

DISSERTATION

Charge Transport Models for Reliability Engineering of Semiconductor Devices

ausgeführt zum Zwecke der Erlangung des akademischen Grades
eines Doktors der technischen Wissenschaften

eingereicht an der Technischen Universität Wien
Fakultät für Elektrotechnik und Informationstechnik

von

MARKUS BINA

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

Wien, im März 2014

Zusammenfassung

Mit der weiter andauernden Skalierung von Halbleiterbauelementen in den Nanometerbereich erhöht sich die Wahrscheinlichkeit eines Ausfalls der Funktionalität aufgrund einzelner Punktdefekte oder Parameterfluktuationen stetig. Daher spielt die Simulation von Bauelementen in der Halbleiterindustrie zur Optimierung von Entwürfen und Prozessen vor der eigentlichen Produktion eine wichtige Rolle in der Vorhersage der Lebensdauer eines Bauelements. In diesem Kontext befasst sich die vorliegende Arbeit mit dem Einfluss der Modellierung des Transports von Ladungsträgern auf die Genauigkeit der Vorhersage von Lebensdauern bei Vorliegen von Bias Temperature Instabilität (BTI) oder Degeneration durch heiße Ladungsträger (HCD).

Zu diesem Zweck wurde ein Vierzustandsmodell basierend auf der Theorie von nicht radiativen mehrphononen (NMP) Übergängen implementiert. Mit Hilfe dieses Modells lässt sich BTI studieren. Weiters muss beachtet werden, dass in skalierten Bauelementen mit Abmessungen von wenigen Nanometern die typischen Dotierkonzentrationen nur mehr von wenigen Dotieratomen hergestellt werden. Dies wiederum führt zu nicht mehr vernachlässigbaren Fluktuationen im elektrostatischen Potential innerhalb des Bauteils. Daher muss die Granularität der Dotierung und die statistische Verteilung von Dotieratomen sowie der Punktdefekte bei der Simulation berücksichtigt werden. Um dies zu bewerkstelligen, wurde ein entsprechendes Modell (random discrete dopant model) in den Drift-Diffusions Simulator Minimos-NT implementiert. In einem ersten Schritt um BTI in p-Kanal MOSFETs bei erhöhter Temperatur besser zu verstehen, wurden Messdaten mit Hilfe der direct-current-current-voltage Methode nach Bias Temperatur Stress aufgenommen und erfolgreich mit dem erwähnten Vierzustandsmodell reproduziert. Dabei wurde selbiges mit dem Modell von Shockley, Read und Hall (SRH) versucht. Es konnte gezeigt werden, dass die Messdaten mit dem SRH Modell nicht vollständig reproduziert werden können. Dieses Ergebnis ist übereinstimmend mit der Literatur, wo gezeigt wurde, dass ein SRH Modell die experimentell gefundenen charakteristischen Zeitkonstanten nach Bias Temperatur Stress nicht reproduzieren kann. In der Literatur wurde unter Anderem auch von einer starken Korrelation zwischen Gate- und Drainstrom nach Bias Temperatur Stress berichtet. Im Zuge dieser Arbeit wurde versucht dies zu erklären. Zuerst wurde versucht diese Korrelation und die beobachtete starke Reduktion des Gatestromes über den Einfluss einer elektrostatischen Interaktion zwischen den statistisch verteilten Dotieratomen und Punktdefekte auf den direkten Tunnelstrom zu erklären. Jedoch konnte die gemessene starke Reduktion des Gatestromes so nicht reproduziert werden. Stattdessen konnte jedoch erfolgreich gezeigt werden, dass ein NMP basiertes Tunnelmodell alle gemessenen Vorgänge sowie die beobachteten Zeitkonstanten erklärt. Dies zeigt, dass ein NMP basiertes Modell ein gute Wahl zur Vorhersage der Bauteillebensdauer ist. Jedoch, muss um eine Vorhersage der Lebensdauer treffen zu können noch die Schwellenspannungsdrift eines einzelnen Punktdefektes erfolgreich reproduziert werden. Dazu wurden MOSFETs mit kleiner Gatefläche statistisch untersucht und die kumulative Verteilung der einzelnen Schwellenspannungsdriften eines Defektes unter Berücksichtigung von diskreten Dotieratomen beobachtet. In diesen Experimenten können ebenso die charakteristischen Zeitkonstanten des Ladungsträgeraustausches, abhängig vom verwendeten Modell, gefunden werden. Es wurde herausgefunden, dass eine Berücksichtigung der Potentialfluktuationen alleine nicht ausreicht um die Verteilung der Schwellenspannungsdriften von einzelnen Defekten zu erklären. Hierfür sind weitere Untersuchungen notwendig.

Für Untersuchungen zur Bauteildegeneration durch heiße Ladungsträger wird eine genaue Kenntnis der Energieverteilung der Ladungsträger benötigt. Dazu muss die Boltzmanntransportgleichung gelöst werden. In dieser Arbeit wurde dazu der Simulator ViennaSHE benutzt. Zum Zwecke des Vergleichs mit Momenten-basierten Simulatoren sowie zur Evaluierung von HCD Modellen wurde dieser Simulator erweitert. Die Erweiterungen bestanden aus einem Quantenkorrekturmodell, Variabilitätsmodellen um Potentialfluktuationen zu berücksichtigen, dem klassischen SRH Modell sowie dem eingangs erwähnten Vierzustandsmodell. Zusätzlich wurde das erweiterte Vecchi Modell zur vollen Bandstrukturabbildung in ViennaSHE implementiert. Mit all diesen Erweiterungen kann ViennaSHE dazu benutzt werden um frische als auch bereits degradierte Bauelemente auf deren Zuverlässigkeit zu untersuchen. Dabei kann sowohl BTI als auch HCD oder später beide zusammen analysiert werden. Um HCD für diverse Bauelemente studieren zu können wurde im Zuge dieser Arbeit ein HCD Modell entwickelt und auch in ViennaSHE implementiert. Dieses Modell wurde erfolgreich mit Hilfe von HCD Messdaten auf n-Kanal MOSFETs verschiedener Kanallängen unter diversen Stressbedingungen mit einem einzigen Satz an Modellparametern validiert. Zusätzlich konnte gezeigt werden, dass Elektron-Elektron Streuung unter Umständen nur in Langkanal MOSFETs vernachlässigt werden kann. Dieses Ergebnis steht damit im Widerspruch zu kürzlich veröffentlichten Arbeiten.

Abstract

With the scaling of semiconductor devices down to the nanometer regime, the probability of functional failure due to single point defects or parameter fluctuations steadily increases. In this respect the simulation of semiconductor devices provides an important ingredient for the optimization of device designs and for the assessment of device lifetimes before production. In this context, this work investigates the influence of the charge carrier transport model on the accuracy of bias temperature instability and hot-carrier degradation models in MOS devices.

For this purpose, a four-state defect model based on a non-radiative multi phonon (NMP) theory is implemented to study the bias temperature instability. However, the doping concentrations typically used in nano-scale devices correspond to only a small number of dopants in the channel, leading to fluctuations of the electrostatic potential. Thus, the granularity of the doping cannot be ignored in these devices. To study the bias temperature instability in the presence of fluctuations of the electrostatic potential, the advanced drift diffusion device simulator Minimos-NT is employed. In a first effort to understand the bias temperature instability in p-channel MOSFETs at elevated temperatures, data from direct-current-current-voltage measurements is successfully reproduced using a four-state defect model. Differences between the four-state defect model and the commonly employed trapping model from Shockley, Read and Hall (SRH) have been investigated showing that the SRH model is incapable of reproducing the measurement data. This is in good agreement with the literature, where it has been extensively shown that a model based on SRH theory cannot reproduce the characteristic time constants found in BTI recovery traces. Upon inspection of recorded recovery traces after bias temperature stress in n-channel MOSFETs it is found that the gate current is strongly correlated with the drain current (recovery trace). Using a random discrete dopant model and non-equilibrium greens functions it is shown that direct tunnelling cannot explain the magnitude of the gate current reduction. Instead it is found that trap-assisted tunnelling, modelled using NMP theory, is the cause of this correlation. This shows that an NMP-based theory of the bias temperature instability can both explain characteristic time constants experimentally found in the drain and the gate current after bias temperature stress as well as the overall threshold voltage shift. These findings imply that for an accurate lifetime prediction an NMP-based theory is a good choice. However, in order to obtain an accurate lifetime prediction information on the threshold voltage shift caused by a single discrete trap created during bias temperature stress needs to be investigated. To this end small area MOSFETs have been investigated on a statistical basis using random discrete doping in order to determine the cumulative distribution function (CFD) of threshold voltage shifts caused by random discrete charged traps as well as their characteristic capture and emission times. It is found that the experimentally observed CFDs of the threshold voltage shifts caused by single charged traps cannot be reproduced using Minimos-NT by considering potential fluctuations alone. Thus further investigations into this subject are needed.

Since the study of hot-carrier degradation requires exact information on the energy distribution of charge carriers, a solution of the Boltzmann transport equation is necessary. For detailed investigations into hot-carrier degradation, ViennaSHE, a device simulator based on a spherical harmonics expansion (SHE) of the Boltzmann transport equation, has been extended in the course of this thesis. To compare SHE to moment-based transport models, quantum correction models, variability caused by random discrete dopants, the classical SRH trapping

theory as well as a four state degradation model based on non-radiative multi-phonon theory are incorporated into the simulator. These additions to ViennaSHE allow to evaluate the device characteristics of virgin as well as degraded devices under hot-carrier or bias temperature stress or both. Additionally, ViennaSHE is extended by the extended Vecchi full-band model in order to accurately model the charge carrier transport in the presence of high electric fields. For the simulation of hot-carrier degradation in MOSFETs, a new hot-carrier model is developed and implemented into ViennaSHE. This hot-carrier model is successfully validated for multiple stress conditions against measurement using a unique set of model parameters. In the discussion of the new model the importance of the various ingredients for hot-carrier modelling are investigated and discussed. Additionally, it is shown that electron-electron scattering is paramount for a successful reproduction of the measurement data for short-channel devices. In this context it is also found that electron-electron scattering may only be neglected in long-channel devices. These results contradict recent findings in the literature, where it was suggested that electron-electron scattering in the context of hot-carrier degradation can be neglected.

Acknowledgement

You live and learn.
At any rate, you live.

(Douglas Adams)

First and foremost I'm indebted to my girlfriend, *Astrid*, for her love, understanding and her endurance in listening to all my rants, I often had throughout the whole time I've been working towards my PhD.

Many thanks go to Prof. *Tibor Grasser* for granting me a position as a PhD candidate. He always provided me with many insights and useful advice. Additionally, he gave me a lot of opportunities to get to know highly ranked scientists and supported my studies as a professor of the TU Wien.

I'm also indebted to *Karl Rupp* for reading through the spherical harmonics centered chapter of this thesis. He also provided great insight into the time-dependent H-transformed spherical harmonics expanded Boltzmann Transport equation, about which I wanted to write a paper. Unfortunately, we could not complete all investigations in time to publish our work elsewhere. I really enjoyed the short time I was allowed to work with him on ViennaSHE. With respect to spherical harmonics expansions many thanks go to Prof. *Christoph Jungemann*. He did not hesitate to provide advice as well as simulation data to verify software I've written whenever necessary. Additionally, he acted as second examiner of my thesis.

To great length I am indebted to *Stanislav Tyaginov*, who heavily supported me in my endeavor into the subject of hot-carrier degradation. When I was pessimistic he was quick to help out with a funny (at least to me and that's the only thing that matters) and to my liking inappropriate joke to break the ice. Working and discussing matters with Stanislav was a true pleasure, although he often would complain about long sentences in my publications, like the last one. Without his help I most likely would not have been able to finish my PhD.

During my thesis I had the pleasure to get to know *Ben Kaczer*, *Maria Toledano* and *Franco Jacopo* from imec, Leuven. I would like to thank them for all their help, hour long fruitful discussions and measurement data. I am especially indebted to Ben Kaczer for inviting me to be a reviewer of scientific papers for journals of the IEEE.

Many thanks go to *Wolfgang Gös*, who, with great endurance, read most of this thesis in order to raise its quality to an acceptable level. He also taught me everything I know about non-radiative multiphonon transitions. Thanks also go to *Franz Schanovsky* who helped me to improve the quality of this thesis.

To my grandfathers, Herbert and Karl

Contents

Der Andersdenkende ist kein Idiot, er hat sich eben eine andere Wirklichkeit konstruiert.

(Paul Watzlawick)

1	Introduction	1
2	Charge Carrier Transport	3
2.1	Band structure	4
2.2	The Boltzmann Transport Equation	7
2.2.1	Scattering	10
2.2.2	Recombination and Generation	11
2.3	The Method of Moments	13
2.3.1	The Drift Diffusion Model	14
2.3.2	The Hydrodynamic Model	15
2.3.3	Higher Order Models	16
2.4	The Monte Carlo Method	16
2.5	The Spherical Harmonics Expansion Method	18
2.6	Quantum Mechanical Effects - Confinement	18
2.7	Non-Equilibrium Greens Functions Approach	20
3	Spherical Harmonics Expansion	21
3.1	Theory	21
3.1.1	Expansion of the Free-Streaming Operator	22
3.1.2	Expansion of the Scattering-Streaming Operator	23
3.2	The H-Transform	24
3.3	Discretization	25
3.4	Full-band Effects	26
3.4.1	The Anisotropic Band Model	27
3.4.2	The extended Vecchi Model	28
3.5	Recombination and Generation	28
3.6	Time-dependent SHE of the BTE	29
3.6.1	Comparison to Drift Diffusion	30
3.6.2	Stability	33
3.6.3	Energy Grid Interpolation	35
3.6.4	Probable Violation of Gauss' Law	37
3.6.5	The Shockley-Haynes Experiment	38
4	Quantum Correction Methods	41
4.1	The Density Gradient Model	41
4.2	The First-Order Quantum Corrected Drift Diffusion Model	42
4.3	The First-Order Quantum Corrected SHE of the BTE	42
4.4	Discretization	43
4.4.1	The Simple Scheme	43
4.4.2	The Full Scheme	44
4.5	Boundary Conditions	44
4.6	Calibration	46
4.6.1	Calibration for the Drift Diffusion Model	46

4.6.2	Calibration for a SHE of the BTE	46
5	Variability	48
5.1	Random Discrete Dopands	48
5.1.1	Random Discrete Dopands Algorithm	49
5.1.2	Screening Charges	50
5.1.3	Simulation Results using the Drift Diffusion Model	52
5.2	Random Discrete Traps	53
5.2.1	Single Trap	54
5.2.2	Multiple Traps	55
5.2.3	Mobility	56
5.3	Random Discrete Doping and a SHE of the BTE	59
6	Bias Temperature Instability	61
6.1	Measurement Techniques	62
6.1.1	Measure Stress Measure Technique	62
6.1.2	On-the-Fly Technique	63
6.1.3	Direct Current Current Voltage	64
6.1.4	Time Dependent Defect Spectroscopy	65
6.2	Models for the Bias Temperature Instability	66
6.2.1	Phenomenological Models	68
6.2.2	Non-radiative Multiphonon Transitions	69
6.2.3	Structural Relaxation	73
6.2.4	The four State NMP Model	73
6.3	Implementation and Requirements	77
6.3.1	Self-Consistent Solutions	78
6.3.2	Parameter Dispersion	79
6.3.3	Suitable Transport Models	80
6.4	Model Evaluation on pMOSFETs using the Direct Current Current Voltage Method	81
6.4.1	Experimental Setup	81
6.4.2	Comparison of SRH and the four State NMP Model	82
6.5	Results on Trap-Assisted Tunneling	87
7	Hot Carrier Degradation	92
7.1	Measurement Technique	92
7.2	Electric Field Dependence	94
7.3	Temperature Dependence	94
7.4	Channel Length Dependence	96
7.5	Models	96
7.5.1	Early Models	97
7.5.2	Latest Model	97
7.6	Requirements for Simulation	100
7.6.1	Importance of the High Energy Tail	100
7.7	Calibration	101
7.7.1	Experiment	101
7.7.2	Results	101
7.7.3	Importance of Electron-Electron Scattering	102
8	Conclusions	107

Abandon all hope, you who enter here.

(Dante Alighieri)

1 Introduction

The ongoing trend in the semiconductor industry towards ever smaller devices has led to ever faster digital circuits at decreasing cost. However, scaling devices down to the nano-meter regime increases the probability of functional failure due to single point defects or parametric fluctuations, such as fluctuations in the doping. Single point defects are studied in reliability, whereas parametric fluctuations are often referred to as variability. The widespread and readily available drift diffusion (DD) model and its first-order quantum corrected versions have often been used to assess both reliability [1] and variability in nano-scale devices, often despite better knowledge or lack of alternatives [2]. The challenge in selecting the appropriate transport model lies in the exact definition of the problem to be assessed and the quantities of interest. In reliability as well as in variability, one is interested in changes in device parameters, such as the absolute threshold voltage shift in MOSFETs, due to charges in the semiconductor and the insulator of the device. Variability, however often needs to be accounted for when investigating reliability. This is the case in sub-100 nm node devices, where fluctuations of the doping cannot be ignored anymore [3, 4]. For example, in nano-scale FinFets the discrete nature of the doping needs to be accounted for since on average there are only very few dopant atoms in the channel of the device. Additionally, in order to avoid unphysically overestimation of carrier concentrations in nano-scale devices, first-order quantum correction models [5-7] or solutions of the Schrödinger-Poisson equation are required. Identifying and capturing essential quantum mechanical effects in a charge transport model, which needs to be computationally efficient and accurate at the same time is challenging.

Considering device degradation, two effects have received wider attention of reliability researchers and engineers alike: the bias temperature instability (BTI) and hot-carrier degradation (HCD). BTI is mainly attributed to oxide defects sensitive to temperature and the electric field. HCD, on the other hand, is attributed to highly energetic carriers, which impinge on the oxide-semiconductor interface in MOSFETs, thereby generating interface defects. Most studies investigating reliability on nano-scale devices are either strongly based on measurement data or on oversimplified device or degradation models. The selection of an appropriate transport model to be coupled with a reliability model is highly dependent on the device design and the degradation effect of interest. For example, when investigating BTI for a nano-scale FinFet [8], carrier confinement and variability of the doping are crucial, whereas, e.g. on a planar 500 nm n-channel MOSFET a classic DD simulation is often sufficient. However, if one is interested in understanding HCD, the drift diffusion transport model has proven to be insufficient in any device, since HCD requires exact modelling of the energy distribution of the charge carriers.

Due to the delicate relation of transport-, reliability-models and variability Chapter 2 gives an overview of charge transport models and solution methods based on the Schrödinger equation and Boltzmann's Transport equation (BTE). In this chapter various transport models, the basic assumptions during their derivation as well as solution methods is be discussed. Since modern

hot-carrier degradation models require a solution of the BTE, the spherical harmonics expansion (SHE) method is used to solve the BTE in a time efficient manner. Chapter 3 describes the spherical harmonics expansion of the Boltzmann transport equation, the derivation of the respective equations for the expansion coefficients and its numeric implementation. In Chapter 4 first-order quantum correction methods, especially the density gradient (DG) model [9], are briefly discussed. The density gradient model is required in order to efficiently assess variability due to random discrete dopants. This is the focus of Chapter 5, which is mainly concerned with the ΔV_{th} variability of single discrete oxide traps in the presence of random discrete dopants. Additionally, it will be shown that a first-order quantum corrected drift diffusion transport model is insufficient to capture the distribution of ΔV_{th} steps caused by single point defects in planar nano-scale MOSFETs. Chapter 6 describes the typical setups for simulation and measurement of the bias temperature instability. Additionally, the latest model for BTI and the appropriate selection of a transport model for BTI are discussed. The chapter ends with an investigation of BTI in p-channel MOSFETs, using an integrated poly-heater technology, at high temperatures. The differences of the current BTI model to the standard SRH model as well as the general deficits of the SRH model will be shown. The last chapter (Chapter 7) is exclusively about HCD. Important characteristics of HCD are discussed, followed by an exhaustive explanation of our current HCD model. Additionally, suitable transport models for the simulation of HCD in MOSFETs are discussed. It is shown that SHE is at present the better BTE solution technique, compared to Monte Carlo approaches, in the context of HCD. Finally the application of the current HCD model using distribution functions obtained by SHE for two different devices and four different stress conditions is be discussed.

Die Möglichkeit einer mechanischen Erklärung der ganzen Natur ist nicht bewiesen, ja, dass wir dieses Ziel vollkommen erreichen werden, kaum denkbar.

(Ludwig Boltzmann)

2 Charge Carrier Transport

Modelling the transport of charge carriers in a crystal, eg. silicon, containing artificially introduced impurities, ie. dopants, to modify the electrical conductance of the crystal is not an easy task. In order to fully describe the many-particle system, where \mathbf{r} denotes the positions of the electrons and \mathbf{R} denotes the positions of the ions, one needs to employ

$$\left(-\frac{\hbar^2}{2m_0} \nabla_{\mathbf{r}}^2 - \frac{\hbar^2}{2M_0} \nabla_{\mathbf{R}}^2 + H_{\text{ion-ion}}(\mathbf{R}) + H_{\text{el-ion}}(\mathbf{r}, \mathbf{R}) + H_{\text{el-el}}(\mathbf{r}) + \dots \right) \Psi(\mathbf{r}, \mathbf{R}) = i\hbar \partial_t \Psi(\mathbf{r}, \mathbf{R}), \quad (2.1)$$

which is the time-dependent Schrödinger equation of the full many-particle system. In the above equation various interactions between the ions and electrons, such as the interactions between ions is given by the operator $H_{\text{ion-ion}}(\mathbf{r}, \mathbf{R})$, between electrons and ions is described by $H_{\text{el-ion}}(\mathbf{r}, \mathbf{R})$ and electron-electron scattering is described by $H_{\text{el-el}}(\mathbf{r}, \mathbf{R})$. At present it is practically impossible to solve the many-particle Schrödinger equation even for only a few atoms. Thus the usual approach to this is to decouple the electron-system and the system of lattice atoms by the separation Ansatz

$$\Psi(\mathbf{r}, \mathbf{R}) = \psi(\mathbf{r})\psi_{\text{ion}}(\mathbf{R}),$$

where the wave function ψ only describes the system of electrons and ψ_{ion} only describes the lattice. Inserting this into Equation (2.1) one obtains

$$\left(-\sum_i \frac{\hbar^2}{2m_0} \nabla_{\mathbf{r}}^2 + \sum_i V_{\text{ion}}(\mathbf{x}_i) + H_{\text{el-ion}}(\mathbf{r}, \mathbf{R}) + H_{\text{el-el}}(\mathbf{r}) \right) \psi(\mathbf{r}, \mathbf{R}) = E_{\text{el}}\psi(\mathbf{r}, \mathbf{R}),$$

$$\left(-\sum_i \frac{\hbar^2}{2M_i} \nabla_{\mathbf{R}}^2 + \sum_i V_{\text{el}}(\mathbf{x}_i) + E_{\text{el}} \right) \psi_{\text{ion}}(\mathbf{R}) = E_{\text{ion}}\psi_{\text{ion}}(\mathbf{R}),$$

where $V_{\text{ion}}(\mathbf{x}_i)$ and $V_{\text{el}}(\mathbf{x}_i)$ are the ionic or electronic background potentials seen by the electrons or ions. To reduce this set of equations further, the ions are assumed to only slowly oscillate around their mean lattice position and the Schrödinger equation for the ions is not solved (jellium model). The main argument of this approach is based on the assumption that the mass of an ion M is much larger than the mass of an electron m , which means that the ions are much slower in their movement than the electrons. The approximation itself is often referred to as the 'adiabatic approximation' or Born-Oppenheimer-Approximation [10]. A thorough justification and investigation of the Born-Oppenheimer-Approximation is beyond the scope of this work and can be found in [11]. In order to reduce the many-particle Schrödinger equation for the electrons to a more tractable single particle equation the Hartree-Fock-Approximation and the

Slater-Approximation are applied [12]. The Schrödinger equation under these approximations for a single electron state in a homogeneous semiconductor reads

$$\left(-\frac{\hbar^2}{2m_0}\nabla^2 + V(\mathbf{x}) + \varphi_{\text{out}}(\mathbf{x}, t) + H_{\text{scatter}}(\mathbf{x}, t) \right) \psi(\mathbf{x}, t) = i\hbar\partial_t\psi(\mathbf{x}, t) \quad (2.2)$$

where $\psi(\mathbf{x}, t)$ is now a single electron wave function representing the whole ensemble of electrons, $\varphi_{\text{out}}(\mathbf{x}, t)$ is the external electrostatic potential and all other particle-particle interactions have been summarized in H_{scatter} . In order to obtain a semi-classical approximation in the special case of a periodic lattice Equation (2.2) has to be transformed such that the potential of the ions does not explicitly appear in the equation. This is done by solving Equation (2.2) without any scattering H_{scatter} and time dependence [12, 13],

$$\left(-\frac{\hbar^2}{2m_0}\nabla^2 + V(\mathbf{x}) \right) \psi(\mathbf{x}, \mathbf{k}) = E(\mathbf{k})\psi(\mathbf{x}, \mathbf{k}), \quad (2.3)$$

and inserting the solution back into Equation (2.2), where the reciprocal wave vector \mathbf{k} and the ionic potential $V(\mathbf{x})$ (jellium-model for semiconductors) have been introduced [12, 14]. A solution $\psi(\mathbf{x}, \mathbf{k})$ to Equation (2.3) is easily obtained by using plane-waves and Bloch-functions [15] to account for the periodicity in the lattice. Additionally the solution to Equation (2.3) gives the relation $E_\nu(\mathbf{k})$ for the ν th band. $E_\nu(\mathbf{k})$ is often referred to as the dispersion relation and is only dependent on the type of semiconductor. By introducing the operator $E_\nu(-i\nabla_{\mathbf{x}})$, where

$$E_\nu(-i\nabla_{\mathbf{x}})\psi(\mathbf{x}, \mathbf{k}) = E_\nu(\mathbf{k})\psi(\mathbf{x}, \mathbf{k})$$

according to [12], one can transform Equation (2.2) into

$$\left(\underbrace{E_\nu(-i\nabla_{\mathbf{x}})}_{-\frac{\hbar^2}{2m_0}\nabla^2 + V(\mathbf{x})} + \varphi_{\text{out}}(\mathbf{x}, t) + H_{\text{scatter}}(\mathbf{x}, \mathbf{k}) \right) \psi(\mathbf{x}, \mathbf{k}, t) = i\hbar\partial_t\psi(\mathbf{x}, \mathbf{k}, t). \quad (2.4)$$

The Hamilton operators used in Equation (2.4) form the basis of all transport models that are described in the course of this thesis. Often Equation (2.4) is arbitrarily separated into two equations, one for all bands (valence bands) below and one for all conduction bands above the center of the band gap. Next a new quasi-particle, the hole, is introduced in the equation for all valence bands. A hole represents an unoccupied state and thus has a positive elementary charge associated with it. This is done, since the valence bands contain many electrons occupying the bands. Thus it is easier to describe the flow of the empty states instead. For a proper definition of holes refer to [12].

2.1 Band structure

In Equation (2.4) the kinetic energy operator and the periodic lattice potential have been replaced by the dispersion relations $E_\nu(\mathbf{k})$ for ν bands. This enables us to treat an electron in a lattice as a free quasi-electron, where the first derivative of $E_\nu(\mathbf{k})$ is related to the momentum \mathbf{p}_ν and therefore the group velocity (cf. Equation (2.15)) and the second derivative of $E_\nu(\mathbf{k})$ is related to the effective mass tensor \mathbf{m}_ν^*

$$\mathbf{p}_\nu = m_0 v_g = \frac{m_0}{\hbar} \nabla_{\mathbf{k}} E_\nu(\mathbf{k}), \quad (2.5)$$

$$(\mathbf{m}_\nu^*)^{-1} = \frac{\partial^2 E_\nu(\mathbf{k})}{\hbar^2 \partial k^2}. \quad (2.6)$$

In order to efficiently compute the electron and hole densities from Equation (2.16) the density of states (DOS) $g_\nu(E)$ is needed. The DOS can be calculated from the dispersion relation $E_\nu(\mathbf{k})$ using [12]

$$g_\nu(E) = \frac{2}{(2\pi)^3} \int_{\mathcal{B}} \delta(E - E_\nu(\mathbf{k})) d^3k, \quad (2.7)$$

where \mathcal{B} denotes the first Brillouin Zone (BZ) [14, 16, 17] (cf. Figure 2.1) in the band structure, shown for silicon in Figure 2.2, of the semiconductor. Most often in the analysis of semiconductors, the bands are separated in the middle of the band gap into conduction and valence bands for electrons and holes respectively. Thus also the DOS has to be separated into a DOS for electrons $g_\nu^n(E)$ and holes $g_\nu^p(E)$. The so obtained DOS for silicon is shown in Figure 2.1.

Upon combining the definitions of the distribution function (Equation (2.16)) and the DOS (Equation (2.7)) one arrives at

$$n = \int f_n^\nu g_\nu^n(E) dE, \quad p = \int f_p^\nu g_\nu^p(E) dE, \quad (2.8)$$

to calculate the macroscopic electron n and hole p concentrations. As will become clearer later, n and p are moments of their respective BTE.

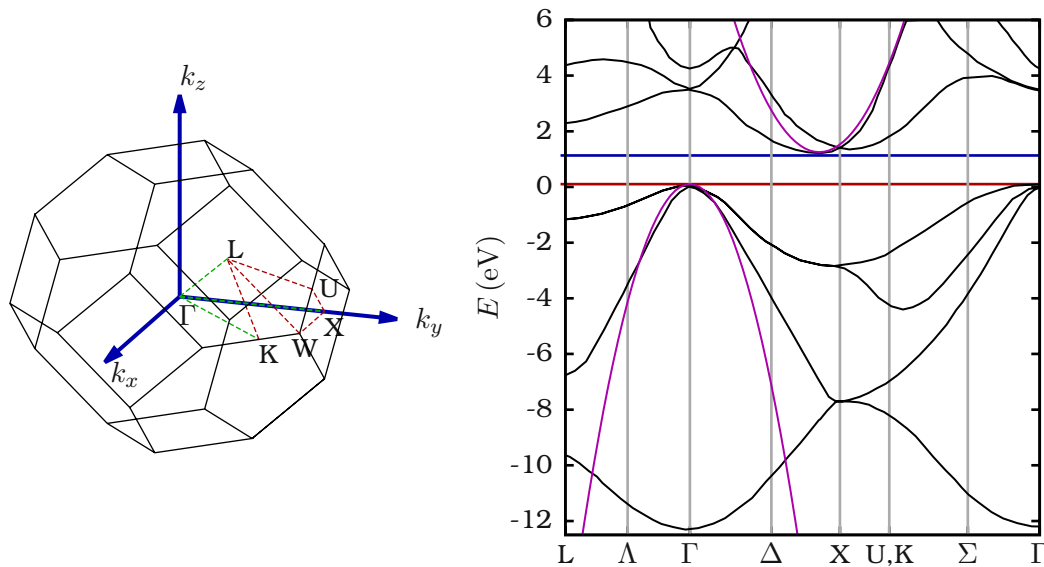


Figure 2.1: **Left:** The first Brillouin zone (BZ) of silicon. The lines and points of symmetry are shown in red (on the edge of the BZ) and green (within the BZ). **Right:** The band structure of silicon from [18] calculated using a non-local pseudopotential method, which neglects spin-orbit interaction [19]. The conduction band edge is indicated by the blue line, whereas the red line indicates the valence band edge. The bandstructure as obtained by a parabolic approximation is also shown (purple).

Parabolic Band Approximation

From Figure 2.1 one can easily deduce that, in the important case of silicon, there is no simple analytic expression for the bandstructure. This poses a fundamental problem for many

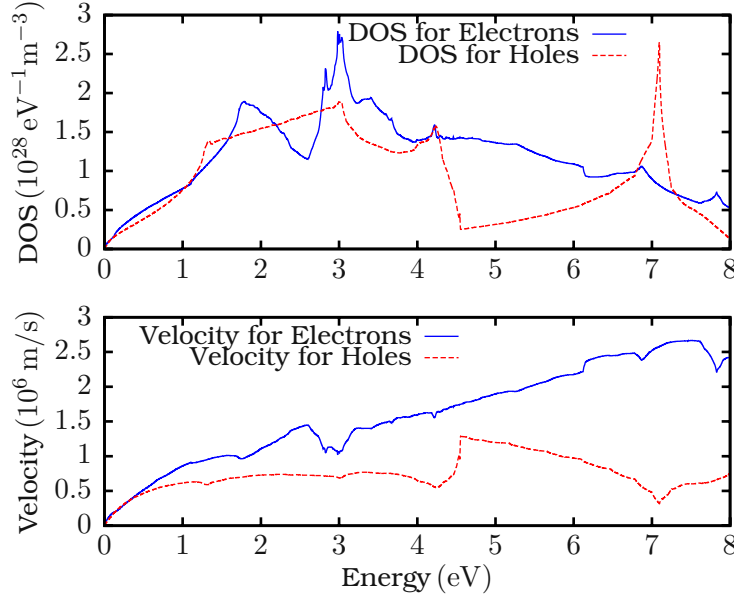


Figure 2.2: The density of states (DOS) and group velocity for relaxed silicon used for the solution of the bipolar BTE.

approximations and solution techniques of the BTE, Monte Carlo methods being a notable exception. Thus analytic approximations for the bandstructure are sought. One of them is the *parabolic band approximation*. The main argument justifying this approximation is that for low and moderate fields the electrons still occupy states close to the band minimum. The parabolic band approximation is obtained by expanding the dispersion relation $E_\nu(\mathbf{k})$ into a Taylor series. Upon evaluating the series for \mathbf{k} at the lowest energy E one obtains

$$E_\nu(\mathbf{k}) \approx \underbrace{E_\nu(\mathbf{k}_{\min(E)})}_{=E_{C/V}} + \frac{1}{2} \frac{\partial^2 E_\nu(\mathbf{k})}{\partial k^2} \Big|_{\min(E)} \|\mathbf{k}\|^2, \quad (2.9)$$

where the first term on the right hand side is a constant energy to account for the distance between the conduction band edge and the choice of the reference energy $E_{\text{ref}} = 0$ (cf. Figure 2.3). If the minimum energy is to be found at $\mathbf{k} = 0$ then the above equation reduces to

$$E_\nu(\mathbf{k}) \approx E_\nu(0) + \frac{\hbar^2 \|\mathbf{k}\|^2}{2m^*}, \quad (2.10)$$

where the effective mass tensor (cf. Equation (2.6)) is assumed to be isotropic and scalar. However, in silicon this is not the case (cf. Figure 2.1) and the effective mass m^* is not isotropic as assumed in the equations above. For the conduction band in silicon, Equation (2.10) is often reduced to

$$E(\mathbf{k}) \approx E_C + \frac{\hbar^2}{2} \left(\frac{k_x^2}{m_y} + \frac{k_y^2}{m_y} + \frac{k_z^2}{m_z} \right), \quad (2.11)$$

where m_x , m_y and m_z are the components of the effective mass tensor \mathbf{m}^* and we restricted ourselves to a single effective band. Often the simple relation Equation (2.10) is favored over Equation (2.11), since a scalar k can be used instead of a vectorial \mathbf{k} thereby simplifying the analysis. This can be achieved by the Herring-Vogt transform [20] without any further assumptions through a linear coordinate transform in \mathbf{k} -space.

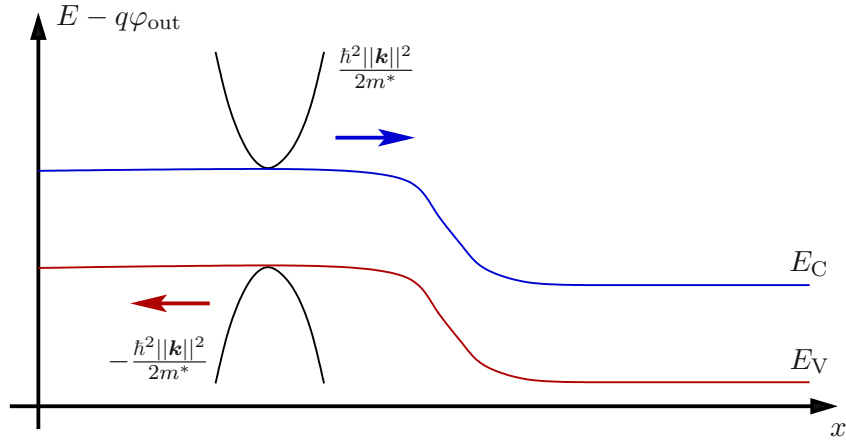


Figure 2.3: Sketch of the parabolic band approximation in a 1D device.

Kane's Model

The parabolic band approximation is often used up to high energies at high fields where it obviously fails to predict the correct band structure (cf. Figure 2.1). In order to improve this the parabolic band approximation has been extended [21] to account for the non-parabolicity in the band structure by introducing the factor α for the first (conduction or valence) band as follows:

$$E(k)[1 + \alpha E(k)] = \frac{\hbar^2 k^2}{2m^*}, \quad (2.12)$$

where $\alpha = 0.5$ is a good choice for electrons in silicon.

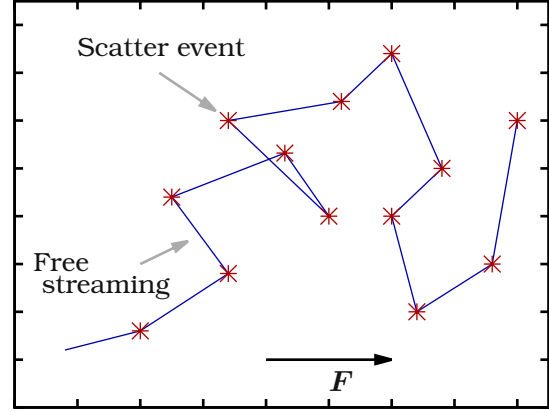
2.2 The Boltzmann Transport Equation

The Boltzmann Transport equation (BTE) has been originally developed by Ludwig Boltzmann to statistically describe transport of atoms and molecules (particles) of an idealized diluted gas. The BTE describes the transport by two processes, namely free streaming and scattering. In order to use the BTE to describe transport of an electron gas in a semiconductor, the BTE for diluted gases needs to be modified. First the free streaming of the charge carriers in the lattice is described by the equations of motion (Newton's law)

$$\hbar \partial_t \mathbf{k} = \mathbf{F} \quad \text{and} \quad \partial_t \mathbf{x} = \mathbf{v}, \quad (2.13)$$

where the relation $\mathbf{p} = \hbar \mathbf{k}$ and the band structure is used. Second, to describe the collisions between electrons in the electron gas, we will use quantum mechanical perturbation theory (Fermi's Golden Rule). It has to be noted that in the classical framework of the BTE, where Heisenberg's uncertainty principle is being neglected, one could directly track position and momentum of each electron. However, this tracking of each single particle in a classical approach is currently infeasible. Instead we will modify the BTE as stated above, by rederiving the BTE using Equation (2.4). Although the derivation of the BTE has been covered in many works before, its derivation will nevertheless be repeated from [12, 13] for the sake of clarity. For this purpose we need to define a distribution function $f(\mathbf{x}, \mathbf{k}, t)$ of the carriers in real- and

Figure 2.4: Illustration of the movement of an electron through a 2D semiconductor. The BTE treats the free streaming (blue lines) classically and the scattering events (red crosses) quantum-mechanically.



reciprocal-space. The distribution function $f(\mathbf{x}, \mathbf{k}, t)$ is defined such that

$$dN = \frac{2}{(2\pi)^3} f(\mathbf{x}, \mathbf{k}, t) d^3x d^3k \quad (2.14)$$

is the number of particles (electrons or holes) in the infinitesimal small volume $d^3x d^3k$ in the six-dimensional phase-space. The distribution function is a semi-classical concept, since it assumes that both position and momentum of a particle can be measured to arbitrary accuracy at the same time. In order to introduce the semi-classical concept into the Hamiltonians of Equation (2.4), we treat the electrons as classical particles governed by classical mechanics in the time intervals they do not scatter with other particles. Thus the external force exerted on a single electron is

$$\mathbf{F}(\mathbf{x}, t) = \partial_t \underbrace{\hbar \mathbf{k}}_{=\mathbf{p}} = -|q| \nabla_{\mathbf{x}} \varphi_{\text{out}}(\mathbf{x}, t).$$

The group velocity $v_g^\nu(\mathbf{k})$ can then be calculated [12] as

$$v_g^\nu(\mathbf{k}) = \frac{\nabla_{\mathbf{k}} E_\nu(\mathbf{k})}{\hbar}.$$

Starting from the definition of the density matrix

$$\rho(\mathbf{x}_1, \mathbf{x}_2, t) = \psi(\mathbf{x}_1, t) \cdot \psi(\mathbf{x}_2, t) \quad (2.15)$$

we define the distribution function as the Fourier transformed density matrix

$$f(\mathbf{x}, \mathbf{k}, t) = \int \exp(-i\mathbf{k} \cdot \mathbf{x}) \rho(\mathbf{x}_1, \mathbf{x}_2, t) d^3k. \quad (2.16)$$

This definition is compatible with the previous definition in Equation (2.14), since the density matrix describes a statistical ensemble of quantum states in real space. In order to transform the density matrix to the phase space, a Fourier transform needs to be carried out. To assess the time evolution of the density matrix we utilize the Liouville-Von-Neumann equation,

$$(H(\mathbf{x}_1) - H(\mathbf{x}_2))\rho(\mathbf{x}_1, \mathbf{x}_2, t) = \partial_t \rho(\mathbf{x}_1, \mathbf{x}_2, t), \quad (2.17)$$

$$\Rightarrow i\hbar \partial_t \rho(\mathbf{x}_1, \mathbf{x}_2, t) + (E_\nu(-i\nabla_{\mathbf{x}_1}) - E_\nu(-i\nabla_{\mathbf{x}_2})) \quad (2.18)$$

$$- \varphi_{\text{out}}(\mathbf{x}_1, t) + \varphi_{\text{out}}(\mathbf{x}_2, t)) \rho(\mathbf{x}_1, \mathbf{x}_2, t) = 0, \quad (2.19)$$

where the simplified and approximated Hamilton-Operators from Equation (2.4) have been used. Scattering will be neglected for the time being, since it will be treated as a perturbation later. After a transformation of coordinates,

$$\mathbf{x} = \frac{1}{2}(\mathbf{x}_1 + \mathbf{x}_2), \quad \mathbf{x}' = \mathbf{x}_1 - \mathbf{x}_2, \quad (2.20)$$

the operators can be linearized by assuming that each electron is sharply distributed in \mathbf{k} -space and that the electrostatic potential changes over \mathbf{x} are small compared to the changes of the electronic wavefunction over \mathbf{x}

$$-\varphi_{\text{out}}(\mathbf{x}_1, t) + \varphi_{\text{out}}(\mathbf{x}_2, t) \approx \nabla_{\mathbf{x}}\varphi_{\text{out}}((\mathbf{x}_2 + \mathbf{x}_1)/2, t) \cdot (\mathbf{x}_2 + \mathbf{x}_1) = \underbrace{-\nabla_{\mathbf{x}}\varphi_{\text{out}}(\mathbf{x}, t)}_{=\mathbf{F}(\mathbf{x})/q} \cdot \mathbf{x}' \quad (2.21)$$

$$E_{\nu}(-i\nabla_{\mathbf{x}_1}) - E_{\nu}(-i\nabla_{\mathbf{x}_2}) \Rightarrow i\hbar\mathbf{v}_{\mathbf{g}}^{\nu}(-i\nabla_{\mathbf{x}'}) \cdot \nabla_{\mathbf{x}} \quad (2.22)$$

Inserting the above approximations into Equation (2.17) one obtains:

$$i\hbar\partial_t\rho(\mathbf{x}, \mathbf{x}', t) + i\hbar\mathbf{v}_{\mathbf{g}}^{\nu}(-i\nabla_{\mathbf{x}'})\nabla_{\mathbf{x}}\rho(\mathbf{x}, \mathbf{x}', t) + \hbar^{-1}\mathbf{F}(\mathbf{x}) \cdot \mathbf{x}' \cdot \rho(\mathbf{x}, \mathbf{x}', t) = 0. \quad (2.23)$$

Finally, a Fourier transformation of the equation above is carried out and one obtains after a few algebraic rearrangements the well known homogeneous BTE for the ν th band

$$\partial_t f^{\nu}(\mathbf{x}, \mathbf{k}, t) + \underbrace{\mathbf{v}_{\mathbf{g}}^{\nu}(\mathbf{x}, \mathbf{k}) \cdot \nabla_{\mathbf{x}} f^{\nu}(\mathbf{x}, \mathbf{k}, t) - \hbar^{-1}\mathbf{F}(\mathbf{x}, t) \cdot \nabla_{\mathbf{k}} f^{\nu}(\mathbf{x}, \mathbf{k}, t)}_{=\mathcal{L}\{f^{\nu}(\mathbf{x}, \mathbf{k}, t)\}} = 0, \quad (2.24)$$

where $\mathcal{L}\{f^{\nu}(\mathbf{x}, \mathbf{k}, t)\}$ is the free-streaming operator. In order to incorporate scattering processes a semi-classical perturbation term $\mathcal{Q}\{f^{\nu}(\mathbf{x}, \mathbf{k}, t)\}$ on the right hand side of the BTE is required

$$\partial_t f^{\nu}(\mathbf{x}, \mathbf{k}, t) + \underbrace{\mathbf{v}_{\mathbf{g}}^{\nu}(\mathbf{x}, \mathbf{k}) \cdot \nabla_{\mathbf{x}} f^{\nu}(\mathbf{x}, \mathbf{k}, t) - \hbar^{-1}\mathbf{F}(\mathbf{x}, t) \cdot \nabla_{\mathbf{k}} f^{\nu}(\mathbf{x}, \mathbf{k}, t)}_{=\mathcal{L}\{f^{\nu}(\mathbf{x}, \mathbf{k}, t)\}} = \mathcal{Q}\{f^{\nu}(\mathbf{x}, \mathbf{k}, t)\}. \quad (2.25)$$

$\mathcal{Q}\{f^{\nu}(\mathbf{x}, \mathbf{k}, t)\}$ is obtained by perturbation theory and modeled using a statistical description of each scattering process. In the BTE the operator $\mathcal{L}\{f^{\nu}(\mathbf{x}, \mathbf{k}, t)\}$ describes the classical free-streaming of the carriers in between scattering events which are described by the scattering operator $\mathcal{Q}\{f^{\nu}(\mathbf{x}, \mathbf{k}, t)\}$ (cf. Figure 2.4).

The force $\mathbf{F}(\mathbf{x}, t)$ due to a gradient of the electrostatic potential or the band edges in the absence of magnetic fields is given by

$$\mathbf{F}(\mathbf{x}, t) = -q\nabla_{\mathbf{x}}\varphi_{\text{out}}(\mathbf{x}, t) = -\nabla_{\mathbf{x}}(\pm E_{C/V} \mp |q|\varphi(\mathbf{x}, t)(\mathbf{x})), \quad (2.26)$$

where $|q|$ is the sign-less elementary charge q , $\varphi(\mathbf{x}, t)$ the electrostatic potential and $E_{C/V}$ the band edge energy for either electrons or holes. Since the electrostatic potential is needed in order to compute the force exerted on each charge carrier, one needs to solve Poisson's equation and the BTE for electrons and holes self-consistently. The full system of equations thus reads,

$$\begin{aligned} \nabla \cdot (\varepsilon(\mathbf{x})\nabla\varphi(\mathbf{x}, t)) &= |q|(n - p + C), \\ \partial_t f_n^{\nu} + \underbrace{\mathbf{v}^n \cdot \nabla_{\mathbf{x}} f_n^{\nu} + \hbar^{-1}\mathbf{F} \cdot \nabla_{\mathbf{k}} f_n^{\nu}}_{=\mathcal{L}\{f_n^{\nu}\}} &= \mathcal{Q}^n\{f_n^{\nu}\} - \Gamma^n\{f_n^{\nu}, f_p^{\nu}\}, \\ \partial_t f_p^{\nu} + \underbrace{\mathbf{v}^p \cdot \nabla_{\mathbf{x}} f_p^{\nu} - \hbar^{-1}\mathbf{F} \cdot \nabla_{\mathbf{k}} f_n^{\nu}}_{=\mathcal{L}\{f_p^{\nu}\}} &= \mathcal{Q}^p\{f_n^{\nu}\} - \Gamma^p\{f_n^{\nu}, f_p^{\nu}\}, \end{aligned} \quad (2.27)$$

where additional scattering terms $\Gamma^{n/p}\{f_n^\nu, f_p^\nu\}$ have been added to account for recombination and generation of electrons and holes. Since each BTE, without further approximations, has to be solved in three spatial and three phase space dimensions in addition to the time dimension, a direct discretization of the full system would lead to prohibitive memory and computational time requirements for many applications. Thus further approximations or alternative discretization schemes have to be employed to solve the system of equations.

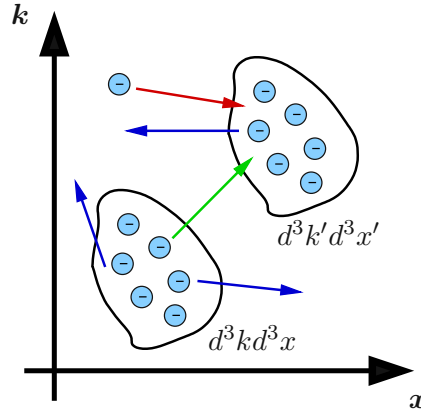
2.2.1 Scattering

In this section a short introduction into modelling the scattering term $Q\{f^\nu(\mathbf{x}, \mathbf{k}, t)\}$ on the right hand side of Equation (2.25) is given. While traveling through the lattice the electrons interact with the lattice or other electrons. Such interactions, where the electrons change their momentum, are termed scattering events. Considering an infinitesimal small volume $d^3k d^3x$ in phase space (cf. Figure 2.5), the electrons can either scatter into or out of this volume. Thus the scattering operator is normally split into an in-scattering and out-scattering operator as follows:

$$Q\{f^\nu(\mathbf{x}, \mathbf{k}, t)\} = Q^{\text{in}}\{f^\nu(\mathbf{x}, \mathbf{k}, t)\} - Q^{\text{out}}\{f^\nu(\mathbf{x}, \mathbf{k}, t)\}.$$

The transitions in and out of an infinitesimal small volume $d^3k d^3x$ in phase space do not occur

Figure 2.5: Illustration of scattering events in phase space. Shown is a group of electrons moving from an infinitesimal volume $d^3k d^3x$ at time t in phase space to $d^3k' d^3x'$ at time $t + d\tau$, where the out-scattering of electrons is depicted using blue arrows, in-scattering is illustrated by red arrows and free streaming is indicated with a green arrow.



instantly. Nevertheless they are often modeled as an instant transition via a rate of expected scattering charge carriers per second. Additionally it is assumed that the scattering occurs at a certain point in real space and is a local event. For each flavor of scattering there is a rate of transitions $S^\nu(\mathbf{k}, \mathbf{k}', t)$ from \mathbf{k} to \mathbf{k}' in the phase space. An incomplete list of interactions considered in this work can be found in Table 2.1. Note that in a homogeneous semiconductor the scattering rate does not depend on the spatial location of the electron. This simplifies the evaluation of the scattering terms. In order for an electron to scatter, the initial state must be occupied by an electron and the final state must be empty by virtue of Pauli's exclusion principle. With all of the above the scattering operator is often modelled as

$$Q\{f^\nu(\mathbf{x}, \mathbf{k}, t)\} = \frac{1}{(2\pi)^3} \sum_\nu \int_B \underbrace{S^\nu(\mathbf{k}', \mathbf{k}, t) f^\nu(\mathbf{x}, \mathbf{k}', t) (1 - f^\nu(\mathbf{x}, \mathbf{k}, t))}_{Q^{\text{in}}\{f^\nu(\mathbf{x}, \mathbf{k}, t)\}} - \underbrace{S^\nu(\mathbf{k}, \mathbf{k}', t) f^\nu(\mathbf{x}, \mathbf{k}, t) (1 - f^\nu(\mathbf{x}, \mathbf{k}', t))}_{Q^{\text{out}}\{f^\nu(\mathbf{x}, \mathbf{k}, t)\}} d^3k', \quad (2.28)$$

where the integral runs over the first Brillouin zone. In thermal equilibrium Equation (2.28) as scattering operator in the BTE yields the *Fermi-Dirac distribution*

$$f^\nu(\mathbf{x}, \mathbf{k}, t) = \frac{1}{1 + \exp\left(\frac{E_\nu(\mathbf{k}) - E_F}{k_B T_L}\right)}, \quad (2.29)$$

where E_F is the Fermi-level and equals the chemical potential μ known from thermodynamics [12]. The computational burden to evaluate the integral in Equation (2.28) can strongly depend on how the bandstructure is resolved and becomes easier to evaluate with the parabolic band approximation. Nevertheless, upon solving the BTE the term $f^\nu(\mathbf{x}, \mathbf{k}', t)(1 - f^\nu(\mathbf{x}, \mathbf{k}, t))$ in Equation (2.28) might prove to be too challenging for a certain numerical BTE solver. Thus the Pauli exclusion principle is often dropped and the scattering operator reduces to

$$\mathcal{Q}\{f^\nu(\mathbf{x}, \mathbf{k}, t)\} = \frac{1}{(2\pi)^3} \sum_\nu \int_{\mathcal{B}} S^\nu(\mathbf{k}', \mathbf{k}, t) f^\nu(\mathbf{x}, \mathbf{k}', t) - S^\nu(\mathbf{k}, \mathbf{k}', t) f^\nu(\mathbf{x}, \mathbf{k}, t) d^3 k' \quad (2.30)$$

and therefore gives the *Maxwell-Boltzmann distribution* as equilibrium distribution function

$$f^\nu(\mathbf{x}, \mathbf{k}, t) = \exp\left(-\frac{E_\nu(\mathbf{k}) - E_F}{k_B T_L}\right). \quad (2.31)$$

The integral in Equation (2.30) is easier to evaluate than Equation (2.28), especially when using the parabolic band approximation and forms the basis of many important approximations, such as the *relaxation time approximation* (RTA) [22]. Dropping Pauli's exclusion principle is clearly justified if $(1 - f^\nu(\mathbf{x}, \mathbf{k}, t)) \approx 1$ and thus $f^\nu(\mathbf{x}, \mathbf{k}, t) \ll 1$, which is only true for low electron (hole) concentrations and non degenerate semiconductors. Thus dropping Pauli's exclusion principle in Equation (2.28) is often termed *low density approximation*.

Interaction	Elastic/Inelastic	References
Acoustic Phonon Scattering	Approx. Elastic	[23-25]
Optical Phonon Scattering	Inelastic	[23-25]
Impurity Scattering	Approx. Elastic	[23, 25, 26]
Impact Ionization	Inelastic	[25, 27]
Electron-Electron Scattering	Inelastic	[28]

Table 2.1: A list of particle interactions considered in this thesis including references to the used models for the respective scattering rates.

2.2.2 Recombination and Generation

The scattering operator as described in Section 2.2.1 does not consider electron-hole recombination (cf. Equation (2.25)). In order to be able to consider recombination/generation processes the BTE needs to be split into a BTE for electrons and holes respectively. Since recombination/generation models can be quite complex [29], we first restrict this section to the most simple two-state defect (cf. Figure 2.6), where only a single function over time, namely the trap occupancy $f_T^\nu(t)$ and trap level E_T are needed to describe the state of the defect. Additionally it will be assumed that there are enough electrons and holes that are either spin up or spin down, such that the trap occupancy is independent of the electron spin. In case

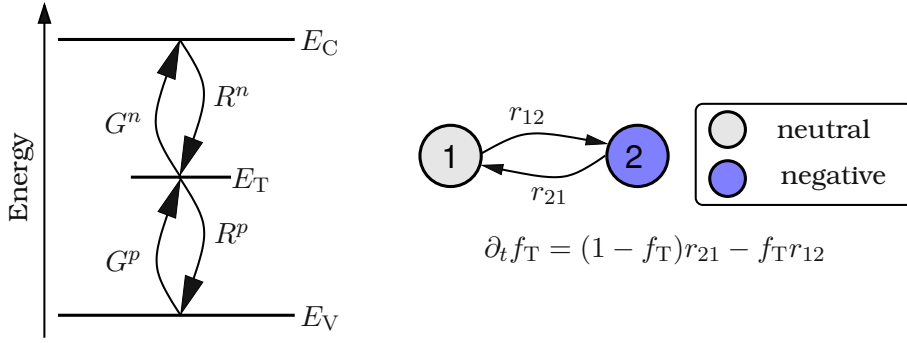


Figure 2.6: Left: The trap occupancy is governed by a first order rate equation and two transition rates r_{12} and r_{21} , which are obtained from the four recombination/generation rates shown in the band diagram. The defect can become charged (state 2) by either capturing an electron from the conduction band or by emitting a hole to the valence band and vice versa. **Right:** The two state model of a defect located within the band gap ($E_V < E_T < E_C$). Only in state two does the defect carry a charge, which needs to be considered upon solving Poisson's equation.

the spin cannot be neglected, the trap occupancy needs to be split into its spin components, $f_T^\nu(t) \Rightarrow f_{T, \text{up}}^\nu + f_{T, \text{down}}^\nu$. With these assumptions the macroscopic rate equation for the trap occupancy reads [12, 30]

$$\begin{aligned} \partial_t f_T^\nu(t) = & \sum_\nu \int_{\mathcal{B}} (1 - f_T^\nu(t)) \underbrace{(R_\nu^n(\mathbf{k}) f_n^\nu - G_\nu^p(\mathbf{k})(1 - f_p^\nu))}_{r_{21}} \\ & + f_T^\nu(t) \underbrace{(R_\nu^p(\mathbf{k}) f_p^\nu - G_\nu^n(\mathbf{k})(1 - f_n^\nu))}_{r_{12}} d^3k, \end{aligned} \quad (2.32)$$

where $G_\nu^n(\mathbf{k})$ and $G_\nu^p(\mathbf{k})$ are the number of generated electrons and holes per second per d^3k , $R_\nu^n(\mathbf{k})$ and $R_\nu^p(\mathbf{k})$ are the number of recombined electrons and holes per second per d^3k . Putting all of the above together, the recombination operators $\Gamma^{\text{p/n}}\{f_n^\nu, f_p^\nu\}$ for electrons and holes, which have been introduced in Equation (2.27), read

$$\Gamma^n\{f_n^\nu, f_p^\nu\} = \frac{N_T}{(2\pi)^3} \sum_\nu \int_{\mathcal{B}} G_\nu^n(\mathbf{k}) f_T^\nu(t) (1 - f_n^\nu) - R_\nu^n(\mathbf{k}) (1 - f_T^\nu(t)) f_n^\nu d^3k, \quad (2.33)$$

$$\Gamma^p\{f_n^\nu, f_p^\nu\} = \frac{N_T}{(2\pi)^3} \sum_\nu \int_{\mathcal{B}} G_\nu^p(\mathbf{k}) (1 - f_T^\nu(t)) (1 - f_p^\nu) - R_\nu^p(\mathbf{k}) f_T^\nu(t) f_p^\nu d^3k, \quad (2.34)$$

where N_T is the trap concentration. Dropping the sum over all bands in the equations above, the recombination rates for electrons and holes are [30, 31]

$$R_\nu^n(\mathbf{k}) = \sigma^n v^n(\mathbf{k}) \quad \text{and} \quad R_\nu^p(\mathbf{k}) = \sigma^p v^p(\mathbf{k}), \quad (2.35)$$

where σ^n and σ^p are experimentally determined capture cross sections and v^n and v^p reaction velocities. From detailed balance [12, 31] the generation rates can be determined

$$G_\nu^n(\mathbf{k}) = \sigma^n v^n(\mathbf{k}) \exp\left(\frac{E_T - E^n(\mathbf{k})}{k_B T_L}\right), \quad (2.36)$$

$$G_\nu^p(\mathbf{k}) = \sigma^p v^p(\mathbf{k}) \exp\left(\frac{E^n(\mathbf{k}) - E_T}{k_B T_L}\right). \quad (2.37)$$

2.3 The Method of Moments

In order to fully analyze semiconductor devices such as metal-oxide-semiconductor field-effect transistors (MOSFET), one needs to solve a coupled system of equations consisting of the Poisson equation and the BTE for electrons and holes. A direct discretization yields a system of equations in seven dimensions. Thus simpler models, which at least capture the essential features of the BTE are often used. One of these methods by which plenty of charge transport models can be obtained is the method of moments [32, 33]. The j th moment of the BTE is defined as

$$\langle \chi_j \rangle = \int_{\mathcal{B}} \chi_j f(\mathbf{x}, \mathbf{k}, t) d^3k = p_j \int_{\mathcal{B}} \mathbf{k}^j f(\mathbf{x}, \mathbf{k}, t) d^3k, \quad (2.38)$$

where χ_j is the j th weight function and p_j is the prefactor of the j th weight function. Table 2.2 lists a few important moments of the distribution function. It is important to note that the moments of the distribution function infer no assumptions and are used to obtain macroscopic quantities often employed in the analysis of the performance of semiconductor devices. To

Moment	χ_j	Formula	Macroscopic Quantity
$\langle \chi_0 \rangle$	$\chi_0 = 1$	$1 \int_{\mathcal{B}} f(\mathbf{x}, \mathbf{k}, t) d^3k$	n or p
$\langle \chi_1 \rangle$	$\chi_1 = \mathbf{k}$	$\hbar \int_{\mathcal{B}} f(\mathbf{x}, \mathbf{k}, t) \mathbf{k} d^3k$	$\langle \mathbf{v} \rangle = \mathbf{J}_{n/p}/q$
$\langle \chi_2 \rangle$	$\chi_2 = E(\mathbf{k})$	$\int_{\mathcal{B}} E(\mathbf{k}) f(\mathbf{x}, \mathbf{k}, t) \mathbf{k}^2 d^3k$	$\langle E_{n/p} \rangle$
...	

Table 2.2: A list of the first few moments of the distribution function, where n and p are the electron and hole concentrations, $\mathbf{J}^{n/p}$ is the current density and $\langle E^{n/p} \rangle$ is the average carrier energy for electrons and holes respectively.

obtain equations for the moments of the BTE, such as the drift diffusion (DD) model, the BTE is multiplied by increasing orders of the wave vector \mathbf{k} and a scalar prefactor p and afterwards integrated over the Brillouin zone

$$\int_{\mathcal{B}} \chi_j (\partial_t f(\mathbf{x}, \mathbf{k}, t) + \mathcal{L}\{f(\mathbf{x}, \mathbf{k}, t)\}) d^3k = \int_{\mathcal{B}} \chi_j \mathcal{Q}\{f(\mathbf{x}, \mathbf{k}, t)\} d^3k + R. \quad (2.39)$$

In order to analytically evaluate the integrals over the free-streaming operator and the scattering operator further assumptions are necessary. Since the integral over the free streaming operator in Equation (2.39) would lead to a tensor equation for odd weight functions, identified by an odd index j , a further assumption is needed. This assumption has to be chosen such that all tensors after integration on the left-hand side of Equation (2.39) are diagonal with equal entries along the diagonal. It turns out that it is sufficient to decompose the distribution function into a symmetric and antisymmetric part,

$$\begin{aligned} f(\mathbf{x}, \mathbf{k}, t) &= f_S(\mathbf{x}, \mathbf{k}, t) + f_A(\mathbf{x}, \mathbf{k}, t), \\ f_S(\mathbf{x}, \mathbf{k}, t) &= f_S(\mathbf{x}, -\mathbf{k}, t), \\ f_A(\mathbf{x}, -\mathbf{k}, t) &= -f_A(\mathbf{x}, \mathbf{k}, t), \end{aligned}$$

and assume that the symmetric part is isotropic,

$$f_S(\mathbf{x}, \mathbf{k}, t) = f_S(\mathbf{x}, \|\mathbf{k}\|, t),$$

as well as

$$|f_A(\mathbf{x}, \mathbf{k}, t)| \ll |f_S(\mathbf{x}, \mathbf{k}, t)|$$

which yields a single unique equation per moment [34]. These assumptions are termed *diffusion approximation*, since it assumes that diffusion of carriers dominates over carrier drift, caused by an electric field. To analytically evaluate the right hand side of Equation (2.39), the *relaxation time approximation* (RTA) is often utilized. The main assumption of the RTA is that due to scattering the distribution function will relax exponentially into the equilibrium distribution function $f_{\text{eq}}(\mathbf{x}, \mathbf{k}, t)$,

$$\int_{\mathcal{B}} \chi_j (\partial_t f(\mathbf{x}, \mathbf{k}, t) + Q\{f(\mathbf{x}, \mathbf{k}, t)\}) d^3k = \frac{f(\mathbf{x}, \mathbf{k}, t) - f_{\text{eq}}(\mathbf{x}, \mathbf{k}, t)}{\tau_s},$$

where the transition is characterized by a scattering process dependent relaxation time constant τ_s [22]. This approximation also implies, as can be shown [22], that the relaxation time constants of various scattering processes can be added up to a single time constant as follows:

$$\frac{1}{\tau_s} = \frac{1}{\tau_1} + \frac{1}{\tau_2} + \dots$$

After employing all of the assumptions, detailed above, one will obtain an infinite enumerable number of equations. The equation for the j th moment will always contain the $j + 1$ moment. Thus the array of equations needs to be terminated by replacing the equation for the $j + 1$ moment with an analytic formula. Such an analytic formula is termed *closure condition*, which is also the major drawback of any transport model obtained through the method of moments.

2.3.1 The Drift Diffusion Model

The drift diffusion (DD) model is one of the simplest and earliest numerically used and investigated charge transport models [35]. It is obtained by employing the parabolic band approximation (cf. Section 2.1) under the additional assumption of a single conduction and a single valence band. Evaluating the first two moments of the BTE one obtains,

$$\langle \chi_0 \rangle \Rightarrow \partial_t n - |q|^{-1} \nabla \mathbf{J}_n - R = 0, \quad (2.40)$$

$$\langle \chi_0 \rangle \Rightarrow \partial_t p + |q|^{-1} \nabla \mathbf{J}_p + R = 0, \quad (2.41)$$

$$\langle \chi_1 \rangle \Rightarrow \tau_m \partial_t \mathbf{J}_n - |q| \tau_m k_B \nabla \cdot (n T_n) + m^* \mathbf{J}_n = 0, \quad (2.42)$$

$$\langle \chi_1 \rangle \Rightarrow \tau_m \partial_t \mathbf{J}_p + |q| \tau_m k_B \nabla \cdot (p T_p) - m^* \mathbf{J}_p = 0, \quad (2.43)$$

where

$$\langle \chi_2 \rangle \Rightarrow \langle E_{n/p} \rangle = \frac{3nk_B T_L}{2} \quad \text{with} \quad T_L = T_n = T_p, \quad (2.44)$$

is used as a closure condition, τ_m is the momentum relaxation time constant, T_n and T_p are the electron and hole temperatures respectively and R is a scalar term accounting for electron/hole recombination and generation. In order to further simplify the DD model, the influence of the time derivative of the current density is often neglected, which is justified if the frequency of the electric field \mathbf{E} is two times lower than the plasma frequency [36]. Thus, including Poisson's

equation to calculate the electric field, the DD model reads

$$\nabla \cdot (\varepsilon(\mathbf{x})\nabla\varphi) = |q|(n - p + C), \quad (2.45)$$

$$|q|\partial_t n - \nabla \mathbf{J}_n = |q|R, \quad (2.46)$$

$$|q|\partial_t p + \nabla \mathbf{J}_p = -|q|R, \quad (2.47)$$

$$\mathbf{J}_n = \underbrace{\frac{|q|\tau_m}{m^*}}_{\mu_n} k_B \left(\nabla(nT_L) + \frac{|q|}{k_B} \mathbf{E} \right) = \underbrace{|q|\mu_n \mathbf{E}}_{\text{drift term}} + \underbrace{|q|D_n \nabla n}_{\text{diffusion term}}, \quad (2.48)$$

$$\mathbf{J}_p = \underbrace{\frac{|q|\tau_m}{m^*}}_{\mu_p} k_B \left(-\nabla(pT_L) + \frac{|q|}{k_B} \mathbf{E} \right) = |q|\mu_p \mathbf{E} - |q|D_p \nabla p, \quad (2.49)$$

where the electron μ_n and hole mobilities μ_p , the thermal voltage $V_T = k_B T_L / |q|$ and the Einstein relation

$$D_{n/p} = V_T \mu_{n/p}$$

have been used. Upon employing this method, one loses exact information regarding the kinetic energy and momentum (\mathbf{k}) of the charge carriers and their distribution. Although various material, field and doping dependent models for carrier mobilities have been developed [35], the DD model still suffers from many of the assumptions needed in the course of its derivation. As such, a consequence of setting the lattice temperature and carrier temperature equal is that carrier diffusion will be underestimated by the model. Additionally, by employing the equipartition theorem or its enhancement, the homogeneous energy balance equation [34], to calculate the average carrier energy it is not possible to describe effects of rapidly changing electric fields. Thus it is for example not possible to explain the phenomenon of velocity overshoot [37] and energy transport phenomena using the DD model. Nevertheless, because of its ease to comprehend and compact formulation, the DD model is the best known and most widely used charge transport model today.

2.3.2 The Hydrodynamic Model

The hydrodynamic model (HD), first developed by [32, 33], is derived like the DD model by employing the method of moments. Instead of using only the first two moments of the BTE, the HD model uses the first three moments in order to incorporate spatial dependencies on the average carrier energy, thereby adding two new unknowns and at least one new parameter. As a closure condition Fourier's law is used for the fourth moment. In addition the initial assumption that

$$|f_A(\mathbf{x}, \mathbf{k}, t)| \ll |f_S(\mathbf{x}, \mathbf{k}, t)|,$$

is dropped in the HD model, which reads

$$\nabla \cdot (\varepsilon(\mathbf{x})\nabla\varphi) = |q|(n - p + C), \quad (2.50)$$

$$\partial_t n - |q|^{-1} \nabla \mathbf{J}_n - R = 0, \quad (2.51)$$

$$\tau_m \partial_t \mathbf{J}_n - |q|\mu_n n \mathbf{E} - \mu_n k_B \nabla(nT_n) - \mathbf{J}_n = \tau_m |q|^{-1} (\mathbf{J}_n \otimes \mathbf{J}_n n^{-1}), \quad (2.52)$$

$$\nabla \cdot (nS_n) + \partial_t(n\omega) - \mathbf{E} \cdot \mathbf{J}_n + n(\omega - \omega_0) * \tau_E^{-1} = 0, \quad (2.53)$$

$$nS_n + |q|^{-1} (\omega + k_B T_n) \mathbf{J}_n + \kappa(T_n) \nabla T_n = 0, \quad (2.54)$$

for electrons. There κ is given by the Wiedemann-Franz law, τ_E is the energy relaxation time, nS_n is the carrier heat flux (third moment), ω is the average carrier energy and ω_0 is the equilibrium average carrier energy. For holes a similar set of equations is obtained. A simplified version of the HD model, the energy transport model, can be obtained by assuming that

$$\tau_m |q|^{-1} (\mathbf{J}_n \otimes \mathbf{J}_n n^{-1}) \approx 0 \quad \text{and} \quad \omega \approx \frac{3k_B T_n}{2}. \quad (2.55)$$

The benefit of the HD model over the DD model is, that it can describe velocity overshoot, although it can only cover processes which can be explained by employing heated Maxwellian distribution functions. Upon investigating this property of the HD model it was found that one can observe an artificial velocity overshoot for decreasing electric fields too [38]. Sadly, it was also found that this is due to the truncation after the 4th moment in the construction of the model and thus an intrinsic property of the HD model [38]. Additionally, the HD model tends to overestimate the number of carriers in the bulk of MOSFETs. In addition to all that the equations of the HD model are strongly hyperbolic, which makes them numerically challenging to solve for arbitrary device geometries.

2.3.3 Higher Order Models

Higher order charge transport models have been developed in hope to overcome the limitations of the hydrodynamic model. The most notable one is the Six Moments model, where the closure condition is obtained through an empirical relation [39]. Although higher order models capture more of the essential physics described by the BTE they also tend to be more complex, have more parameters and are harder to understand, derive and implement. The increased complexity in these models often, despite successfully describing many transport effects observed in semiconductors, trades off the benefit of the more accurate description.

2.4 The Monte Carlo Method

The Monte Carlo method itself is a stochastic algorithm to solve integral equations [25, 40, 41]. Since the BTE can be formally transformed into an integral equation, the Monte Carlo (MC) method provides a way to solve the BTE. The main idea of the MC method for the BTE is to simulate the flow (cf. Figure 2.4) of a statistically representative ensemble of particles by following the path of each single particle. In between scattering events each particle propagates freely according to

$$\hbar \partial_t \mathbf{k} = \mathbf{F} \quad \text{and} \quad \partial_t \mathbf{x} = \mathbf{v}, \quad (2.56)$$

while scattering events are determined stochastically and occur instantly. In order to determine the instance of a scattering event one starts from the formally integrated BTE. From this form one can, through algebraic transformations, obtain the conditional probability density $p(\mathbf{x}, \mathbf{k}, \nu, t_0 + T | \mathbf{x}_0, \mathbf{k}_0, \nu_0, t_0)$ to find a particle (electron or hole), without being scattered, in state $(\mathbf{x}, \mathbf{k}, \nu)$ at time $t_0 + T$ after it has been at time t_0 in state $(\mathbf{x}_0, \mathbf{k}_0, \nu_0)$. Now $p(\mathbf{x}_0, \mathbf{k}_0, \nu_0, t_0 | \mathbf{x}, \mathbf{k}, \nu, t_0 + T)$ is evaluated using random numbers to determine whether or not a particle is scattered, giving the method its name and characteristics (cf. Figure 2.7). It is important to state that by using the MC method it is possible to directly solve the uncoupled (no recombination or generation) BTEs for electrons and holes self-consistently with Poisson's equation (cf. Equation (2.27)). By virtue of the method it is not necessary to approximate the

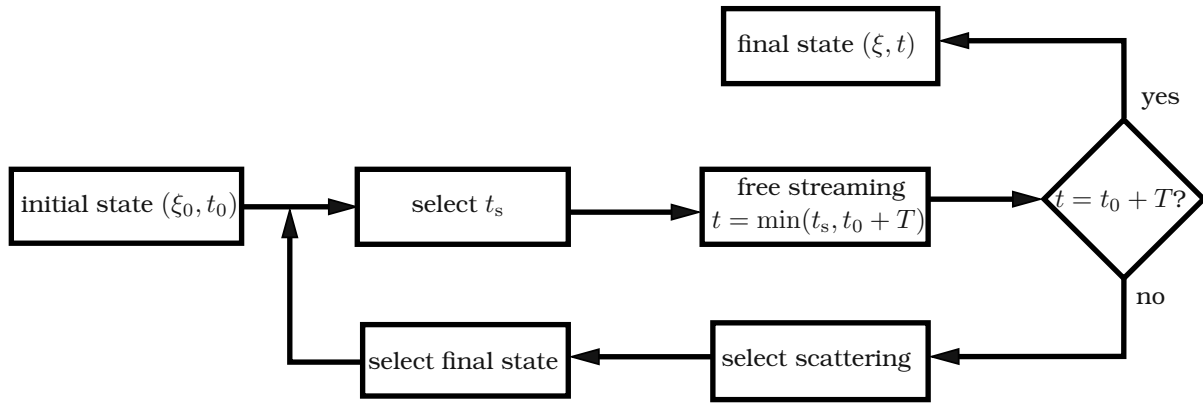


Figure 2.7: A flowchart of the Monte Carlo method for solving the BTE in time T , where $\xi = (\mathbf{x}, \mathbf{k}, \nu)$ and t_s is the time at which the next scattering event occurs. Reproduced from [25].

bandstructure $E_\nu(\mathbf{k})$ of the semiconductor and it is indeed straightforward to incorporate the full dispersion relation into a simulator using the MC method. Additionally, one can directly obtain f_n^ν and f_p^ν from any MC simulator by averaging over all particles and thus calculate the moments by evaluating the respective integrals numerically (cf. Section 2.3). This is indeed very satisfying, since a maximum of physically relevant and accurate information can in principle be obtained by a minimum of approximations. Nevertheless, approximations to the BTE are necessary. In order to be able to apply the MC method, the BTE needs to be linearized [25]. This linearization in turn requires one to drop Pauli's exclusion principle in the scattering operator and also severely impacts the ability to rigorously implement non-linear scattering operators such as electron-electron scattering. However, electron-electron scattering in Monte Carlo based device simulators, using a plethora of approximations, has been implemented for 1D device geometries [42, 43]. Additionally, in order for the MC method to yield useful results very large numbers of particles are required in a simulation, since the accuracy in $f^\nu(\mathbf{x}, \mathbf{k}, t)$ exhibits a square root dependence on the number of particles [44]. Large ensembles of particles in turn lead to a large computational burden, such that significant MC simulations can take from a few hours up to several workdays or even months to complete. The reason for this unfortunate property is due to the stochastic nature of any MC method, which can even be easily observed in simple simulations. In Figure 2.9, the electron distribution function f_n for a n+-n-n+ structure (cf. Figure 2.8) obtained by using the MC simulator MONJU [45] is presented. The numeric noise towards higher kinetic energies is clearly visible.

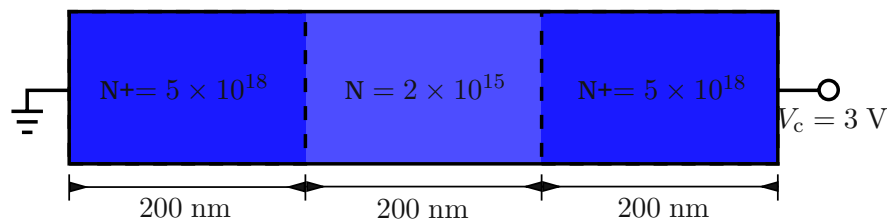


Figure 2.8: The exemplary 2D n+-n-n+ structure simulated using MONJU [45]. The structure is 100 nm wide.

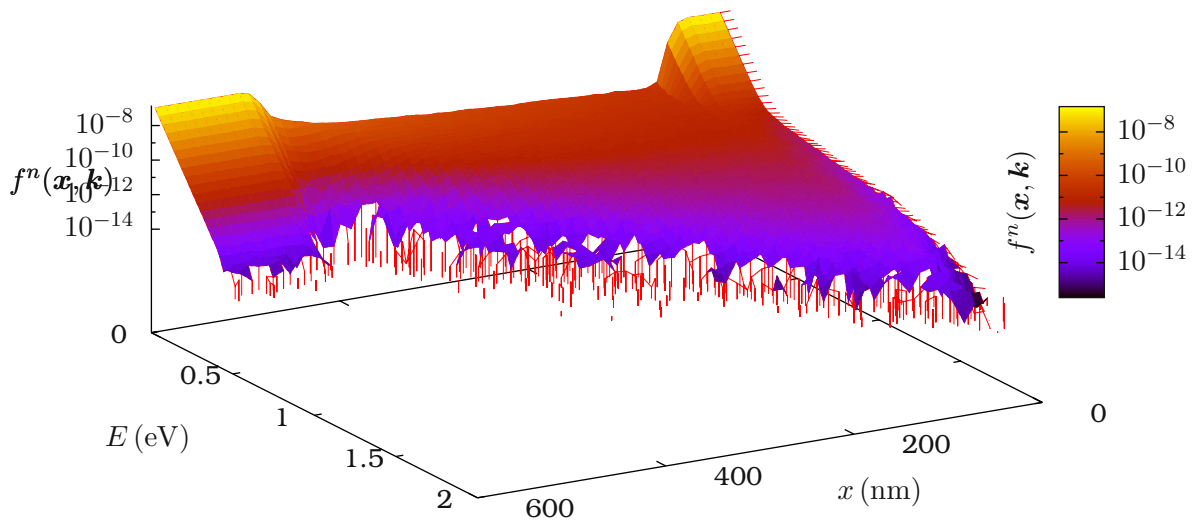


Figure 2.9: The electron distribution function plotted over the long side of the n⁺-n-n⁺ diode, obtained by full-band MC simulation using MONJU [45]. The numeric noise towards higher kinetic energies is clearly visible. In the simulation, which took one workday to complete, five million particles have been used.

2.5 The Spherical Harmonics Expansion Method

Inherent numerical noise in the Monte Carlo solutions of the Boltzmann Transport Equation, as well as the $\mathcal{O}(N^{-1/2})$ dependence of the integration error inherent in Monte Carlo methods has motivated the development of new techniques to solve Boltzmann's Transport Equation. Thus special attention has been devoted to deterministic efforts for solving Boltzmann's Equation. A very attractive deterministic method is to expand the distribution function into spherical harmonics and to project the k-space onto a single energy axis. This way a seven dimensional system is reduced to a five dimensional system, which does not have the limitations of a Monte Carlo approach. Nevertheless, a spherical harmonics expansion (SHE) of the BTE is still challenging [46] and it is, compared to Monte Carlo, difficult to consider as many full-band effects in the simulation [47] as possible. However, a quite satisfying approach to consider full-band effects has been found by [48], where current of diodes could be predicted using a fifth order expansion within an error margin of 8% in saturation compared to the results of a full-band Monte Carlo simulation. In addition to being free of stochastic noise, a spherical harmonics expansion allows to straightforward and self-consistently solve the BTE with any other partial differential equation, especially Poisson's Equation using a Newton-Raphson solver [49]. Thus special attention will be paid to the SHE method in Chapter 3 and for hot-carrier degradation modelling in the last chapter of this thesis.

2.6 Quantum Mechanical Effects - Confinement

Until now purely quantum mechanical effects, such as charge carrier confinement, have not been addressed. In MOS structures, considered in this work, carrier confinement emerges

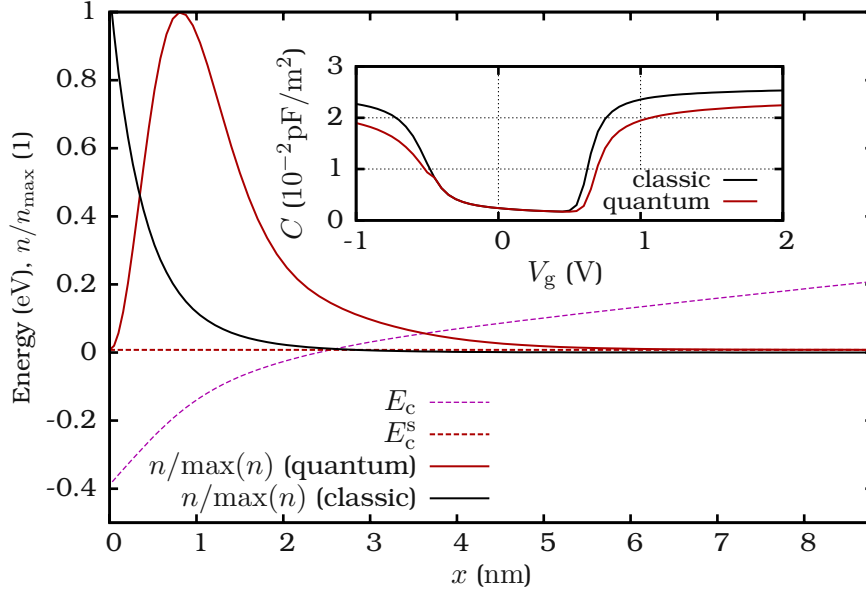


Figure 2.10: The confinement of electrons and the emergence of sub-bands in a exemplary 1D nMOS structure, where $x = 0$ nm corresponds to the insulator-semiconductor interface. Shown are the conduction band edge E_c and the first sub-band E_c^s at 7.7 meV as well as the normalized electron concentrations as obtained by a classical calculation and by a quantum mechanical one. The model error of a classical calculation can be well seen in the Capacitance-Voltage curve (inset). For pMOS analog results are obtained. For the calculation the Vienna Schrödinger Poisson Solver (VSP) [51] has been used.

whenever local potential wells are formed [50]. Deep potential wells cause discrete quantized energy levels and confine the charge carriers in one or more directions. One of the most widely recognized case of carrier confinement occurs in all MOSFETs at the semiconductor-insulator interface, whenever the MOSFET is driven into inversion as shown in Figure 2.10. In this case the charge carriers are confined in the direction perpendicular to the interface and free to move in the other two directions, thus forming a two dimensional electron gas [50]. When deriving the BTE in Section 2.2, the wave character of the electrons was assumed to be negligible in favor of a classical description of the physical system. In MOS structures with thin insulators in the range of nanometers, e.g. $t_{\text{ox}} = 1$ nm, charge carrier confinement needs to be included in order to yield an accurate description of the device. For an accurate quantum mechanical description of carrier confinement the Schrödinger equation

$$-\left(\frac{\hbar^2}{2}(\mathbf{m}^*)^{-1}\nabla + qV'(\mathbf{x})\right)\psi^s(\mathbf{x}) = \epsilon^s(\mathbf{x})\psi^s(\mathbf{x}) \quad (2.57)$$

for a given confining electrostatic potential $V'(\mathbf{x})$ must be solved (self-consistently) with Equation (2.27), where s is the sub-band index and ϵ^s is the energy eigenvalue. In order to solve the above equation boundary conditions need to be set for an isolated device. It is natural to assume vanishing electron wave functions at the boundaries of the semiconductor. If one is interested in the penetration of the wave function into an oxide in order to calculate direct tunneling currents [52] this can be achieved by setting a Dirichlet boundary condition at the boundaries of the oxide not interfacing with a semiconductor, although the method itself is dependent on the oxide thickness and only justified for thick oxides.

2.7 Non-Equilibrium Greens Functions Approach

In the course of the derivation of Boltzmann's transport equation, it was assumed that the wave nature of the electrons can be neglected. The concept of the BTE, as shown in Figure 2.4, is that during free streaming the electrons are treated as classical particles embedded in the lattice of fixed positively charged ions. As long as the electrons scatter frequently while traveling, the phase of their wave function and initial speed is lost, which is true for channels longer than the coherence length and slowly spatially varying electrostatic potentials. This is for example not the case when the channel only consists of a few atoms or a single molecule (biomolecule) [53,54]. In this case a non-equilibrium Schrödinger equation for charge transport needs to be solved. The formalism most commonly used to solve the Schrödinger equation in non-equilibrium is referred to as *Non-Equilibrium Greens Functions Approach* (NEGF) [55]. The channel, of a MOSFET for example, is then modelled as consisting of a few slices (2D case) of atoms with contacts in thermal equilibrium on the left and right of the channel (cf. Figure 2.11). The one-particle Schrödinger equation in NEGF formalism then reads

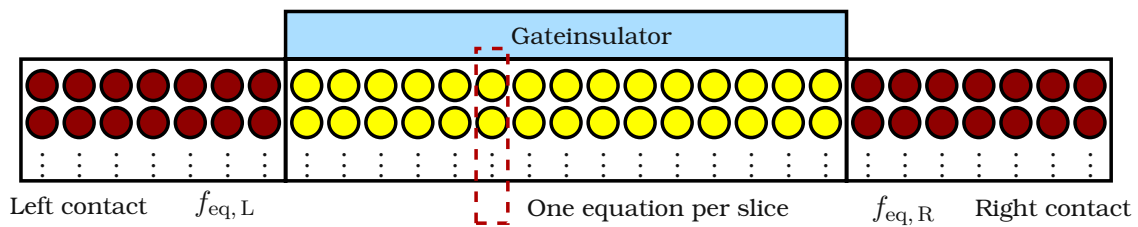


Figure 2.11: Schematic of a 2D nanoscale MOSFET, consisting of a few atoms, simulated using NEGF. For every 1D slice (yellow atoms) a Schrödinger equation is assembled and solved. It is assumed that the contact (red atoms) are in thermal equilibrium.

$$(EI - H_C - \Sigma_L - \Sigma_R - \Sigma_S)G = BG = I, \quad \text{and} \quad BG^< = \Sigma^<G^A, \quad (2.58)$$

where E is the electron energy, H_C is the Hamiltonian for the channel, Σ_L and Σ_R are the self-energies accounting for the left and right contacts, Σ_S is the self-energy due to scattering, B is to account for the boundary conditions at the contacts, G is the retarded Green's function, $G^<$ is the lesser Green's function and G^A is the advanced Green's function [55,56]. From $G^<$ macroscopic quantities, such as the electron concentration and the current density can be obtained. The main advantage of NEGF is that the wave character of the electrons is preserved, which leads to a highly accurate description of nanoscale ($\sim 10\text{nm}$) devices. But as soon as the devices become larger, the matrices in Equation (2.58) become too large for any numerical solver. In large devices, scattering occurs more frequently, which increases the number of off-diagonals in Equation (2.58). In addition it is not possible to rigorously consider electron-electron scattering in the NEGF method. Thus special attention has to be paid to the calculation of the scattering rates. Nevertheless, NEGF is one of the most accurate methods to describe charge carrier transport in nanoscale devices as well as direct tunneling in MOS structures [56]. In the course of this thesis NEGF was used to estimate direct tunneling currents, where only phonon scattering was considered.

3 Spherical Harmonics Expansion

Recent advances in the spherical harmonics expansion (SHE) method of the Boltzmann transport equation (BTE) allow for the accurate solution of the BTE on arbitrary three-dimensional device geometries [57], full-band effects [48, 58, 59], rare scattering events (e.g. impact ionization) [60], carrier-carrier scattering [61], charge carrier recombination/generation [62], small signal analysis [47, 63], quantum mechanics [49, 64], and hot carrier degradation [46]. The SHE method has been developed to a point where it is now an attractive alternative to the common Monte Carlo (MC) method, which has a square root dependence of its accuracy on CPU time [44]. As discussed in Section 2.4, the Monte Carlo method suffers from inherent stochastic noise in the solution and the requirement of sufficiently small time steps to achieve self-consistency with Poisson's equation. An expansion of the BTE using spherical harmonics does not impose such restrictions, since it is a deterministic approach. In the course of this work the SHE simulator ViennaSHE [65] has been extended and used to solve BTE (cf. Equation (2.27)).

3.1 Theory

The main idea of the SHE method is to expand the distribution function $f^\nu(\mathbf{x}, \mathbf{k}, t)$ into spherical harmonics (SH)

$$Y_{l,m}(\theta, \varphi) = \begin{cases} N_{l,m} P_l^m(\cos(\theta)) \cos(m\varphi), & \text{if } m \geq 0, \\ N_{l,m} P_l^{-m}(\cos(\theta)) \sin(-m\varphi), & \text{if } m < 0, \end{cases} \quad (3.1)$$

where $P_l^m(x)$ are the associated Legendre polynomials, $l \geq 0$ is the order, m is the sub-order index bounded by $-l \leq m \leq l$, and $N_{l,m}$ are normalization factors given by

$$N_{l,m} = \begin{cases} \sqrt{\frac{2l+1}{4\pi}}, & \text{if } m = 0, \\ \sqrt{\frac{(2l+1)(l-|m|)!}{2\pi(l+|m|)!}}, & \text{if } m \neq 0. \end{cases}$$

The spherical harmonics form an orthonormal basis since

$$\underbrace{\int_{\theta=0}^{\pi} \int_{\varphi=0}^{2\pi} Y_{l,m} Y_{l',m'} \sin(\theta) d\varphi d\theta}_{\equiv \oint} \underbrace{\equiv d\Omega} = \delta_{l,l'} \delta_{m,m'},$$

where the Kronecker delta $\delta_{ab} = 1$ if and only if $a = b$, else it equates to zero. Additionally, the shorthands \oint and $d\Omega$ have been introduced. With the above definition of the spherical

harmonics, the truncated expansion of the distribution function $f(\mathbf{x}, \mathbf{k}, t)$ is

$$f(\mathbf{x}, \mathbf{k}, t) \simeq \sum_{l=0}^L \sum_{m=-l}^l f_{l,m}(\mathbf{x}, \epsilon, t) Y_{l,m}(\theta, \varphi), \quad (3.2)$$

where the order l is bounded by the maximum expansion order ($l < L$), ϵ is the energy, θ and φ are the spherical angles of the spherical harmonics $Y_{l,m}$ [47, 59, 64, 66]. Additionally, the expansion of the distribution function in spherical harmonics on equi-energy surfaces implicitly assumes a bijective mapping $E(\mathbf{k})$ (cf. Chapter 2) between energy and wave vector. An exhaustive treatment of the derivation of the BTE expanded in spherical harmonics can be found in [47]. For the sake of clarity the derivation will be briefly summarized in here. In order to avoid notational clutter, the band index ν will be dropped in this chapter entirely, since the projection on spherical harmonics does not change anything regarding the band index. With this, any physical quantity X can be expanded into SH on equi-energy surfaces using

$$X \simeq \sum_{l=0}^L \sum_{m=-l}^l X_{l,m}(\epsilon) Y_{l,m}(\theta, \varphi). \quad (3.3)$$

The expansion coefficients for any quantity X on equi-energy surfaces are then found as [47, 67]

$$\begin{aligned} X_{l,m} &= \frac{1}{(2\pi)^3} \int_{\mathcal{B}} Y_{l,m}(\theta(\mathbf{k}), \varphi(\mathbf{k})) X(\mathbf{x}, \mathbf{k}) \delta(\epsilon - E(\mathbf{k})) d^3k \\ &= \oint Y_{l,m}(\theta, \varphi) X(\mathbf{x}, \mathbf{k}(\epsilon, \theta, \varphi)) Z(\epsilon, \theta, \varphi) d\Omega, \end{aligned} \quad (3.4)$$

where δ is the Dirac delta. Projecting the distribution function $f(\mathbf{x}, \mathbf{k}, t)$ using Equation (3.4) one obtains the expansion coefficients $f_{l,m}$ as

$$\begin{aligned} f_{l,m} &= \frac{1}{(2\pi)^3} \int_{\mathcal{B}} Y_{l,m}(\theta(\mathbf{k}), \varphi(\mathbf{k})) f(\mathbf{x}, \mathbf{k}, t) \delta(\epsilon - E(\mathbf{k})) d^3k \\ &= \oint Y_{l,m}(\theta, \varphi) f(\mathbf{x}, \mathbf{k}(\epsilon, \theta, \varphi), t) Z(\epsilon, \theta, \varphi) d\Omega, \end{aligned} \quad (3.5)$$

The *generalized density of states* $Z(\epsilon, \theta, \varphi)$ for a single spin direction in the absence of magnetic fields transformed to spherical coordinates is

$$Z(\epsilon, \theta, \varphi) = \frac{2\|\mathbf{k}\|^2 \partial\|\mathbf{k}\|}{(2\pi)^3 \partial\epsilon}. \quad (3.6)$$

Due to the integration over the spherical angles the generalized density of states differs by a factor of 4π from the conventional density of states.

3.1.1 Expansion of the Free-Streaming Operator

In order to obtain a set of equations for the expansion coefficients up to order L , the BTE is multiplied by the generalized density of states (cf. Equation (3.6)), $Y_{l,m}$, and afterwards integrated over the unit sphere. Hence, a spherical harmonics expansion of the BTE,

$$\partial_t f(\mathbf{x}, \mathbf{k}, t) + \underbrace{\mathbf{v}_g^\nu(\mathbf{x}, \mathbf{k}) \cdot \nabla_{\mathbf{x}} f(\mathbf{x}, \mathbf{k}, t) - \mathbf{F}(\mathbf{x}, t) \cdot \nabla_{\mathbf{k}} f(\mathbf{x}, \mathbf{k}, t)}_{\mathcal{L}\{f(\mathbf{x}, \mathbf{k}, t)\}} = Q\{f(\mathbf{x}, \mathbf{k}, t)\} - \Gamma, \quad (3.7)$$

to find the unknown expansion coefficients $f_{l,m}$, reads

$$\oint Y_{l,m}(\theta, \varphi) \{ \text{BTE} \} Z(\epsilon, \theta, \varphi) d\Omega. \quad (3.8)$$

More precisely,

$$\begin{aligned} & \int_{\mathcal{B}} Y_{l,m} \delta_t f(\mathbf{x}, \mathbf{k}, t) \delta(\epsilon - E(\mathbf{k})) d^3k \\ & + \int_{\mathcal{B}} Y_{l,m} \mathcal{L}\{f(\mathbf{x}, \mathbf{k}, t)\} \delta(\epsilon - E(\mathbf{k})) d^3k \\ & = \int_{\mathcal{B}} Y_{l,m} (\mathcal{Q}\{f(\mathbf{x}, \mathbf{k}, t)\} - \Gamma) \delta(\epsilon - E(\mathbf{k})) d^3k, \end{aligned} \quad (3.9)$$

where the recombination term Γ will be treated separately in Section 3.5 and the time-dependent term will be discussed in detail later. In a second step the expansion 3.2 as well as a spherical harmonics expansion for the density of states are used in Equation (3.9) in order to find the equations per spatial location \mathbf{x} , energy ϵ , order l and m [47, 67]. Assuming that $Z(\epsilon, \theta, \varphi) = Z(\epsilon)$ the equations for the expansion coefficients fulfilling the BTE read term by term [67],

$$\partial_t f(\mathbf{x}, \mathbf{k}, t) \Rightarrow \partial_t \oint Y_{l,m} f(\mathbf{x}, \epsilon, t) Z(\epsilon) d\Omega, \quad (3.10)$$

$$\mathbf{v}_g^\nu(\mathbf{x}, \mathbf{k}) \cdot \nabla_{\mathbf{x}} f(\mathbf{x}, \mathbf{k}, t) \Rightarrow \nabla_{\mathbf{x}} \cdot \underbrace{\oint Y_{l,m} f(\mathbf{x}, \epsilon, t) v_g(\epsilon) Z(\epsilon) d\Omega}_{\mathbf{j}_{l,m}(\mathbf{x}, \epsilon, t)}, \quad (3.11)$$

$$\mathbf{F}(\mathbf{x}, t) \cdot \nabla_{\mathbf{k}} f(\mathbf{x}, \mathbf{k}, t) \Rightarrow \mathbf{F}(\mathbf{x}, t) \cdot \left(\frac{\partial \mathbf{j}_{l,m}(\mathbf{x}, \epsilon, t)}{\partial \epsilon} - \mathbf{A}_{l,m} \right), \quad (3.12)$$

$$\mathbf{A}_{l,m} = \oint \frac{1}{\hbar \|\mathbf{k}\|} \left(\frac{\partial Y_{l,m}}{\partial \theta} \mathbf{e}_\theta + \frac{\partial Y_{l,m}}{\sin(\theta) \partial \varphi} \mathbf{e}_\varphi \right) f Z(\epsilon) d\Omega, \quad (3.13)$$

where \mathbf{e}_θ and \mathbf{e}_φ are the unit vectors in the space spanned by the spherical harmonics. Whilst the time-dependent term can be projected in a straight-forward manner, the free streaming operator $\mathcal{L}\{f(\mathbf{x}, \mathbf{k}, t)\}$ is commonly split into two separate contributions. These two contributions are then also separately transformed. The first contribution (cf. Equation (3.12)) is often referred to as the diffusion term, since it can be expressed as the divergence of a generalized current density $\mathbf{j}_{l,m}(\mathbf{x}, \epsilon, t)$. The definition of the generalized current density is very convenient since the current density of charge carriers is obtained by

$$\mathbf{J}(\mathbf{x}, t) = \frac{1}{Y_{0,0}} \int \mathbf{j}_{0,0}(\mathbf{x}, \epsilon, t) d\epsilon. \quad (3.14)$$

The second contribution to the transformed free streaming operator is referred to as the drift term, since this term contains the force \mathbf{F} and no spatial derivative.

3.1.2 Expansion of the Scattering-Streaming Operator

The scattering operator $\mathcal{Q}\{f(\mathbf{x}, \mathbf{k}, t)\}$ is split into an in-scattering and an out-scattering term,

$$\mathcal{Q}\{f(\mathbf{x}, \mathbf{k}, t)\} = \mathcal{Q}^{\text{in}}\{f(\mathbf{x}, \mathbf{k}, t)\} - \mathcal{Q}^{\text{out}}\{f(\mathbf{x}, \mathbf{k}, t)\}. \quad (3.15)$$

Each term is then transformed to spherical coordinates on equi-energy surfaces by

$$\oint Y_{l,m}(\theta, \varphi) \mathcal{Q}\{f(\mathbf{x}, \mathbf{k}, t)\} Z(\epsilon, \theta, \varphi) d\Omega. \quad (3.16)$$

Utilizing the low-density approximation [68] for the scattering operators in Equation (2.30) the transformation simplifies to

$$\mathcal{Q}^{\text{in}}\{f(\mathbf{x}, \mathbf{k}, t)\} \Rightarrow \oint Y_{l,m} Z(\epsilon) \oint \sigma(\mathbf{k}(\epsilon', \theta', \varphi'), \mathbf{k}(\epsilon, \theta, \varphi)) f(\mathbf{x}, \epsilon', t) Z(\epsilon') d\Omega' d\Omega, \quad (3.17)$$

$$\mathcal{Q}^{\text{out}}\{f(\mathbf{x}, \mathbf{k}, t)\} \Rightarrow \oint Y_{l,m} Z(\epsilon) \oint \sigma(\mathbf{k}(\epsilon, \theta, \varphi), \mathbf{k}(\epsilon', \theta', \varphi')) f(\mathbf{x}, \epsilon, t) Z(\epsilon') d\Omega' d\Omega, \quad (3.18)$$

where it was assumed that the charge carrier is scattered from energy ϵ' to ϵ . The physics of the scattering process itself are summarized in the function σ . In this framework, the scattering process is considered to be elastic if $\epsilon' = \epsilon$, else it is an inelastic process. If the scattering operators are assumed to be *velocity randomizing* [68], i.e. the projected rates σ do not depend on the angles, the projection can be simplified further to yield

$$\mathcal{Q}^{\text{in}}\{f(\mathbf{x}, \mathbf{k}, t)\} \Rightarrow \sigma(\epsilon', \epsilon) \oint Y_{l,m} Z(\epsilon) \oint f(\mathbf{x}, \epsilon', t) Z(\epsilon') d\Omega' d\Omega, \quad (3.19)$$

$$\mathcal{Q}^{\text{out}}\{f(\mathbf{x}, \mathbf{k}, t)\} \Rightarrow \sigma(\epsilon, \epsilon') \oint Y_{l,m} Z(\epsilon) \oint f(\mathbf{x}, \epsilon, t) Z(\epsilon') d\Omega' d\Omega. \quad (3.20)$$

This can be considerably simplified by inserting the projection of the generalized density of states,

$$Z_{l,m}(\epsilon) = \oint Y_{l,m}(\theta, \varphi) Z(\epsilon, \theta, \varphi) d\Omega, \quad (3.21)$$

into Equation (3.22). After a few algebraic operations the scattering operators finally can be reduced to read,

$$\begin{aligned} \mathcal{Q}^{\text{in}}\{f(\mathbf{x}, \mathbf{k}, t)\} &\Rightarrow \sigma(\epsilon', \epsilon) \oint Y_{l,m} Z(\epsilon) \oint f(\mathbf{x}, \epsilon', t) Z(\epsilon') d\Omega' d\Omega \\ &= \frac{1}{Y_{0,0}} \sigma(\epsilon', \epsilon) Z_{l,m}(\epsilon) Z(\epsilon') f_{l',m'}(\epsilon', t) \delta_{0,l'} \delta_{0,m'} = \mathcal{Q}_{l',m',l,m}^{\text{in}}, \end{aligned} \quad (3.22)$$

$$\begin{aligned} \mathcal{Q}^{\text{out}}\{f(\mathbf{x}, \mathbf{k}, t)\} &\Rightarrow \sigma(\epsilon, \epsilon') \oint Y_{l,m} Z(\epsilon) \oint f(\mathbf{x}, \epsilon, t) Z(\epsilon') d\Omega' d\Omega \\ &= \frac{1}{Y_{0,0}} \sigma(\epsilon', \epsilon) Z_{0,0}(\epsilon') Z(\epsilon) f_{l',m'}(\epsilon, t) \delta_{l,l'} \delta_{m,m'} = \mathcal{Q}_{l',m',l,m}^{\text{out}}. \end{aligned} \quad (3.23)$$

3.2 The H-Transform

The spherical harmonics expansion applied to the BTE yields a set of equations for the unknown expansion coefficients $f_{l,m}$ in real space and energy. For the assembly of this set of equations, numeric stabilization techniques, such as the H-transform [68, 69], are needed [67]. The H-transform is a linear transformation of coordinates, where the kinetic energy ϵ is translated to the total energy H by

$$H = \begin{cases} \epsilon - |q| \varphi(\mathbf{x}, t), & \text{for electrons,} \\ \epsilon + |q| \varphi(\mathbf{x}, t), & \text{for holes.} \end{cases} \quad (3.24)$$

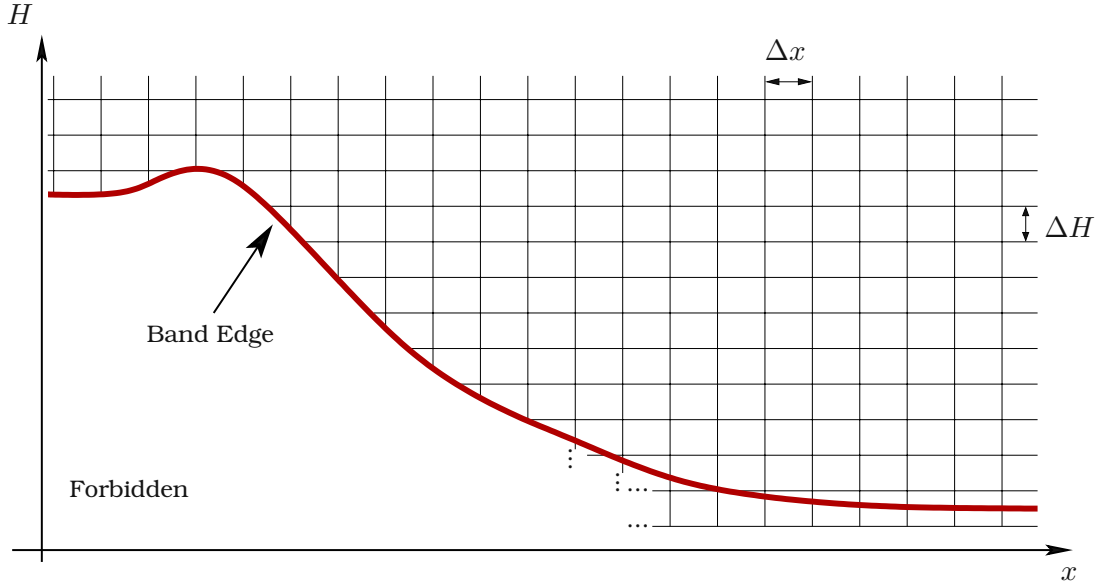


Figure 3.1: Sketch of the H-grid after an H-transform. The grid starts (black lines) at the band edge (thick red line) and is uniformly spaced.

The electrostatic potential $\varphi(\mathbf{x}, t)$ used in the transformation is obtained from a solution of Poisson's equation (cf. Figure 3.1). In a SHE of the BTE, the H-transform is used to eliminate the derivative with respect to energy in Equation (3.12), rendering the set of equations numerically stable. Although the H-transform simplifies the free streaming operator, it unfortunately results in a potential-dependent energy grid, which needs to be recalculated during each iteration of the self-consistent solution process. Thus, with the H-transform and Equation (3.2) the full Boltzmann Transport Equation by collecting each term from Equation (3.10) to Equation (3.13) for either electrons or holes at (\mathbf{x}, H, t) reads

$$\begin{aligned} & \sum_{l',m'} \left(Z_{l',m'}^{l',m'} \partial_t f_{l,m} + q Z_{l',m'}^{l',m'} \frac{\partial \varphi}{\partial t} \frac{\partial f_{l,m}}{\partial H} + \nabla_{\mathbf{x}} \mathbf{j}_{l,m}^{l',m'} f_{l,m} - \mathbf{F} \cdot \mathbf{A}_{l,m}^{l',m'} f_{l,m} \right) \\ & = \frac{1}{Y_{0,0}} \sum_{l',m'} \left(Z_{l',m'}^{l',m'}(H) \sigma(H', H) Z_{l,m}^{0,0} f_{l,m} - Z_{0,0}(H') \sigma(H, H') Z_{l,m}^{l',m'} f_{l,m} \right), \end{aligned} \quad (3.25)$$

where the shorthands

$$Z_{l,m}^{l',m'}(\epsilon) = \oint Y_{l',m'} Y_{l,m} Z(\epsilon) d\Omega, \quad (3.26)$$

$$\mathbf{j}_{l,m}^{l',m'}(\epsilon) = \oint Y_{l',m'} v_g Y_{l,m} Z(\epsilon) d\Omega, \quad (3.27)$$

$$\mathbf{A}_{l,m}^{l',m'}(\epsilon) = \oint \frac{1}{\hbar \|\mathbf{k}\|} \left(\frac{\partial Y_{l',m'}}{\partial \theta} \mathbf{e}_\theta + \frac{\partial Y_{l',m'}}{\sin(\theta) \partial \varphi} \mathbf{e}_\varphi \right) Y_{l,m} Z(\epsilon) d\Omega, \quad (3.28)$$

from [47, 67] have been used.

3.3 Discretization

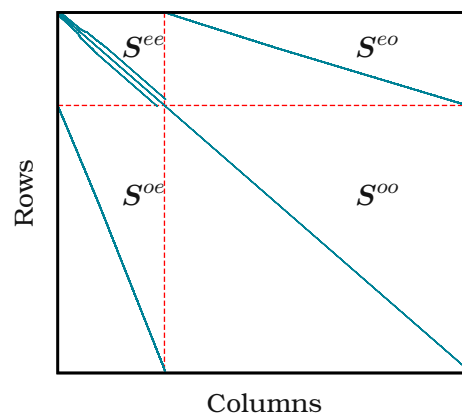
To numerically solve the set of equations presented in Section 3.1 they need to be discretized. In the course of this work the finite volume method (FVM) [35] has been used for the space

spanned by the spatial coordinates and the total energy. Although it is possible to discretize the energy space by non-equidistant grids, equidistant grids are used here for simplicity. Upon discretization the expansion coefficients (cf. Equation (3.2)) are usually split into an even and odd part

$$f_{l,m}(\mathbf{x}, H, t) = \begin{cases} f_{l,m}^e(\mathbf{x}, H, t), & \text{if } l \text{ is even,} \\ f_{l,m}^o(\mathbf{x}, H, t), & \text{if } l \text{ is odd,} \end{cases} \quad (3.29)$$

where the even part $f_{l,m}^e$ yields densities, such as the electron concentration, and the odd part $f_{l,m}^o$ yields fluxes, such as the current density (cf. Section 2.3). This is necessary since the continuity equations, which involve the odd parts of the distribution function, need proper stabilization [70].

Figure 3.2: The sparsity pattern of the system matrix, where dashed red lines have been added to visualize the submatrices. The sparsity pattern of the system matrix with all terms excluding the time derivative assembled is shown. As can be seen, the submatrix S^{oo} has no off-diagonals, since no odd unknown couples with another odd unknown. This allows to eliminate the odd-order unknowns from the equation system as demonstrated in [71].



When splitting the BTE (cf. Equation (3.25)) into equations for odd and even unknowns an interesting pattern in the system matrix of the SHE equations emerges (cf. Figure 3.2). The system matrix for the BTE can be split into four submatrices, depending on the coupling they describe. The matrix which contains couplings between even unknowns is S^{ee} . The submatrix that couples even with odd unknowns is S^{eo} and so forth. It is interesting to note that, since elastic and inelastic scattering operators only couple even unknowns with even unknowns, the submatrix S^{oo} , which couples odd unknowns with odd unknowns, is a diagonal matrix. This allows for the use of the Schur-complement to compress the matrix to the size of the submatrix S^{ee} [71]. This dramatically reduces the number of unknowns to be solved for to the size of S^{ee} , which makes the system of equations much smaller and thus easier to solve on a common workstation. The structure of the system matrix is shown in Figure 3.3 and clearly visualizes the dependence of the SHE equations via the H-grid on the electrostatic potential. An exhaustive explanation of the discretized set of equations can be found in [67].

3.4 Full-band Effects

In Monte Carlo simulators full-band effects can be integrated without further approximations in a straight-forward fashion [25]. For a SHE of the BTE this is not possible, since the generalized density of states is only well-defined if there is a bijective mapping between k and energy per band (cf. Equation (3.6)). For silicon there is no straight-forward bijective mapping between energy and wave vector (cf. Figure 2.1) and thus approximations need to be applied. The simplest approximation for silicon is the parabolic band approximation or the Modena model

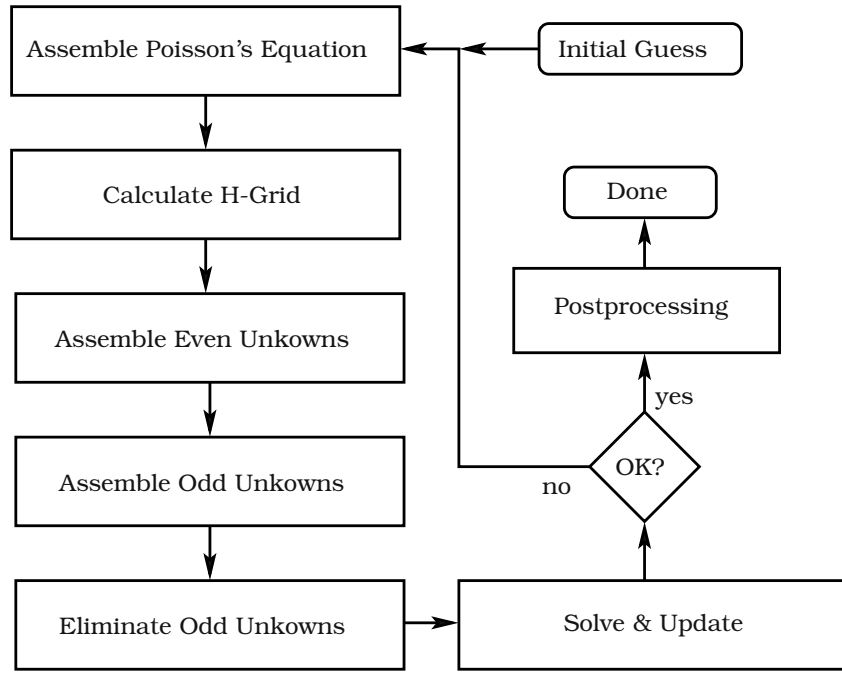


Figure 3.3: A flowchart of the SHE method. In every iteration Poissons equation is assembled first. After this the H-grid is built for the assembly of the equations for even and odd unkowns. Before solving the system of equations and reevaluating the potential and carrier concentrations, the size of the system matrix is reduced by elimination of odd-order unkowns [71]. This is repeated until changes in the potential are smaller than a given value.

together with the Herring-Vogt transform, both of which are bijective. Aside from these two approximations, other ways to include full-band effects have been investigated and are summarized in [47]. In this work the anisotropic band model and the extended Vecchi model will be explained briefly. Both band models deliver the same amount of accuracy, whilst the extended Vecchi model requires less computational resources. A third technique to cover full-band effects is described in [47]. In this model the first conduction band an anisotropic model is used, whilst for higher bands an isotropic model is employed. During the derivation of said model, the assumption of isotropic vallies for higher bands, generally throughout the derivation of the extended Vecchi model, is delayed to the last moment of the derivation. However, this third technique is more complex then the extended Vecchi model and will not be further discussed in this work in favour of the simpler extended Vecchi model.

3.4.1 The Anisotropic Band Model

The anisotropic band model, developed by [72], was an attempt to increase the accuracy of the Modena model by taking the anisotropic nature of the bandstructure into account. The main idea of the anisotropic band model is to expand the inverse dispersion relation $k(\epsilon, \theta, \varphi)$ for all bands per valley n into spherical harmonics,

$$k_{\text{fit}}^n(\epsilon, \theta, \varphi) = \sum_{l=0}^L \sum_{m=-l}^l k_{l,m}^n Y_{l,m}(\theta, \varphi), \quad (3.30)$$

where the expansion coefficients $k_{l,m}^n$ are obtained by Least-Squares fitting [73, 74] to the real bandstructure. This expansion can then be inserted into Equation (3.6) to obtain the anisotropic density of states for the SHE equations. Although being numerically quite challenging, the model only delivers moderate accuracy as shown in [47] and was thus not used in the course of this thesis.

3.4.2 The extended Vecchi Model

Building upon the Herring-Vogt transformed elliptic dispersion, Vecchi et al. [58] modified this model by using the full density of states and group velocity (cf. Figure 2.2), but still assuming an isotropic and parabolic dispersion relation. This model has been subsequently extended by Jin et al. [48], such that the resulting quantities, e.g. drift velocity, match better with the results from a full-band Monte Carlo simulator than the original band model [58]. The derivation of the extended Vecchi model starts from the Herring-Vogt transformed density of states and group velocity

$$g_\nu(E) \quad \text{and} \quad v'_g(E) = \sqrt{\frac{m_c^*}{m_d^*}} v_g^\nu, \quad (3.31)$$

where m_c^* is the conductivity effective mass and m_d^* is the DOS effective mass. Now the generalized density of states Z is approximated such that $\partial||\mathbf{k}||/\partial\epsilon$ disappears

$$Z = \frac{2||\mathbf{k}||^2}{(2\pi)^3} \frac{\partial||\mathbf{k}||}{\partial\epsilon} \approx \frac{\partial(g_\nu(E)v'_g(E))}{\partial\epsilon}. \quad (3.32)$$

This assumption is only justified by numerical results and shows the same quality as the anisotropic band model, but requires significantly less computational time. Due to its minimal runtime penalty, this full-band model for SHE has been employed in the course of this thesis.

3.5 Recombination and Generation

In Section 2.2.2 recombination/generation has been discussed for a single two-state defect located energetically within the bandgap, better known as the SRH model [31]. This model has been incorporated into the bipolar spherical harmonics expansion of the BTE [62]. Thus the electron and hole BTE are now coupled by a recombination term Γ and the solution $f_T^\nu(t)$ of Equation (2.32) per trap level E_T . The recombination operators $\Gamma^{p/n}\{f_n, f_p\}$ for electrons and holes are

$$\Gamma^n\{f_n, f_p\} = \frac{N_T}{(2\pi)^3} \int_{\mathcal{B}} G_\nu^n(\mathbf{k}) f_T^\nu(t) (1 - f_n) - R^n(\mathbf{k}) (1 - f_T^\nu(t)) f_n d^3k, \quad (3.33)$$

$$\Gamma^p\{f_n, f_p\} = \frac{N_T}{(2\pi)^3} \int_{\mathcal{B}} G_\nu^p(\mathbf{k}) (1 - f_T^\nu(t)) (1 - f_p) - R^p(\mathbf{k}) f_T^\nu(t) f_p d^3k, \quad (3.34)$$

where N_T is the trap concentration. The expansion into spherical harmonics using Equation (3.3), Equation (3.4) as well as the H-transform gives

$$\Gamma^n\{f_n, f_p\}_{l,m} = N_T Z^n(H) \left(G^n(H) f_T (1 - f_{l,m}^n) - R^n(H) \left(\frac{1}{Y_{0,0}} - f_T \right) f_{l,m}^n \right) \delta_{l,0} \delta_{l,m}, \quad (3.35)$$

$$\Gamma^p\{f_n, f_p\}_{l,m} = N_T Z^p(H) \left(G^p(H) \left(\frac{1}{Y_{0,0}} - f_T \right) (1 - f_{l,m}^p) - R^p(H) f_T f_{l,m}^p \right) \delta_{l,0} \delta_{l,m}. \quad (3.36)$$

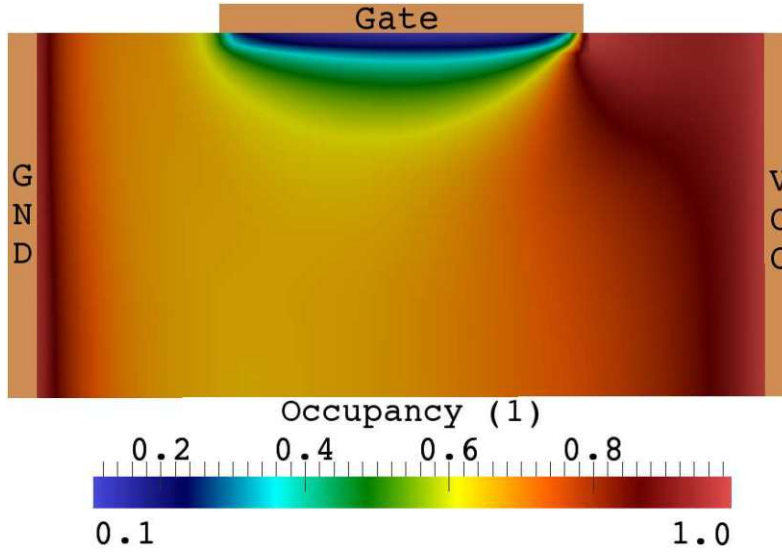


Figure 3.4: Trap occupancy in a NPN-transistor, where the gate potential is 0.7 V and $V_{cc} = 0.1$ V. Note that the base of the transistor has been named ‘Gate’.

To provide an example of the implementation of the above equations, a 2D NPN-transistor in steady state has been simulated with a trap density of 10^{19} cm^{-3} and a single trap level at 0.1 eV away from the center of the bandgap. The domain has been discretized using an orthonormal grid with 3600 vertices. The doping was set to $N_D = 10^{16} \text{ cm}^{-3}$ in both n-doped regions and $N_A = 10^{16} \text{ cm}^{-3}$ in the p-doped center. The gate potential was set to 0.7 V and V_{cc} to 0.1 V. The resulting trap occupancy is shown in Figure 3.4.

3.6 Time-dependent SHE of the BTE

First results for a time-dependent SHE using the H-transform have been reported in [75]. However, the additional derivative,

$$\begin{aligned} \partial_t f(\mathbf{x}, E, t) &= \frac{\partial f(\mathbf{x}, H, t)}{\partial t} + \frac{\partial f(\mathbf{x}, H, t)}{\partial H} \frac{\partial H}{\partial t} \\ &= \frac{\partial f(\mathbf{x}, H, t)}{\partial t} \pm |q| \frac{\partial \varphi}{\partial t} \frac{\partial f(\mathbf{x}, H, t)}{\partial H}, \end{aligned} \quad (3.37)$$

resulting from the H-transform was not considered in [75] and first reported in [49]. As stated, in a SHE of the BTE the H-transform is used to eliminate the derivative with respect to energy in the free streaming operator.

When applying the finite volume method in the energy space, the additional coupling terms read

$$\begin{aligned} &\pm \int_{H_n^-}^{H_n^+} |q| \frac{\partial \varphi}{\partial t} \frac{\partial f_i(H_n, t_{k+1}) Z_i(H_n)}{\partial H} dH \\ &\Rightarrow \pm |q| \frac{\Delta \varphi_i}{\Delta t} \int_{H_n^-}^{H_n^+} \frac{\partial f_i(H_n, t_{k+1}) Z_i(H_n)}{\partial H} dH \\ &= \pm |q| \frac{\Delta \varphi_i}{\Delta t} \frac{f_i(H_{n+1}, t_{k+1}) Z_i(H_{n+1}) - f_i(H_{n-1}, t_{k+1}) Z_i(H_{n-1})}{2}, \end{aligned} \quad (3.38)$$

where the subscript i denotes the i th spatial grid point and n is the n th energy grid point (cf. Figure 3.7). The energy space is discretized using equidistant sampling points, that is $H_n - H_{n-1} = \Delta H$, where ΔH is the distance between two adjacent energy grid points. For integration H_n^+ and H_n^- are used and defined as

$$H_n^+ = (H_n + H_{n+1})/2, \quad (3.39)$$

$$H_n^- = (H_n + H_{n-1})/2. \quad (3.40)$$

Although it is possible to discretize the energy space by non-equidistant grids here, we use equidistant grids for simplicity. The additional term on the right hand side of Equation (3.38) couples neighboring energies for even and odd-order unknowns in the system matrix S

$$\underbrace{\begin{pmatrix} S^{ee} & S^{eo} \\ S^{oe} & S^{oo} \end{pmatrix}}_S \begin{pmatrix} f^e \\ f^o \end{pmatrix} = \begin{pmatrix} b^e \\ b^o \end{pmatrix}, \quad (3.41)$$

where $f^{e/o}$ are the even and odd unknowns respectively, $b^{e/o}$ is the right hand side for even and odd unknowns, S^{ee} is the upper left sub-matrix coupling even unknowns to even unknowns, and S^{oo} is the lower right sub-matrix coupling odd-order unknowns with odd-order unknowns, and so forth. In a SHE of the stationary BTE the sub-matrix S^{oo} is a diagonal matrix. This makes it possible to reduce the number of unknowns considerably using the Schur-complement $(S^{ee} - S^{eo}(S^{oo})^{-1}S^{oe})x^e = b^e - S^{eo}(S^{oo})^{-1}b^o$ as shown in [71]. When assembling the time derivative the coupling of energies introduced by Equation (3.38) for even-order unknowns appearing in S^{ee} is similar to any coupling an elastic scattering operator would introduce. But the coupling of the energies for the odd-order unknowns from Equation (3.38), showing up in S^{oo} , destroys the diagonal sub-matrix structure of S^{oo} . This in turn increases the effort to eliminate the odd-order unknowns [71], since one has to carry out multiple line operations to restore the diagonal structure of S^{oo} in order to reduce the number of unknowns as it is done in a SHE for the time independent BTE (cf. Figure 3.5). To account for the time derivative of the expansion coefficients in the expanded BTE, we first assemble the stationary BTE and include Equation (3.38). Thus, the additional derivative for the total energy H in Equation (3.37) is accounted for. Now, assuming a known solution f_k of the SHE equations at time t_k , the fully assembled SHE equations for the next time $t_{k+1} = t_k + \Delta t$ without $\partial_t f^{n/p}$ are given as

$$S_{k+1}^j f_{k+1}^j = b_{k+1}^j, \quad (3.42)$$

where j denotes the j th iteration of the non-linear solver (e.g. Newton-Raphson). In order to assemble the time derivative for t_{k+1} , we use the implicit Euler scheme, although other more complex higher-order methods are available [76]. After a few algebraic transformations this yields

$$\Delta t \left(S_{k+1}^j + \frac{I}{\Delta t} \right) = \Delta t b_{k+1}^j + f_k, \quad (3.43)$$

where I is the unit matrix. Thus, one can assemble the system of SHE-BTE equations in each iteration of any non-linear solver as before and then include the time derivative by simple algebraic manipulations of the assembled stationary system of equations.

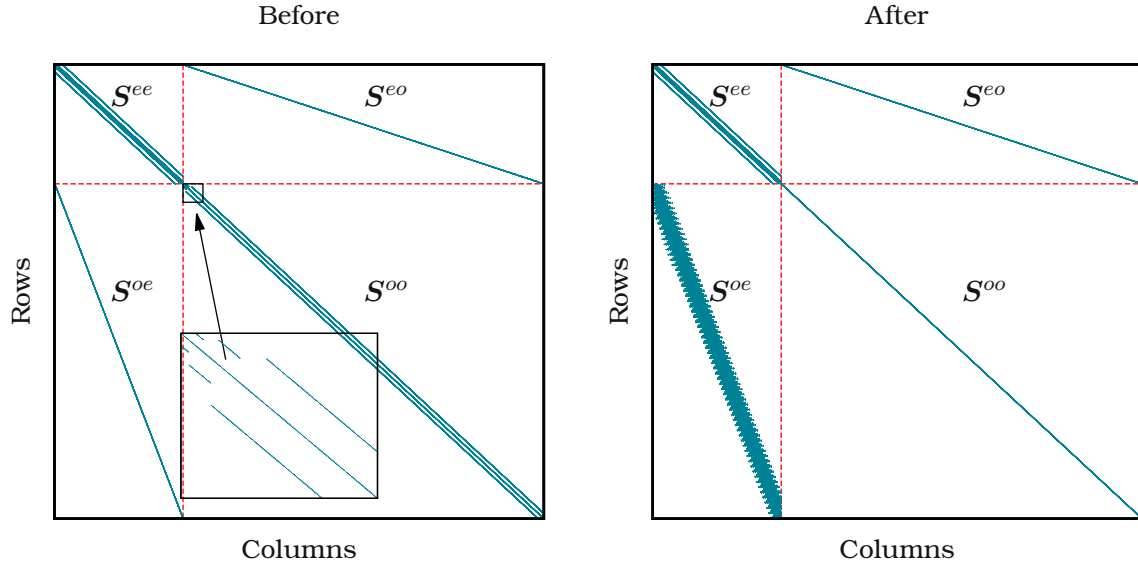


Figure 3.5: The sparsity pattern of the system matrix before and after diagonalizing S^{oo} , where dashed red lines have been added to visualize the submatrices. On the left hand side the sparsity pattern of the system matrix with all terms (left figure), including the time derivative, assembled is shown. As can be seen the submatrix S^{oo} has non-zero offdiagonals, which prohibit the elimination of the odd-order unknowns. Through simple line operations the system matrix can be transformed such that S^{oo} is a diagonal matrix again. This allows to eliminate the odd-order unknowns from the equation system as demonstrated in [71].

3.6.1 Comparison to Drift Diffusion

A comparison of the time derivative in the BTE and in the drift diffusion equations can be made by reproducing and exploring the origins of plasma oscillations. In this section the mathematical coherences between plasma oscillations as described by the BTE and by the drift diffusion model are compared.

Oscillations of the whole electron plasma in a semiconductor occur when electrons are shifted out of their equilibrium position around the fixed ions. Upon relaxation back into their respective equilibrium positions, the electrons oscillate around their equilibrium position. The characteristic frequency of plasma oscillations in a semiconductor is

$$\omega_p = \sqrt{\frac{nq^2}{\kappa m_e}},$$

where n is the carrier concentration, m_e is the electron mass and in this section here κ is used to denote the dielectric constant. For silicon ω_p is usually on the order of 1 THz. We expect to see a peak in the real part of the admittance of a semiconductor device at the plasma frequency [36]. In the following we will use this fact to asses various assumptions for simulations using SHE and a stability analysis and show why the plasma frequency is important to test these assumptions. As stated, upon solving the time-dependent BTE using SHE and the H-transform one has to include an additional term in the time derivative (cf. Equation (3.37)) due to the H-transform. This term couples neighboring energies for even and odd-order unknowns. In

particular, the coupling of the energies for the odd-order unknowns destroys the diagonal sub-matrix structure. One possible way to solve this is to diagonalize the sub-matrix for odd-order unknowns by a suitable algorithm before applying the algorithm from [71]. Nevertheless, this transformation renders a required stability analysis intractable. Another possibility to solve this problem is to assume that the potential change over time (cf. Equation (3.37)) as well as the gradient of the distribution function over energy are sufficiently small, effectively eliminating the couplings of the energies for the even and odd-order unknowns. As will be shown, a physically sound way to solve the problem of additional couplings for odd-order unknowns is to assume that the time derivative in the equations for the odd-order unknowns is sufficiently small. More precisely,

$$\Delta t^{-1} (f_{i,k+1}^o - f_{i,k}^o) \pm \frac{|q| \Delta \varphi_i}{2\Delta H \Delta t} (f_{i,k+1}^o(H_{n+1}) - f_{i,k+1}^o(H_{n-1})) \simeq 0.$$

In order to assess the physical meaning of the above assumption we take a look at the drift diffusion model and show that neglecting the time derivative for odd-order unknowns is equivalent to neglecting the acceleration term

$$\tau_m \partial_t \mathbf{J} \quad (3.44)$$

in the drift diffusion model, where \mathbf{J} is the current density and τ_m is the moment relaxation time. This assumption is equal to neglecting plasma oscillations and was shown to be valid for frequencies up to the plasma frequency [36]. For the drift diffusion model only the first two moments of the BTE are used, where closure is obtained by assuming equivalence of carrier temperature T_n and lattice temperature T_L . The moments of the distribution function are calculated using

$$\langle \psi_j \rangle = \int_{\mathcal{B}} \psi_j f d^3 k = w_j \int_{\mathcal{B}} \mathbf{k}^j f d^3 k, \quad (3.45)$$

where j is the j th moment, ψ_j denotes the j th weight function and w_j is the j th prefactor. To obtain equations for the moments of the BTE, such as the drift diffusion model, the BTE is multiplied by increasing integer exponent of the wave vector \mathbf{k} and a scalar prefactor w and afterwards integrated over the Brillouin zone. The first two moments of the BTE read

$$\begin{aligned} \psi_0 = 1 &\Rightarrow \langle \psi_0 \rangle = \int_{\mathcal{B}} f d^3 k = n \\ \psi_1 = \hbar \mathbf{k} &\Rightarrow \langle \psi_1 \rangle = \hbar \int_{\mathcal{B}} \mathbf{k} f d^3 k \\ &= \hbar \left(\underbrace{\int_{\mathcal{B}} \mathbf{k} f^e d^3 k}_{=0} + \int_{\mathcal{B}} \mathbf{k} f^o d^3 k \right) = \mathbf{J}/q, \end{aligned}$$

where n is the electron density. The integral over the even part cancels out due to the symmetry of the first Brillouin zone. Thus, the current density is associated with the odd part or asymmetry of the distribution function. Applying the method of moments, assuming parabolic bands, and approximating the scattering operator $Q\{f\}$ using the relaxation time approximation (RTA), one obtains the drift diffusion model [22]. Repeated from chapter 2 Section 2.3.1,

the drift diffusion model with acceleration term, excluding Poisson's equation, reads

$$\psi_0 \Rightarrow \partial_t n - |q|^{-1} \nabla \cdot \mathbf{J}_n = -R, \quad (3.46)$$

$$\psi_1 \Rightarrow \tau_m \partial_t \mathbf{J}_n - |q| \tau_m k_B \nabla \cdot (n T_L) + m^* \mathbf{J}_n = 0, \quad (3.47)$$

$$\text{with } \partial_t \mathbf{J}_n = \int_{\mathcal{B}} \mathbf{k} \partial_t f d^3 k = \partial_t \int_{\mathcal{B}} \mathbf{k} f^o d^3 k \quad (3.48)$$

with the mobility $\mu_n = |q| \tau_m / m^*$, R the scalar recombination term from $R^n\{f^n, f^p\}$, m^* is the effective mass and k_B the Boltzmann constant. As derived above the current density is associated with the odd part of the distribution function and thus also with the odd expansion orders in a spherical harmonics expansion. Thus neglecting the acceleration term in the DD model is equivalent to neglecting the time derivative of the odd part of the distribution function.

3.6.2 Stability

Even if the time derivative has been correctly implemented, numerical stability issues might still arise and lead to numeric artifacts in the simulation results [77]. At least a guideline, for how small Δx for spatial, ΔH for energy space and Δt for temporal discretization have to be in order to avoid any numeric artifacts, is required. To investigate the numerical stability of the time-dependent SHE-BTE a von Neumann analysis for electrons is carried out. The analysis for holes is done in the same way, yielding the same results and thus not shown here. Since the full numeric system is too complex and analytically intractable, we need to simplify matters. Thus, in the following we neglect plasma oscillations in the calculations to avoid the time derivative of the odd unknowns in the SHE. To keep the equations in the stability analysis as simple as possible, the finite volume method (FVM) in the 1D case under bulk conditions for silicon (cf. Figure 3.7) is used. Thus the force \mathbf{F} , the generalized density of states Z as well as the group velocity $\mathbf{v}(\mathbf{k})$ are assumed to be spatially constant. To avoid notational clutter we use a superscript (o) to mark odd-order unknowns, indices and variables. The same is done for even-order unknowns (superscript e). We also drop the indices l, m, l', m' including the summation over these indices as well as the arguments to the distribution function. Additionally, we consider only elastic, velocity randomizing scattering processes and neglect the Pauli principle to further simplify the analysis. To summarize, we condense the scattering operator to

$$\frac{1}{Y_{0,0}} \left(\underbrace{s(\mathbf{x}, E_i, E_f) Z(E_f)_{l,m} \delta_{l',0} \delta_{m',0}}_{\text{in-scattering}} - \underbrace{s(\mathbf{x}, E_f, E_i) Z(E_i)_{0,0} \delta_{l',l} \delta_{m',m}}_{\text{out-scattering}} \right) = \begin{cases} c^e, & \text{for even-orders} \\ c^o, & \text{for odd orders,} \end{cases}$$

where $Y_{0,0}$ is the first spherical harmonic, s is the scattering rate. Inelastic processes are not considered in the stability analysis. This assumption is needed to decouple the equations with respect to the total energy, which in turn leads to a tractable number of equations. Following [47, 66] the expanded and discretized BTE in a single spatial dimension (cf. Figure 3.7) for odd expansion orders reads

$$\underbrace{\partial_t f^o}_{\simeq 0} + \frac{A^{e,+}}{\Delta x} (f_{i+1,k+1}^e - f_{i,k+1}^e) + \frac{B^{e,+}}{2} (f_{i,k+1}^e + f_{i+1,k+1}^e) = c^o f_{i+1/2,k+1}^o,$$

$$\underbrace{\partial_t f^o}_{\simeq 0} + \frac{A^{e,-}}{\Delta x} (f_{i,k+1}^e - f_{i-1,k+1}^e) + \frac{B^{e,+}}{2} (f_{i,k+1}^e + f_{i-1,k+1}^e) = c^o f_{i-1/2,k+1}^o,$$

where $f_{i,k}^e$ is the even unknown expansion coefficient for l, m at vertex i in the k th time step and $f_{i,k}^o$ is the same for odd unknowns for l', m' . Note that in the equations above we have neglected the acceleration term by setting the time derivative for odd unknowns to zero. For even-orders the expanded BTE in 1D reads

$$\Delta t^{-1} (f_{i,k+1}^e - f_{i,k}^e) - \underbrace{\frac{1}{2\Delta H} |q| \Delta \varphi_i}_{=\gamma} \frac{1}{\Delta t} \underbrace{(f_{i,k+1}^e(H_{n+1}) - f_{i,k+1}^e(H_{n-1}))}_{\simeq f_{i,k+1}^e x} + \frac{A^{e,+}}{\Delta x} f_{i+1/2,k+1}^o - \frac{A^{e,-}}{\Delta x} f_{i-1/2,k+1}^o - \frac{B^{e,+}}{2} f_{i-1/2,k+1}^o - \frac{B^{e,-}}{2} f_{i+1/2,k+1}^o = c^e f_{i,k+1}^e.$$

Here it was assumed that $f_{i,k+1}^e(H_{n\pm 1})$ can be written as $f_{i,k+1}^e \chi^\pm$ and $\chi = \chi^+ - \chi^-$, where χ^\pm are measures of how strong the distribution function increases/decreases over energy. In the above equations the shorthands

$$A^{e/o,\pm} = \int_{H_n^-}^{H_n^+} \mathbf{j}_{l,m}^{l',m'}(x_{i\pm 1/2}, H_n) dH, \quad (3.49)$$

$$B^{e/o,\pm} = \int_{H_n^-}^{H_n^+} \mathbf{A}_{l,m}^{l',m'}(x_{i\pm 1/2}, H_n) dH, \quad (3.50)$$

have been used for even and odd-orders, where \mathbf{A} and \mathbf{j} are the shorthands from Equation (3.26) for integrals over spherical harmonics. In this discretization $A^{e/o,\pm}$ accounts for the projected diffusion term in the free streaming operator of the BTE, whereas $B^{e/o,\pm}$ is the projected drift term and thus dependent on the driving force (electric field). Since we are assuming a homogeneous material and a constant driving force, we have $A^{e,+} = A^{e,-}$, $B^{e,+} = B^{e,-}$, $A^{o,+} = A^{o,-}$ and $B^{o,+} = B^{o,-}$. With the simplified system given above, a von Neumann stability analysis is possible. For this we

- (i) eliminate the odd unknowns in the equation for the even unknowns,
- (ii) Fourier transform the obtained equation in space to transform spatial offsets into phase factors and
- (iii) express the gain G as a function of Δx and Δt .

The Fourier transforms, where j is the imaginary unit, of $f_{i,k}^e$ and $f_{i,k}^o$ read $G^k F^e \exp(-ji\theta)$ and $G^k F^o \exp(-ji\theta)$ respectively, where G is the gain. Thus one obtains, after a few algebraic transformations, the gain

$$|G|^2 = \frac{s - (a^o b^e + a^e b^o) \cos(\theta) + (a^e b^o - a^o b^e) \sin(\theta)}{(y - z)^2}, \quad (3.51)$$

where

$$\begin{aligned} s &= c^o c^o \Delta t^{-2} (a^e a^o + b^e b^o + c^o (c^e - \Delta t^{-1} + \gamma \chi \Delta t^{-1})) \\ y &= a^{o2} b^{e2} + a^{e2} (a^{o2} + b^{o2}) \\ &\quad + 2a^e a^o (b^e b^o + c^o (c^e - \Delta t^{-1} + \gamma \chi \Delta t^{-1})) \\ &\quad + (b^e b^o + c^o (c^e - \Delta t^{-1} + \Delta t^{-1} \gamma \chi))^2, \\ z &= 2(a^o b^e + a^e b^o) (a^e a^o + b^e b^o + c^o (c^e - \Delta t^{-1} + \gamma \chi \Delta t^{-1})) \\ &\quad \cos(\theta) + 2a^e a^o b^e b^o \cos(\theta), \end{aligned}$$

and

$$a^e = \frac{A^{e,\pm}}{\Delta x} + \frac{B^{e,\pm}}{2}, \quad a^o = \frac{A^{o,\pm}}{\Delta x} + \frac{B^{o,\pm}}{2}, \quad (3.52)$$

$$b^e = \frac{A^{e,\pm}}{\Delta x} - \frac{B^{e,\pm}}{2}, \quad b^o = \frac{A^{o,\pm}}{\Delta x} - \frac{B^{o,\pm}}{2}. \quad (3.53)$$

In order to have numeric stability of the time-discretized SHE-BTE in the von Neumann sense [78], the following relation must be fulfilled

$$|G| \leq 1 \Leftrightarrow |G|^2 \leq 1, \quad (3.54)$$

which states that any frequencies propagated by the BTE must not be amplified. For the lowest expansion order $L = 1$ the stability condition reduces to $\Delta t > 0$, since for the lowest order we have $c = 0$ and $B = B' = 0$ due to the H-transform. For $L > 1$, the stability condition reduces to

$$\Delta t \leq \frac{c^o(2 - \gamma\chi)}{a^o(a^e - a^e) + b^o(b^e - a^e) + c^e c^o}, \quad \text{if } \gamma\chi < 2, \quad (3.55)$$

$$\Delta t > 0, \quad \text{otherwise}, \quad (3.56)$$

provided that

$$c^e > 0 \text{ and } c^o < 0. \quad (3.57)$$

These conditions are naturally fulfilled due to the M-matrix property of the SHE equations for the lowest expansion order $L = 1$. In the limit of no driving force,

$$B^{e\pm} = 0 \text{ and } B^{o\pm} = 0 \text{ and } \gamma = 0, \quad (3.58)$$

or sufficiently small driving force

$$|B^{e\pm}| \ll \frac{|A^{e\pm}|}{\Delta x} \text{ and } |B^{o\pm}| \ll \frac{|A^{o\pm}|}{\Delta x}, \quad (3.59)$$

the stability condition reduces to

$$\Delta t \begin{cases} \leq (2 - \gamma\chi)/c^e, & \text{if } \gamma\chi \leq 2, \\ > 0, & \text{else,} \end{cases} \quad (3.60)$$

where c^e is usually of the order of the density of states. Even though the above relation has been derived using a number of simplifying assumptions, it can be used as a rough guideline to choose Δt .

3.6.3 Energy Grid Interpolation

Aside from the usual stability concerns regarding hyperbolic partial differential equations discretized using the finite volume method, the H-transform poses another restriction to the maximum Δt . When considering the time-dependent BTE, the old solution $f(x, H, t_k)$ from time step t_k has to be transferred from (x, H^k) to a new grid (x, H^{k+1}) (cf. Figure 3.6). Convergence problems and numerical artifacts arise whenever a value of the old distribution function on (x, H^k) cannot be transferred/interpolated to a point on the new grid (x, H^{k+1}) , because there

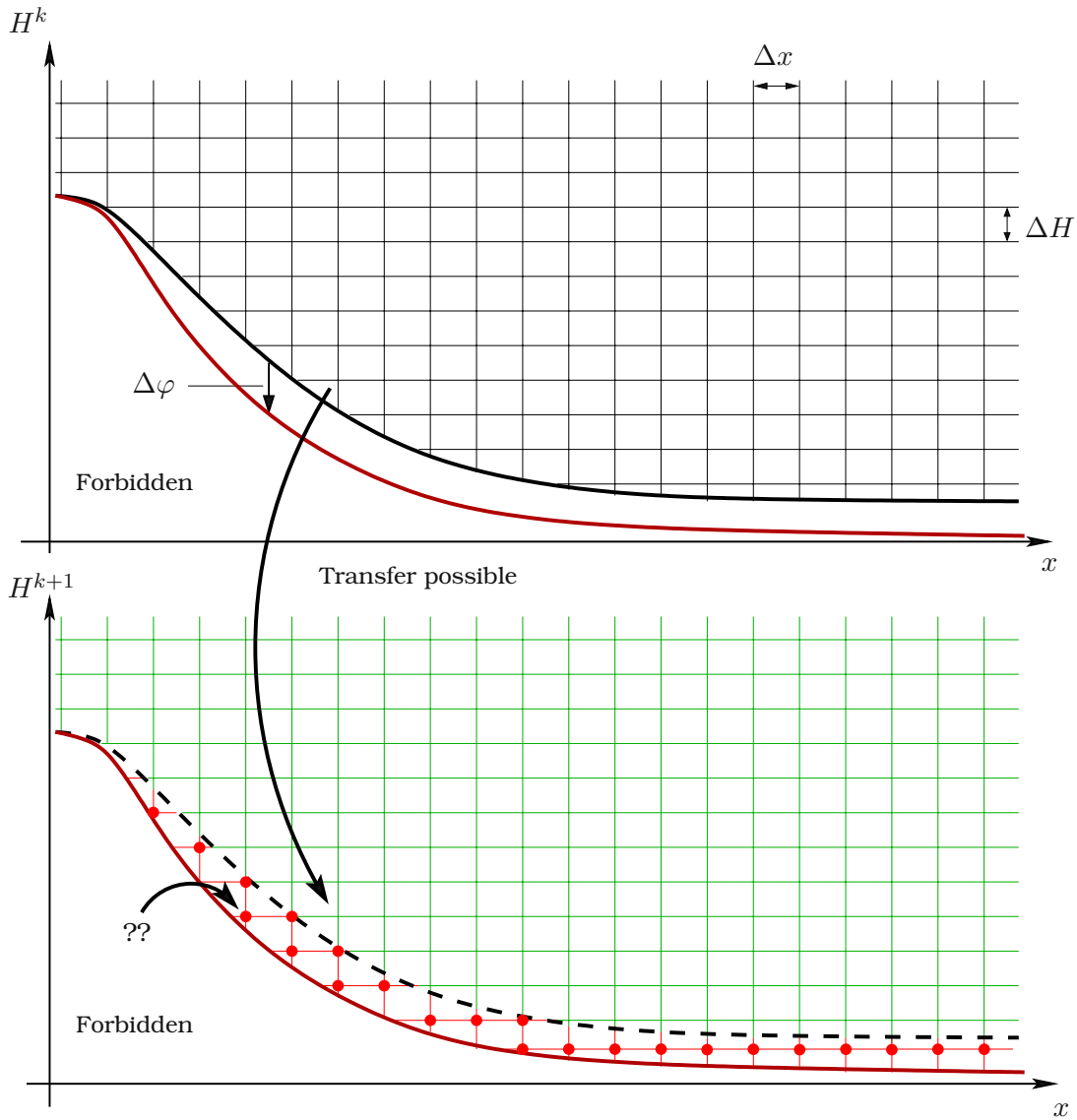


Figure 3.6: The transfer of the old distribution function at time t_k from (x, H^k) (top graph) to a new grid (x, H^{k+1}) (bottom graph) during the iterations of the non-linear solver. If the change in the electrostatic potential φ is larger than $\Delta H / |q|$ there are grid points (red dots) in the new grid (red grid lines) for which there is no value of the old distribution function at time t_k available to be transferred from the old grid (question marks).

is no value of the old distribution function available at that particular energy. This problem is always present, whenever the potential from time t_k to $t_{k+1} = t_k + \Delta t$ changes by more than

$$\frac{\Delta H}{|q|} \geq \Delta\varphi, \tag{3.61}$$

where ΔH has to be carefully chosen. One can now either choose ΔH large enough to accommodate a predicted potential change $\Delta\varphi$ within a time step and accept the inaccuracies in the distribution function introduced by this choice, or choose the boundary conditions for the Poisson equation and the subsequent time steps carefully in order not to violate Equation (3.61). This is very unsatisfying, since it is for example not possible to apply step functions

as boundary conditions, with steps larger than $\Delta\varphi$. However, in case one is interested in the small signal response, a linearization of the free streaming operator of the BTE around the bias point [47] should be used instead of the time-dependent BTE to avoid the condition in Equation (3.61).

Even if the condition in Equation (3.61) is not violated, the distribution function $f(\mathbf{x}, H^k, t_k)$ needs to be transferred to the new grid such that the macroscopic quantities, that is charge carrier concentration and charge carrier current, are not modified by the interpolation of f from (\mathbf{x}, H^k) to (\mathbf{x}, H^{k+1}) . Thus, the charge carrier density and current density are not modified by the interpolation of f from (\mathbf{x}, H^k) to (\mathbf{x}, H^{k+1}) . Since the even part of the distribution function completely defines the charge carrier concentration and the odd part defines the charge carrier current, the even part is transferred such that

$$f^e(\mathbf{x}, H^k, t_k)Z(H^k) = f^e(\mathbf{x}, H^{k+1}, t_k)Z(H^{k+1}) \quad (3.62)$$

holds for the conservation of the carrier density. Likewise the new odd part $f^o(\mathbf{x}, H^{k+1}, t_{k+1})$ is independently renormalized such that

$$f^o(\mathbf{x}, H^k, t_k)Z(H^k)v_g(H^k) = f^o(\mathbf{x}, H^{k+1}, t_k)Z(H^{k+1})v_g(H^{k+1}) \quad (3.63)$$

is fulfilled, in order to keep the current density constant. However, interpolation errors will occur during the transfer of the old solution onto the new H-grid, even if the condition in Equation (3.61) is fulfilled. Since the band edge is shifted by the potential during a single timestep, the first energy grid point ($H^k = 0$) closest to the band edge will be shifted under the new band edge or a grid point below the old band edge ($H^k = -1$) will be shifted above the new band edge. In the first case a sample point for the distribution function is lost. In the second case a new energy grid point for which there is no old distribution function available will be obtained. This effect cannot be mitigated when the H-transform is employed. The error introduced by this is quantified by

$$\frac{\hat{f}^{e/o}(\mathbf{x}, H^{k+1}, t_k)Z(H^{k+1})}{\Delta t}. \quad (3.64)$$

3.6.4 Probable Violation of Gauss' Law

As stated in the previous section, it is currently not possible to avoid errors during energy grid interpolation. To better understand the implications of this, the discretized SHE-BTE equations per energy grid point, for a three-point stencil are investigated. In Figure 3.7 a homogeneous, uniformly discretized three-point stencil for a single spatial dimension is shown. The system of equations for this simple three point stencil will be derived and evaluated in the following. Assuming that the even unknowns on the left and right point of this three point stencil are fixed by dirichlet boundary conditions, only the even expansion coefficient $f_{l,m}^e$ in the center and the two odd unknowns on the two edges $f_{l,m}^o$ need to be determined. To simplify matters, a first-order expansion ($L = 1$) and elastic scattering is assumed. Thus we are left with three unknowns, the even unknown $f_{0,0}^i$, the odd unknown on the left edge $f_{1,0}^{i-1/2}$ and the odd unknown $f_{1,0}^{i+1/2}$ on the right edge. For these unknowns the equations of any point on the H-grid read

$$VZ^i\partial_t f_{0,0}^i - \frac{A\Delta x}{2}v^{i-1/2}Z^{i-1/2}f_{1,0}^{i-1/2} + \frac{A\Delta x}{2}v^{i+1/2}Z^{i+1/2}f_{1,0}^{i+1/2} = 0, \quad (3.65)$$

$$VZ^{i-1/2}\partial_t f_{1,0}^{i-1/2} - \frac{A\Delta x}{2}v^{i-1/2}Z^{i-1/2}f_{0,0}^i + \frac{A\Delta x}{2}v^{i-1/2}Z^{i-1/2}f_{0,0}^i = S f_{1,0}^{i-1/2}V, \quad (3.66)$$

$$VZ^{i+1/2}\partial_t f_{1,0}^{i+1/2} - \frac{A\Delta x}{2}v^{i+1/2}Z^{i+1/2}f_{0,0}^i + \frac{A\Delta x}{2}v^{i+1/2}Z^{i+1/2}f_{0,0}^i = S f_{1,0}^{i+1/2}V, \quad (3.67)$$

where V is the box volume, Z is the generalized density of states, v is the group velocity, S is the elastic scattering term and the interface area A between two points connected by an edge is assumed to be equal to unity. Introducing $a = (A\Delta x)/(2V)$, neglecting the time derivatives of the odd-order unknowns for the sake of argument and using a Backward-Euler scheme one obtains

$$\begin{aligned}
 & \frac{Z^i f_{0,0}^i - Z^{i,\text{old}} f_{0,0}^{i,\text{old}}}{\Delta t} - \underbrace{\frac{Z^{i,\text{old}} \hat{f}_{0,0}^{i,\text{old}}}{\Delta t}}_{\text{Energy grid interpolation error}} \\
 & + q \frac{\Delta\varphi_i}{\Delta t} \frac{f_i(H_{n+1}, t_{k+1})Z_i(H_{n+1}) - f_i(H_{n-1}, t_{k+1})Z_i(H_{n-1})}{2} \\
 & - \underbrace{a^2 S^{-1} v^{i-1/2} Z^{i-1/2} v^{i-1/2} Z^{i-1/2}}_A \left(f_{0,0}^i - f_{0,0}^{i-1} \right) \\
 & + \underbrace{a^2 S^{-1} v^{i+1/2} Z^{i+1/2} v^{i+1/2} Z^{i+1/2}}_B \left(f_{0,0}^{i+1} - f_{0,0}^i \right) = 0, \tag{3.68}
 \end{aligned}$$

where the quantities from the previous timestep are highlighted by the superscript old and the energy grid interpolation error has been considered. Since it will be important when considering the time derivative, it is shown how to obtain the familiar charge conservation law (Gauss Law). Integration of the above equation over energy from $H = 0$ to infinity yields term by term,

$$\partial_t n_i - (D_1 + D_2) \hat{n}_i + 0 + D_2 n_{i+1} + 2(D_1 + D_2) n_i - D_1 n_{i-1} = 0, \tag{3.69}$$

where D are transport coefficients and

$$\int_0^\infty \frac{Z^{i,\text{old}} \hat{f}_{0,0}^{i,\text{old}}}{\Delta t} dH = \hat{n}_i, \tag{3.70}$$

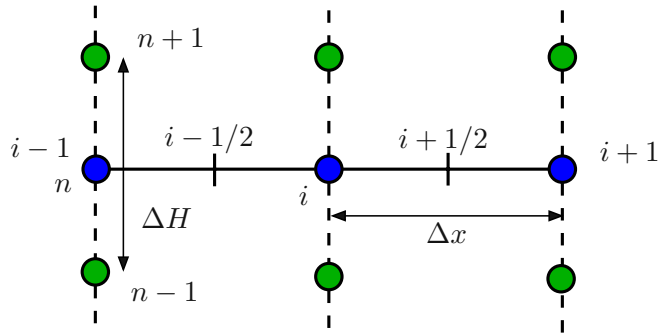
$$q \int_0^\infty \frac{\Delta\varphi_i}{\Delta t} \frac{f_i(H_{n+1}, t_{k+1})Z_i(H_{n+1}) - f_i(H_{n-1}, t_{k+1})Z_i(H_{n-1})}{2} dH = 0. \tag{3.71}$$

If there is no interpolation error at all, i.e. $\hat{n}_i = 0$, one would obtain the familiar charge conservation law

$$\partial_t n^i - D_1 n^{i-1} + (D_1 + D_2) n^i - D_2 n^{i+1} = 0, \tag{3.72}$$

instead. From this it is clear that any energy grid interpolation error leads to artifacts (\hat{n}) in the charge carrier density.

Figure 3.7: Illustration of a staggered three point stencil in a single spatial dimension. The even unknowns are assembled on the vertices (blue circles), whereas the odd unknowns are assembled on the edges. Additionally for each vertex there is an equidistant grid (green circles) for the total energy of the charge carriers over which the distribution function is resolved.



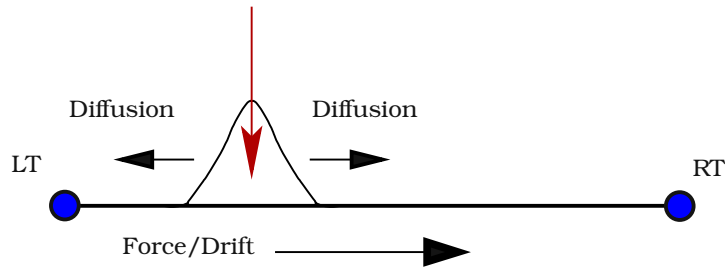


Figure 3.8: A schematic of the 1D domain used. In this experiment a Gaussian disturbance in the electron density was introduced close to the left terminal (LT) by raising the energy distribution function uniformly over energy. The carriers then drift, caused by the electric field, towards the right terminal (RT) and diffuse in all directions.

3.6.5 The Shockley-Haynes Experiment

To assess the derived results we added the time derivative of the BTE to the open source simulator ViennaSHE [65]. In the first numerical experiment, similar to the famous Shockley-Haynes experiment in [79], the drift and diffusion of minority carriers in a p-type silicon resistor $5\ \mu\text{m}$ long under carrier-phonon and impurity scattering were investigated. The Shockley-Haynes experiment was selected, since throughout the whole simulation the potential and thus the H-grid remain virtually unchanged, provided that the distortion in the carrier concentration remains small. Assuming symmetry in two axis, the simulation was carried out in a single dimension using a Δx of 10 nm. Additionally, to have the low-field conditions required for the estimation of the low-field mobility, we chose an uniform electric field of 0.5kV/cm . In this initial configuration we artificially introduced, at time zero, a Gaussian disturbance in the electron density (cf. Figure 3.8), such that changes in the electric field over time can be safely neglected. This was accomplished by uniformly raising the electron distribution function over the energy such to reach the desired electron density. Then a time-dependent simulation with a stepping of $\Delta t = 10\ \text{ps}$ was carried out. Since changes in the electric field have been kept neglected, the displacement current was neglected too. In the first few hundred picoseconds the high energy carriers diffuse strongly. Thus, at first most electrons diffuse out of the left terminal before being accelerated towards the right terminal by the small electric field (cf. Figure 3.9). To see most of the diffusion in the current, the disturbance has been placed close to the left terminal (cf. Figure 3.10). After 700 ps the carrier drift dominates and the current through the right terminal peaks. One can also calculate the low field mobility from this experiment by observing the velocity of the electron peak towards the right contact. In the presented experiment we obtained an electron mobility of $1430\ \text{cm/Vs}$ with an uncertainty of $\pm 20\ \text{cm/Vs}$ due to the discretization.

Figure 3.9: The electron current over time in a p-type resistor (cf. Figure 3.8) for the left (LT) and the right terminal (RT). In good approximation it can be said that the current at the left terminal corresponds to the diffusion and the current at the right terminal corresponds to the drift of the electrons. It is also noticeable that the electron current at the left terminal is higher for drift diffusion compared to a SHE solution of the BTE.

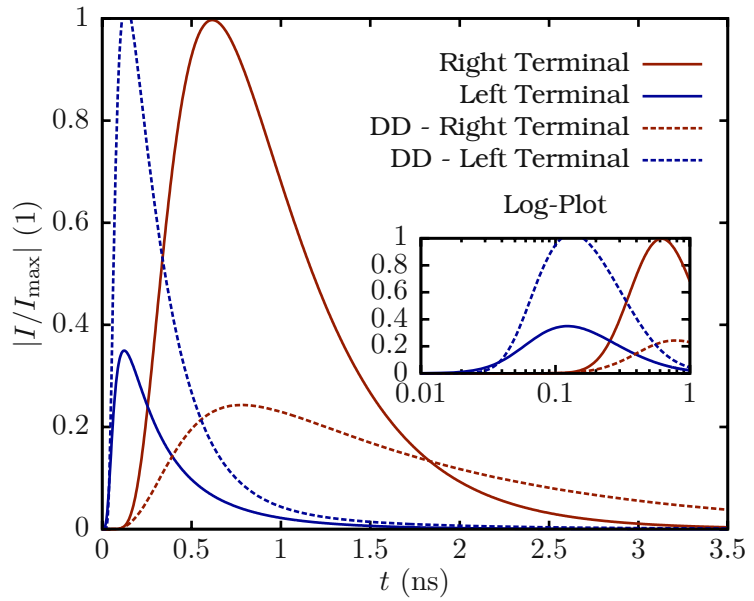
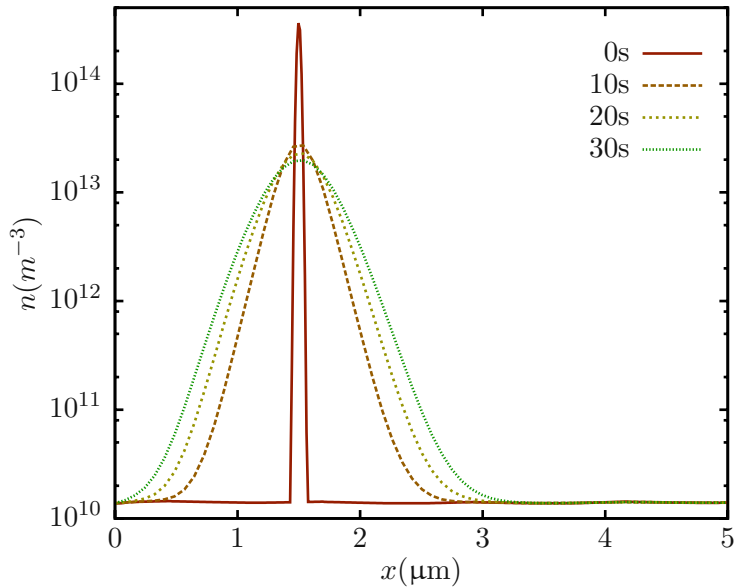


Figure 3.10: The electron concentration in a p-type resistor for various times. At time $t = 0$ s a Gaussian disturbance is introduced and observed over time. In the first few time steps ($\Delta t = 10$ ps) the carriers strongly diffuse, before gaining enough momentum towards the right terminal ($5\mu\text{m}$).



Das ist nicht nur nicht richtig, es ist
nicht einmal falsch.

(Wolfgang Pauli)

4 Quantum Correction Methods

In Section 2.6 and Section 2.7 purely quantum mechanical effects and quantum transport have been briefly discussed. During the derivation of the BTE, purely quantum mechanical effects such as confinement have been ignored. As stated in Section 2.6, a Schrödinger equation for the confined electron gas has to be solved in order to obtain the correct electron density. Solving the Schrödinger equation (cf. Equation (2.57)) self-consistently with the BTE for electrons and holes as well as Poisson's equation might be computationally too demanding. Thus, approximations are sought, which at least capture the most important characteristics of confinement and quantum transport. Since having the physically correct electron concentration at the interface of any MOS structure is of highest importance, the effect of carrier confinement should be correctly reproduced by any approximation. One of the first approximations to be used was the improved modified local density approximation (IMLDA) [7] and its predecessor MLDA [80]. Both methods are based on solutions of the Schrödinger equation for an infinitely wide square well potential, where IMLDA also takes the electric field normal to the oxide-semiconductor interface into account. Nevertheless, IMLDA can only be used to obtain a corrected electron or hole density at the interface of MOS structures, where carrier confinement occurs. As will be shown later in this work, the important case of screening of a single fixed point charge in a semiconductor or insulator needs to be correctly and generically described. The model to describe carrier confinement considered in this work is the density gradient approximation.

4.1 The Density Gradient Model

A method, other than IMLDA, to describe carrier confinement in arbitrary potential wells is called 'density gradient' model (DG) [5,81–83], which is a first-order quantum-correction model. Density gradient uses, as its name suggests, the gradient of the carrier densities to describe carrier confinement by locally modifying the electrostatic potential through a correction potential γ . Nevertheless, neither the charge carrier wave functions nor the sub band structure can be obtained by this approach, although it might be possible to assess direct tunneling using DG [83]. The equations for the correction potential are derived from Wigner's equation, where it is assumed that any effects associated with Fermi-Dirac statistics and many-body effects can be safely neglected. Additionally, the effective mass and parabolic band approximations are used and it is assumed that the described electron gas has an infinite extent. The correction potential $\gamma(\mathbf{x}, t)$ reads [82, 83]

$$\gamma(\mathbf{x}, t) = \frac{\hbar^2}{12\lambda k_B T_L m^*} \left(\nabla_{\mathbf{x}}^2 \varphi(\mathbf{x}, t) - \frac{1}{2k_B T_L} (\nabla_{\mathbf{x}} \varphi(\mathbf{x}, t))^2 \right), \quad (4.1)$$

where λ is a fitting parameter, which is determined by comparing the carrier density in a MOS structure to the carrier density obtained by the solution of Equation (2.57). Since the electric

field $\mathbf{E} = -\nabla_{\mathbf{x}}\varphi$ can be undefined at abrupt potential steps, the electrostatic potential φ is replaced, in the above equation, by $\varphi + \gamma$, yielding

$$\gamma(\mathbf{x}, t) = \frac{\hbar^2}{12\lambda k_B T_L m^*} \left(\nabla_{\mathbf{x}}^2 \varphi(\mathbf{x}, t) + \nabla_{\mathbf{x}}^2 \gamma(\mathbf{x}, t) - \frac{1}{2k_B T_L} (\nabla_{\mathbf{x}} \varphi(\mathbf{x}, t) + \nabla_{\mathbf{x}} \gamma(\mathbf{x}, t))^2 \right), \quad (4.2)$$

which is fine for this first order model, since the error in γ is of second order. This correction potential has to be calculated separately for each carrier type and added to the electrostatic potential for and only for the respective charge carrier transport equation.

4.2 The First-Order Quantum Corrected Drift Diffusion Model

The density gradient model is first applied to the drift diffusion model (cf. Section 2.3.1), where in thermal equilibrium the electron concentration is given by

$$n = N_C \exp\left(\frac{E_c - \varphi - \gamma_n}{k_B T_L}\right). \quad (4.3)$$

Inserting the above relation into Equation (4.2) yields

$$\gamma_n = -\frac{\hbar^2}{12\lambda_n k_B T_L m_n^*} \left(\nabla_{\mathbf{x}}^2 \ln(n^2) + \frac{1}{2} (\nabla_{\mathbf{x}} \ln(n))^2 \right) = -\frac{\hbar^2}{6\lambda_n k_B T_L m_n^*} \frac{\nabla_{\mathbf{x}}^2 \sqrt{n}}{\sqrt{n}}, \quad (4.4)$$

which gave the model its name ‘density gradient’.

Applying the model to the drift diffusion equations, the quantum corrected drift diffusion equations now read:

$$\nabla \cdot (\varepsilon \nabla \varphi) = |q| (n - p + C), \quad (4.5)$$

$$|q| \partial_t n - \nabla \mathbf{J}_n = |q| R, \quad (4.6)$$

$$|q| \partial_t p + \nabla \mathbf{J}_p = -|q| R, \quad (4.7)$$

$$\mathbf{J}_n = -|q| \mu_n \nabla_{\mathbf{x}} (\varphi + \gamma_n) + |q| D_n \nabla n, \quad (4.8)$$

$$\mathbf{J}_p = -|q| \mu_p \nabla_{\mathbf{x}} (\varphi + \gamma_p) - |q| D_p \nabla p, \quad (4.9)$$

$$\gamma_n = \frac{\hbar^2}{12\lambda_n k_B T_L m_n^*} \left(\nabla_{\mathbf{x}}^2 \varphi + \nabla^2 \gamma_n - \frac{1}{2k_B T_L} (\nabla_{\mathbf{x}} \varphi + \nabla_{\mathbf{x}} \gamma_n)^2 \right), \quad (4.10)$$

$$\gamma_p = \frac{\hbar^2}{12\lambda_p k_B T_L m_p^*} \left(\nabla_{\mathbf{x}}^2 \varphi + \nabla^2 \gamma_p - \frac{1}{2k_B T_L} (\nabla_{\mathbf{x}} \varphi + \nabla_{\mathbf{x}} \gamma_p)^2 \right). \quad (4.11)$$

4.3 The First-Order Quantum Corrected SHE of the BTE

The density gradient model has been previously successfully introduced into a full-band Monte Carlo simulator [84]. Introducing quantum correction potentials in a SHE of the BTE influences the step in which the H-grid is calculated, shown in Figure 3.3, quite strongly. First the discretization of the H-space needs to be split for electrons and holes, since different correction potentials are applied for electrons and holes respectively. When evaluating the recombination terms (cf. Section 3.5), care must be taken to only use the distribution function directly or the charge carrier concentrations, but not the quantum corrected electrostatic potential in order

to avoid mistakes. Incorporating the quantum correction potentials

$$\gamma_n = \frac{\hbar^2}{12\lambda_n k_B T_L m_n^*} \left(\nabla_{\mathbf{x}}^2 \varphi + \nabla^2 \gamma_n - \frac{1}{2k_B T_L} (\nabla_{\mathbf{x}} \varphi + \nabla_{\mathbf{x}} \gamma_n)^2 \right), \quad (4.12)$$

$$\gamma_p = \frac{\hbar^2}{12\lambda_p k_B T_L m_p^*} \left(\nabla_{\mathbf{x}}^2 \varphi + \nabla^2 \gamma_p - \frac{1}{2k_B T_L} (\nabla_{\mathbf{x}} \varphi + \nabla_{\mathbf{x}} \gamma_p)^2 \right), \quad (4.13)$$

for electrons and holes respectively leads to a modified H-transform

$$H = \begin{cases} \epsilon - |q| [\varphi(\mathbf{x}, t) + \gamma_n(\mathbf{x}, t)], & \text{for electrons} \\ \epsilon + |q| [\varphi(\mathbf{x}, t) + \gamma_p(\mathbf{x}, t)], & \text{for holes,} \end{cases} \quad (4.14)$$

where the force \mathbf{F} is calculated per charge carrier using

$$\mathbf{F}_n(\mathbf{x}, t) = -\nabla_{\mathbf{x}}(\pm E_C \mp |q| [\varphi(\mathbf{x}, t) + \gamma_n(\mathbf{x}, t)]), \quad (4.15)$$

$$\mathbf{F}_p(\mathbf{x}, t) = -\nabla_{\mathbf{x}}(\pm E_V \mp |q| [\varphi(\mathbf{x}, t) + \gamma_p(\mathbf{x}, t)]). \quad (4.16)$$

4.4 Discretization

Various authors [85–89] reported numeric instabilities, when directly applying the finite volume method to Equation (4.2) using drift diffusion and thus developed various discretization schemes for DG and developed higher resolution schemes. An almost complete list of schemes is given in [9]. In this work only two discretization schemes, suggested by [9], will be discussed. The first discretization is called the simplified scheme, which simply neglects the second order term in Equation (4.2) and exhibits numeric stability. The ‘full scheme’, is the second discretization of Equation (4.2). In this scheme Equation (4.4) is used instead of Equation (4.2). In the course of this thesis, both schemes, have been implemented in ViennaSHE [65]. The simplified and full scheme have been implemented into the device simulator MinimosNT by [90]. In all implementations convergence of the non-linear solver has been achieved without any problems for the full and the simplified scheme. Although the simplified scheme often lead to faster convergence for the non-linear solver, when used in a Gummel-loop with the spherical harmonics expanded BTE.

4.4.1 The Simple Scheme

Following [9], the second order terms in Equation (4.2) are neglected yielding,

$$\gamma(\mathbf{x}, t) = \frac{\hbar^2}{12\lambda k_B T_L m^*} (\nabla_{\mathbf{x}}^2 \varphi(\mathbf{x}, t) + \nabla_{\mathbf{x}}^2 \gamma(\mathbf{x}, t)), \quad (4.17)$$

in a homogeneous semiconductor. Directly applying a finite volume discretization gives

$$\gamma_i = \frac{\hbar^2}{12\lambda k_B T_L m^*} \frac{1}{V_i} \sum_j \frac{A_{ij}}{d_{ij}} (\varphi_i + \gamma_i - \varphi_j - \gamma_j), \quad (4.18)$$

where V_i is the volume of the box around vertex i , A_{ij} are the interfaces and d_{ij} are the distances between vertices i and j . Additionally it has been assumed that the semiconductor is homogeneous.

4.4.2 The Full Scheme

In the full scheme, Equation (4.2) is replaced by Equation (4.4). Thus the equations for electrons and holes to be discretized are

$$\gamma_p(\mathbf{x}, t) = -\frac{\hbar^2}{6\lambda_n k_B T_L m_n^*} \frac{\nabla_{\mathbf{x}}^2 \sqrt{n(\mathbf{x}, t)}}{\sqrt{n(\mathbf{x}, t)}}, \quad (4.19)$$

$$\gamma_n(\mathbf{x}, t) = \frac{\hbar^2}{6\lambda_p k_B T_L m_p^*} \frac{\nabla_{\mathbf{x}}^2 \sqrt{p(\mathbf{x}, t)}}{\sqrt{p(\mathbf{x}, t)}}. \quad (4.20)$$

Applying the finite volume method yields

$$\gamma_{n,i} = -\frac{\hbar^2}{6\lambda_n k_B T_L m_n^*} \frac{1}{V_i} \sum_j \frac{A_{ij}}{d_{ij}} \left(\frac{\sqrt{n_j}}{\sqrt{n_i}} - 1 \right), \quad (4.21)$$

$$\gamma_{p,i} = \frac{\hbar^2}{6\lambda_p k_B T_L m_p^*} \frac{1}{V_i} \sum_j \frac{A_{ij}}{d_{ij}} \left(\frac{\sqrt{p_j}}{\sqrt{p_i}} - 1 \right), \quad (4.22)$$

which can result in convergence problems, since changes in the electrostatic potential result in exponential changes in the charge carrier concentrations. To mitigate this one can approximate the charge carrier concentrations using Boltzmann statistics and thus obtains

$$\gamma_{n,i} = -\frac{\hbar^2}{6\lambda_n k_B T_L m_n^*} \frac{1}{V_i} \sum_j \frac{A_{ij}}{d_{ij}} \left(\exp\left(\frac{\varphi_i + \gamma_i - \varphi_j - \gamma_j}{2}\right) - 1 \right), \quad (4.23)$$

$$\gamma_{p,i} = \frac{\hbar^2}{6\lambda_p k_B T_L m_p^*} \frac{1}{V_i} \sum_j \frac{A_{ij}}{d_{ij}} \left(\exp\left(\frac{\varphi_i + \gamma_i - \varphi_j - \gamma_j}{2}\right) - 1 \right), \quad (4.24)$$

where it was assumed that the semiconductor is homogeneous.

4.5 Boundary Conditions

There are at least three possible choices of boundary conditions. Systematic research was, for example, conducted in this area by [9, 90], where all kinds of conditions, such as

$$\gamma = f \quad \text{Dirichlet}, \quad (4.25)$$

$$\frac{\partial \gamma}{\partial \mathbf{d}} = f \quad \text{von Neumann and} \quad (4.26)$$

$$\alpha \gamma + \beta \frac{\partial \gamma}{\partial \mathbf{d}} = f \quad \text{Robin} \quad (4.27)$$

boundary conditions have been tested, where \mathbf{d} is the normal vector of the oxide semiconductor interface, α , β and f need to be reasonably chosen. At the boundaries of the device or semiconductor segments which do not border to an insulator segment Dirichlet conditions with $\gamma_n = \gamma_p = 0$, have always been employed throughout the literature. Whereas various types of boundary conditions at semiconductor-insulator interfaces have been tested, including Dirichlet, von Neuman and Robin conditions. It is necessary to find boundary conditions for semiconductor-insulator interfaces with which it is possible to correct the electrostatic potential such, that the carrier concentrations calculated by the transport model yield the same as predicted by the Schrödinger equation (cf. Section 2.6). It was shown that Robin boundary

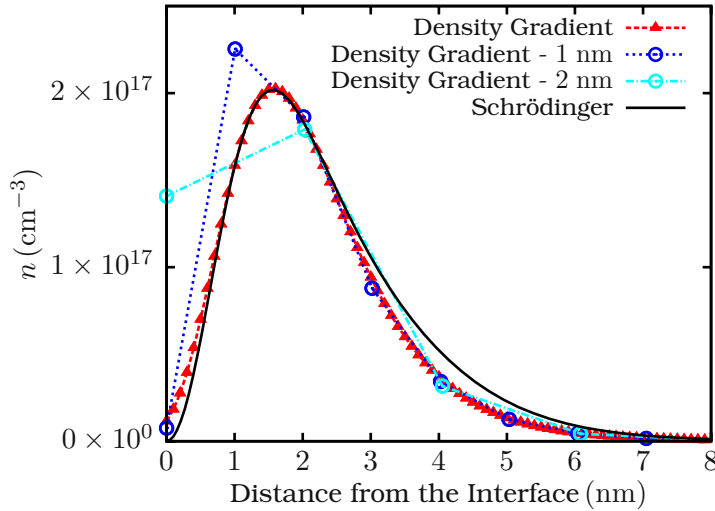


Figure 4.1: Calibration results for a density gradient model for quantum corrected drift diffusion, implemented in MinimosNT, to a solution of the Schrödinger-Poisson equation, using VSP [51]. In this case the fit has been obtained using a 1D NMOS structure, Robin boundary conditions at the silicon-silicon-dioxide-interface, an acceptor doping of $3 \cdot 10^{17} \text{ cm}^{-3}$, an oxide thickness of 1 nm and a uniform grid spacing (orthonormal grid) of 0.1 nm. A second fit using a grid spacing of 1 nm and 2 nm are shown to illustrate the grid spacing dependence of density gradient. The parameters have been previously published in [90].

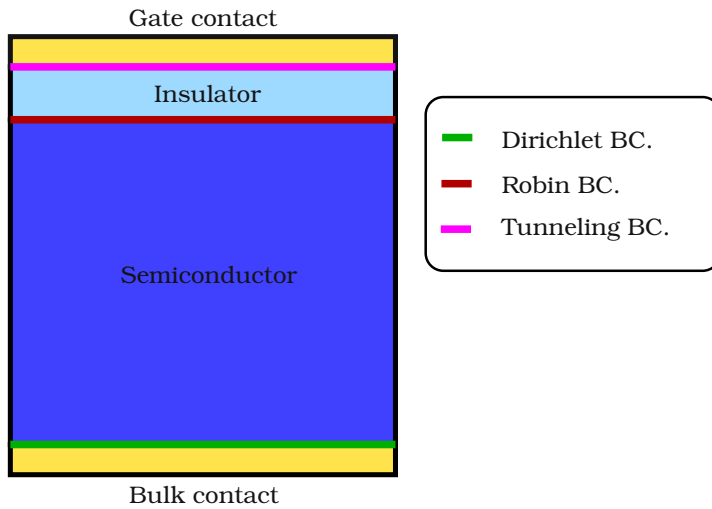


Figure 4.2: The boundary conditions applied for density gradient in a 2D MOS structure. The Dirichlet conditions, where the correction potential vanishes, are drawn in green. If tunneling is not considered, Robin conditions are used at the interface (shown in red). In case tunneling is assessed using density gradient, the Robin condition is replaced by a Dirichlet condition (termed Tunneling in the graph) at the insulator-gate interface.

conditions prove to be the best choice at semiconductor-insulator interfaces, since they allow the best fit with the carrier concentration obtained from a solution of the Schrödinger equation [90]. Subsequently Robin boundary conditions at semiconductor-insulator interfaces for Equation (4.2) are used throughout this thesis (cf. Figure 4.2). However, it was found on [91] that upon solving Equation (4.2) for insulator segments too, it has been suggested that it is possible to assess direct tunneling. Nevertheless, in the course of this thesis this feature of the DG model was not used since a more powerful technique, namely NEGF, is available to assess direct tunneling currents through oxides in MOS structures.

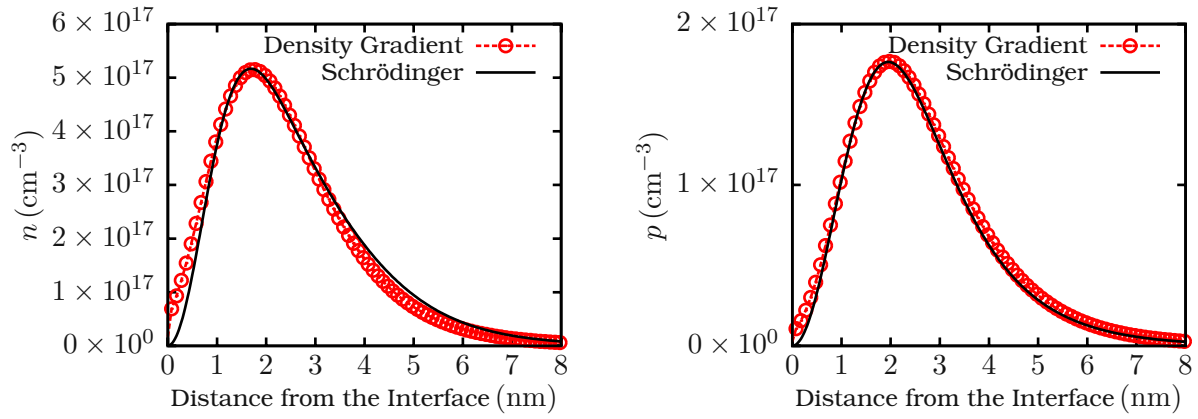


Figure 4.3: Results of the calibration of the DG model to a first-order SHE [65] using a solution of the Schrödinger-Poisson equation obtained via VSP [51]. **Left:** The fit for electrons has been obtained using a 1D NMOS structure, Robin boundary conditions at the silicon-silicon-dioxide-interface, an acceptor doping of $3 \times 10^{17} \text{ cm}^{-3}$, an oxide thickness of 1 nm and a uniform grid spacing (orthonormal grid) of 0.1 nm. **Right:** The fit for holes has been obtained using exactly the same device as for electrons but instead of an acceptor doping a donor doping of $3 \times 10^{17} \text{ cm}^{-3}$ has been used. The parameters are given in Table 4.1

4.6 Calibration

Whether or not DG reproduces confinement correctly largely depends on the boundary conditions, the mesh spacing and the parameters λ_n for electrons and λ_p for holes. As stated in the last section, Robin boundary conditions at insulator-semiconductor interfaces have proven to be best suited since they deliver the best fit with the carrier concentrations obtained from a solution of the Schrödinger equation (cf. Section 2.6).

4.6.1 Calibration for the Drift Diffusion Model

In Figure 4.1 the results of a fit, with a density gradient quantum corrected drift diffusion model, for electrons are shown for various grid spacings. This figure clarifies the need for a finely spaced grid for simulation, when using a quantum correction model. In Figure 4.4 Capacitance-Voltage curves, obtained by density gradient and a Schrödinger-Poisson solver, for n- and p-channel 1D MOS structures for various dopings are shown and compared [90]. Since density gradient does not exhibit a free, doping dependent parameter, a set of parameters only works for a certain bulk doping [5]. This is the main disadvantage of density gradient. Another requirement of density gradient, as for any Schrödinger-Poisson solver, is that the grid needs to be in the sub-nanometer regime in order to fully refine the 2D electron gas in a MOS structure, as demonstrated in Figure 4.1.

4.6.2 Calibration for a SHE of the BTE

This section is devoted to the calibration of density gradient in ViennaSHE [65]. To this end Robin-Boundary conditions and the simple scheme have been implemented and a comparison with VSP [51] has been carried out in weak inversion. The parameters of the fit are given in Table 4.1 and the resulting calibration is shown in Figure 4.3.

Carrier Type	α	β	f
Electrons	-61.3V/m	$-11.4 \cdot 10^{-5}$ V	0.0
Holes	-36.9V/m	$-8.3 \cdot 10^{-5}$ V	0.0

Table 4.1: Robin boundary condition parameters for density gradient and SHE. The parameters have been obtained by manual optimization using VSP [51].

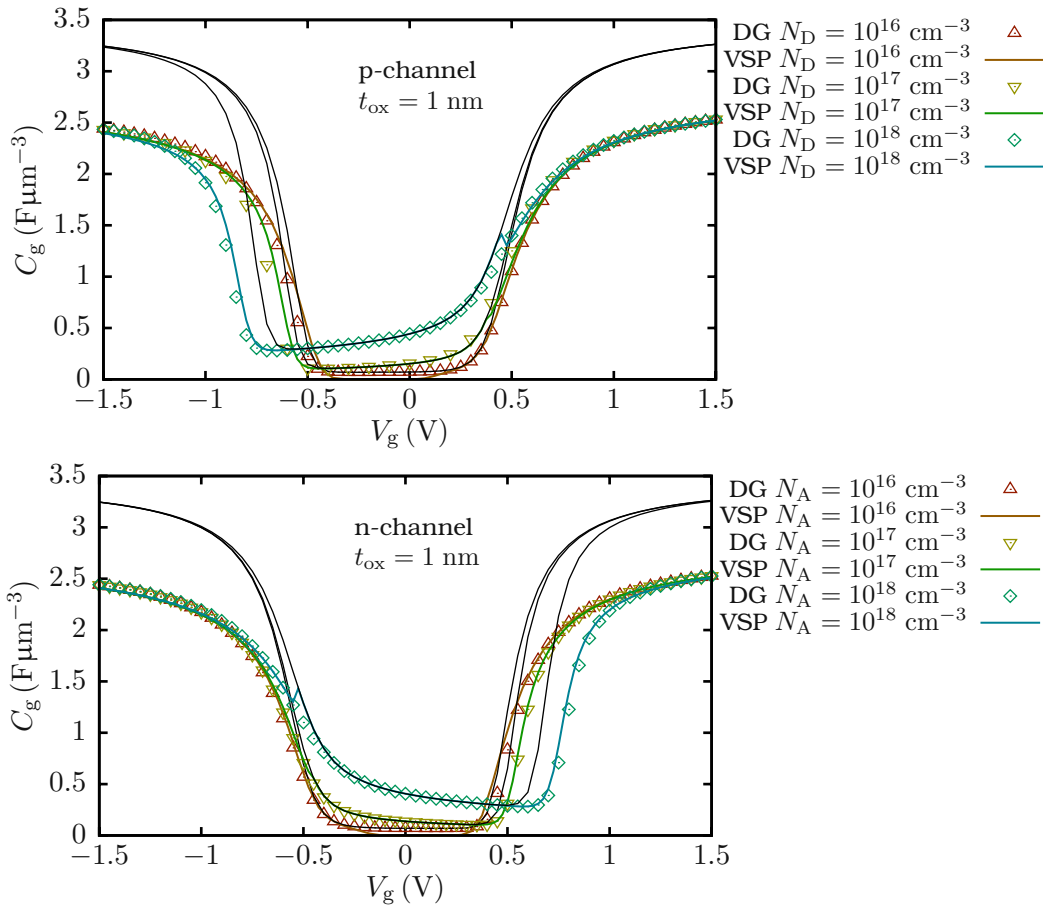


Figure 4.4: Capacitance-Voltage curves for various channel dopings in 1D MOS structures, obtained by a DG solution compared to a solution obtained using VSP [51]. Well visible is the increasing discrepancy between DG and a real Schrödinger-Poisson solution for dopings above and below $N_D = 10^{17}$ cm^{-3} , for which the Robin boundary condition coefficients have been fitted to the Schrödinger-Poisson solution. For comparison the CV-curves obtained using plain drift diffusion are shown in black. The fitting parameters used with MinimosNT have been previously published in [90].

5 Variability

In the full system of equations (cf. Equation (2.27) in Chapter 2) the fixed charges expressed as a charge density on the right hand side of Poisson's equation are modelled as macroscopic densities. All charges, whether they arise from doping, the free carriers or traps have so far been viewed as macroscopic quantities, neglecting their granular nature. This approach is well-applicable as long as the number of charged particles is large enough, which is the case for semiconductor devices with characteristic lengths larger than devices of the 100 nm node [3]. The influence of the granular, random nature of the donor and acceptor atoms in sub-100 nm node field effect devices has been well investigated by [3, 92, 93] and many others. Often, the study of this kind of granularity is called the study of random discrete dopants (RDD). The granularity itself stems from the device fabrication, during which the dopants are implanted into the silicon substrate resulting in a Poisson distribution of dopants per unit volume. In deca-nanometer devices the natural random, non-smooth distribution of dopants can lead to significant inter-device variability of important parameters such as the threshold voltage V_{th} . Nevertheless, random discrete dopants are not the only source of variability. In sub-100 nm field effect devices, trapped charges also have to be viewed as discrete, since they can electrostatically interact with the dopants in the channel [4], giving rise to the different variability in ΔV_{th} in degraded MOSFETs. Yet another source of variability emerges in small area metal contacts [94], since the granularity of the metal (MGG) becomes apparent: a volume of metal in a semiconductor device consists, due to fabrication, of multiple grains. The work function of the grains is statistically distributed and is usually assumed to follow a normal distribution. In field effect devices, this is important if a metal gate instead of a highly doped poly-silicon gate is used, which is the case for the sub-32 nm nodes [94].

In the course of this thesis RDD and random discrete traps have been investigated since both are essential to describe the 'step heights' seen in ΔV_{th} recovery traces [4, 95], especially after bias temperature stress. MGG is important to correctly assess the variability of device parameters of sub-100 nm metal-gate MOSFETs and has thus been implemented into the simulator MinimosNT for evaluation purposes. It was found that MGG has negligible impact on the threshold voltage shifts caused by discrete oxide traps close to the semiconductor interface. Therefore, MGG is not considered in the remainder of this chapter.

5.1 Random Discrete Dopands

As stated above, in sub-100 nm field effect devices, the discreteness of the dopant atoms cannot be ignored anymore. In modern semiconductor devices, the semiconductor doping is introduced by 'ion implantation'. In this technique the semiconductor material is bombarded by ionized dopant atoms. Upon hitting the target material the dopant atoms are randomly scattered by the lattice atoms of the semiconductor. Thus, to obtain the positions of the dopant atoms, in an

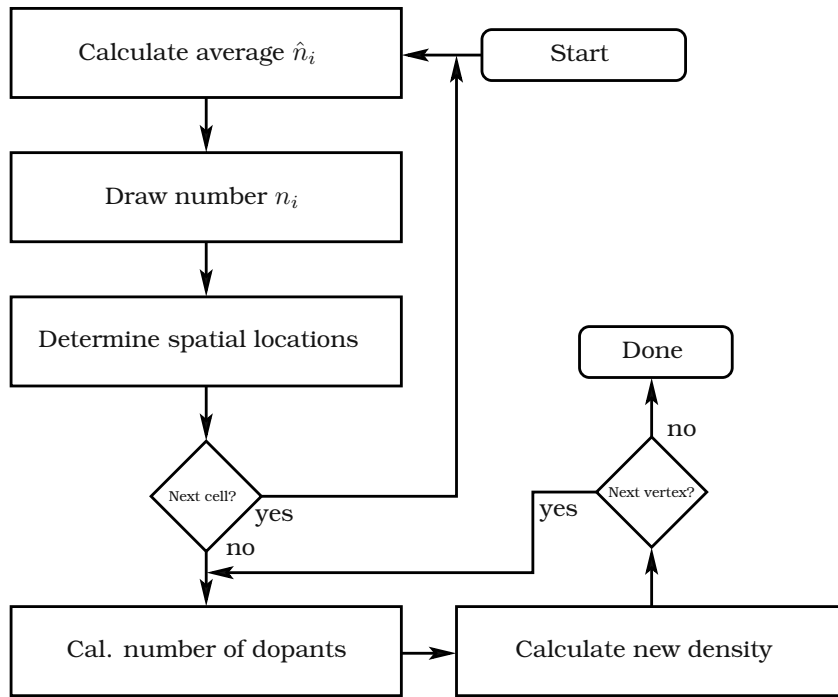


Figure 5.1: A flowchart of the Monte Carlo algorithm for random discrete dopants detailed in the text.

ion-shell model of the atom, one can either directly use the data from a process simulation or discretize the given macroscopic doping density (N_A and N_D) and calculate the dopant positions by a straight-forward Monte Carlo algorithm [95, 96].

5.1.1 Random Discrete Dopands Algorithm

In this thesis the Monte Carlo approach to the placement of random discrete dopants is chosen. The algorithm works as follows (cf. Figure 5.1). First the expected number of dopants \hat{n}_i is calculated per cell i using $\hat{n}_i = V_i \times N_i$, where N_i is the dopant concentration and V_i is the volume of the cell. Next, the actual number of dopants of the i th cell is obtained by drawing a Poisson distributed (\hat{n}_i being the mean value) random number n_i , which is the total number of dopants in that cell. In a third step the positions of each of the n_i dopants are determined by drawing random numbers for each spatial direction under the condition that each position must be located within the boundaries of the cell. Since in Poisson's equation charge densities are needed, equivalent charge densities have to be calculated from the positions of the dopants. When using a finite volume discretization for Poisson's equation this is done by finding the vertex with the shortest distance to each dopant and counting the number of dopants (j) associated with each vertex (cf. Figure 5.2). Then, after the dual grid has been calculated, the new donor/acceptor concentrations per volume in the dual grid is calculated by dividing the number of dopants j associated with the vertex at the center of the finite volume by the volume. The precision of this algorithm strongly depends on the resolution of the grid. In [93] it has been shown that a grid with a spacing below 1 nm is sufficient to capture the mean value of the threshold voltage variability due to RDD in field effect devices. This results in a 'jittered', finely resolved doping, which is suitable for simulation of RDD with conventional device simulators employing Poisson's equation.

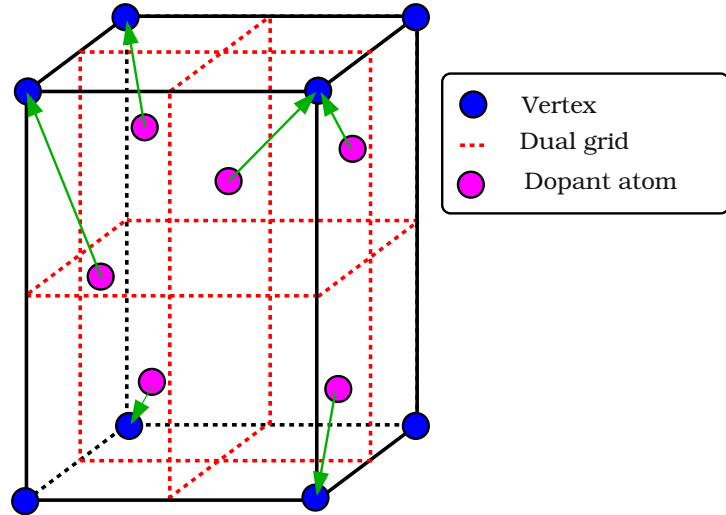


Figure 5.2: Illustration of the randomization algorithm, in a single cell, detailed in the text. The dual grid is shown in red. The green arrows depict the assignment of the dopants to the respective vertices.

5.1.2 Screening Charges

Considering single point charges in the numerical solution of the Poisson equation a problem emerges. Since the potential of a point charge located at \mathbf{r}_0 is $q\delta(\mathbf{r} - \mathbf{r}_0)$, this would result in an infinite number of charge carriers screening the point charge (cf. Figure 5.3). Certainly this behavior is unphysical and an artifact of a classical or semi-classical system description, since the point charge serves as a potential well for charge carriers and thus can be screened by a few electrons or holes occupying discrete eigenenergies as described by the Schrödinger equation. Additionally, a semi-classical description of point charges is highly dependent on the grid spacing. Thus in a semi-classical system description correction methods are required. In the literature two approaches can be found to eliminate the artificial screening effect and to make the results of the simulation fairly independent of the grid spacing. In the first approach the Fourier transformed charge density of the point charge is calculated and formally split into a long range and a ‘short range’ part per finite volume i :

$$\rho(r) = q\delta(r) = qV_i^{-1} \left(\sum_{k < k_c} \exp(-ik \cdot r) + \sum_{k > k_c} \exp(-ik \cdot r) \right) = \rho_{\text{short}}(r) + \rho_{\text{long}}(r), \quad (5.1)$$

where k_c is the cut-off radius in reciprocal space. Using the screening length of the Conwell-Weisskopf model [97] k_c is determined using

$$k_c = \kappa N^{1/3}, \quad (5.2)$$

where N is either the donor or acceptor concentration and $\kappa = 2$ is empirically chosen. This model is termed ‘Long-Range’ model since only the long range part $\rho_{\text{long}}(r)$ is considered on the right hand side of Poisson’s equation, thus eliminating the artificial screening [6]. The Long-Range model works well [6, 98], but the question remains whether $\kappa = 2$ is a reasonable choice. Additionally, the Long-Range model does not feature any quantum correction, which is needed for sub-100 nm field effect devices. Thus the Long-Range model will not be further used or discussed in this thesis. A good replacement, which features first-order quantum correction (carrier confinement) for semi-classical transport models, is the density gradient model (DG) (cf. Section 4.1) [93]. A density gradient first-order quantum correction features a set of parameters (λ , α , β) which need to be determined. Calibration of the DG model is usually

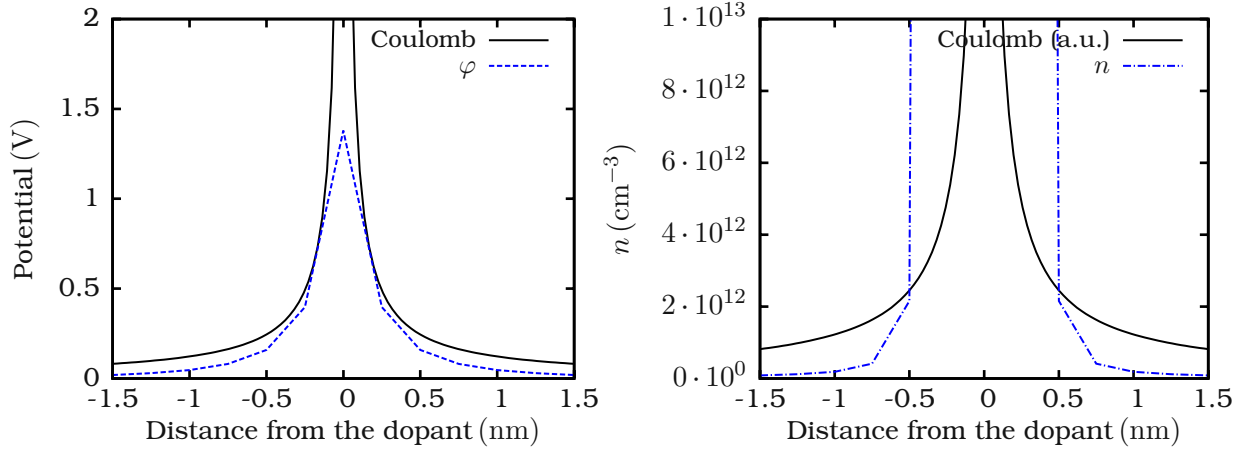


Figure 5.3: The potential ϕ (left figure) and the electron concentration n (right figure) around a single point charge embedded in pure silicon obtained by a DD or SHE simulation in equilibrium are shown. Without any correction of the coloumb potential a semi-classical simulator will predict an infinite amount of electrons around the point charge.

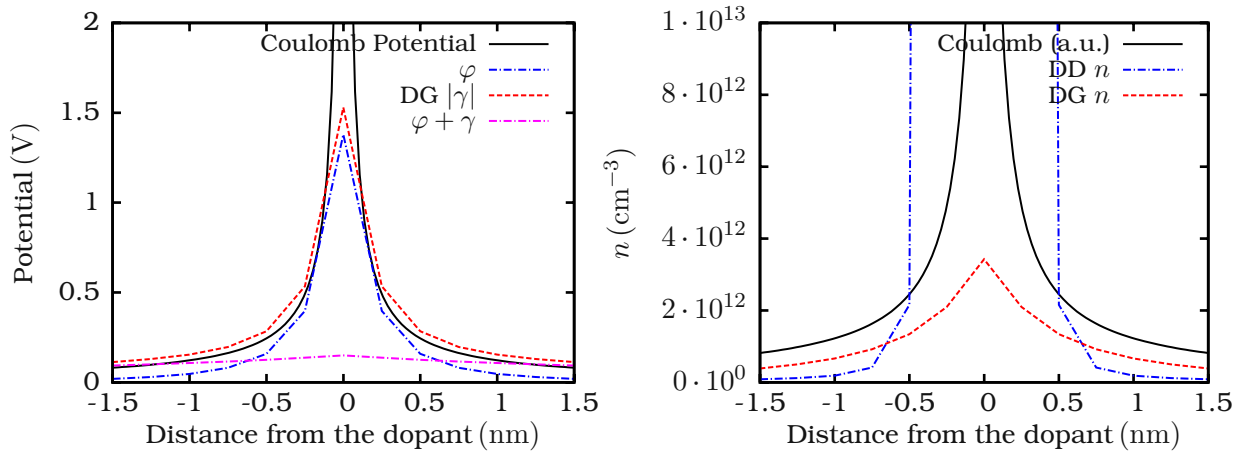


Figure 5.4: The potential ϕ (left figure) and the electron concentration n (right figure) around a single point charge embedded in pure silicon obtained by a quantum-corrected DD or SHE simulation in equilibrium are shown. A grid independent correction of the coloumb potential is achieved via the density gradient model, such that only the long range part $\rho_{\text{long}}(r)$ of the coloumb potential enters Poisson's equation. In the figure γ denotes the quantum correction potential and $\phi + \gamma$ is the quantum corrected potential.

done using an accurate description of the carrier confinement by the Schrödinger equation. This eliminates the need for empirically determined parameters. When employing the density gradient model, the results also become independent of the grid spacing. Nevertheless, the grid needs to be resolved finer than 1 nm in order to capture the effects of carrier confinement and discrete dopant screening (cf. Figure 5.3).

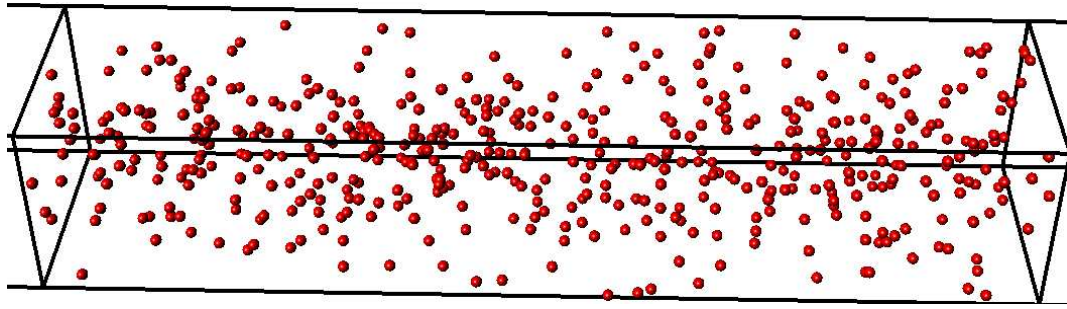


Figure 5.5: The discrete doping in an n-doped 50 nm long, 10 nm times 10 nm wide resistor with idealized ohmic contacts. The initial doping $N_D = 10^{20} \text{cm}^{-3}$ concentration was assumed to be uniform.

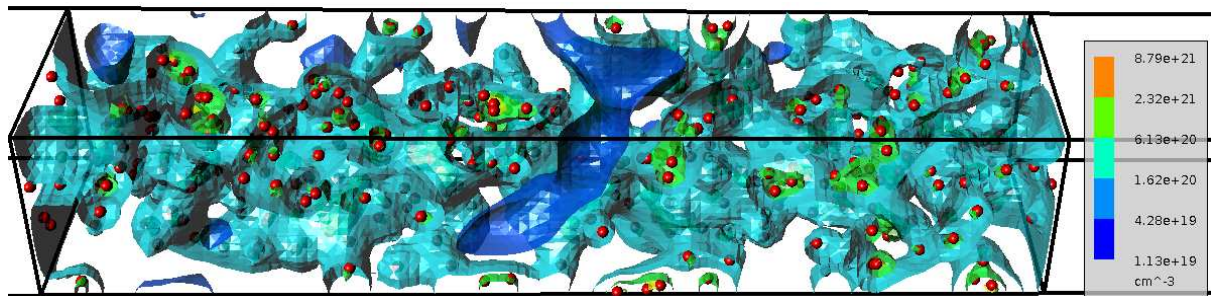


Figure 5.6: A surface plot of equal electron concentrations in equilibrium (all terminals are grounded). The simulation has been carried out at room temperature using MinimosNT. The dopant distribution is shown in Figure 5.5. The grid is orthogonal and regular with a spacing of 0.5 nm.

5.1.3 Simulation Results using the Drift Diffusion Model

The algorithms laid out in the sections before have been implemented into the simulator MinimosNT [90]. First, in order to investigate the grid dependence of the results, a simple n-type $10 \text{ nm} \times 10 \text{ nm} \times 100 \text{ nm}$ nanoscale resistor is simulated (cf. Figure 5.5 and Figure 5.6). Due to the fluctuations of the doping, the potential is expected to be spatially fluctuating as well, resulting in charge carrier rich and charge carrier deprived areas. This leads to fluctuations in device parameters (resistivity) from device to device, although macroscopically they are exactly the same. But due to the small characteristic lengths the doping cannot be viewed macroscopically. Thus hundreds of microscopically different devices have to be investigated in order to assess the inter-device distribution of parameters in any nano-scale devices.

Percolation Paths

Upon investigation of the influence of random discrete dopants with the density gradient model as quantum correction for the drift diffusion model to a MOSFET a new effect emerges. In sub-100 nm channel length MOSFETs, there are only few dopants in the channel leading to considerable potential fluctuations. Thus the inversion condition becomes a function of the spatial locations of the discrete dopants and leads to the formation to current percolation paths (cf. Figure 5.7). This in turn leads to substantial fluctuations in the $I_d - V_g$ -characteristic and consequently to fluctuations of the threshold voltage from device to device (cf. Figure 5.8).

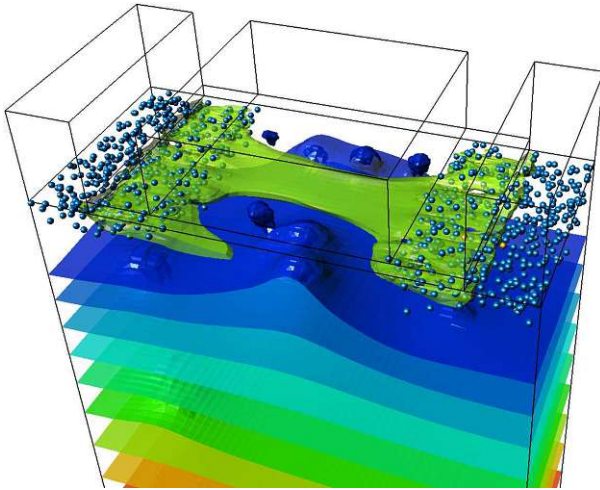


Figure 5.7: The percolation path (in green) in a MOSFET in weak inversion and equipotential surfaces. Due to the potential fluctuations the conditions for charge carrier inversion is only fulfilled in certain regions of the channel, causing dominant current paths. Donor dopants are rendered as blue balls and acceptor dopants as golden balls.

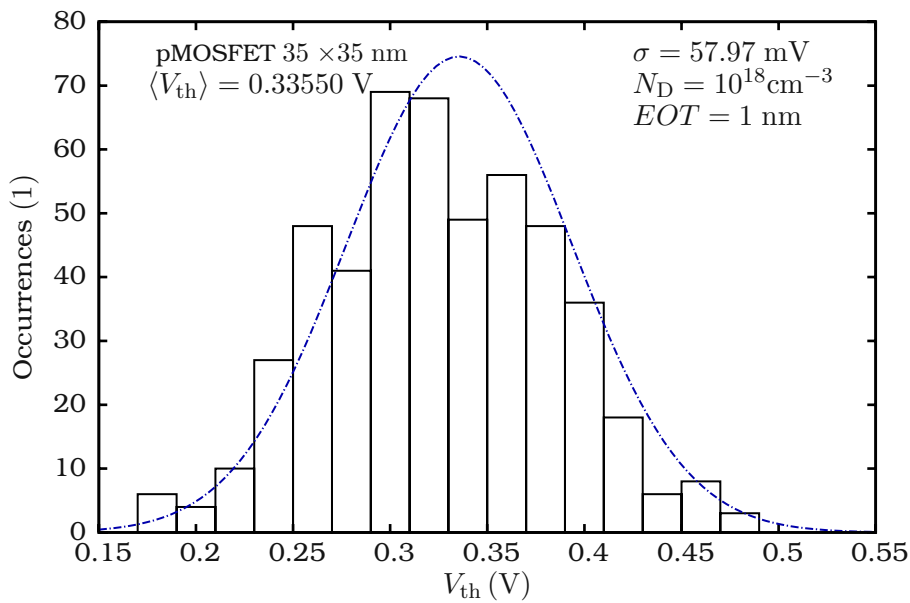


Figure 5.8: Histogram of the distribution of the threshold voltage in a MOSFET composed of 500 microscopically different devices. A Gaussian distribution of V_{th} has been fitted to the data. The standard deviation is above 30 mV. In order to circumvent this large deviation in the threshold voltage the utilization of a delta-doping has been put forward [99].

5.2 Random Discrete Traps

Recent studies have demonstrated that in modern deca-nanometer devices the variability due to random discrete dopants (RDD) and oxide defects has become critical in the context of reliability [100–103]. For instance, discrete charges in the oxide or at the interface lead to the occurrence of potentially huge ΔV_{th} shifts due to the non-uniform current flow in the channel (Figure 5.10). RDD needs to be considered when investigating the ‘step heights’ in the drain current and threshold voltage observed in field effect devices under bias temperature or hot carrier stress. Most studies have considered either fixed positive/negative charges or described charge capture into a fixed number of defects by standard SRH trapping kinetics [104], where the goal is to explain the statistical distribution of ΔV_{th} shifts in an ensemble of

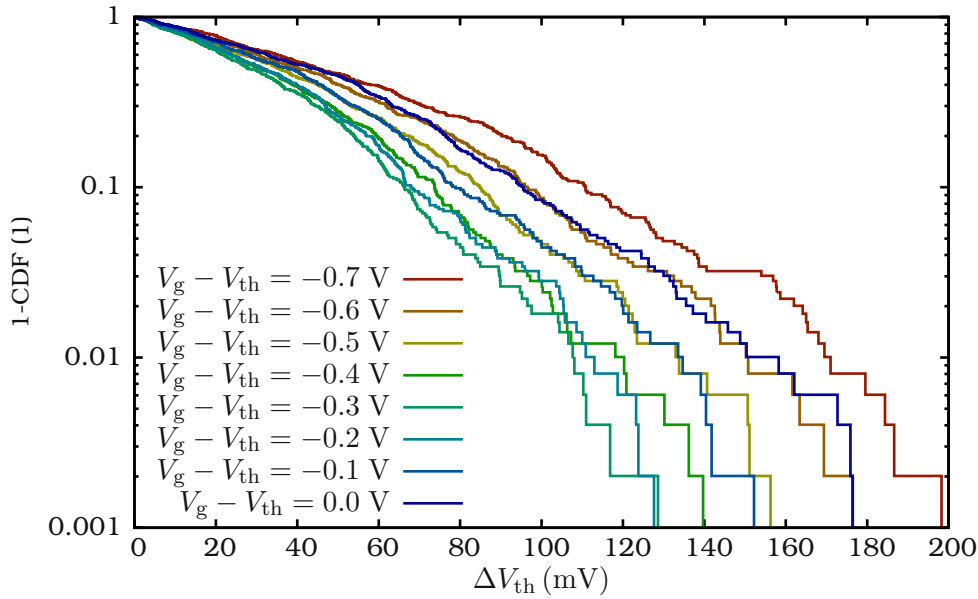


Figure 5.9: The cumulative distribution functions (CDF) of ΔV_{th} over gate overdrive for 500 different microscopic p-channel MOSFETs due to RDD only (no traps). Clearly visible is the large deviation in maximum step-heights of the CDFs over gate overdrive. This is due to the emergence of new percolation paths with higher gate overdrive.

devices (cf. Figure 5.9).

In this section the mere electrostatic effects of random discrete charges (traps) and random dopants on the drain current are discussed. A discussion on the actual number of available defects in a time-dependent manner and their activation process itself is delayed to Chapter 6.

5.2.1 Single Trap

The effect of a single charged oxide defect on the drain current in a MOSFET can be explained without considering the time dependence of its occupancy and has been studied in depth using a first-order quantum corrected drift diffusion model [103]. Depending on the spatial location of the trap the resulting ΔV_{th} , that is its ‘step height’, varies. In the course of this thesis, together with the author of [90], this was investigated for planar devices.

In order to obtain the spatial dependence of the ΔV_{th} for a single defect for a single microscopic device (constant dopant configuration), first a discrete random dopant configuration is determined (cf. Section 5.1.1) for the planar MOSFET under investigation. Then for each position of a single discrete elementary oxide charge in a predefined grid a single device simulation is carried out to determine the $I_d - V_g$ curve. At last these $I_d - V_g$ curves are compared to the $I_d - V_g$ curve for the device without any oxide charges. In Figure 5.11 the ΔV_{th} maps for a planar p-channel MOS, with a channel doping of $N_D = 10^{18} \text{ cm}^{-3}$, a channel length and width of 35 nm and an effective oxide thickness (EOT) of 1 nm silicon-dioxide are shown. The ΔV_{th} maps are a practical tool to show the trap position dependence of the ΔV_{th} caused by a single elementary charge, however they are not suitable to compare simulation and measurement. For this the cumulative distribution function, obtained by considering a statistical representative ensemble of devices, of the individual step heights is used. In order to obtain the cumulative distribution of the step-heights of a statistically representative sample of microscopic devices

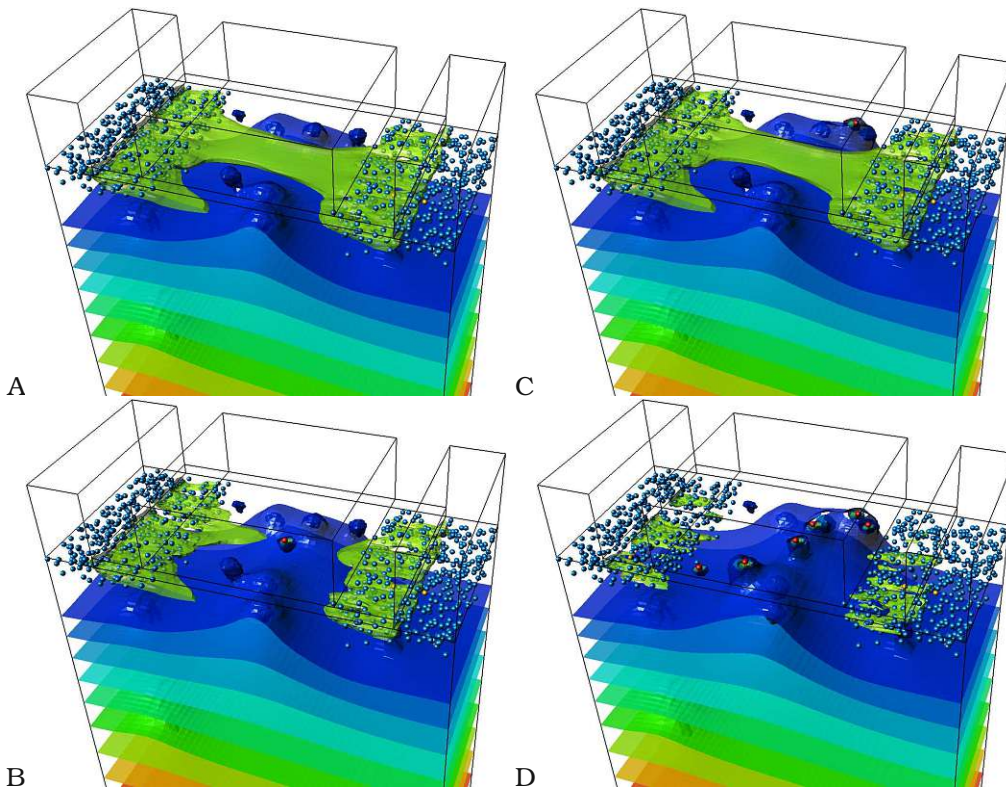


Figure 5.10: The percolation path in a selected microscopic MOSFET due to random discrete dopants and equi-potential surfaces. In sub-figure A there are no oxide traps and a single percolation path dominates. Upon the the formation of a fully charged trap right above the dominant percolation path, the device is switched off and the dominant percolation path vanishes (sub-figure B) leading to a huge ΔV_{th} . When an oxide trap is being charged next to the percolation path (sub-figure C), the trap has a negligibly small influence on the current flow and causes only a small ΔV_{th} . Sub-figure D shows the formation of six fully charged traps perpendicular to the current flow. In such a case the device becomes much harder to switch on.

has to be simulated. In Figure 5.12 the cumulative distribution function of the ‘step heights’ for a single trap in a p-channel MOSFET is plotted. From the plot the dependence of the step-height maximum on gate bias and channel doping is clearly visible. Figure 5.13 finally shows that in order to describe the occurrence of large steps in V_{th} a microscopic description of the doping is necessary by comparing the CDFs for RDD and continuous doping.

5.2.2 Multiple Traps

In real devices there are often a number of oxide traps. In sub-deca micrometer devices the oxide volumes are so small that only a few traps, if any, can be observed in a single device [105]. When considering multiple traps and RDD the question arises whether or not two discrete traps have an influence on the ΔV_{th} caused by one another [4]. It was found by [106] that two or more traps act independent of each other, at least electrostatically. The CDFs of ΔV_{th} for multiple traps are comparatively plotted in Figure 5.14.

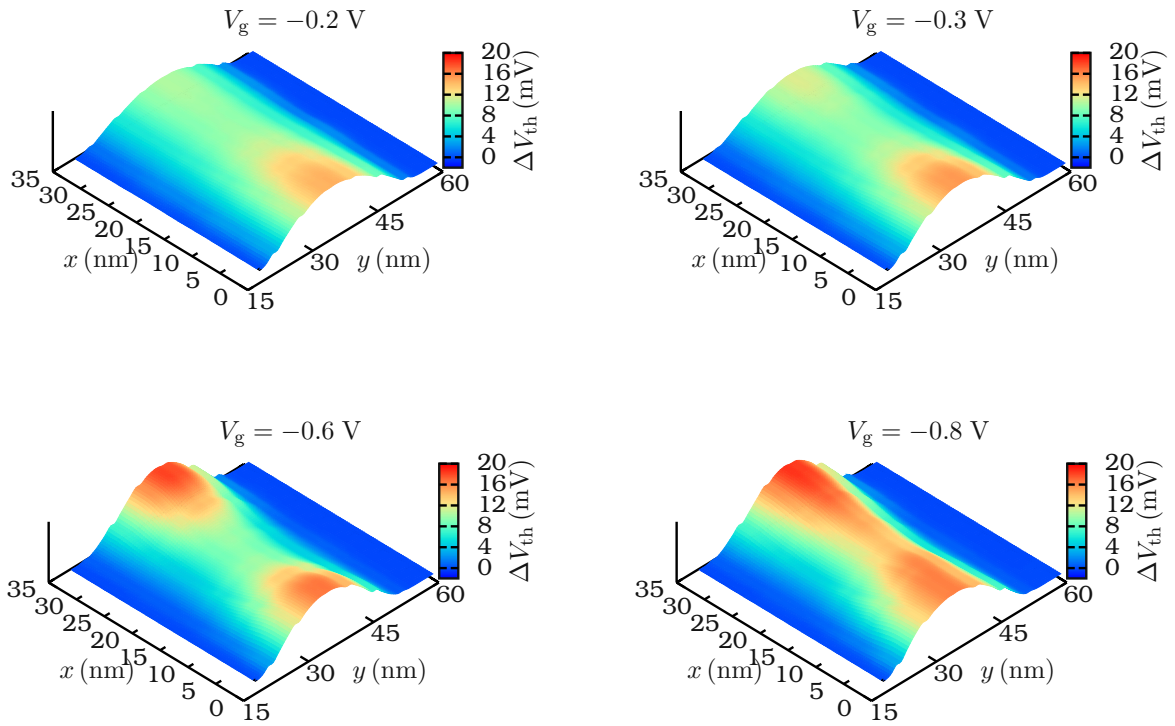


Figure 5.11: Spatial distribution of ΔV_{th} in $35 \text{ nm} \times 35 \text{ nm}$ p-channel MOSFET, with a channel doping of $N_D = 10^{18} \text{ cm}^{-3}$. To generate these maps 1500 simulations of $I_d - V_g$ curves with $V_{ds} = 1 \text{ mV}$ have been carried out, where in each simulation the elementary point charge has been moved along the channel surface and was located directly at the interface.

5.2.3 Mobility

Comparing simulation data (DG and DD), to measurement data it was observed that the doping dependent mobility model needs to be slightly adjusted. Most authors have used constant homogeneous mobilities in their random dopant studies, see for example [107]. It was later found in [98, 108] through classical Monte Carlo simulations that random discrete dopants not only cause potential fluctuations, but also spatial fluctuations in the scattering rates. For moments based charge transport models it was shown by [109] that it is sufficient to adjust the doping dependent mobility model parameters such that the bulk-mobility of a resistor is the same when simulated with continuous and discrete doping. In the course of this thesis it was found that this is not quite correct. It is possible to reproduce measured mean and standard deviation in V_{th} due to random discrete dopants only [110]. However, it was found impossible to reproduce measurements taken on an ensemble of p-channel MOSFETs by simulation with quantum corrected drift diffusion taking into account discrete traps and random discrete dopants. The devices were p-channel MOSFET with a channel doping of $N_D = 10^{18} \text{ cm}^{-3}$ produced by imec. The measurements have been taken with great care to ensure that only a single trap is created per device. The microscopic device for simulation was calibrated to the mean $I_d - V_g$ curve of the undegraded imec devices. From both sets of data, namely measurement and simulation, the cumulative distribution functions of distinct step heights are shown in

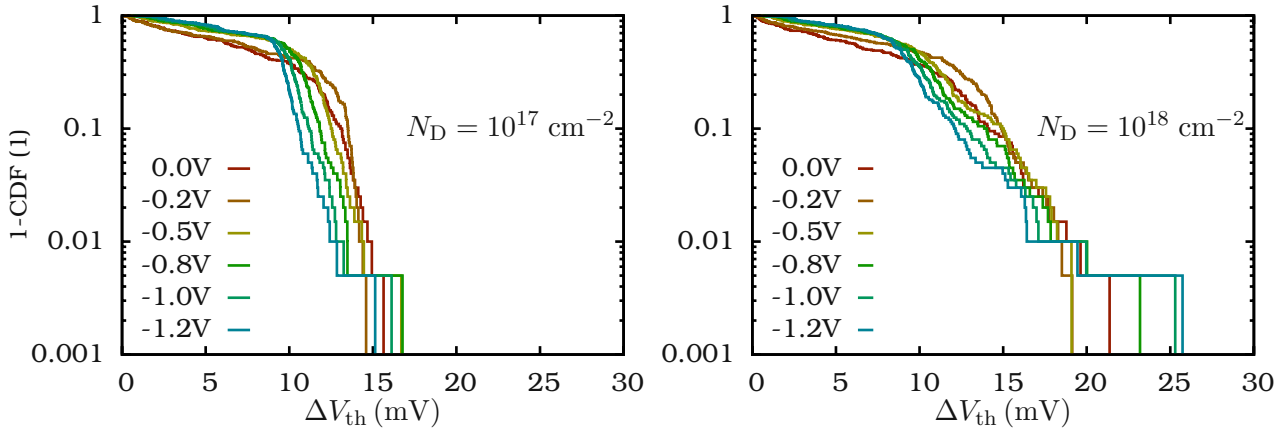


Figure 5.12: The CDF of the ΔV_{th} in a $35 \text{ nm} \times 35 \text{ nm}$ p-channel MOSFET due to a single trap for two distinct channel dopings. It can be seen that a higher channel doping results in bigger ΔV_{th} steps for a single charge at the interface. This is due to the increasing number of dopants in the channel, when the average channel doping is higher. Each graph is composed from a thousand different simulations using MinimosNT to have a statistically meaningful sample size.

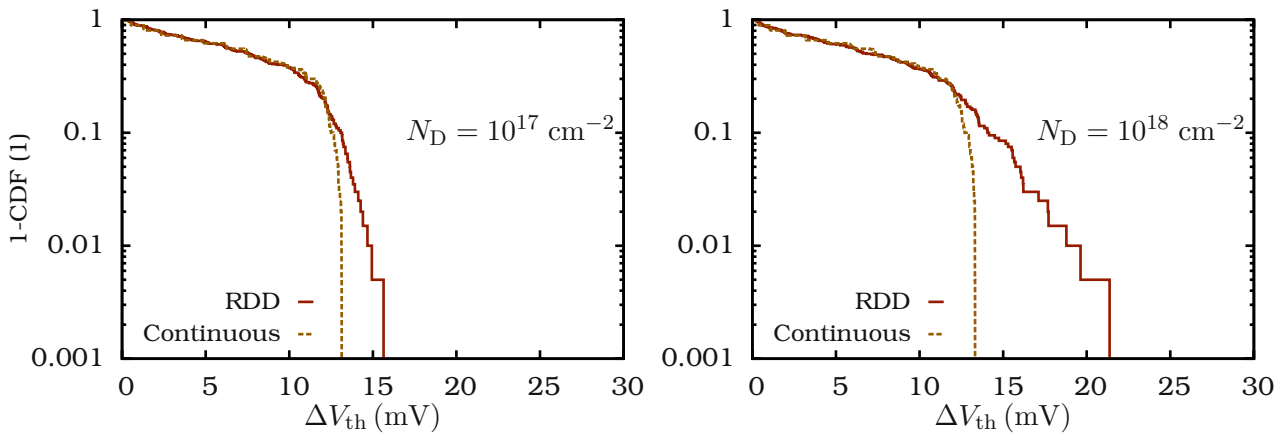


Figure 5.13: The data from Figure 5.12 at $V_g = V_{th}$ compared to the step heights obtained if the doping is continuous and not discrete. Clearly visible is the need for RDD in order to correctly describe the large dispersion in step heights, which cannot be obtained using a continuous doping. It is also remarkable that the macroscopic doping concentration has virtually no influence on the maximum step height for continuous channel doping. Each graph is composed from a thousand different simulations using MinimosNT.

Figure 5.15. To highlight the amplification of the defect impact due to the random dopants and to be able to compare MOSFETs of slightly different geometries, we normalize the step height by the theoretical value obtained through the charge-sheet approximation assuming that each

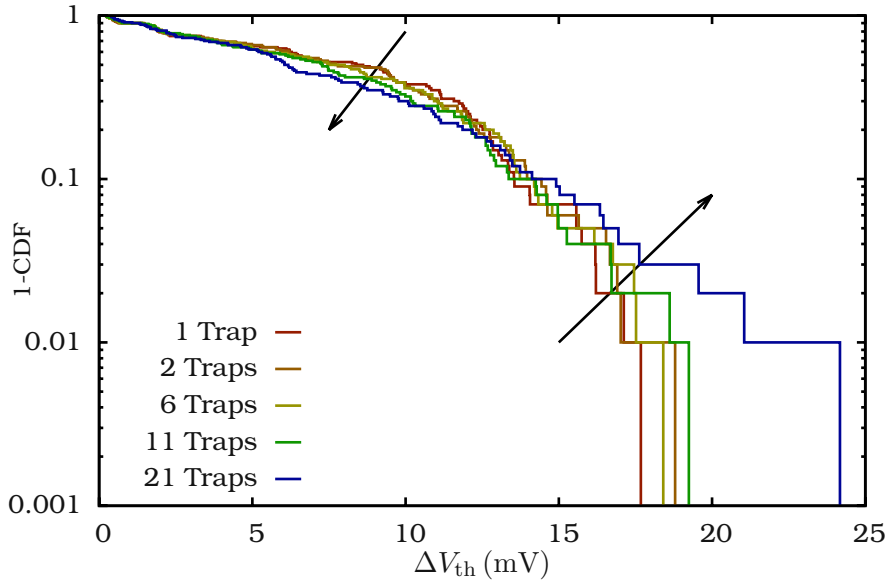


Figure 5.14: The cumulative distribution of step heights for multiple evenly distributed traps in a 35 nm p-channel MOSFET with a channel doping of $N_D = 10^{18} \text{ cm}^{-3}$. It is obvious that the maximum step height increases with the number of traps. More interestingly no ΔV_{th} step above 30 mV could be observed even for 21 oxide traps, which corresponds to a trap concentration of $1.7 \cdot 10^{19} \text{ cm}^{-3}$.

trap is located directly at the interface:

$$\eta_r = \Delta V_{th} \eta_0 \quad \text{with} \quad \eta_0 = \frac{q}{C_{ox}}, \quad (5.3)$$

where C_{ox} is the nominal gate oxide capacitance. Surprisingly the simulation cannot reproduce the measurement by a factor of 5 in η_r . This suggests that the purely electrostatic picture to determine step heights is not entirely correct as previously reported, for example in [98]. Not even the largest step heights, caused by a single trap, can be reproduced. Large step heights are, in the simulation, due to huge fluctuations in the potential leading to a dominant percolation, which can be ‘blocked’ by a single trap causing a huge step in drain current. Recalling that ΔV_{th} is estimated by comparing $I_d - V_g$ curves or by calculating ΔV_{th} from measured steps in the drain current ΔI_d by a SPICE level 1 model [111], the only parameter that has not been accounted for is charge carrier mobility μ . As stated initially, in [98] it was shown that a quantum corrected drift diffusion simulation underestimates the standard deviation in current fluctuations caused by random discrete dopants or random discrete charges by at least 20%, when compared to solutions obtained from a calibrated Monte Carlo simulator incorporating RDD. Thus deviations seen in Figure 5.15 can be minimized by incorporating fluctuations in mobility on a macroscopic level or ionized impurity scattering on a microscopic level. This deviation of 20% explains the difference between measurement and simulation using DD and DG models.

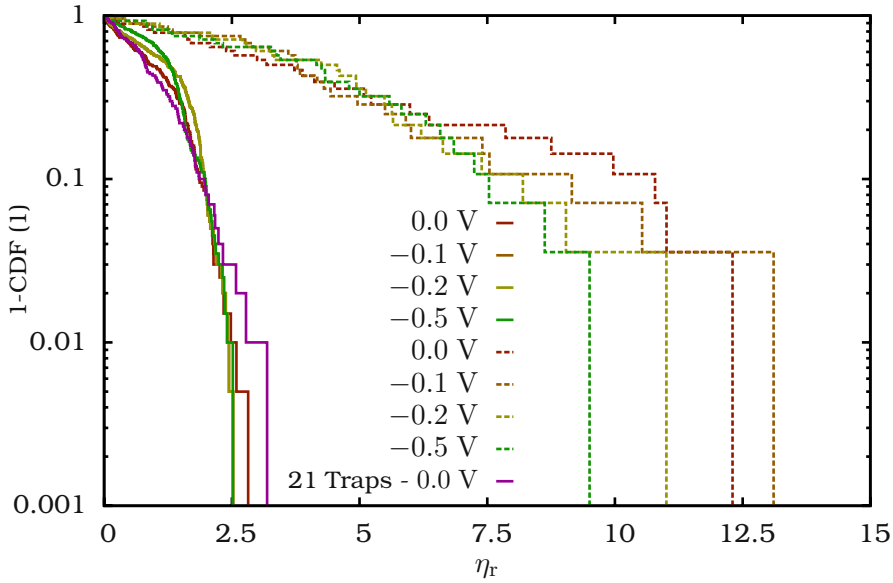


Figure 5.15: Comparison of simulation (solid lines) and measurement data (dashed lines) for a p-channel MOSFET with a single occupied trap, where ΔV_{th} has been normalized using the charge sheet approximation. Various CDFs as a function of gate overdrive as well as the CDF for 21 traps from Figure 5.14 are plotted. None of the simulation data, even the one for 21 traps, are close to the measurement and are off by a factor 5. This is a hint that the purely electrostatic picture to determine step heights is not entirely correct. The measurement data has been provided by imec.

5.3 Random Discrete Doping and a SHE of the BTE

To test the hypothesis from Section 5.2.3 the density gradient quantum correction model as well as the random discrete dopant algorithm detailed in Section 5.1.1 have been employed. As a proof of concept a n-channel MOSFET in inversion has been simulated using ViennaSHE [65] (cf. Figure 5.16).

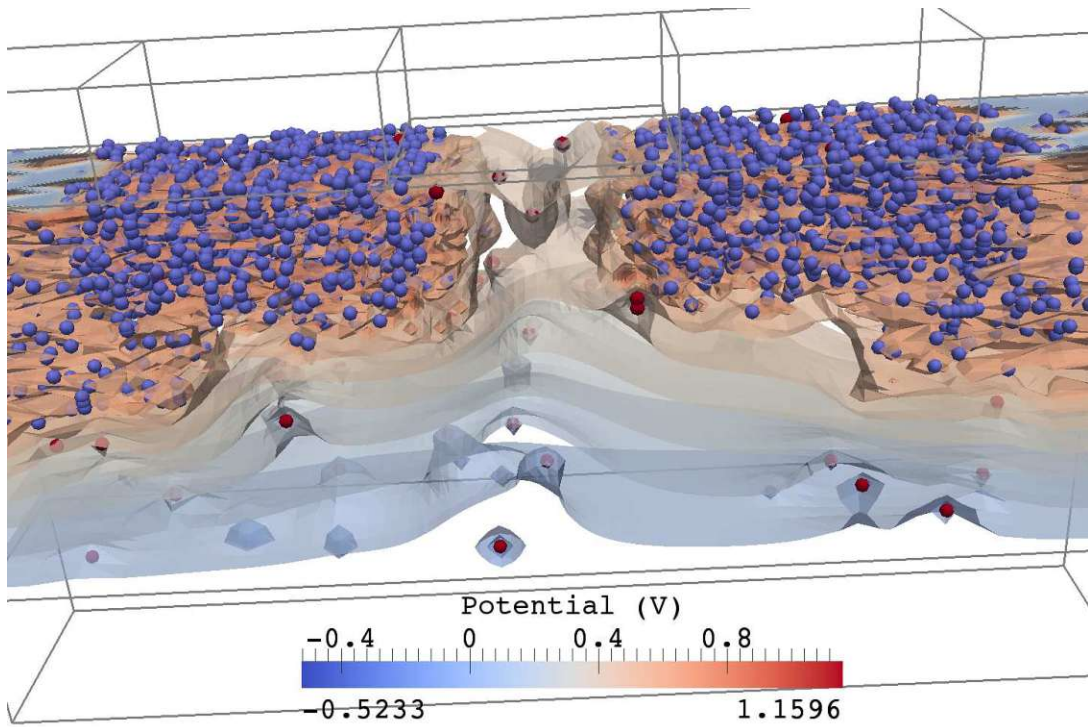


Figure 5.16: A 25 nm n-channel MOSFET simulated using SHE, density gradient and random discrete doping in inversion. The MOSFET featured a channel doping of $N_D = 5 \times 10^{18} \text{ cm}^{-3}$ and an effective oxide thickness of 1 nm. The potential fluctuations are clearly visible.

Es hat doch keiner eine Ahnung was da wirklich im Oxid vor sich geht.

(Franz Schanovsky)

6 Bias Temperature Instability

The Bias Temperature Instability (BTI) is a time, temperature and electric field dependent effect encountered in metal-oxide-semiconductor (MOS) field effect devices, leading to a drift in the threshold voltage V_{th} [29, 112, 113] (cf. Figure 6.1). The drift in threshold voltage is usually measured as a drift in drain I_d or source I_s current over time. Depending on the gate voltage V_g , it is either referred to as negative BTI (NBTI), if $V_g \leq 0$, or as positive BTI (PBTI), if $V_g > 0$. It has been established that BTI is due to the formation of chargeable defects inside the gate insulator or directly at the interface of the gate insulator in MOS devices, especially MOSFETs [29, 114, 115]. The exact influence of temperature, oxide field and stress time on the time evolution of ΔV_{th} has been established by carefully designed experiments, which will be briefly introduced in Section 6.1. Any differences in the setup lead to a misinterpretation of the measurement data and unreproducibility (by other researchers) of the experiment [116]. A typical BTI experiment (cf. Figure 6.2) involves a temperature-controlled environment, fast voltage and if possible temperature transients as well as fast and highly accurate measurement equipment [117]. For simulation of BTI, it is important to design the simulation such that the input (voltages, temperature, stress time, etc.) over time is as close as possible to the design of the experiment in order to avoid any errors. In this chapter the intricacies involved in numerically assessing BTI and numerically reproducing measurement results will be laid out.

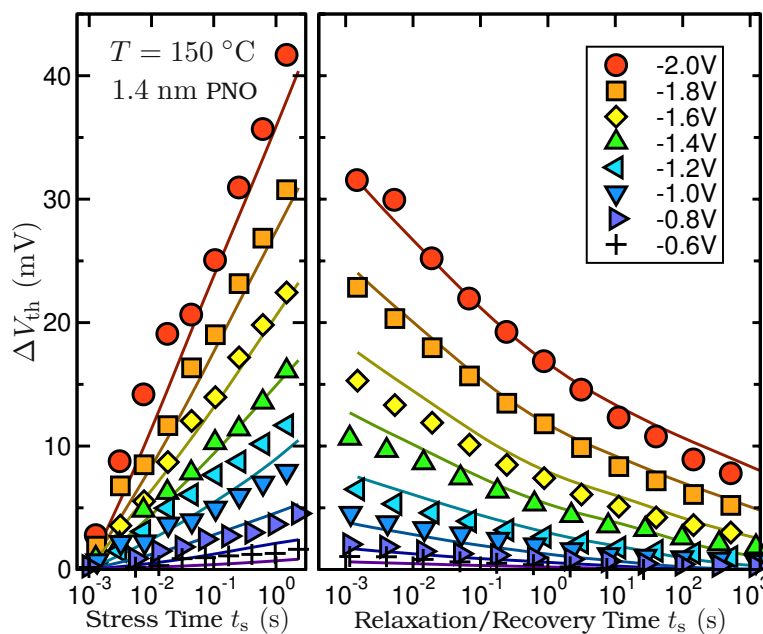


Figure 6.1: Threshold voltage drifts at elevated temperatures depending on the oxide field. The plot shows the result of a measure-stress-measure experiment (symbols) at 150 °C for various stress gate voltages. During recovery the gate voltage was kept at the nominal threshold voltage. Between the end of stress and the first measurement point in recovery is a delay of 1 ms. For comparison a fit of the Two-Stage model to the data (lines) is also shown. Most evident is the asymmetry in time between stress and recovery. Data are taken from [118].

The origins and physical descriptions for BTI available so far will be briefly summarized in Section 6.2. For an exhaustive discussion of the physical background, historic BTI models and mathematical modelling the reader is referred to [29, 112, 113]. Historically the role of hydrogen in the gate oxide is of high importance. Nevertheless, this work will not cover the influence of hydrogen in the gate oxide on the degradation. Details on the influence of the hydrogen concentration in the oxide on BTI is for example given in [119, 120].

6.1 Measurement Techniques

For an explanation of many BTI experiments it is instructive to introduce one often employed basic element of a BTI experiment first. In this work this basic element is termed a single ‘stress-relax’ cycle for BTI (cf. Figure 6.2). Before the experiment starts, the fresh device is characterized by measuring the $I_d - V_g$, $I_d - V_d$ or capacitance-voltage curves, while taking great care not to significantly change the device characteristics by prematurely BTI stressing the device. However, stressing the pristine device in this initial characterisation stage is often unavoidable. For the experiment itself the drain voltage V_d is often regulated to be as small as possible to guarantee low field conditions. Nevertheless, constraints imposed by the measurement setup often require a higher V_d , to for example, minimise noise in the measurement data. Whenever one is assessing homogenous BTI and not interested in the V_{th} shift either during stress or relaxation, then V_d is often chosen to be at zero volts during the respective phase. While the gate voltage V_g is precisely controlled to cycle between stress and relaxation, the drain current can be recorded to measure the threshold voltage deviation ΔV_{th} . If a fast heater, such as a poly-heater device [121], is available it is also possible to cycle the temperature in the same fashion. In a setup where the temperature can be cycled, the device temperature during stress T_s is usually much lower than the relaxation temperature T_r . The relaxation gate voltage V_r is normally chosen to be equal to the nominal threshold voltage V_{th0} . Next the gate voltage is set to the stress value, where V_s normally corresponds to strong inversion for a well defined time t_s (stress). After bias temperature stress the gate voltage is, ideally instantly, set back to its relaxation value (relax/measurement) and kept there for a given time t_r until the experiment ends. During this stress-relax cycle the drain current is recorded and compared to the initial measurement taken before the experiment. Due to the fast transient nature of BTI often a logarithmic time-stepping in the recording of the drain current is chosen to capture the transient behavior of the drain current right after bias or temperature changes. Additionally, most often one is restricted by the measurement range and bandwidth of the equipment and can only measure the drain current over the recovery time (recovery trace). In this work, BTI is defined as the set of gate voltages V_g and device temperatures T , which cause a change in the threshold voltage V_{th} in a given time. Thus we define bias temperature stress as an increase in threshold voltage $|\Delta V_{th}|$ and drain current shift $|\Delta I_d|$ over time. Relaxation is defined as a decreasing threshold voltage $|\Delta V_{th}|$ and drain current shift $|\Delta I_d|$ over time bringing the actual threshold voltage V_{th} closer to its nominal value V_{th0} .

6.1.1 Measure Stress Measure Technique

The measure stress measure (MSM) technique is a succession of multiple stress and $I_d - V_g$ measurement cycles [122]. First an initial $I_d - V_g$ curve is measured on the fresh device. Then the device is subjected to bias temperature stress for t_s seconds. Right after stress a final $I_d - V_g$ curve is taken. Note that during the measurement of the final $I_d - V_g$ curve the device unavoidably relaxes. This cycle can be repeated many times in order to measure

'Stress-Relax' Cycle

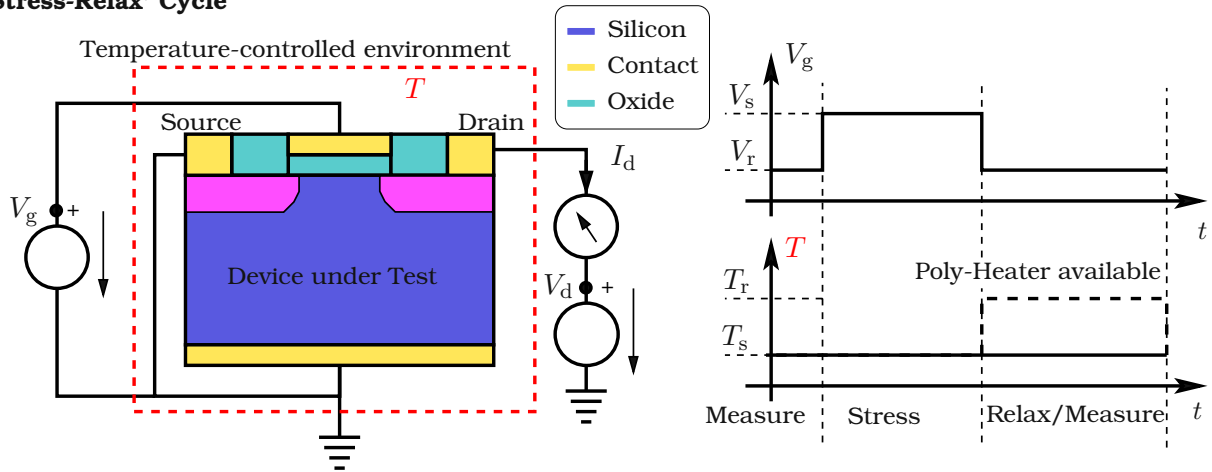


Figure 6.2: A typical setup for a single stress-relax cycle at constant drain voltage and temperature. In such an experiment, the drain potential V_d is usually regulated to be as small as possible (low field conditions). The gate voltage V_g is regulated to cycle between stress, recovery and measurement phases. If a fast heater, such as an integrated poly-heater [121], is available the temperature can also be switched reliably and fast. Initially, V_g is kept at the relaxation voltage, V_r , to measure the undegraded drain current I_d (measure).

the degradation over various stress times. In an extended MSM measurement one records several relaxation phases after single exponentially growing stress phases, where the device temperature is kept constant during the whole experiment. The actual extended experiment, for bias stress, is shown in Figure 6.3. In the MSM technique the threshold voltage shift is obtained by comparing the measured drain current at a certain gate voltage against an initially measured $I_d - V_g$ curve. This is possible since the gate voltage for relaxation is chosen to be equal close to the nominal threshold voltage. Additionally, it is also a possibility to record a fast $I_d - V_g$ curve just before or during switching the gate voltage from V_s to V_r . In [123] it was shown that the MSM technique is quite insensitive to mobility changes induced by stress. Nevertheless, it was also shown that the mobility variations induced are linearly dependent on temperature. This dependence has to be taken into account, when comparing MSM measurements taken at different device temperatures.

6.1.2 On-the-Fly Technique

The on-the-fly (OTF) technique is a method of extracting the threshold voltage shift from the recorded drain current with different levels of accuracy and not a separate measurement setup. In OTF the first recorded drain current under stress conditions at a fixed drain voltage is used as a reference to determine the threshold voltage shift. However, due to an inherently unavoidable delay between the onset of stress and the first recorded drain current there will always be an error in the reference drain voltage. Thus one is obliged to keep this delay in the first measurement point as small as possible in order to minimize this systematic error. This is also the major drawback of any OTF method. To extract the threshold voltage shift induced by operating the device under stress conditions, usually a SPICE Level 1 model is used [116, 124, 125]. In the simplest method (OTF1), the drain voltage V_d and the SPICE parameter θ are assumed to be small and the effective mobility μ_{eff} (SPICE parameter) is assumed to be constant throughout the experiment. With these assumptions the threshold voltage shift ΔV_{th}

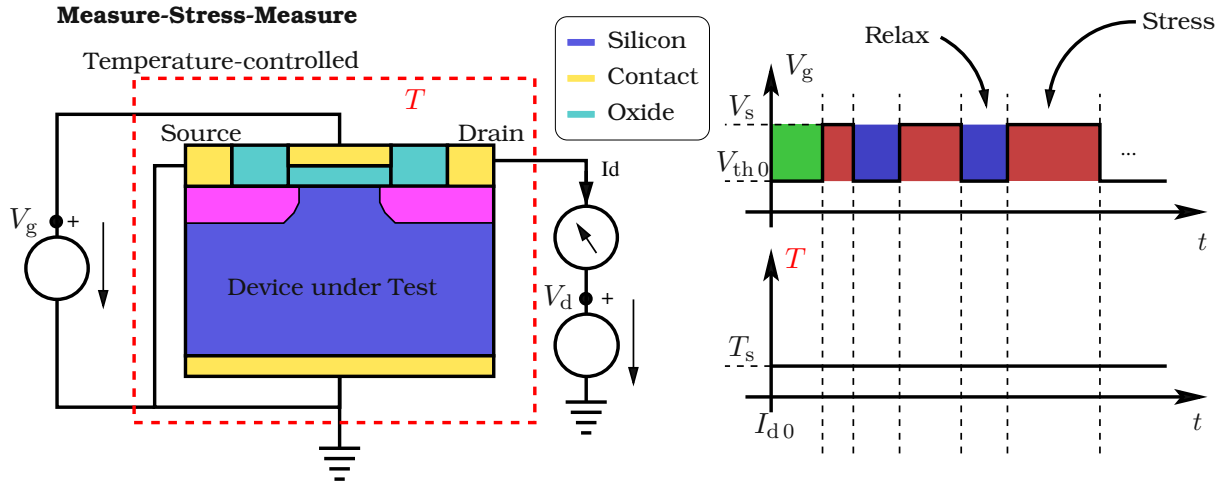


Figure 6.3: The basic setup of a measure stress measure experiment. Initially the device is characterized and at least an $I_d - V_g$ curve at a fixed drain voltage is measured (green). Then the device is periodically switched between stress (red) and relax (blue) conditions, while the drain current is being measured. With each new stress cycle the stress time t_s is increased exponentially. If a poly-heater device is available, it is possible to accelerate the recovery by switching to higher temperatures during recovery, else the device temperature is kept at a predefined stress level throughout the experiment.

in the OFT1 method can be obtained by

$$\Delta V_{th} \approx \frac{I_d - I_{d0}}{I_{d0}} (V_g - V_{th0}), \quad (6.1)$$

where I_{d0} is the reference drain current and V_{th0} is the threshold voltage corresponding to I_{d0} . Since the OFT1 method cannot, due to the assumptions made, predict or at least compensate for mobility changes, other OTF methods [126] have been developed. Nevertheless, all of them suffer from the unavoidable measurement delay between stress and the first measurement point I_{d0} . In addition, all OTF methods also feature the inherent modelling error of the employed SPICE models to determine ΔV_{th} . An example of recorded stress and recovery traces using the OTF method is shown in Figure 6.1. However, due to the inherent errors in OTF data the MSM-method and variants thereof are often used, especially when one is only interested in the recovery traces.

6.1.3 Direct Current Current Voltage

First introduced in [127] and [128], the direct-current-current-voltage (DCIV) method is used to directly monitor the defect density by measuring the bulk current I_b , which is the result of the carrier recombination in the oxide and at the silicon-oxide interface (cf. Figure 6.4). In order to monitor the stress induced degradation, DCIV experiments [127, 128] are performed on fresh devices before and after stress using a drain voltage V_d high enough to forward bias the pn junctions. When assessing bias temperature stress, a DCIV experiment is conducted before and after stress to compare the defect densities before and after stress.

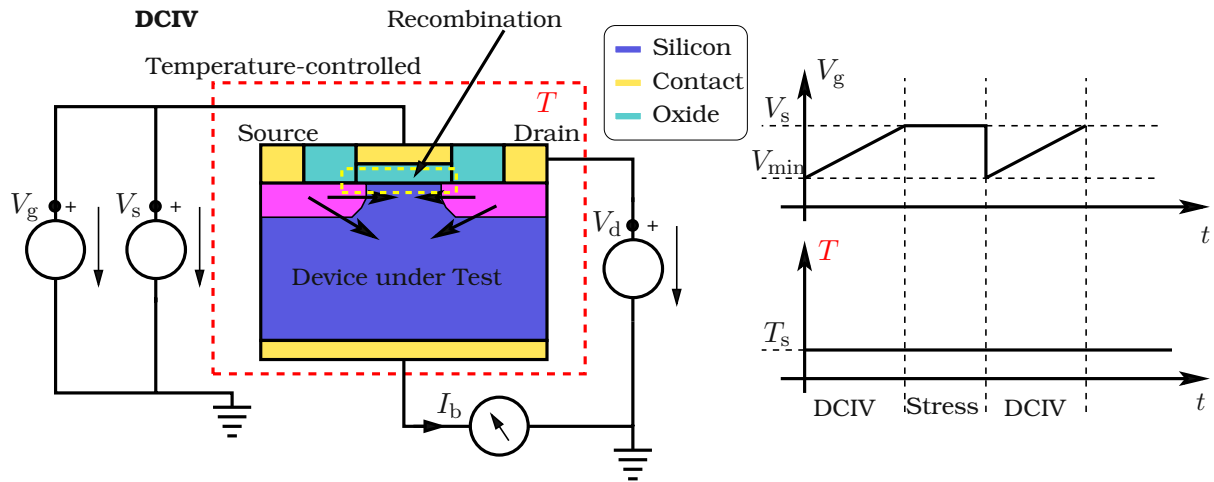


Figure 6.4: In this example, the DCIV method is used to assess the threshold voltage shift caused by bias temperature stress in a single stress pulse. Before stress a DCIV curve is recorded. In a single DCIV measurement the device is swept from accumulation to inversion by changing the gate voltage, while the source/drain pn junctions are forward biased. This forward bias allows the injection of minority carriers into the space charge region, where they may recombine with an oxide trap, causing a measurable bulk current. Then the drain and source voltages are switched to low field conditions. At the same time the gate voltage is kept at the stress level. Right after stress another DCIV measurement is carried out. By comparing the post-stress and the pre-stress DCIV curves one can, using a suitable model, obtain the concentration of interface, oxide or border traps.

6.1.4 Time Dependent Defect Spectroscopy

Time dependent defect spectroscopy (TDDS) is a data analysis technique to assess border traps in the oxides of MOSFETs, where the devices need to be sufficiently small in order to be able to discriminate different traps [105]. Figure 6.5 shows the ΔV_{th} recovery traces recorded after bias temperature stress on a small area ($L \times W = 2 \mu\text{m} \times 160 \text{ nm}$) n-channel MOSFET. The only assumption TDDS does require is that the step-height and emission time of a single defect/charge carrier trap in the oxide (cf. Chapter 4) uniquely characterizes a particular trap. The technique itself works as follows: First a statistical significant number of subsequent ‘stress-relax’ cycles on a single device at a certain but fixed drain voltage are recorded for a certain but fixed stress temperature. The recovery traces can then be compared by accounting for the residual degradation from the previous relax phase. Inspecting ΔV_{th} over recovery time from each relaxation phase and employing the initial assumption that each trap can be uniquely identified by a characteristic step-height one is able to calculate the characteristic time constant, e.g. the emission time τ_e [129]. In Figure 6.6 the extraction technique is illustrated. Furthermore, TDDS can be used to produce so called discrete Capture-Emission-Time (CET) maps [29]. To this end, the TDDS method shown in Figure 6.6 is not only applied to the recovery traces but also to the ΔV_{th} recorded during the stress phase of the stress-relax experiments. Then by identifying the various traps by the individual step heights they cause it is possible to combine these two TDDS maps to a single discrete CET map.

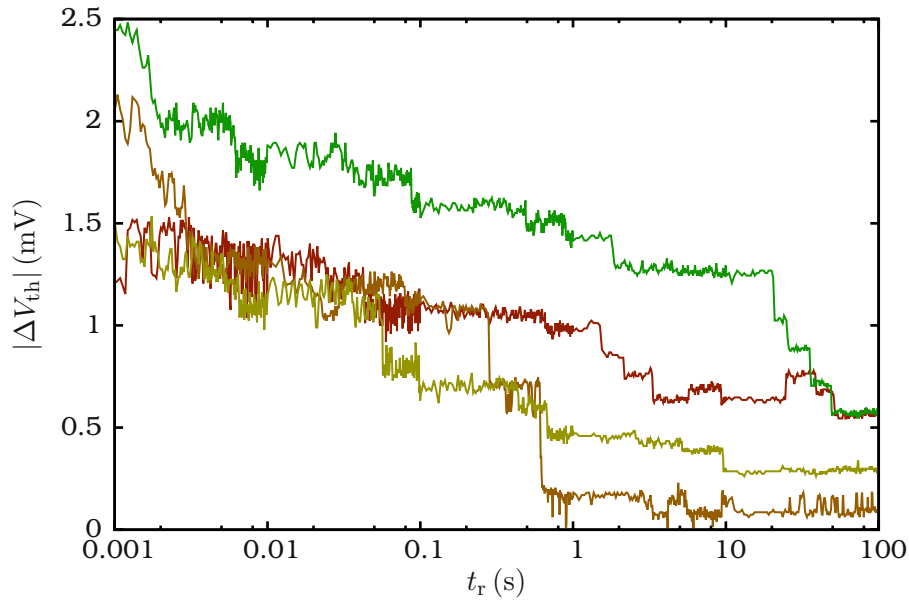


Figure 6.5: Five recovery traces from a stress-relax experiment. The experiment has been carried out in a temperature controlled environment on a n-channel MOSFET, where the ΔV_{th} has been calculated from ΔI_d using an initial $I_d - V_g$ curve. Multiple of these traces are used to identify the capture and emission times of single traps. The data has been provided by Michael Waltl.

6.2 Models for the Bias Temperature Instability

As stated in the beginning of this chapter, BTI is, as its name suggests, a gate voltage, device temperature and time dependent threshold voltage shift ΔV_{th} , which indicates a charge buildup in the oxide of the investigated MOS structures. As the magnitude of ΔV_{th} over stress time t_s can be expressed by a power law [130, 131], this spawned a debate over the correct power law exponent. At this stage of research recovery after bias temperature stress has only been superficially investigated and the reaction-diffusion (RD) models [130, 132] could explain the published measurement data. All flavours of the RD models assume a charge buildup at the semiconductor-oxide-interface through hydrogen-based reaction and diffusion. It was soon noticed that the measurement techniques and all time transients differ from publication to publication and that time delays and transients need to be well defined. Additionally, measurement and ΔV_{th} extraction techniques for BTI have been investigated [123]. Utilizing well defined experiments the quantitative relations between stress time, recovery time, device temperature and gate bias or oxide field can be identified. As such it was, for example, found that ΔV_{th} depends quadratically on the magnitude of the oxide field. However, the authors of [115] had to introduce a permanent component of BTI, due to the limited recovery time in the experiments. This permanent component is still under debate and is necessary to reproduce recovery traces, due to the stark asymmetry of stress and recovery over time as can also be seen in Figure 6.1. By asymmetry we mean that the recovery of 1 mV of ΔV_{th} after BTI stress takes longer than the buildup of 1 mV of ΔV_{th} during stress, provided the device temperature is kept constant. Thus when applying a model to experimentally recorded ΔV_{th} data, the data outside the measurement window is accounted for by a permanent non-recoverable component although a plateau of recovery has to date never been, reproducibly, observed [133]. With the

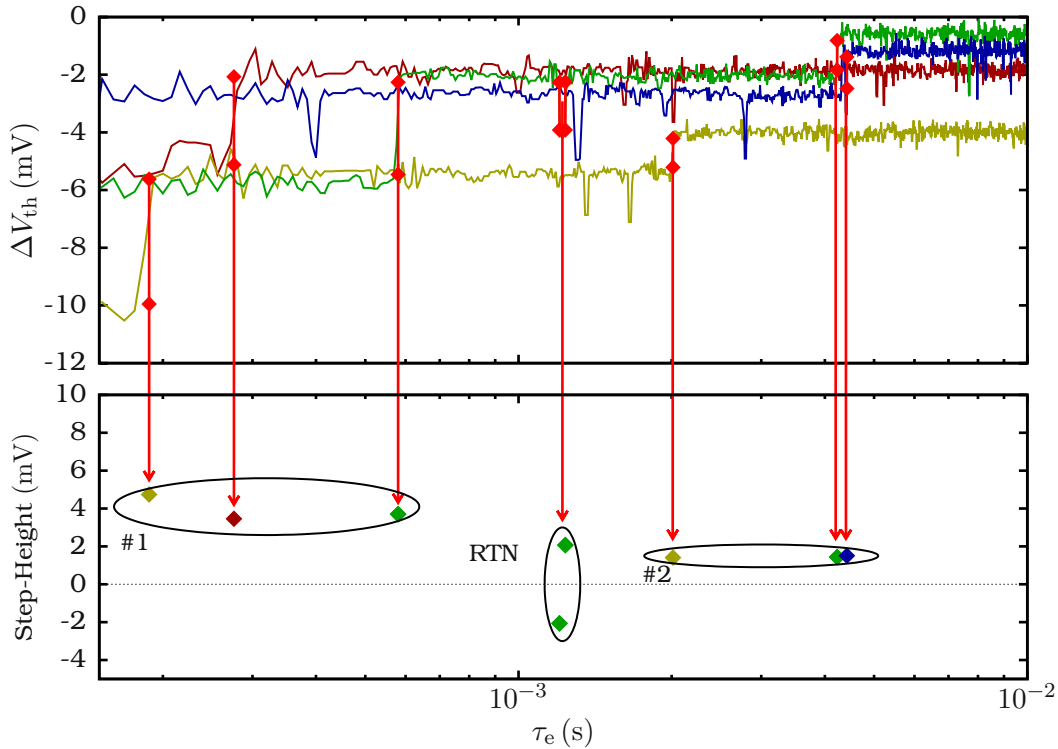


Figure 6.6: Assembly of a TDDS map from drain current measurements during device relaxation. All recorded recovery traces are plotted on the same time axis, where 0 s corresponds to the onset of relaxation. Then the characteristic step-heights are identified by an algorithm published in [129]. Each step-height corresponds to a capture/emission event and its time of occurrence and step-height are marked (diamonds) in a map (lower figure). Clusters of points in this map most likely correspond to the same trap. This way single traps and their time constants can be identified. Taken from [129].

published amount of data, it was finally conclusively found that any reaction diffusion model, particularly on the microscopic scale, cannot possibly explain all the observed BTI characteristics [112]. Finally, with the emergence of time dependent defect spectroscopy (TDDS), it was possible to investigate the charge capture and emission of single defects in small-area MOSFETs and to identify a multiphonon process as the charge exchange process between the substrate and the defects in the oxide.

In the course of this thesis three different models for BTI have been implemented into the drift diffusion simulator MinimosNT. First a simple two well model, second the Two-Stage-Model [134] and last a four state non-radiative multiphonon (NMP) model [4]. In contrast to the four-state NMP model alone, which can reproduce the recoverable component, the Two-Stage-Model predicts both the recoverable and the permanent component of BTI. Thus, the four-state NMP model is often combined with a two well model, since the two well model is used to reproduce the permanent component of BTI. The permanent degradation is usually modelled as a buildup of interface states, whereas the recoverable component is modelled as formation and annealing of oxide traps. Thus the two well model describes the time evolution of N_{it} , whilst the four-state NMP model describes the time evolution of N_{ot} . The Two-Stage model, in contrast to the other two models, includes a description of N_{ot} (recoverable) and N_{it} (permanent) over time. Since the four state NMP model combined with the two well model can

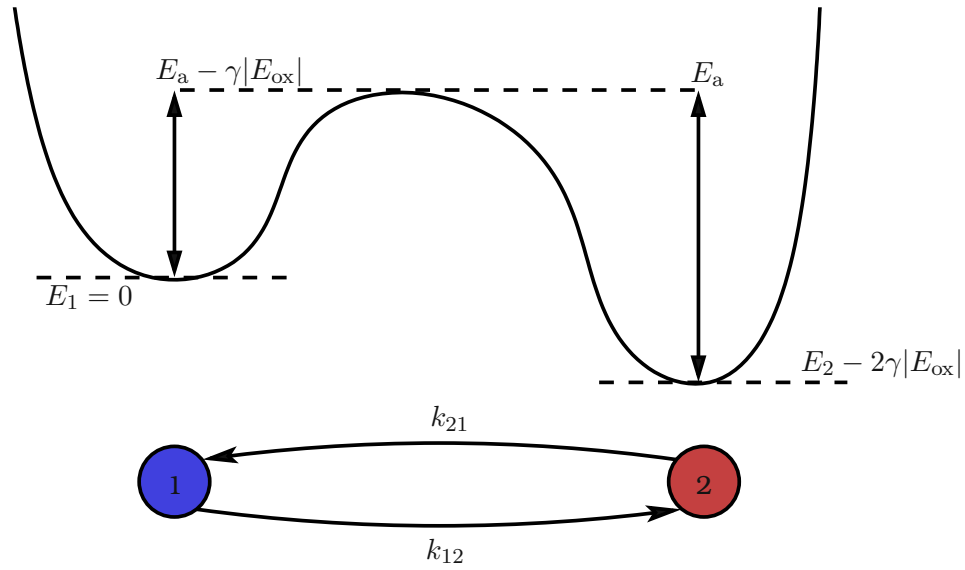


Figure 6.7: A phenomenological model for BTI. In the precursor state 1 the defect is neutral and can, by being thermally activated, undergo a transition into the active trapping state 2, where it can capture and emit charge carriers. The energy barrier which must be overcome by the defect can, in this model, be lowered by an oxide field. In this model the parameters γ and ν must be found experimentally.

explain more BTI characteristics than the other two models alone, this is the model which will be explained in detail.

6.2.1 Phenomenological Models

In order to model the weak temperature and the quadratic oxide field dependence of the charge capture process causing the threshold voltage shifts many phenomenological models have been developed. In addition to that many authors, e.g. [135], tried to explain the charge capture process by extending the successful SRH [136] trapping model with a tunneling coefficient to account for the fact that a charge trapping defect is not necessarily located directly at the oxide interface. However, it was shown that these models cannot sufficiently explain the physics involved in BTI [113], especially the capture and emission time constants found using TDDS. Nevertheless, a double-well model serves as a good example how interface defects caused by BTI can be modelled. Starting from the SRH model introduced in Section 2.2.2, one can explain the formation of new defects causing a shift in threshold voltage by introducing a model for the trap concentration N_{it} . To capture the oxide field dependence and the temperature dependence of the formation and annealing of defects an Arrhenius-type law is often employed. In a two-well model for N_{it} , a defect can be in two states. It can either be active, where trapping and detrapping of charge carriers is governed by SRH statistics (charge trapping), or inactive, where transitions are modelled by defect reaction rates. The barrier height E_a determining the defect reactions rates is assumed to be oxide field dependent and temperature activated. For the model in Figure 6.7 the defect reaction rates are

$$k_{12} = \nu_0 \exp(-(E_a - \gamma|E_{ox}|)/(k_B T_L)), \quad (6.2)$$

$$k_{21} = \nu_0 \exp(-(E_a - E_2 + 2\gamma|E_{ox}|)/(k_B T_L)), \quad (6.3)$$

where $\nu_0 \approx 10^{13} \text{ s}^{-1}$ is the attempt frequency.

Another often employed approach is to explain the defect buildup by initially electrical inactive traps and hole trapping instead. To this end the trap is assumed to reside in the oxide, a distance x_T away from the semiconductor-oxide interface and the trapping kinetics are described by the SRH model and elastic tunneling [137]. The modified SRH rates per trap read

$$k_{12} = k_{12}^{\text{SRH}} \lambda(x_T, E_{\text{ox}}) \quad \text{and} \quad k_{21} = k_{21}^{\text{SRH}} \lambda(x_T, E_{\text{ox}}), \quad (6.4)$$

where the function λ denotes the elastic tunneling coefficient, usually a modified WKB approximation [138–140], depending on the depth of the trap and on the oxide field energy barrier. In [113] it has been shown that these simple models are insufficient to describe all the observed characteristics of BTI, especially those found by time dependent defect spectroscopy on nanoscale devices.

6.2.2 Non-radiative Multiphonon Transitions

The first researchers to extend the SRH theory [136] formulated a charge trapping theory which was loosely based on non-radiative multiphonon transitions (NMP) [141]. These models considered charge carrier tunneling and a thermally activated, oxide field dependent process as demonstrated in the previous section [142]. In this model [142] the capture and emission time constants, which are obtained from the reaction rates using Markov-Chain theory [143], read

$$\tau_c = \tau_0 \underbrace{\exp\left(\frac{x_T}{x_0}\right)}_{\lambda(x_T)} \frac{N_v}{p} \exp\left(\frac{E_a - E_t + |q| x_T E_{\text{ox}}}{k_B T_L} \theta(E_C - E_t)\right), \quad (6.5)$$

$$\tau_e = \tau_0 \underbrace{\exp\left(\frac{x_T}{x_0}\right)}_{\lambda(x_T)} \exp\left(\frac{E_a + E_t - |q| x_T E_{\text{ox}}}{k_B T_L} \theta(E_t - E_V)\right), \quad (6.6)$$

where E_t is the trap level, E_C and E_V are the respective band edges and θ denotes the Heaviside function [113]. When fitting this model to capture and emission times obtained through TDDS it was realized that one can either obtain a fit for τ_c or τ_e but not for both [113]. Since none of the above models can fully capture the findings obtained by TDDS experiments a new model was required [29]. This new model has been fully based on non-radiative multiphonon theory, which shall be explained in the following.

Basic Theory

The non-radiative multiphonon theory is based on the possible scattering between multiple phonons with a single electron, which eventually gets trapped in the process. It is theorized that this particular electron can, depending on the configuration and bonds of the surrounding atoms, not only emit photons to loose this gained energy, but now has sufficient energy to occupy a previously unoccupied state (charge trapping). What the atomic configuration at the defect site is, is still not completely clear [112]. In order to derive the reaction rates necessary to describe the charge trapping process a plenty of approximations to the full Schrödinger equation (cf. Equation (2.1)) describing the many particle problem are required. First the Born-Approximation is applied to separate the equation into two loosely coupled equations: one for

the system of electrons and one for the system of nuclei. Per electron i and nucleus j the system of equations is

$$(E_{\text{el}} + H_{\text{ion-ion}}(\mathbf{R}) + H_{\text{el-ion}}(\mathbf{r}, \mathbf{R}) + H_{\text{el-el}}(\mathbf{r})) \psi_i(\mathbf{r}, \mathbf{R}) = V_i(\mathbf{R}) \psi_i(\mathbf{r}, \mathbf{R}), \quad (6.7)$$

$$(E_{\text{ion}} + V_i(\mathbf{R})) \phi_{iJ}(\mathbf{R}) = E_{iJ} \phi_{iJ}(\mathbf{R}), \quad (6.8)$$

where $V_i(\mathbf{R})$ denotes the adiabatic potential, which corresponds to the energy of a certain atomic configuration and ϕ_{iI} is the vibrational wave function corresponding to the states i and I . Next, to describe the complex scattering process, first order time-dependent perturbation theory and the Franck-Condon approximation [144, 145] are applied. Putting all these approximations together the reaction rates between the electronic state i , the electronic state j and the respective states I and J of the nuclei can be written as [146]

$$k_{ij} = A_{ij} f_{ij} \quad (6.9)$$

where

$$A_{ij} = 2\pi\hbar^{-1} |\langle \psi_i | V' | \psi_j \rangle|^2 \quad (6.10)$$

is the electronic matrix element, V' is the adiabatic perturbation operator, $\langle \psi_i |$ is the electronic wave function corresponding to state i . In Equation (6.9) f_{ij} is the so-called lineshape function which formally equates to,

$$f_{ij} = \text{avg}_I \left(\sum_J |\langle \phi_{iI} | \phi_{jJ} \rangle|^2 \delta(E_{iI} - E_{jJ}) \right), \quad (6.11)$$

where the operator avg denotes the thermal average operator over all initial states of the nuclei I . The lineshape function describes the likelihood of a transition from the nuclei state I to J and the electronic matrix element describes the electronic transition probability. In the NMP theory both need to be non-vanishing in order to have a defect reaction. For the lineshape function to be non-vanishing the energies of the final and initial state need to be very close. In Equation (6.11) this has been approximated by the Dirac-Delta. Thus the Franck-Condon factor $|\langle \phi_{iI} | \phi_{jJ} \rangle|^2$ is interpreted as a transition probability, determined by the overlap of the two vibrational wave functions ϕ_{iI} and ϕ_{jJ} (cf. Figure 6.8). The lineshape function, its formal derivation, numeric evaluation using data from density functional theory (DFT) calculations and subsequent approximations to Equation (6.11) have been extensively discussed previously [112]. For the following discussion the classic approximation of the line shape function is briefly summarized. Parameterizing the adiabatic potentials (cf. Figure 6.8) using the reaction coordinate concept [147] yields

$$V_i(q) = c_i(q - q_i)^2 + d_i, \quad (6.12)$$

$$V_j(q) = c_j(q - q_j)^2 + d_j, \quad (6.13)$$

where c is the curvature, d is the potential offset, q_i and q_j denote the positions of the adiabatic potentials (cf. Figure 6.8). Additionally, $d_i - d_j$ is usually expressed as a function of $\Delta E = E - E_t$ to fit into the conventions used when describing the energy bands in semiconductors. Using the parameterized potentials to find the intersection points (IP) q_1 and q_2 (cf. Figure 6.8) of the adiabatic potentials and the classic approximation of the line shape function in the high temperature limit [112] one obtains

$$f_{ij} = \frac{\sqrt{c_i}}{2\sqrt{\pi k_B T_L}} \left(\frac{\exp\left(-\frac{c_i q_1^2}{k_B T_L}\right)}{|c_i q_1 - c_j(q_1 - q_j - q_i)|} + \frac{\exp\left(-\frac{c_i q_2^2}{k_B T_L}\right)}{|c_i q_2 - c_j(q_2 - q_j - q_i)|} \right), \quad (6.14)$$

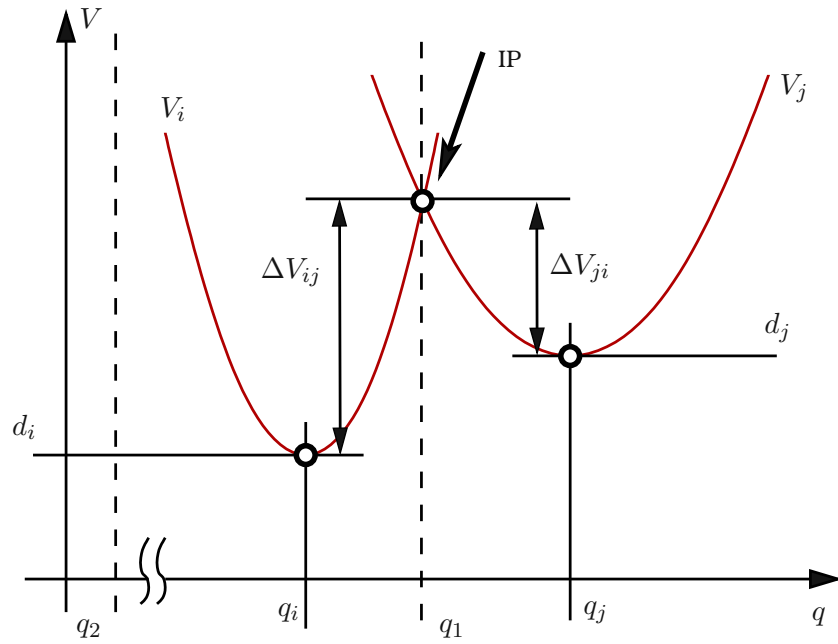


Figure 6.8: Definition of the parameterized adiabatic potentials V_i and V_j . Upon parameterizing the adiabatic potential $V_i(\mathbf{R})$, it is possible to analytically calculate the classic transition barriers ΔV_{ij} and the intersection points q_1 and q_2 , where usually one intersection point dominates the lineshape function (here q_1).

where q_1 and q_2 are the intersection points between $V_i(q)$ and $V_j(q)$. Depending on the actual parameters of the adiabatic potentials V_i and V_j , multiple further approximations are possible to obtain simpler formulae for the lineshape function. The barrier heights from the parameterized potentials, which have been used in the previous equation are

$$\Delta V_{ij} = \frac{c_i(q_j - q_i)^2}{(c_i/c_j - 1)^2} \left(1 \pm \sqrt{\frac{c_i}{c_j} + \frac{(d_i - d_j)(c_i/c_j - 1)}{c_j(q_j - q_i)^2}} \right)^2 \quad (6.15)$$

$$= \frac{S\hbar\omega}{(R^2 - 1)^2} \left(1 \pm R \sqrt{\frac{S\hbar\omega + (E - E_t)(R^2 - 1)}{S\hbar\omega}} \right)^2, \quad (6.16)$$

where the parametrisation with $c_i/c_j = R^2$ and the Huang and Rhys factor $S\hbar\omega = c_i(q_i - q_j)^2$, often found in the literature, has been employed.

Depending on the location of the intersection between V_i and V_j two cases of phonon coupling can be distinguished. When the dominant intersection point q_1 (cf. Figure 6.8) is located between the minima of the potentials ($q_i < q_1 < q_j$) the process is referred to as strong phonon coupling, otherwise it is referred to as weak phonon coupling [113].

Evaluation of Transition Rates in Semiconductors

To obtain useful transition rates for semiconductors, where a continuum of energy levels must be considered, an integration over the electronic energy is necessary. Employing the definition of the density of states (cf. Equation (2.7)) the rates for capturing a hole (or emitting an electron)

from/to the valence or conduction band reads

$$\begin{aligned}
k_{12} = & \nu^p \int_{-\infty}^{E_v} g(E)(1 - f^p(\mathbf{x}, E, t))A_{12}(E, x_T) \underbrace{f_{12}(E - E_t)}_{\Delta E} dE \\
& + \nu^n \int_{E_c}^{\infty} g(E)(1 - f^n(\mathbf{x}, E, t))A_{12}(E, x_T) \underbrace{f_{12}(E - E_t)}_{\Delta E} dE.
\end{aligned} \tag{6.17}$$

The rate for absorbing an electron (or emitting a hole) from/to the valence or conduction band reads

$$\begin{aligned}
k_{21} = & \nu^p \int_{-\infty}^{E_v} g(E)f^p(\mathbf{x}, E, t)A_{21}(E, x_T) \underbrace{f_{21}(E_t - E)}_{\Delta E} dE \\
& + \nu^n \int_{E_c}^{\infty} g(E)f^n(\mathbf{x}, E, t)A_{21}(E, x_T) \underbrace{f_{21}(E_t - E)}_{\Delta E} dE.
\end{aligned} \tag{6.18}$$

For repetitive evaluation it is often beneficial to approximate the energy barrier (Equation (6.15)) by expanding Equation (6.15) into a Taylor series about ΔE up to the quadratic term,

$$\Delta V_{12} \approx \frac{S\hbar\omega}{(1+R)^2} + \frac{R\Delta E}{1+R} + \frac{R\Delta E^2}{4S\hbar\omega}. \tag{6.19}$$

Now, in order to evaluate the transition rates numerically, the electronic matrix elements A_{ij} need to be evaluated. The evaluation of these functions strongly depends on the quantities available from the employed transport model. Thus when using a Schrödinger-Poisson solver together with the transport model one can directly evaluate integrals of the electronic wave functions. For semi-classic transport models, such as SHE or the drift diffusion model, one needs to suitably approximate the electronic matrix elements. Assuming that the wave function of the trapped charge carrier is strongly localized around the defect site, one can approximate the electronic matrix element by an energy dependent tunneling coefficient and a prefactor, usually a WKB-approximation of charge carrier tunneling [113]. All factors, which are not dependent on energy can be pulled out of the integrand and are summarized in the parameter σ_0 . Having a solution of the Boltzmann Transport equation, the integrals (cf. Equation (6.17) and Equation (6.18)) can be directly evaluated. However, for moment based transport models further assumptions regarding the density of states and the energy distribution functions are necessary. For drift diffusion models Maxwellian distribution functions and a parabolic band are usually assumed. With these approximations the non-radiative multiphonon transition rates read

$$\begin{aligned}
k_{12} = & \sigma_0 \exp\left(-\frac{x}{x_T}\right) v_{th}^p p \exp\left(-\frac{\Delta V_{12}}{k_B T_L}\right) \\
& + \sigma_0 \exp\left(-\frac{x}{x_T}\right) v_{th}^n n \exp\left(-\frac{\Delta V_{12}}{k_B T_L}\right) \exp\left(-\frac{E_t - E_c + E_F}{k_B T_L}\right)
\end{aligned} \tag{6.20}$$

$$\begin{aligned}
k_{21} = & \sigma_0 \exp\left(-\frac{x}{x_T}\right) v_{th}^p p \exp\left(-\frac{\Delta V_{12}}{k_B T_L}\right) \exp\left(-\frac{E_t + E_v - E_F}{k_B T_L}\right) \\
& + \sigma_0 \exp\left(-\frac{x}{x_T}\right) v_{th}^n n \exp\left(-\frac{\Delta V_{12}}{k_B T_L}\right)
\end{aligned} \tag{6.21}$$

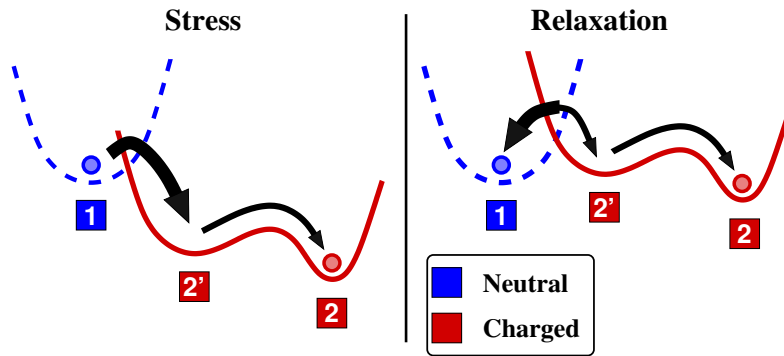


Figure 6.9: Schematic representation of the adiabatic potentials used to define transition rates for a defect with three states. During stress an NMP transition from state 1 to state 2' is very likely (bold arrow). In relaxation, when there is a small oxide field, the reverse reaction (from 2' to 1) is more likely. However, the likelihood for the defect to undergo a structural relaxation (red potentials) does not change with the oxide field (relaxation/stress). Once the defect has undergone a transition from state 2' to state 2 (structural relaxation) it cannot, in this model, directly exchange charges with the substrate.

6.2.3 Structural Relaxation

In [121] the strong temperature dependence of BTI was explicitly shown. Also in TDDS measurements an oxide field independent, but temperature dependent regime was found [105]. To account for these findings it was theorized that this is due to a structural relaxation of the nuclei forming the defect site after the non-radiative multiphonon transition took place. Due to structural relaxation the effective energy barriers for a subsequent NMP transition change. Such a mechanism can also explain the strongly temperature dependent recovery, where a high temperature accelerates the recovery. It was theorized and investigated by [148] that a dimer configuration or a hydrogen bridge can be broken by capturing a hole (or emitting an electron) by an NMP transition and that the resulting configuration can structurally relax such that the inverse NMP transition becomes less likely (cf. Figure 6.9). It is assumed that the process of structural relaxation is solely dependent on the temperature. Thus the transition rates are modelled using an Arrhenius law,

$$k_{2'2} = \nu_0 \exp\left(-\frac{E_{2'2}}{k_B T_L}\right) \quad \text{and} \quad k_{22'} = \nu_0 \exp\left(-\frac{E_{22'}}{k_B T_L}\right), \quad (6.22)$$

where $E_{2'2}$ and $E_{22'}$ are the constant barriers and $\nu_0 \approx 10^{13} \text{ s}^{-1}$. Structural relaxation was shown to be essential to describe DCIV experiments [149, 150] and the possible field independence of charge emission times τ_e found in TDDS data.

6.2.4 The four State NMP Model

NMP as a charge capture and emission process as well as structural relaxation have been combined by [151] into a four state model (cf. Figure 6.11 and Figure 6.10), which was shown to successfully describe the capture and emission time constants found in TDDS experiments. The model is constructed such that a defect can in each state either exchange charges with the substrate or undergo structural relaxation. However, it is noteworthy that this model does not attempt to explain the step-heights of single defects. Instead the ΔV_{th} contribution of

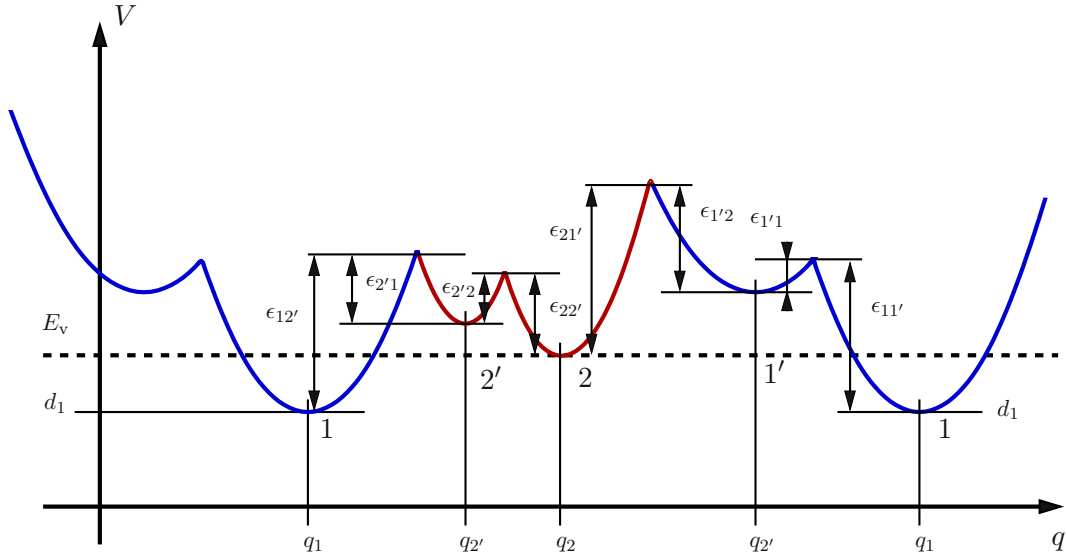


Figure 6.10: Definition of the adiabatic potentials for the oxide trap model from [151]. The sketch shows a finite state diagram for a single defect close to the valence band. State 1 is the stable and electrically neutral precursor state. Via an NMP transition (intersection between blue and red parabolas) the defect goes into state 2' and the defect becomes charged but remains meta-stable. A defect in state 2' can, upon the emission of a hole, undergo a transition back to the neutral and stable precursor state 1. The charge carrier exchange processes are modeled using NMP theory.

each defect is attributed to the interaction between the charged defect and the random discrete dopants in the channel underneath the defect (cf. Chapter 4). In order to determine whether or not a defect is charged, a system of equations describing the state transitions per defect are required. To this end, the framework of first-order continuous-time Markov-Chains [143] can be directly applied [29]. A single defect can only be in one state at a time. More precisely $X_i(t)$ is the random variable for state i at time t , which is exactly 1 if the defect is in the i th state and 0 otherwise. The condition that any defect has to be in any of its N states, where N is a finite integer number, can be expressed by

$$\sum_i^N X_i(t) = 1. \quad (6.23)$$

Next, the probabilities for state transitions need to be defined. Since a transition from i to j depends on the occupation probabilities $X_i(t)$ and $X_j(t)$, where $X_i(t) = 1$ and $X_j(t) = 0$, one can only define conditional probabilities. Formally, for an infinitesimal small time step dt the conditional transition probability from state i to j reads,

$$P(X_j(t + dt) = 1 | X_i(t) = 1) = k_{ij}dt + \mathcal{O}(dt), \text{ where } \lim_{dt \rightarrow 0} \mathcal{O}(dt) = 0. \quad (6.24)$$

In the above equation k_{ij} are the transition rates, which have been covered in the preceding sections. Inserting Equation (6.24) into Equation (6.23) and defining that

$$p_i(t) = P(X_i(t + dt) = 1 | X_i(t) = 1) \quad (6.25)$$

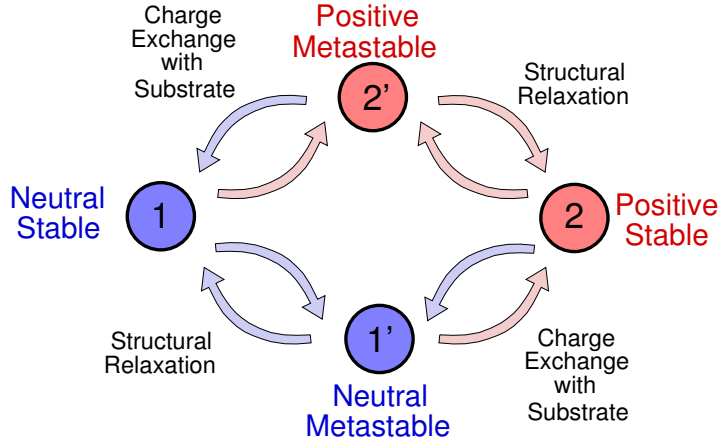


Figure 6.11: The oxide trap model from [151]. The figure shows a finite state diagram for a single defect, in which state 1 is the stable and electrically neutral precursor state. The charge carrier exchange processes with the substrate are modeled using NMP theory. Upon hole capture (red arrow from state 1 to state 2') the defect becomes positively charged but remains meta-stable. A defect in state 2' can, upon the emission of a hole, undergo a transition back to the neutral and stable precursor state 1. Alternatively, a defect in state 2' can undergo the slow process of structural relaxation, become stable and stay positively charged (state 2). In state 2 the defect can either go back into state 2' or can emit a hole, thus becoming electrically neutral and return to state 1'. In state 1' the defect is neutral and can either undergo a transition into the stable precursor state 1 by structural relaxation or can capture a hole and thus change into the stable positively charged state 2.

is the probability for the defect to stay in a certain state, one obtains

$$p_j(t + dt) = k_{ij}dt p_i(t) - k_{ji}dt p_j(t). \quad (6.26)$$

Equation (6.26) can now be written as

$$\partial_t p_j(t) = k_{ij}p_i(t) - k_{ji}p_j(t). \quad (6.27)$$

In addition Equation (6.23) can be reexpressed to

$$\sum_i^N p_i(t) = 1. \quad (6.28)$$

Equation (6.27) and Equation (6.28) yield $N + 1$ equations for N unknowns $p_i(t)$. Since the system is overdetermined one can omit a single equation from the equation system. The $p_i(t)$ are occupancy probabilities, which determine the probability of the defect to be in state i . Thus the $X_i(t)$ are at a single instance in time either 1 or 0 with the probability $p_i(t)$. Sometimes one is only interested in the expectation values of the occupancies. Consequently, applying the expectation operator $E\{\}$ to Equation (6.27) and Equation (6.28) one obtains

$$\partial_t p_j(t) = k_{ij}p_i(t) - k_{ji}p_j(t) \Rightarrow \partial_t f_j(t) = k_{ij}f_i(t) - k_{ji}f_j(t), \quad (6.29)$$

$$E \left\{ \sum_i^N X_i(t) = 1 \right\} \Rightarrow \sum_i^N f_i(t) = 1, \quad (6.30)$$

where $E\{X_i(t)\} = f_i(t)$, meaning that $f_i(t)$ can have any value between zero and one. It is usually referred to as occupancy of state i (cf. Chapter 2).

Evaluation of the Four State Model for the Drift Diffusion Model

Putting all of the above together, one is able to write down the system of equations for the four state NMP model from Figure 6.11 and Figure 6.10. The full, time averaged system for the drift diffusion model reads,

$$k_{12'} = \sigma_0 \lambda(x_T) v_{\text{th}}^p p \exp\left(-\frac{\epsilon_{12'}^v}{k_B T_L}\right) + \sigma_0 \lambda(x_T) v_{\text{th}}^n N_c \exp\left(-\frac{\epsilon_{12'}^c}{k_B T_L}\right), \quad (6.31)$$

$$k_{2'1} = \sigma_0 \lambda(x_T) v_{\text{th}}^p N_v \exp\left(-\frac{\epsilon_{2'1}}{k_B T_L}\right) + \sigma_0 \lambda(x_T) v_{\text{th}}^n n \exp\left(-\frac{\epsilon_{2'1}}{k_B T_L}\right), \quad (6.32)$$

$$k_{2'2} = \nu_0 \exp\left(-\frac{\epsilon_{2'2}}{k_B T_L}\right), \quad k_{22'} = \nu_0 \exp\left(-\frac{\epsilon_{22'}}{k_B T_L}\right), \quad (6.33)$$

$$k_{11'} = \nu_0 \exp\left(-\frac{\epsilon_{11'}}{k_B T_L}\right), \quad k_{1'1} = \nu_0 \exp\left(-\frac{\epsilon_{1'1}}{k_B T_L}\right), \quad (6.34)$$

$$k_{1'2} = \sigma_0 \lambda(x_T) v_{\text{th}}^p p \exp\left(-\frac{\epsilon_{1'2}^v}{k_B T_L}\right) + \sigma_0 \lambda(x_T) v_{\text{th}}^n N_c \exp\left(-\frac{\epsilon_{1'2}^c}{k_B T_L}\right), \quad (6.35)$$

$$k_{21'} = \sigma_0 \lambda(x_T) v_{\text{th}}^p N_v \exp\left(-\frac{\epsilon_{21'}^v}{k_B T_L}\right) + \sigma_0 \lambda(x_T) v_{\text{th}}^n n \exp\left(-\frac{\epsilon_{21'}^c}{k_B T_L}\right), \quad (6.36)$$

$$1 = f_1 + f_{2'} + f_2 + f_{1'}, \quad (6.37)$$

$$\partial_t f_1 = k_{2'1} f_{2'} + k_{1'1} f_{1'} - (k_{12'} + k_{11'}) f_1, \quad (6.38)$$

$$\partial_t f_{2'} = k_{12'} f_1 + k_{22'} f_2 - (k_{2'1} + k_{2'2}) f_{2'}, \quad (6.39)$$

$$\partial_t f_2 = k_{2'2} f_{2'} + k_{1'2} f_{1'} - (k_{22'} + k_{21'}) f_2, \quad (6.40)$$

$$\partial_t f_{1'} = k_{11'} f_1 + k_{21'} f_2 - (k_{1'1} + k_{1'2}) f_{1'}, \quad (6.41)$$

where $\nu_0 \approx 10^{13}/\text{s}$, $\lambda(x_T)$ denotes the tunneling coefficient and superscripts v and c have been added to account for electron and hole dependent parameters, respectively. This system of equations has more than ten parameters per band edge. One parameter set is the position of the defect in the oxide, where only the trap depth x_T directly enters the equations, which together with the dopant positions determines the step-height by electrostatic interaction. The parameters mainly determining the capture and emission times are the energy barriers (cf. Figure 6.11), the trap depth x_T , the Huang-Rhys factors $S\hbar\omega$, $S'\hbar\omega'$ as well as the curvature ratios R and R' . All other quantities in the equation system are either fixed values, such as the band weights N_c and N_v , or directly obtained from the transport model.

Evaluation of Trap Charge and Recombination Terms

The trapping model is also tightly coupled to the transport model via the trapped charge and charge recombination. The trapped charge Q_t is computed via

$$Q_t = \pm |q| (X_{2'} + X_2) \text{ or } E\{Q_t\} = \pm |q| (f_{2'} + f_2), \quad (6.42)$$

where the sign of the charge depends on the type of trap, which can either be donor- or acceptor-like. The recombination per defect can be straightforwardly generalized from Section 2.2.2 and read

$$\Gamma^n \{f_n^\nu, f_p^\nu\} = \sum_{\text{defects}} \frac{1}{(2\pi)^3} \sum_{\nu} \int_{\mathcal{B}} k_{12'} p_1 (1 - p_{2'}) (1 - f_n^\nu) g_n - k_{2'1} p_{2'} (1 - p_1) f_n^\nu g_n d^3 k, \quad (6.43)$$

$$\Gamma^p \{f_n^\nu, f_p^\nu\} = \sum_{\text{defects}} \frac{1}{(2\pi)^3} \sum_{\nu} \int_{\mathcal{B}} k_{2'1} p_{2'} (1 - p_1) (1 - f_p^\nu) g_p - k_{12'} p_1 (1 - p_{2'}) f_p^\nu g_p d^3 k, \quad (6.44)$$

for an acceptor-like defect. For a moment based transport model the above relations simplify to

$$R^n = \sum_{\text{defects}} k_{12'} p_1 (1 - p_{2'}) n - k_{2'1} p_{2'} (1 - p_1) n, \quad (6.45)$$

$$R^p = \sum_{\text{defects}} k_{2'1} p_{2'} (1 - p_1) p - k_{12'} p_1 (1 - p_{2'}) p. \quad (6.46)$$

Evaluation of Capture and Emission Times

For comparison with experiment, the capture and emission times τ_c and τ_e need to be calculated. Since the four state NMP model has been derived using first-order continuous-time Markov-Chain theory one can use the concept of first passage times from the theory. In the four state model, there are two pathways for charge capture and emission. Charge capture either proceeds via state $1'$ or $2'$. Thus, following [113], we define the capture and emission times as

$$\tau_c = \frac{\tau_c(1') \tau_c(2')}{\tau_c(1') + \tau_c(2')}, \quad \tau_e = \frac{\tau_e(1') \tau_e(2')}{\tau_e(1') + \tau_e(2')}, \quad (6.47)$$

where the first passage times $\tau_c(1')$, $\tau_e(1')$ and $\tau_c(2')$, $\tau_e(2')$ are

$$\tau_c(1') = \frac{k_{11'} + k_{1'1} + k_{1'2}}{k_{11'} k_{1'2}}, \quad \tau_c(2') = \frac{k_{12'} + k_{2'1} + k_{2'2}}{k_{12'} k_{2'2}}, \quad (6.48)$$

$$\tau_e(1') = \frac{k_{1'1} + k_{1'2} + k_{21'}}{k_{1'1} k_{21'}}, \quad \tau_e(2') = \frac{k_{2'1} + k_{2'2} + k_{22'}}{k_{2'1} k_{22'}}. \quad (6.49)$$

6.3 Implementation and Requirements

Each defect model, which can be derived based on first-order continuous-time Markov-Chains [143], can be stochastically interpreted. Thus two possibilities to solve the system of equations for the four-state NMP model as well as the SRH model [29] (cf. Chapter 2) are investigated here. The first possibility is to use a kinetic Monte Carlo algorithm to directly solve for the stochastic occupancies $X_i(t)$ via the occupancy probabilities $p_i(t)$, whereas the second one is to solve for the occupancies $f_i(t)$ (cf. Section 6.2.4). One can choose, depending on the particular task, between those two methods independent of the choice for a particular transport model or BTE solving technique. Irrespective of the actual (non-linear) solver used for the transport model, the program flow stays essentially the same and is shown in Figure 6.12. In case the time-averaged occupancies are of interest a straight forward assembly of the equations and the application of a linear solver suffice. For the kinetic Monte Carlo method, one first calculates the transition rates k_{Ij} between the currently occupied state I , ie. $X_I(t_0) = 1$, to all other possible states j . Then two uniformly distributed random numbers r_1 and r_2 between zero and one are drawn. Using random number r_1 the time to the next state transition event,

$$\Delta t = \frac{1}{a_I} \ln \left(\frac{1}{r_1} \right) \quad \text{with} \quad a_I = \sum_j k_{Ij}, \quad (6.50)$$

is calculated such that $X_I(t_0 + \Delta t) = 0$ and $X_J(t_0 + \Delta t) = 1$, where J is the final state. Afterwards, the final state J needs to be selected. For this one iterates, in arbitrary order, over all possible transitions j , where k_{Ij}/a_I are summed up until the random number r_2 is smaller

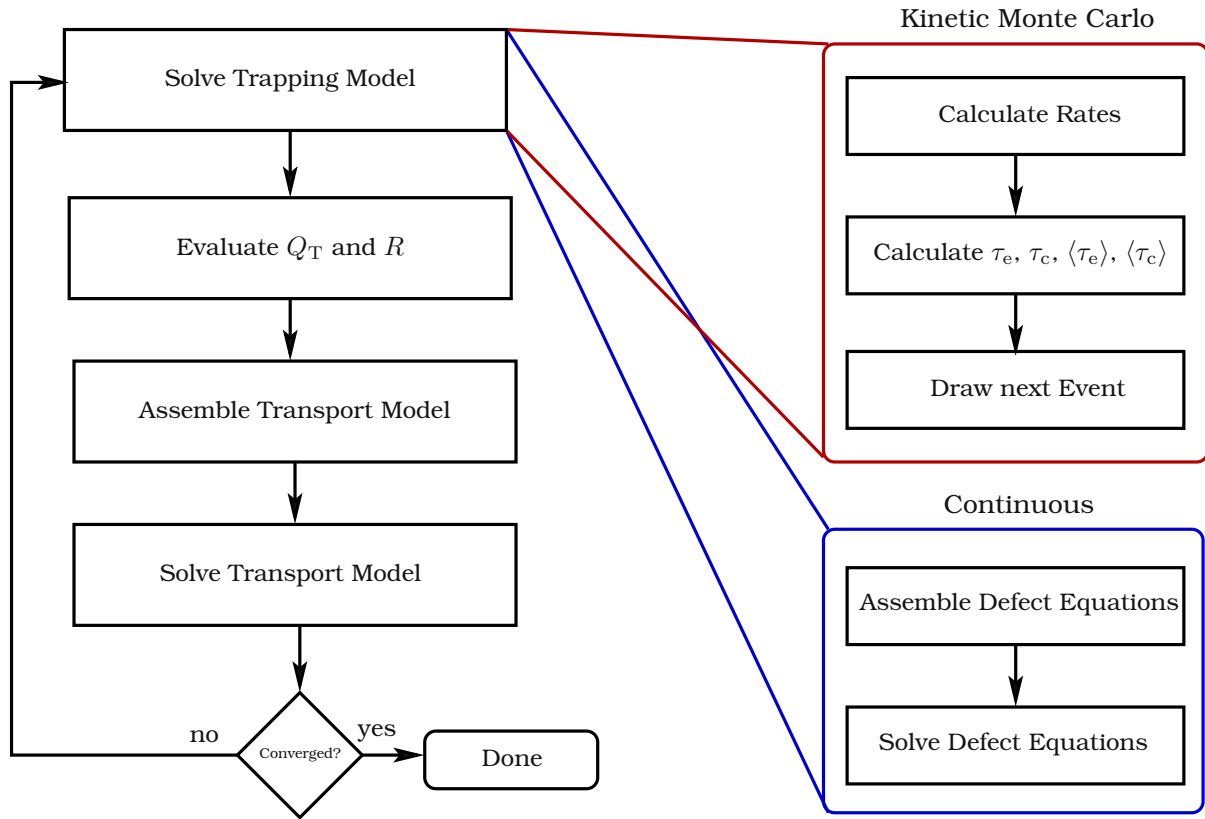


Figure 6.12: A flowchart for the implementation of NMP trapping models with suitable transport models. The transport model is, with the exception of the trapped charge and the recombination term, independent of the trapping model. This allows to use different techniques to solve the trapping model for the occupancies. When one solves for the stochastic occupancies a kinetic Monte Carlo algorithm can be used.

than the current sum. The final rate k_{IJ}/a_I in the summation determines the final state J . The advantage of solving for the stochastic occupancies lies in the greater depth of information on the actual trapping event. By virtue of the algorithm trap self-interaction can be mitigated. This advantage however is then traded-off against the possibility to effectively use the simulator in a mixed-mode simulator and the possibility to carry out a small-signal analysis.

6.3.1 Self-Consistent Solutions

The set of equations shown in the last sections are exponentially dependent on the electrostatic potential, which can be obtained by solving Poisson's equations. Additionally, Poisson's equation is tightly coupled to the transport model from which the quantities to calculate the NMP transition rates are obtained. The most pressing problem in the case of NMP transitions is the strong potential dependence of the NMP model, which results in self-interaction of the trap when solving for the occupancies $f_i(t)$. Since the occupancies $f_i(t)$ can take any value between zero and one, the trapped charge can also take all values between $\pm |q|$ and zero. These fractional charges are added to the charge density in the Poisson equation and thus change the local electric field at the defect site, which in turn can cause a significant change in the transition rates and thus the capture and emission times. This tight coupling of quantities

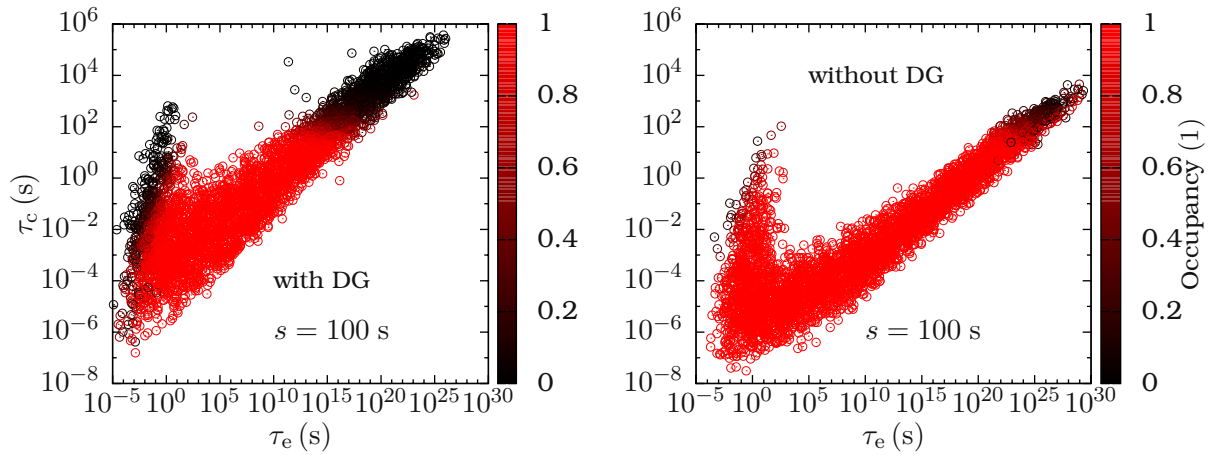


Figure 6.13: For each trap in one of 200 simulated microscopic devices with random discrete doping (left figure) and continuous doping (right figure), the occupancy after 100 s of bias temperature stress is shown (color coded). It can be seen that traps with large emission time constants, which correspond to a more permanent degradation, are occupied. Here the capture time constants have been chosen to be within the stress time in order to demonstrate saturation. Additionally, it can be seen in the left figure that with random discrete doping the capture and emission time constants are more spread out, compared to continuous doping.

for self-consistency results in convergence problems of the non-linear solver, which is in the course of this thesis a Gummel-Iterator or a Newton-Raphson algorithm. However, solving for the time averaged occupancies does allow to effectively use the simulator in mixed-mode and the possibility to carry out a small-signal analysis, provided the full Jacobian is known. To demonstrate this, MinimosNT with a first-order quantum corrected drift diffusion model and the four-state NMP model (solving for $f_i(t)$) have been used to simulate the time evolution of 2600 microscopically different defects [4]. The quantum correction has been accounted for by utilizing the density gradient model (cf. Chapter 4). To demonstrate the influence of potential fluctuations on the time constants predicted by the model, random discrete dopants have been considered too (cf. Figure 6.13). At last, Figure 6.14 shows the probabilities of self-interaction. However, as we have shown in [4] (cf. Figure 6.14) the error due to trap self-interaction is small and the validity of the simulation can be assessed. Nevertheless, if one is only interested in the evaluation of the capture and emission time of a single defect, self-consistency is often neglected and the trapping model can be evaluated after the transport model has been evaluated, without any coupling between trapping and transport model. Thus the system of equations for each defect described by an NMP based model is usually solved as a post processing step to the transport model whenever possible. Nevertheless, for example in the case of inhomogenous BTI this is not possible and one might need to utilize a KMC technique to solve the set of governing equations per trap.

6.3.2 Parameter Dispersion

All bonds between atoms in an amorphous structure, such as the gate oxides used in MOSFETs, are of different configuration and exhibit, for example, great disparity in bond lengths. Thus the parameters of the trapping models used to describe BTI are not fixed but statistically

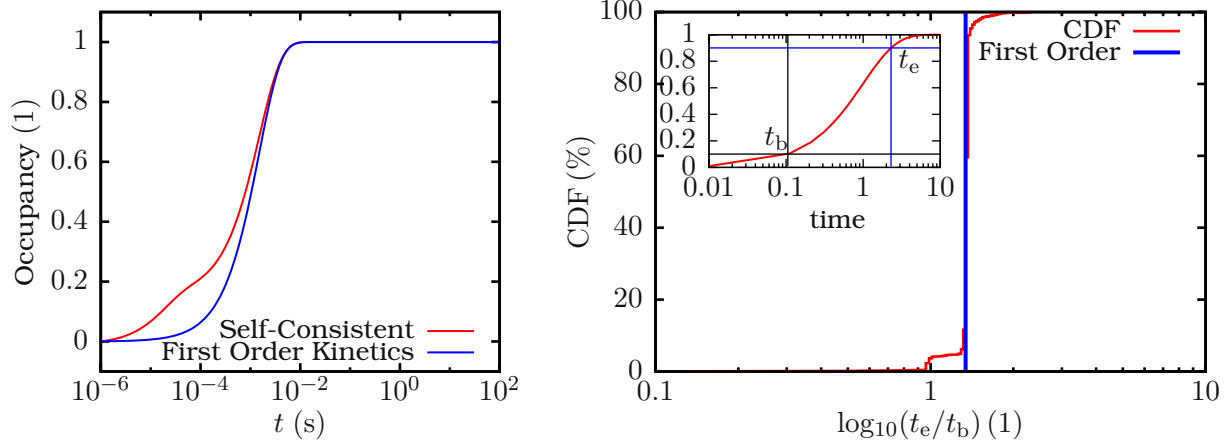


Figure 6.14: Left: The probability of a single self interacting trap to be occupied over stress time. The red curve is the occupancy for a self-consistent solution entering fractional charges leading to a trap self interaction during the charging process. As a reference the blue curve shows the expected charging behavior of the trap according to first-order kinetics. **Right:** The cumulative distribution function for the time it takes each trap to change its occupancy from 10% of its initial value to 90% of its final value on a log-scale (inset) for the whole ensemble of discrete traps in 200 simulations of a 35 nm pMOSFET with an EOT of 1 nm. Since significantly more than 90% of all 2600 traps exhibit the expected first-order behavior, the influence of the self interacting traps is small. For simulation the drift diffusion simulator MinimosNT with density gradient for quantum correction and random discrete traps (four-state model) and dopants have been used.

distributed. Without any prior knowledge of the defect site configuration and its precursor configuration many parameters have to be estimated within physically reasonable bounds. An example for the four-state NMP model is given in Table 6.1. These parameters have also been used for the self-interaction example in the previous section. It is generally assumed that all parameters follow a Gaussian distribution, which is fully characterized by its respective mean value and standard deviation. For the simulation of an ensemble of traps, the same algorithm as for random discrete dopants is used to estimate the number of traps via a Poisson distribution, their spatial location via an uniform distribution and the parameter set for each defect via a Gaussian distribution. This approach, however, is not taken to describe TDDS data [105].

6.3.3 Suitable Transport Models

Having a generalized algorithm for first-order time continuous Markov-Chain based BTI model (cf. Figure 6.12) the question remains: Which transport model is best suited to study BTI? Straightforwardly, any physical model which to sufficient accuracy describes the physical process underlying the unperturbed (no BTI) operation of a particular device is best suited to study BTI in that particular device. This is due to the nature of the NMP based models, which are, up to the inherent model error due to the assumptions made in the derivation, only as accurate as the physical device model, which delivers the necessary quantities, such as the electric field, the energy distribution function, charge carrier concentrations, lattice temperature and so on.

Parameter	Mean Value μ	Standard Deviation σ
E_t	0.6 eV	0.1 eV
E'_t	0.75 eV	0.1 eV
R	0.5	0.01
S	0.6	0.01
R'	0.8	0.01
S'	0.6	0.01
ϵ_{T2}	0.5 eV	0.05 eV
$\epsilon_{1'1}$	0.8 eV	0.05 eV
$\epsilon_{2'2}$	0.5 eV	0.05 eV

Table 6.1: Example mean values and standard deviations for the Gaussian distributed parameters of the four-state NMP model (Figure 6.11) used for demonstration purposes. The trap levels E_t (that is d_1) and E'_t (that is $d_{1'}$) have the largest standard deviations, which results in a few traps being more likely in state $1'$ than in state 1 before stress.

Quantum corrected drift diffusion simulators, like MinimosNT, are well suited to study BTI in MOSFETs where the density gradient model is sufficient to describe carrier confinement, provided the drain voltage is low enough to guarantee low field transport [4]. However, for devices with pronounced quantum mechanical effects, such as small diameter nanowires, SiGe MOSFETs [152] or nanoscale MOS structures, often Schrödinger-Poisson solvers are best suited if there is no transport, since they most accurately describe the device. Fortunately, Boltzmann Equation solvers, such as SHE, are only necessary in non-equilibrium if the charge carrier energy distribution functions for the device under BTI stress cannot be suitably approximated by a Maxwellian distributions anymore. In research this however is seldom the case, since BTI can be well studied, when there is little to no current flow in the device. Nevertheless, for device engineering the effect of BTI during usual use conditions, e.g. inhomogenous BTI, is of utmost importance to make lifetime predictions.

6.4 Model Evaluation on pMOSFETs using the Direct Current Current Voltage Method

In this section a comparison of the SRH model and the four state NMP model using direct current current voltage (DCIV) experiments at various temperatures, conducted using the polyheater technology [153], will be presented. To this end p-type metal-oxide-semiconductor field-effect transistors (pMOSFETs) are studied using the DCIV method before and after bias temperature stress. The ability of the SRH model and the four state NMP model to meaningfully reproduce the acquired DCIV data is compared. It is demonstrated that the SRH model cannot capture the detailed features of the data and that the more detailed four state NMP model is required.

6.4.1 Experimental Setup

For the measurement pMOSFETs with 30 nm thick silicon-dioxide as gate dielectric have been used. All pMOSFETs have been integrated with the polyheater technology presented in [153] in order to be able to locally heat the devices up to 500 °C. To monitor the stress induced

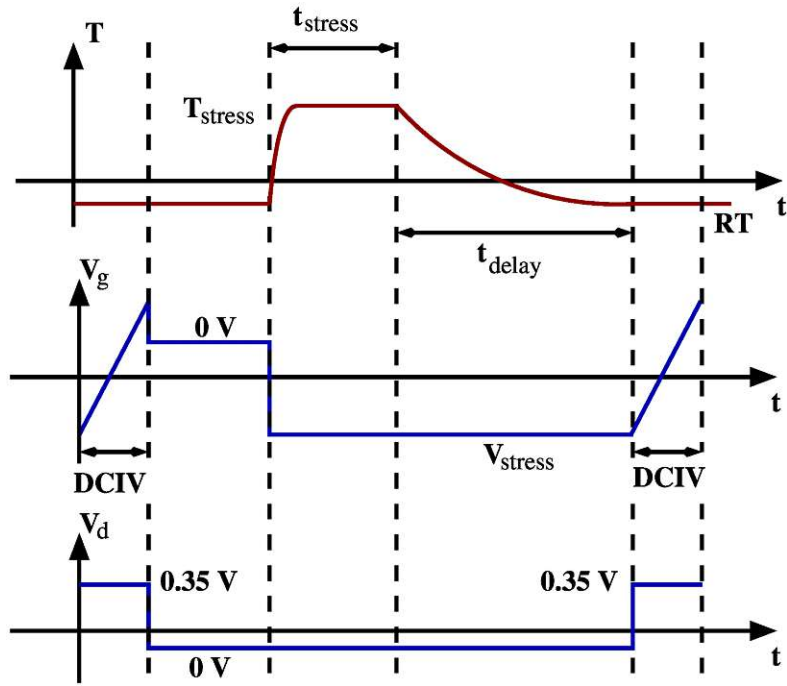


Figure 6.15: Time line of a single BTI experiment involving DCIV. From top to bottom the timeline of the temperature, gate voltage and drain voltage is shown. In each plot the abscissa has been placed such that it does not interfere with the plot and thus does not correspond to zero on the ordinate. For each stress temperature T_s a fresh device was used. Prior to stress an initial DCIV curve was recorded at room temperature. Afterwards the device was heated by integrated polyheaters almost instantly to the stress temperature T_s , while a gate voltage of -20 V was applied. With the end of the bias temperature stress phase the device was cooled down to room temperature (RT). At the end of the cool down phase a DCIV experiment on the degraded device was performed at room temperature (cf. right side of the lower part of the figure).

degradation, DCIV experiments [127, 128] were performed on fresh devices before and after stress using a drain voltage V_d of 0.35 V to forward bias the pn junctions. For each stress temperature a fresh device was stressed for $t_s = 10$ s by applying a gate voltage of -20 V ($V_s = -20$ V and $E_{ox} \approx 6.7$ MV/cm). After 10 s of stress the devices were cooled down for $t_{delay} = 200$ s to room temperature at a gate voltage of -20 V (cf. Figure 6.15) in order to minimize the relaxation during cool down [154] (degradation quenching). With the end of the cool down phase a DCIV curve for the stressed device was recorded.

6.4.2 Comparison of SRH and the four State NMP Model

The standard SRH model for interface traps and the previously introduced four state NMP model for BTI were used to describe the measurement data in order to compare their ability to reflect the (N)BTI stress dependent DCIV data. For the extraction all recombination centers, i.e. stress induced defects, were assumed to be at or near the silicon-oxide interface. During DCIV experiments the bulk current is directly proportional to the number of recombination events [127], simplifying the analysis considerably. It is important to note that any carrier recombination in the bulk, especially at the pn junctions, would cause a constant bulk current

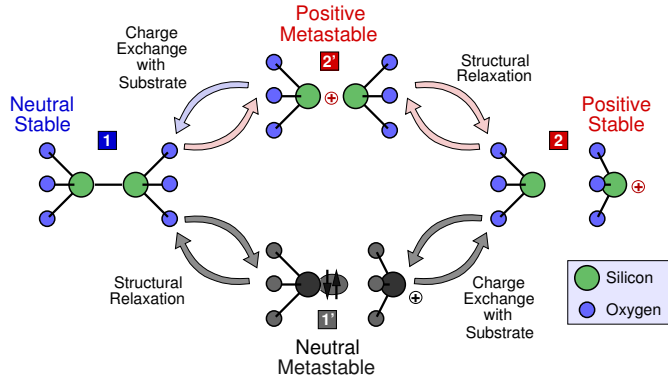


Figure 6.16: The four state NMP model from [151]. As a first approximation we omitted state 1' and transitions from/to this state for the DCIV experiments (gray shaded).

during the DCIV experiment. Since we did not observe a shift of the measured DCIV curves along the ordinate within the accuracy of the measurement equipment we can safely assume negligible carrier recombination in the bulk. Geometrical effects could be safely neglected, since large devices with a 30 nm thick gate dielectric and a nominal gate length larger than 100 nm were used. To assess the ability of the SRH model to reflect the DCIV measurement data, the formula originally derived in [155] was used. It reads

$$R = \frac{0.5\sqrt{\sigma_p\sigma_n}n_i (\exp(\beta(E_{F_p} - E_{F_n})) - 1)}{\exp(\beta(E_{F_p} - E_{F_n})/2) \cosh(\beta U_S^*) \cosh(\beta U_{t1}^*)}, \quad (6.51)$$

$$U_S^* = \phi_s + \ln(\sqrt{\sigma_p\sigma_n}) - (E_{F_p} + E_{F_n})/2, \quad (6.52)$$

$$U_{t1}^* = (E_t - E_i) \ln(\sqrt{\sigma_p\sigma_n}), \quad (6.53)$$

where n_i is the intrinsic carrier concentration, E_{F_n} and E_{F_p} are the quasi Fermi levels for electrons and holes respectively, E_i is the intrinsic energy, σ_p and σ_n are the constant hole and electron capture cross sections, ϕ_s is the surface potential, and $\beta = (k_B T_L)^{-1}$. In the analysis transitions from/to state 1' are neglected to obtain a simplified approximation to the full model. Further it was assumed that the DCIV experiment itself does not stress the device any further. This assumption was experimentally justified by comparing DCIV curves for various measurement durations, i.e. different slopes of the gate voltage applied, whereas the measurements yielded the same DCIV curves. With the stated simplification the derivation of a compact analytical version of the carrier recombination rate for the multistate NMP model can be undertaken. By defining effective rates the three remaining defect states have been reduced to only two equivalent states. The effective rates are [156],

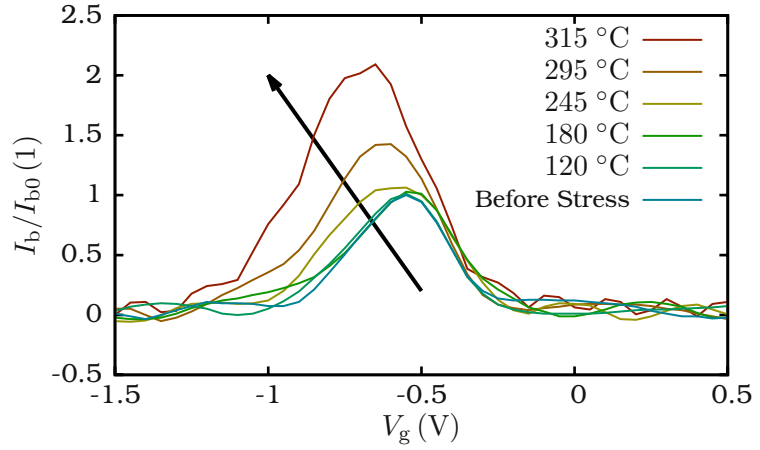
$$k_{12} = \frac{k_{12'}k_{2'2}}{k_{2'1} + k_{2'2}} \quad \text{and} \quad k_{21} = \frac{k_{2'1}k_{22'}}{k_{2'1} + k_{2'2}}. \quad (6.54)$$

Employing Maxwell-Boltzmann statistics and applying the assumptions discussed above to the first-order differential for the four state NMP model one obtains

$$\partial_t f = k_{21}(1 - f) - k_{12}f = 0, \quad (6.55)$$

for the occupancy of a trap. Using the definition of the recombination rate for holes R^p in steady state and inserting the solution of (6.55), gives an analytical formula in the framework

Figure 6.17: For higher stress temperatures the peak in I_b is larger and the curve is broader towards negative gate voltages (black arrow), whereas the slope for $V_g > -0.5$ V remains nearly constant. Also noteworthy is the reproducible shoulder towards negative gate voltages for a stress temperature of 315°C .



of a multistate NMP model for carrier recombination

$$R = -\frac{k_{22'}k_{2'2}(n_i^2 - np)\sigma_p\sigma_n v_{th}^n v_{th}^p}{N} \quad (6.56)$$

$$N = \exp(\beta\epsilon_{12'})k_{22'}n\sigma_n v_{th}^n + \exp(\beta\epsilon_{2'1})k_{2'2}p\sigma_p v_{th}^p \\ + \exp(\beta(-E_i + \epsilon_{12'} + E_t - \epsilon_{T2'}))k_{2'2}n_i\sigma_n v_{th}^n \\ + \exp(\beta(E_i + \epsilon_{2'1} - E_t - \epsilon_{T2'}))k_{22'}n_i\sigma_p v_{th}^p, \quad (6.57)$$

where v_{th}^n and v_{th}^p are the electron and hole carrier velocities, σ_n and σ_p are the constant electron and hole capture cross sections. For all simulations the constant capture cross sections σ_n and σ_p were fixed to a value of $2.0 \times 10^{-16} \text{ cm}^2$ for the SRH and the multistate NMP model.

DCIV curves measured for various stress temperatures, which have been normalized to the peak value for $T_{\text{stress}} \approx 245^\circ\text{C}$ for comparison, are shown in Figure 6.17. The maximum value of the bulk current I_b increases with higher stress temperatures as expected. Also noteworthy is the broadening of the bell-shaped DCIV curve towards negative gate voltages, whereas there is almost no broadening towards positive gate voltages (cf. Figure 6.17 for $V_g > -0.5$ V). This indicates that traps with higher effective activation energies are becoming active trapping centers at higher stress temperatures. A fit of the SRH model [136] for the post-stress measurement data is shown in Figure 6.18. It can be seen that the SRH model can reproduce the DCIV curve only for certain stress temperatures (in this case $T_{\text{stress}} \approx 240^\circ\text{C}$), but not for a wide range of stress temperatures. Especially for stress temperatures above 315°C , when the DCIV bell-shaped curve develops a shoulder towards negative gate voltages, the SRH model cannot reproduce the experimental data as shown in Figure 6.19. Noteworthy is the fact, as seen in Figure 6.19, that the post-stress DCIV curve changes its shape for stress temperatures above 315°C . This is why the SRH model, for which the recombination current exhibits a \cosh^{-1} shape (cf. (6.51)) for all temperatures [155], cannot reproduce the experimental data anymore. In contrast to the SRH model the four state NMP model can give excellent fits to the data for all stress temperatures as shown in Figure 6.20. This can be attributed to the fact that the reduced four state NMP model additionally considers structural relaxation (cf. Figure 6.16 and Figure 6.22). To understand why structural relaxation can explain the additional shoulder above a stress temperature of 315°C in our data, we reformulated the carrier recombination formulae in the four-state NMP model (6.56) and for the SRH model (6.52) such that both formulas have the same structure. The carrier recombination rate for both models has the

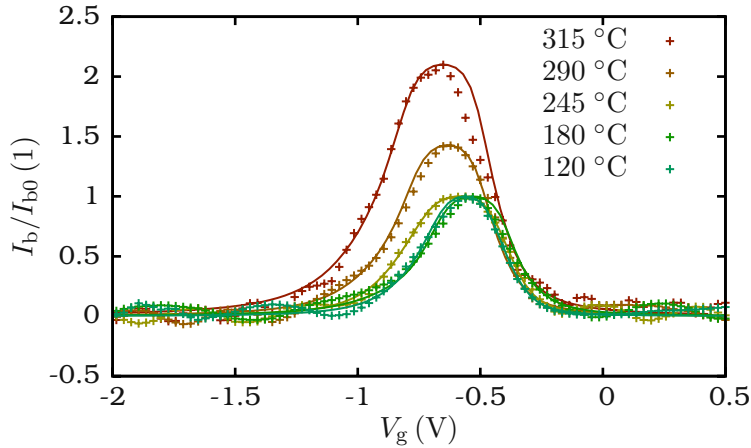


Figure 6.18: Fits of the standard SRH model to the post-stress DCIV curves for various stress temperatures. Especially for higher stress temperatures, such as 315°C, the standard SRH model cannot reproduce the shape of the measurement data anymore.

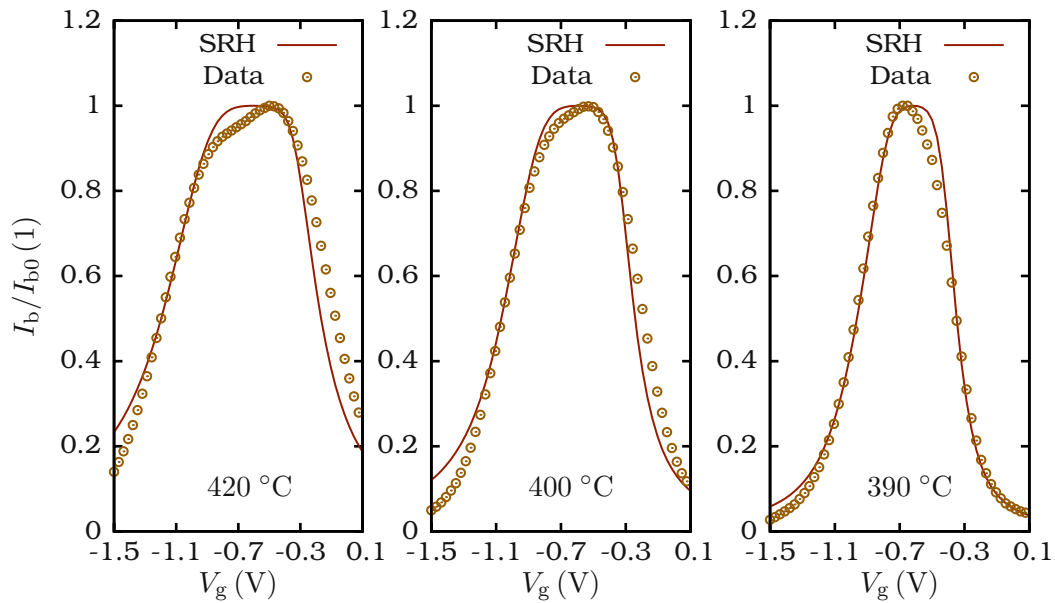


Figure 6.19: Same as in Figure 6.18 but for higher stress temperatures. At very high stress temperatures the post-stress DCIV curve develops a characteristic shoulder. The SRH model cannot predict any of the three DCIV curves, since it cannot reproduce the characteristic shoulder.

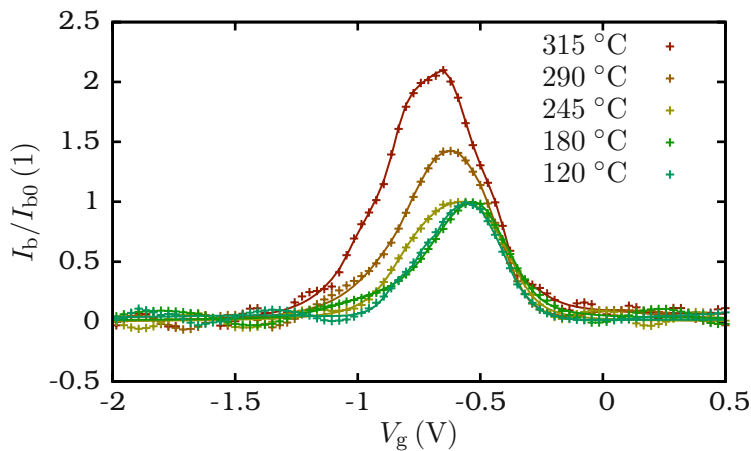


Figure 6.20: In contrast to the fit for the SRH model (cf. Figure 6.18), the NMP model can explain the post-stress DCIV data for all stress temperatures used. Remarkable is the ability of the NMP model to reproduce the DCIV curve for a stress temperature of 315 °C with high accuracy.

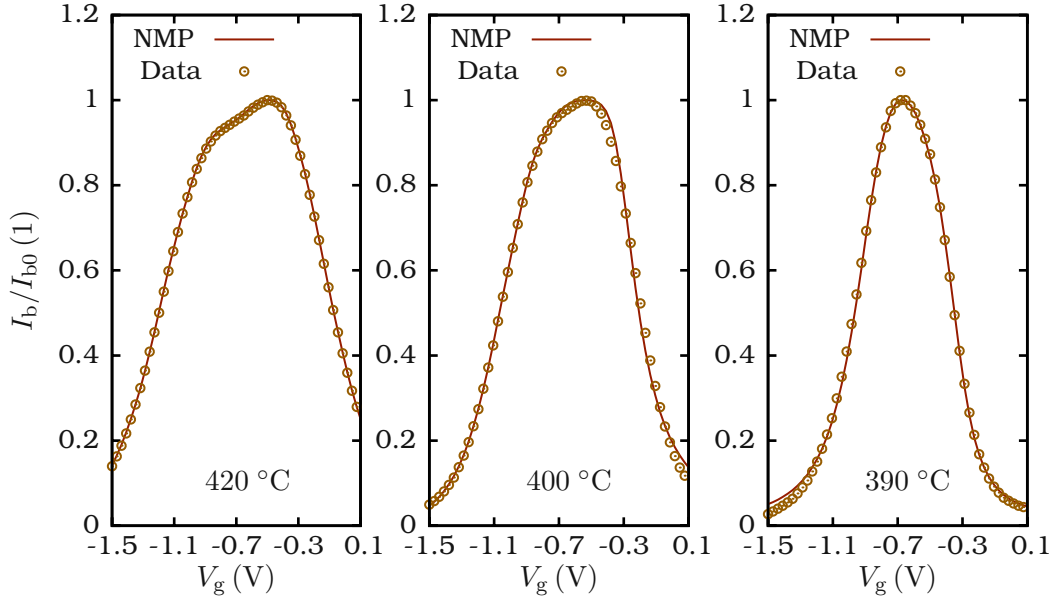


Figure 6.21: For high stress temperatures it is possible to reproduce the post-stress DCIV data with good accuracy. Even the characteristic shoulder, best seen for $T_{\text{stress}} = 420 \text{ }^\circ\text{C}$, can be fitted with excellent accuracy.

following structure

$$R = -\frac{(n_i^2 - np) \sigma'_p \sigma'_n v_{\text{th}}^n v_{\text{th}}^p}{\tau'_n(n + n_1) + \tau'_p(p + p_1)}, \quad (6.58)$$

where τ'_n and τ'_p are the carrier lifetimes for electrons and holes and σ'_n and σ'_p are the capture cross sections for electrons and holes, respectively. These parameters are model-dependent. For the SRH model the capture cross sections have fixed values, where

$$\sigma'_n = \sigma_n = 2.0 \times 10^{-16} \text{ cm}^2, \quad (6.59)$$

$$\sigma'_p = \sigma_p = 2.0 \times 10^{-16} \text{ cm}^2. \quad (6.60)$$

In contrast, the four-state NMP model gives

$$\sigma'_n = \frac{\exp(\beta\epsilon_{2'1})}{k_{22'}k_{2'2}} \sigma_n = \frac{\exp(\beta\epsilon_{2'1})}{\nu^2 \exp(-\beta\epsilon_{22'}\epsilon_{2'2})} \sigma_n, \quad (6.61)$$

$$\sigma'_p = \frac{\exp(\beta\epsilon_{12'})}{k_{22'}k_{2'2}} \sigma_p = \frac{\exp(\beta\epsilon_{12'})}{\nu^2 \exp(-\beta\epsilon_{22'}\epsilon_{2'2})} \sigma_p. \quad (6.62)$$

Furthermore, the carrier lifetimes differ strongly between the SRH model and the multistate NMP model [151]. Thus it can be stated that the effective capture cross sections of the multistate NMP model are device temperature, oxide field and parameter dependent, while those of the SRH model are always constant. By adjusting the energy barriers ($\epsilon_{22'}$ and $\epsilon_{2'2}$) describing the structural relaxation (cf. Figure 6.9 and Figure 6.16) it is possible to perfectly fit the DCIV curves measured after high temperature stress (cf Figure 6.21). Figure 6.22 depicts how the shoulder, for stress temperatures above $315 \text{ }^\circ\text{C}$, can be described using a multistate NMP model. If structural relaxation is neglected (cf state 2 in Figure 6.16) one obtains a standard

NMP model [157, 158]. For the standard NMP model the energy barriers ($\epsilon_{22'}$ and $\epsilon_{2'2}$) describing the structural relaxation are zero. Thus the effective capture cross sections of the standard NMP model reduce to

$$\sigma'_n = \frac{\exp(\beta\epsilon_{2'1})}{\underbrace{k_{22'}}_1 \underbrace{k_{2'2}}_1} \sigma_n = \exp(\beta\epsilon_{2'1}) \sigma_n \quad \text{and} \quad \sigma'_p = \frac{\exp(\beta\epsilon_{12'})}{\underbrace{k_{22'}}_1 \underbrace{k_{2'2}}_1} \sigma_p = \exp(\beta\epsilon_{12'}) \sigma_p. \quad (6.63)$$

This reduction of the multistate NMP model results in a less pronounced temperature dependence [29] and in the loss of two degrees of freedom for parameter extraction.

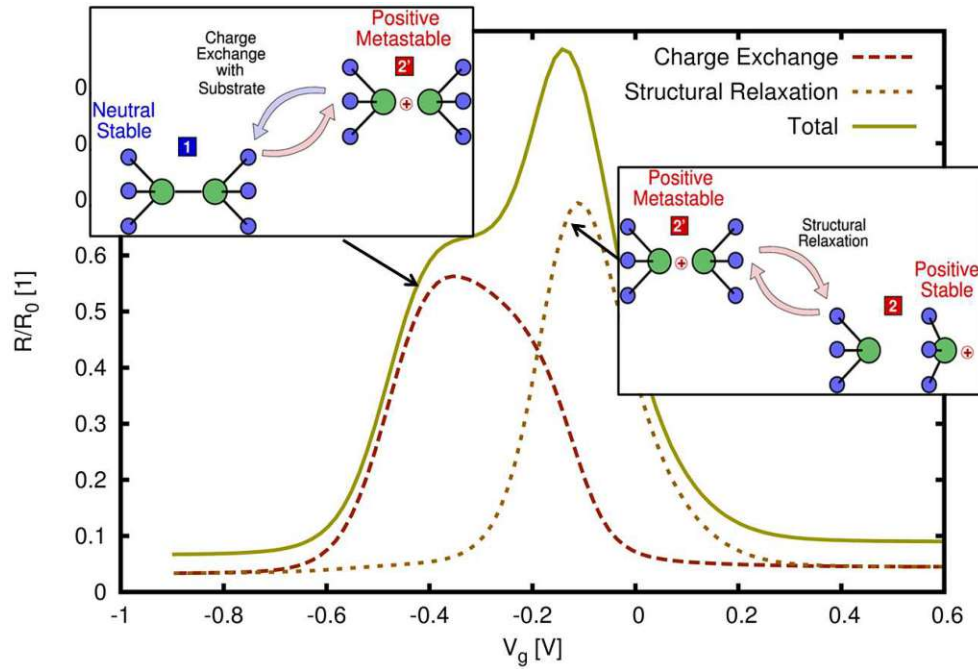


Figure 6.22: When the exchange of charge carriers with the substrate dominates the recombination current exhibits a bell-shaped curve towards negative gate voltages. Whereas when structural relaxation dominates the recombination current exhibits a bell-shaped curve towards positive gate voltages. The weighted sum of these two partial recombination currents gives a bell shaped curve with an additional shoulder [149].

The investigation of the capability of the SRH and NMP model to explain DCIV measurements of pMOS devices after NBTI stress at various stress temperatures shows the importance of structural relaxation for a proper description of recombination currents. A moderate agreement between model and measurement data could be obtained with the conventional SRH model.

6.5 Results on Trap-Assisted Tunneling

In 2012 Toledano et al. [159] showed that there is a correlation between gate and drain current fluctuations in MOSFETs. Figure 6.23 shows a portion of the experimental data. Interestingly the difference between high and low gate current is 80% of the gate current maximum and temperature independent. Due to the correlation with the drain current it is concluded that the charged trap significantly changes the tunneling current. A possible explanation is that the

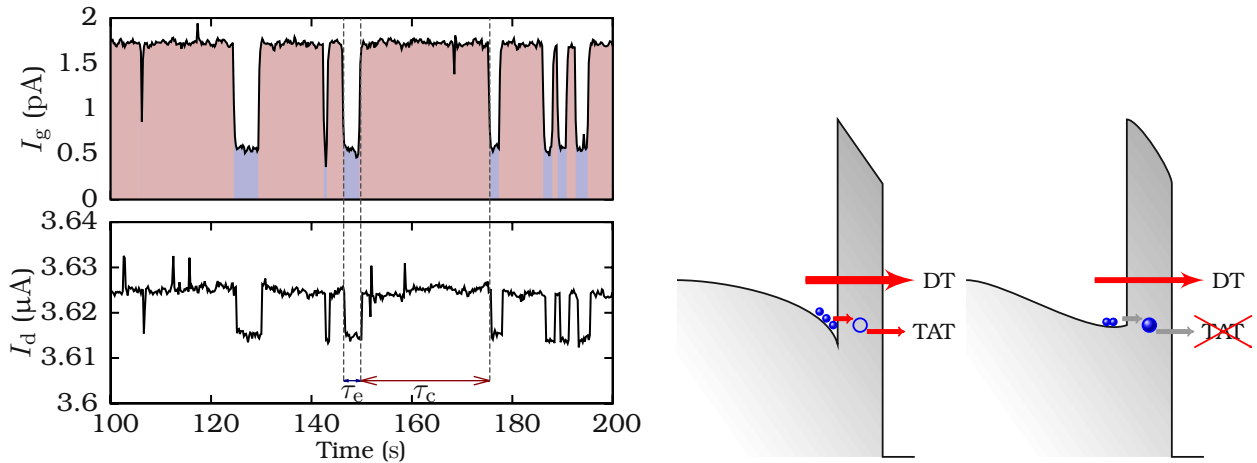


Figure 6.23: Left: Gate I_g and drain current I_d simultaneously recorded on a n-channel MOSFET with an EOT of 2.3 nm [159]. It is clear that gate and drain current are strongly correlated. Whenever a trap becomes occupied (low drain current) the gate current is decreased as well. In total the gate current is reduced by nearly 80%. **Right:** A schematic of the two tunneling processes to explain the experimental data (left). The total tunneling gate current is composed of direct tunneling (DT) and trap-assisted tunneling (TAT). On the far right, the assumption of a purely electrostatic interaction between the direct tunneling current and the trap occupancy is depicted.

charged trap changes the local potential such that the number of charge carriers, which tunnel through the oxide, is reduced by 80%. This would also explain the temperature independence of the reduction, inherent in direct tunneling. Another explanation is that the tunneling current is a trap assisted NMP process. In [160] we explored whether the experimental data could be reproduced by assuming that an occupied trap reduces the direct tunneling current by purely electrostatic interaction. To assess whether or not this is correct, we first simulated the electrostatics of the n-channel MOSFET. For this we used the 3D density gradient and drift diffusion simulator MinimosNT and additionally considered random dopants. The simulations were done per device, once without and once with a single discrete trapped charge in the oxide (cf. Chapter 5). Our sample size consisted of 201 microscopically different devices, i.e. different dopant positions. Then 1D potential profiles (normal to the semiconductor-oxide interface) were extracted in 1 nm increments. These profiles were then loaded into our NEGF (cf. Chapter 2) simulator, VSP [51], in order to calculate the relative change in the gate current (direct tunneling) $\Delta I_g/I_g$. The relative tunneling current change $\Delta I_g/I_g$ was obtained by comparing the tunneling currents for the devices without trap to the devices with trap. In Figure 6.24 (left) a normalized map of the local $\Delta I_g/I_g$ over the whole gate area is shown. The map shows high local changes in the vicinity of the trap. However, as the histogram in Figure 6.24 (right) shows, we could hardly find a dopant-trap configuration for which the total change in tunneling current exceeded 1%. This is in stark contrast with the experimentally reproducibly found decrease of 80% in I_g .

Thus trap-assisted tunneling in the context of the presented four-state NMP model has been (successfully) explored [161]. Extending the four-state NMP model (cf. Figure 6.25) by taking interactions with the poly-gate into account, the experimental data could be reproduced (cf. Figure 6.26). In this extended model one has to differentiate whether the charge carrier to be

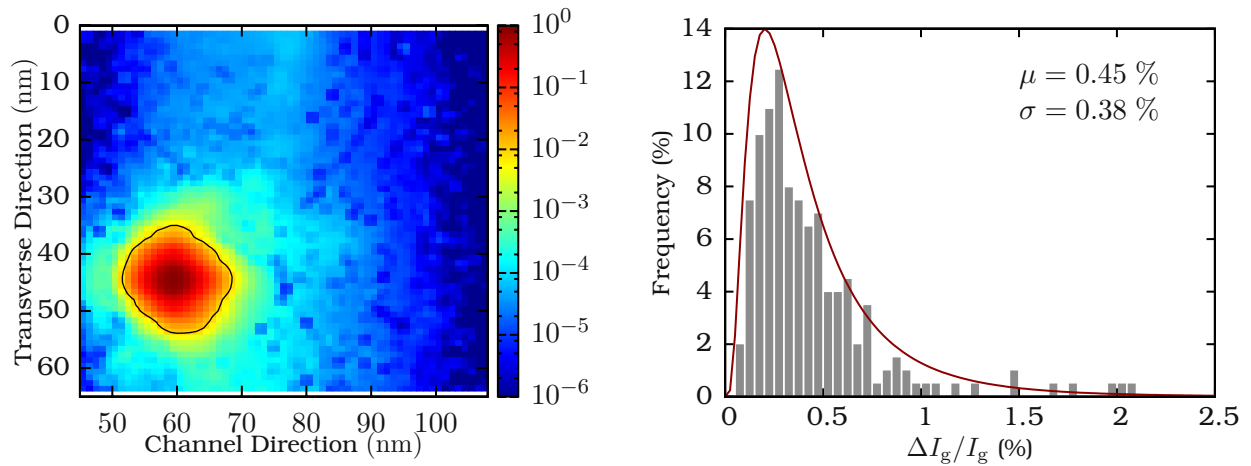


Figure 6.24: Left: Map of *local* $\Delta I_g/I_g$ (1 nm increments) due to the charging of a single trap for a particular random discrete doping. High changes in the direct tunneling current in the vicinity of the trap are visible. The contour line corresponds to a one percent change in the local gate current. Nevertheless, the total gate current does not change by more than two percent [160]. **Right:** Histogram of $\Delta I_g/I_g$ for active and inactive traps exhibiting a log-normal distribution (solid line). From this data it is clear that the reduction of the direct tunneling current due to active traps is unlikely to be decreased by more than 2% [160].

trapped originated from the poly-gate or the substrate. Nevertheless, the basic rate equations do not change. However, the source of the quantities and the tunneling distances are different. In order to correctly evaluate the capture τ_c and emission times τ_e for the trap-assisted tunneling current the charge state (whether or not the defect is charged) in the modified model had to be taken into account. In the course of the evaluation of the five-state NMP model it was found that the weak temperature dependence of the time constants is due to the weak electron-phonon coupling (small NMP transition barriers) needed to reproduce the measurement data (cf. Figure 6.26 right top). For details on the model, the reader is referred to [161].

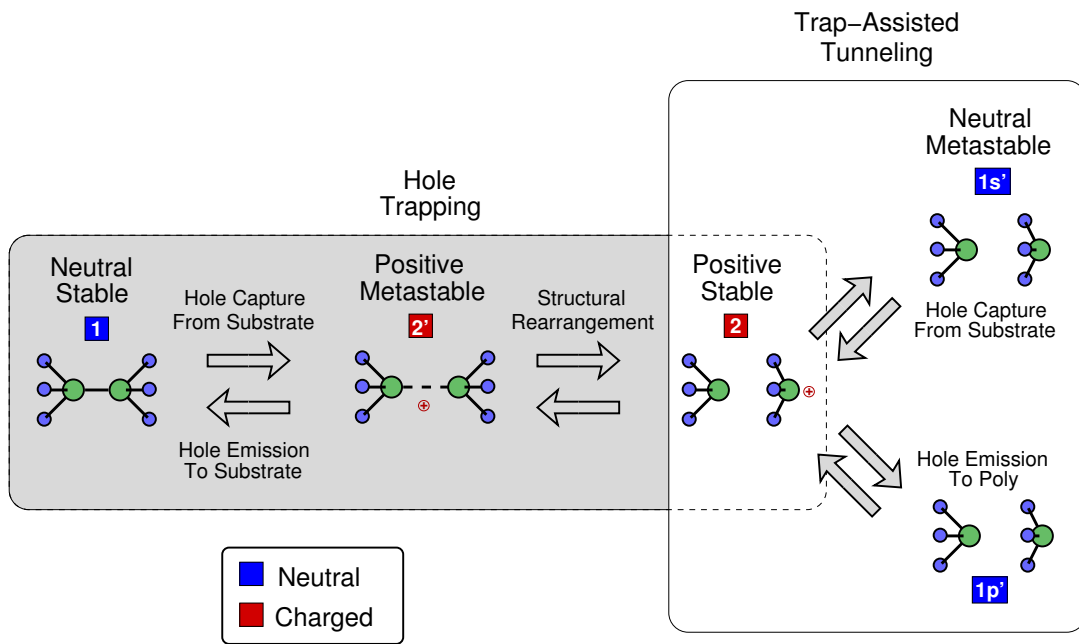


Figure 6.25: A five-state NMP model to explain trap-assisted tunneling in MOSFETs. The first three states (1, 2' and 2) are the same as in the four-state model. All other states have been replaced by the remaining two states. These states describe the charge exchange via an NMP process with the poly-gate.

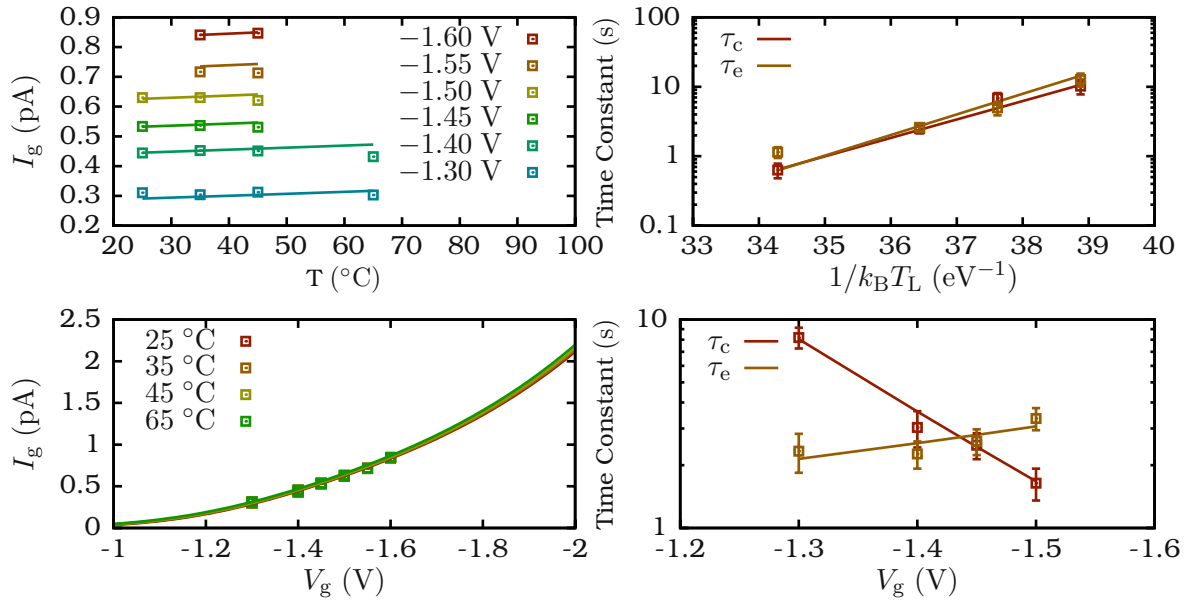


Figure 6.26: Left: A fit of the five-state NMP model for trap-assisted tunneling (cf. Figure 6.25) to the experimental data from [159]. The top figure illustrates the weak temperature dependence of the trap-assisted tunneling process, due to weak electron-phonon coupling. The lower figure shows the exponential dependence of the tunneling current on the gate voltage. **Right:** A fit of the model to the characteristic time constants of the tunneling process (cf. Figure 6.23). The capture and emission time constants can be reproduced quite well. The model can thus represent the temperature and oxide field dependence of the tunneling current as well as the experimentally observed characteristic time constants.

Let's perfect the current HCD model
before we mate it with the BTI model.

(Stanislav Tyaginov)

7 Hot Carrier Degradation

Hot-carrier degradation (HCD) is a change in the device parameters, such as drain current ΔI_d , threshold voltage V_{th} and on-resistance R_{on} , in MOSFETs operating at high lateral (along the channel) electric fields. HCD was initially only attributed to a defect buildup in the oxide caused by hot charge carriers. However, it was found that colder charge carriers, that are carriers with a low average energy, can also cause hot-carrier degradation [162]. Nevertheless, our latest [163] as well as older HCD models [164] include the assumption that hot charge carriers, which have gained sufficient kinetic energy, scatter with passivated dangling bonds at the interface, causing the bond to break. However, newer models include the additional assumption that multiple colder carriers can also cause these bonds to be dissociated. This rupture in turn creates P_b centers at the semiconductor-oxide interface [165, 166], which capture and emit charge carriers, thereby distorting important device characteristics such as e.g. drain current and threshold voltage (cf. Figure 7.1). In contrast to bias temperature instability (BTI), however, hot-carrier degradation is best observed at high electric fields along the channel. Whereas BTI is best observed at high gate voltages and negligibly small lateral electric fields, such that low field conditions in the direction of charge carrier transport are met. It is adopted that oxide defects and border traps play a crucial role in BTI modeling, where the charge trapping kinetics are described by non-radiative multiphonon theory [29]. In HCD modeling, however, it is assumed that P_b centers created by hot carriers can be described by standard SRH theory and that oxide defects can be safely neglected. Nevertheless, oxide defects were shown to be present in devices subjected to hot-carrier stress [167]. For instance, in [168, 169] it has been demonstrated that charging of different types of traps can lead to a change of sign in $\Delta I_{d,lin}$. Thus, in [169] first the $\Delta I_{d,lin}$ increases followed by a reduction in $\Delta I_{d,lin}$ at longer stress times. Such a turned-around effect of the threshold voltage, which is attributed to the interplay of hole trapping by bulk oxide traps and electron capture at the interface has been reported in [167].

7.1 Measurement Technique

The measurement techniques employed to assess hot-carrier degradation are very similar to those used for bias temperature instability. The degradation is often measured as a change in drain current ΔI_d or as a threshold voltage shift ΔV_{th} [170]. Additionally, recovery of hot-carrier stressed devices is currently of increasing scientific interest [171]. Measuring the recovery after hot-carrier stress is very similar to BTI recovery experiments. Recovery can be observed by switching all voltages to their respective relaxation levels and recording the drain current as it is done for stress. However, often it is necessary to accelerate recovery to collect more data points in less time. This can be achieved by switching the device temperature to higher temperatures while measuring recovery. A typical stress or stress-measure technique

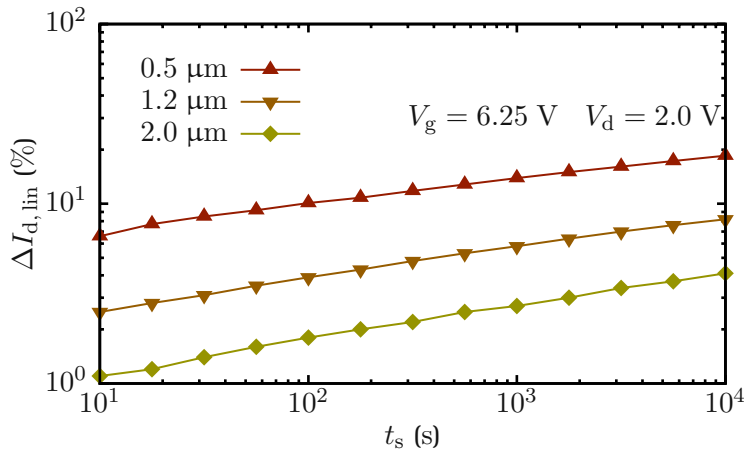


Figure 7.1: Hot-carrier stress manifested as ΔI_d at room temperature. The data has been recorded for n-channel MOSFETs with an EOT of 16 nm. It can be seen that for smaller channel lengths the degradation becomes more severe.

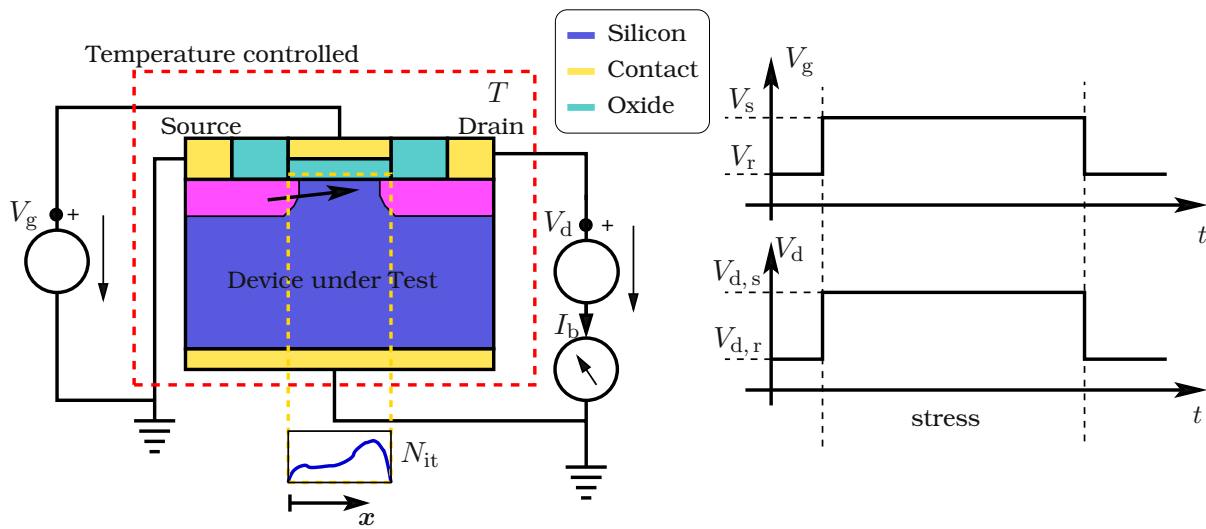


Figure 7.2: A typical hot-carrier stress cycle. In each hot-carrier stress cycle the device gate voltage and the drain voltage are chosen such that charge carriers can be sufficiently accelerated (no low field conditions), while the device temperature is kept constant. These stress conditions lead to an increase in N_{it} along the oxide interface, measurable as a decrease in drain current. To obtain the drain current degradation ΔV_{th} the drain current during stress is compared to an $I_d - V_g$ curve taken before stress.

to assess HCD is shown in Figure 7.2, where the drain current is recorded for a preset pair of gate V_g and drain voltages V_d at a fixed temperature. Before the hot-carrier experiment starts, the fresh device is characterized by measuring for example the $I_d - V_g$, $I_d - V_d$ or CV curves. During this initial measurement great care not to BTI stress the device in this phase has to be taken. Throughout the hot-carrier experiment the drain current is constantly recorded, while gate and drain voltage are regulated by a controller to cycle between the various stress conditions. Hot-carrier stress is defined as the set of voltages (V_g , V_d) at constant device temperature, which cause a prominent drain current shift $|\Delta I_d|$ or an equivalent absolute threshold voltage shift $|\Delta V_{th}|$ in a given time t_s .

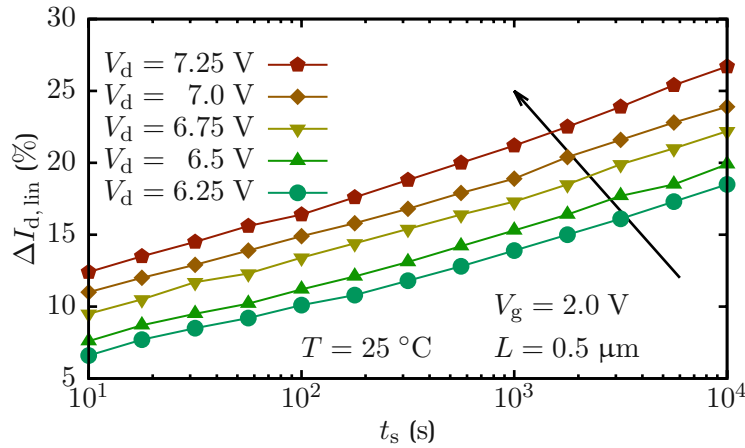


Figure 7.3: Hot-carrier stress assessed as ΔI_d at room temperature for various drain voltages. The data has been recorded for n-channel MOSFETs with an EOT of 16 nm at a gate voltage of 2.0 V. With increasing drain voltage and thus channel electric field, the hot-carrier induced damage is increasing (arrow) [174].

7.2 Electric Field Dependence

Experimentally it was confirmed that hot-carrier degradation in MOSFETs is highly electric field dependent (cf. Figure 7.3) [164, 170, 172], for both short- and long-channel devices. Thus the influence of the electric field and channel length on HCD is uncorrelated. This is due to the constant field scaling applied by the industry [173] and can be attributed to the fact that even a few nanometers of silicon are sufficient for charge carriers to be accelerated or otherwise high drain currents in scaled devices could not be sustained. The significant lateral field dependence inspired the first lucky electron model [164], which assumed that the lateral electric field is the driving force of hot-carrier degradation (field driven paradigm) [164]. Systematic investigations based on Fowler-Nordheim [175] tunneling stress, hot substrate and channel charge carriers revealed that the carrier energy rather than the electric field is the driving force behind hot-carrier degradation (energy driven paradigm) [176, 177]. However, charge pumping experiments revealed that neither the peak of the electric field nor the peak of average kinetic carrier energy coincides with the experimentally found maximum of the created interface defect density N_{it} [178] (cf. Figure 7.4). Thus, the information on the electric field and average carrier energy is not enough to adequately evaluate the the hot-carrier induced N_{it} profile.

7.3 Temperature Dependence

Contrary to bias temperature instability, hot-carrier stress becomes less detrimental with increasing device temperature (cf. Figure 7.5) in longer channel devices. This relation is inverted for short channel devices, where hot-carrier stress becomes more severe at elevated temperatures [179]. This channel length dependent influence of temperature on the degradation can be explained in terms of carrier kinetic energy distribution and scattering. As laid out in Chapter 2, charge carriers are accelerated by the electric field in the semiconductor and undergo scattering events. These events lead to changes in momentum and possibly in a loss/gain of particle energy. At higher temperatures the intensity of all scattering processes increases, which results in depopulation of the high energetical fraction of the carrier ensemble. As a

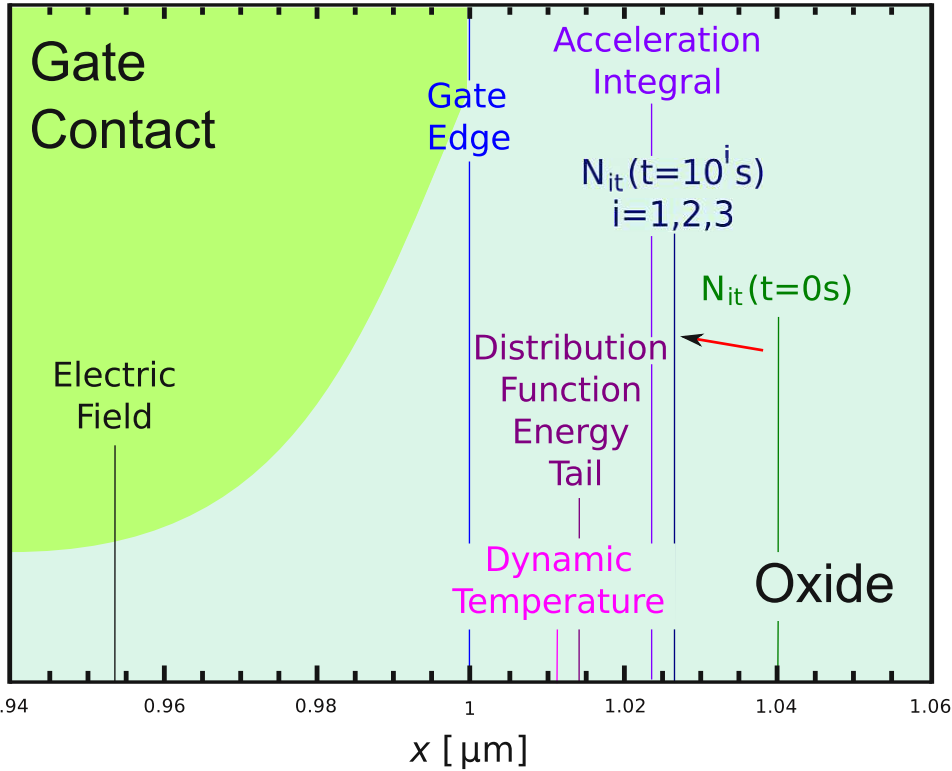


Figure 7.4: The peak positions of various quantities used to model hot-carrier degradation in comparison to the extracted N_{it} peak at the drain side of a MOSFET channel. The figure shows that neither the electric field nor the average carrier energy maximum coincide with the N_{it} peak and thus cannot be used as the driving force behind HCD. Adopted from [170].

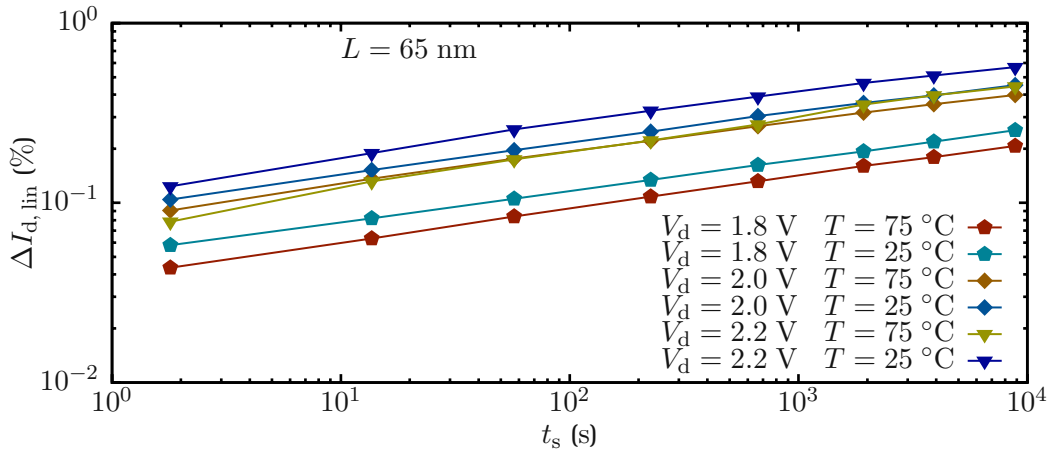


Figure 7.5: Hot-carrier degradation expressed as $\Delta I_{d,lin}$ at room temperature and for device stress temperatures of 25 °C and 75 °C. The data have been recorded for n-channel MOSFETs with an EOT of 2.5 nm at various gate and drain voltages, where $V_g = V_d$. With increasing drain voltage and thus channel electric field, the hot-carrier induced damage increases. However, the induced damage decreases with increasing temperature.

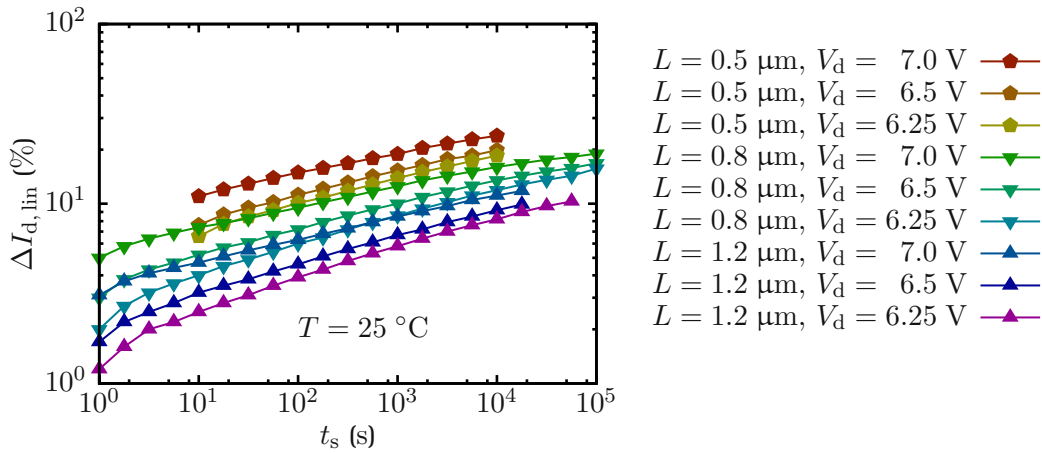


Figure 7.6: Hot-carrier degradation expressed as $\Delta I_{d, \text{lin}}$ at room temperature for various drain voltages and channel lengths. The device under test was an n-channel MOSFETs with an EOT of 16 nm at a gate voltage of 2.0 V. From the plot it can be seen that the general trend over drain voltage is the same for various channel lengths. Nevertheless, the distance between the curves grows with decreasing channel lengths [174].

consequence, increasing lattice temperature causes weakening of hot-carrier degradation.

7.4 Channel Length Dependence

In Figure 7.6 the $\Delta I_{d, \text{lin}}$ over time during hot-carrier stress in the linear regime of various n-channel MOSFETs with different channel lengths is shown. From this figure and the literature [180], it can be concluded that hot-carrier stress is a severe problem in short channel devices too. This can be explained in terms of the charge carrier energy distribution and is still an active area of research. As discussed for the temperature dependence, in shorter channels the charge carriers can accelerate without being scattered as much as in a long channels. This reduction in the number of scattering events per second in scaled MOSFETs reinforces the cumulative probability of the carrier ensemble to dissociate Si-H bonds at the semiconductor-oxide interface, since there are more carriers with sufficient kinetic energy. Additionally, electron-electron scattering is more pronounced in shorter channel devices, where more carriers exceed the threshold to significantly elevate the high energy tail of the distribution function than in long channel devices. In summary the average number of hot electrons is increased by electron-electron scattering, which has a significant influence especially for shorter channel devices, since the carriers scatter less with phonons and impurities.

7.5 Models

To model HCD understanding of the underlying physical mechanisms is naturally required, where the electric field dependence was explored first. Thus, it is not surprising that early models were based on the magnitude of the electric field in the channel (field-driven paradigm). Until today most of the models explained the influence of hot-carriers as a buildup of interface defects and charges along the semiconductor-oxide interface of MOSFETs, which are often modeled using SRH kinetics.

7.5.1 Early Models

One of the first HCD models was the so-called Lucky-Electron model [164]. The model is based on the assumption that a single carrier (called 'lucky electron') is accelerated by the electric field such that it gains sufficient energy to overcome the dissociation energy barrier of a Hydrogen saturated bond at the interface. Additionally, when the lucky electron impinges on the semiconductor-oxide interface it does so without further collisions and without being scattered back into the channel. In this field-driven model the hot-carrier degradation is described as a buildup of interface states ΔN_{it} in addition to interface states that have been active before hot-carrier stress.

$$\Delta N_{it} = K \left(\frac{tI_d}{W} \exp \left(- \frac{E_a}{|q| \lambda \max\{E(\mathbf{x})\}} \right) \right), \quad (7.1)$$

where K is a prefactor, t is time since the onset of stress, W is the channel width, E_a is the activation energy, λ is the electron free mean path and $\max\{E(\mathbf{x})\}$ is the maximum electric field at the semiconductor-oxide interface. However, it has been shown (cf. Section 7.2) that HCD is not field driven and that the Lucky-Electron model fails to explain HCD in short channel devices [181]. From charge pumping experiments [182] the maximum N_{it} has been found not to coincide with the maximum of the electric field along the channel (cf. Figure 7.4). This, however, contradicts the prediction of the Lucky-Electron model, where maximum N_{it} and the peak of the electric field coincide.

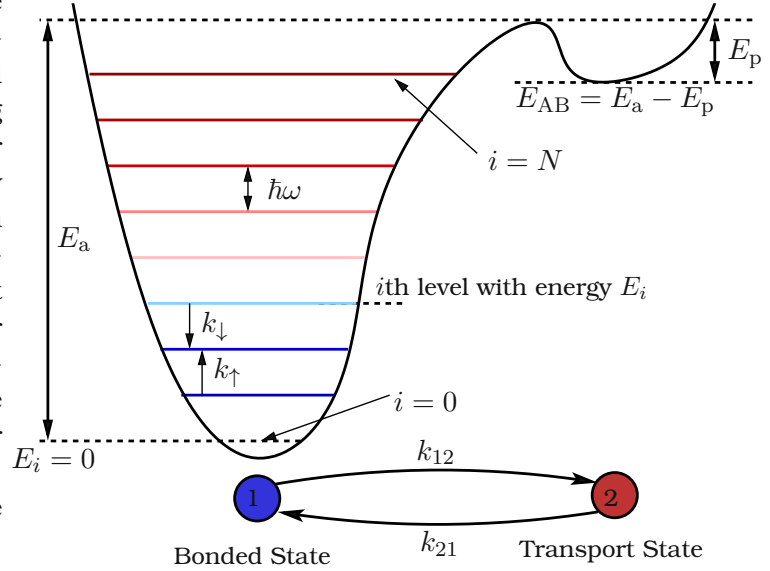
7.5.2 Latest Model

While the first hot-carrier degradation models were based on the channel electric field as the driving force, it has long been realized that the phenomenon is energy- rather than field-driven [172, 176, 177]. Thus, a physics based model requires information on the energy distribution of the charge carriers. Such a model has first been proposed by the group of Hess [183, 184], improved by the group of Bravaix [162, 185] and has been constantly further developed in [174, 186], resulting in the latest carrier energy distribution based model [163], which will be presented and discussed here.

Bond Breakage Mechanisms

Our physics-based model, as many other models, describes the drain current and the degradation of other device parameters (e.g. R_{on}) as the result of a stress induced buildup of interface defects. Thus, the bond breakage mechanism needs to be addressed first. Contrary to the Lucky-Electron model the silicon-hydrogen bonds, present in every forming gas annealed device, are modeled as truncated harmonic oscillators and not by a single bond energy (cf. Figure 7.7). In the truncated harmonic oscillator model of the Si-H bond, charge carriers are assumed to be able to scatter at the Si-H bond, thereby exchanging energy with the bond and scattering back into the substrate. If a sufficient amount of energy is transferred to the Si-H bond from the channel carrier, the bond is ruptured and N_{it} increases. This bond breakage mechanism is termed anti-Bonding (AB) mechanism and corresponds to bond rupture by a single highly energetic carrier. If the charge carrier, however, does not have enough energy to break the Si-H bond, it is assumed that this carrier at least increases the vibrational energy of the Si-H bond, thereby contributing to the multi-vibrational excitation (MVE) process. This situation corresponds to bond dissociation by a series of 'cold' carriers [184]. The MVE process is especially

Figure 7.7: The truncated harmonic oscillator model of the Si-H bond used in the derivation of the HCD model. In this model, the Si-H bond is intact in state 1 (blue circle) and in the anti-bonding state 2 (red circle) it is broken. Impinging carrier can induce a vibrational mode of the bond, thereby exciting the bond. Once excited to a higher level, the bond can be more easily dissociated (i.e. the hydrogen can overcome the potential barrier separating the bonded and transport configurations) either thermally or by a carrier bombarding the interface. This is because the more the bond is vibrating, that is the higher the oscillator level that is occupied, less energy is needed to rupture the Si-H bond.



important for scaled devices, where only few charge carriers reach kinetic energies beyond the dissociation energy of the Si-H bond (≈ 1.5 eV).

The truncated harmonic oscillator can be described by a set of rate equations, which read

$$\partial_t n_i = \begin{cases} k_{\downarrow}(n_{i+1} - n_i) - k_{\uparrow}(n_i - n_{i-1}) - r_i n_0 + p_i N_{it}^2 & \text{if } 0 < i < N - 1, \\ k_{\downarrow} n_1 - k_{\uparrow} n_0 - r_0 n_0 + p_0 N_{it}^2 & \text{if } i = 0, \\ -k_{\downarrow} n_{N-1} + k_{\uparrow} n_{N-2} - r_{N-1} n_{N-1} + p_{N-1} N_{it}^2 & \text{if } i = N - 1, \end{cases} \quad (7.2)$$

where n_i is the occupation number of the i th level, N is the number of levels, k_{\downarrow} and k_{\uparrow} are the bond deexcitation and excitation rates, p_i is the Si-H passivation rate for the i th level and r_i is the Si-H rupture rate for the i th level. In order to solve this rate equation it is assumed that there is a large difference in the timescales of the harmonic oscillator transitions and the passivation/depassivation processes. This assumption is justified by the huge disparity between phonon and Si-H bond lifetimes. Thus, the rate equation above can be reduced to

$$\partial_t N_{it} = (N_{it, \max} - N_{it})R - N_{it}^2 P, \quad (7.3)$$

$$\text{where } R = C \sum_m^{N-1} r_m \left(\frac{k_{\uparrow}}{k_{\downarrow}} \right)^m, \quad P = \sum_m^{N-1} p_m, \quad \text{with } C = \sum_m^{N-1} \left(\frac{k_{\downarrow}}{k_{\uparrow}} \right)^m, \quad (7.4)$$

which is solved analytically with the initial condition $N_{it}(0) = 0$, $N_{it}(t) \geq 0$ to read

$$N_{it}(t) = \frac{N_{it, \max} R}{R + \sqrt{R^2 + N_{it, \max} R P} + \coth(t \sqrt{R^2 + N_{it, \max} R P})}. \quad (7.5)$$

To complete the model, the bond passivation rate p_i is modeled using an Arrhenius law with $p_{N-1} = \nu_p \exp\left(-\frac{E_p}{k_B T_L}\right)$ and $p_i = 0$ for all other levels, where $\nu_p \approx 10^{13}$ /s is an attempt

frequency. The dissociation rate for the i th level with energy E_i is

$$r_i = I_{AB,i} + \nu_r \exp\left(-\frac{E_a - \mathbf{d} \cdot \mathbf{E}_{ox} - E_i}{k_B T_L}\right), \quad (7.6)$$

where ν_r is an attempt frequency, while \mathbf{d} is the bond dipole moment. This dipole moment and the dispersion of the activation energy E_a are discussed in the next section. I_i is the carrier acceleration integral [183, 184, 187] which represents the cumulative ability of the carrier ensemble to break the bonds, whereas the second term models interface state generation via thermal excitation. In the same fashion, the bond excitation and deexcitation rates are defined as

$$k_{\uparrow} = I_{MVE} + \omega' \exp\left(-\frac{\hbar\omega}{k_B T_L}\right) \quad \text{and} \quad k_{\downarrow} = I_{MVE} + \omega' \quad (7.7)$$

where ω' is the oscillator frequency. The acceleration integral for the anti-bonding process (AB) is defined as

$$I_{AB,i} = \sigma_0 \int_{E_{th}}^{\infty} f^{n/p}(\mathbf{x}_{it}, E, t) g^{n/p}(E) v_g^{n/p}(E) \left[\underbrace{(E - E_a + \mathbf{d} \cdot \mathbf{E}_{ox} + E_i)}_{=-E_{th}} / E_{ref} \right]^{11} dE, \quad (7.8)$$

where E_{th} is a threshold energy for electrons and holes respectively and the integral runs to infinity for electrons and to negative infinity for holes and $E_{ref} = 1$ eV. Note that the exponent of 11 in the equation above is an empirically found value. For multi-vibrational excitation the acceleration integral is defined as

$$I_{MVE} = \sigma_0 \int_{E_{th}}^{\infty} f^{n/p}(\mathbf{x}_{it}, E, t) g^{n/p}(E) v_g^{n/p}(E) \left[\underbrace{(E - \hbar\omega)}_{=E_{th}} / E_{ref} \right] dE, \quad (7.9)$$

and not dependent on the current state the oscillator is in. A graphical representation of the acceleration integral is given in Figure 7.8.

Activation Energy Dispersion and Bond Dipole Moment

The defect precursors (passivated Si-H bonds) at the semiconductor-oxide interface, even in modern semiconductor devices, are not of the same atomic and energetic configuration. This is due to the amorphous nature of the employed dielectrics and lattice mismatches at the interface. Thus, the energy needed to break Si-H bonds is naturally dispersed [188], which is in our model reflected by a normal distribution of E_a . The mean value of the activation energy was chosen to be 1.5 eV with a standard deviation of 0.15 eV, in good agreement with experimental data [171]. The importance of the activation energy dispersion was shown in [163]. Additionally, the dipole moment \mathbf{d} of the bond needs to be accounted for when dissociation occurs from the i th level [185]. This is done by subtracting $\mathbf{d} \cdot \mathbf{E}_{ox}$ from the energy $E_a - E_i$ needed to break the bond (cf. Equation (7.6)). To account for the dispersion of the activation energy E_a a Monte Carlo method, similar as for random discrete dopants, is used. Since any Monte Carlo method adds numeric noise to the solution another, deterministic approach can be taken: For example, assuming that the activation follows a Gaussian distribution one can sample this distribution using a sufficient number of sample points and calculate weight factors using the Gaussian distribution. Then for each sample point the model equations are being solved, where every N_{it} value is multiplied by the weight factor of the sample point such that the maximum N_{it} cannot be exceeded (normalization).

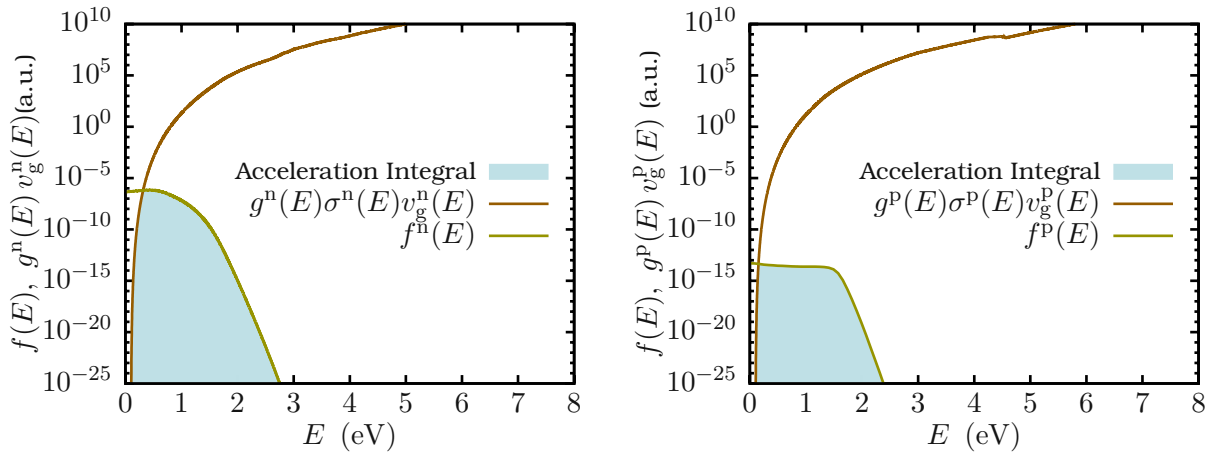


Figure 7.8: Exemplary distribution functions and acceleration integrals (shaded area) for electrons and holes from the middle of an artificial short channel (25 nm) n-channel MOSFET simulated with ViennaSHE [65] at hot-carrier stress with phonon and impurity scattering only. Well visible is the importance of the high energy tail for the calculation of the acceleration integral.

7.6 Requirements for Simulation

For the application of the present hot-carrier model, the evaluation of the acceleration integral (cf. Equation (7.8)) is necessary. To achieve this, the energy distribution of the charge carriers is required. In order to obtain the energy distribution, the Boltzmann Transport equation (BTE) has to be solved, which is challenging in its own right. Unfortunately, HCD is highly sensitive to the high-energy tail of the distribution, and therefore the modeling of the scattering operator requires special attention. In particular, impact ionization scattering as well as electron-electron interactions have to be incorporated. For example, it has been shown that the accuracy of the BTE solution ignoring electron-electron scattering can be seriously impacted [189]. Furthermore, it has been shown that the majority carriers can, in some cases, significantly contribute to the damage, requiring a coupled solution of the BTE for electrons and holes [174]. Finally, since an accurate resolution of the energy distribution at high energies is required, information about the full band structure has to be included into the model. Traditionally, this complicated problem is approached by using the Monte Carlo method [25], which is computationally- and time-intensive, particularly when the high-energy tails of the distribution function have to be resolved in detail [44]. Until a fully functional full-band spherical harmonics expansion (SHE) simulator with electron-electron scattering incorporated has been available the BTE had to be solved using the Monte Carlo approach [25]. The main advantages of the MC method are, that it is a well-known and investigated method, as well suited for the incorporation of the full dispersion relation (cf. Figure 2.1). However, the MC method is capable of resolving the whole phase space only at the cost of a square root dependence on the number of particles per magnitude of resolution in the distribution function. Aside from a limited time-scale, Monte Carlo simulators are not easily extendable to self-consistently solve the BTE, Poisson's Equation and the HCD model equations. Additionally, it has been shown that the use of strongly simplified models derived from approximations of the acceleration integral in moment based simulators leads to unsatisfactory results [190].

7.6.1 Importance of the High Energy Tail

Due to enormous computational burden of the MC method to resolve the high energy tail, a spherical harmonics expansion of the BTE to assess HCD is attractive. Thus, SHE based BTE solvers, such as ViennaSHE [65], have been developed in hopes to overcome the inadequacies of the Monte Carlo method. A SHE based simulator can easily be extended to numerically solve any defect equation self-consistently with the set of SHE-BTE equations. However, the greatest advantage of a SHE based simulator is that there is no numeric noise whatsoever in the energy distribution function. Nevertheless, SHE is still a fairly ‘young’ method. This means that not all possible numeric schemes have been sufficiently explored yet. Additionally, it is quite challenging to incorporate full-band effects into a SHE based simulator [47]. Nevertheless, as a solution technique to the BTE, a spherical harmonic expansion delivers the better noise/performance trade-off compared to Monte Carlo methods [44, 46].

7.7 Calibration

7.7.1 Experiment

In the course of this thesis our model was calibrated using HCD measurements provided by imec on 65 nm gate length and 150 nm gate length n-channel MOSFETs with an EOT of 2.5 nm [163]. The SiON gate oxide has been fabricated using a decoupled plasma nitridation followed by a post-nitridation anneal. For simulation we employed a device structure generated by the Sentaurus process simulator and subsequent calibration using MinimosNT and ViennaSHE. The 65 nm nMOSFET was stressed at $V_g = V_d = 1.8$ V and $V_g = V_d = 2.2$ V, while the 125 nm counterpart was subjected to HCD at $V_g = 0.9$ V, $V_d = 1.8$ V and $V_g = 1.1$ V, $V_d = 2.2$ V. These combinations of voltages correspond to worst-case conditions typical for short-channel (65 nm) and relatively long-channel (150 nm) MOSFETs, respectively. In both cases the ambient temperature was 298 K. To assess degradation, the linear drain current I_d in inversion under low field conditions, i.e. small channel electric field, was measured.

7.7.2 Results

In Figure 7.9 the fit of the model to the measurement data for the 65 nm nMOSFET are plotted disregarding various important ingredients of the model. If any of the mechanisms, such as MVE or EES, are neglected the measurement data cannot be reproduced by the model. This is also reflected in the corresponding N_{it} profiles (cf. Figure 7.10). From the acceleration integrals for the ground state (cf. Figure 7.9) it can be seen that impact ionization has a negligible effect on the HCD for this particular device. However, the influence of electron-electron scattering is massive, as expected for a short channel device. The effect of EES is naturally reflected also in the N_{it} profiles, where it causes an increase of N_{it} in the center of the channel right under the gate (cf. Figure 7.10). Interesting features are also the impact of the interaction between the electric field and the dipole moment as well as the contributions of the single hot carriers (anti-bonding process) into HCD. Contrary to previous speculations [46], the increase of N_{it} shifted towards the source as compared to the ‘classical’ N_{it} maximum near the drain, for the devices considered here, cannot be explained by a prominent majority carrier (hole) contribution. Instead, the electric field-dipole interaction and the activation energy dispersion seem to be responsible for this tendency. The later is highly important for a physics based description of HCD and Figure 7.10 clearly shows the consequence of neglecting the activation energy dispersion. Ignoring multi-vibrational excitation or activation energy dispersion does not drastically

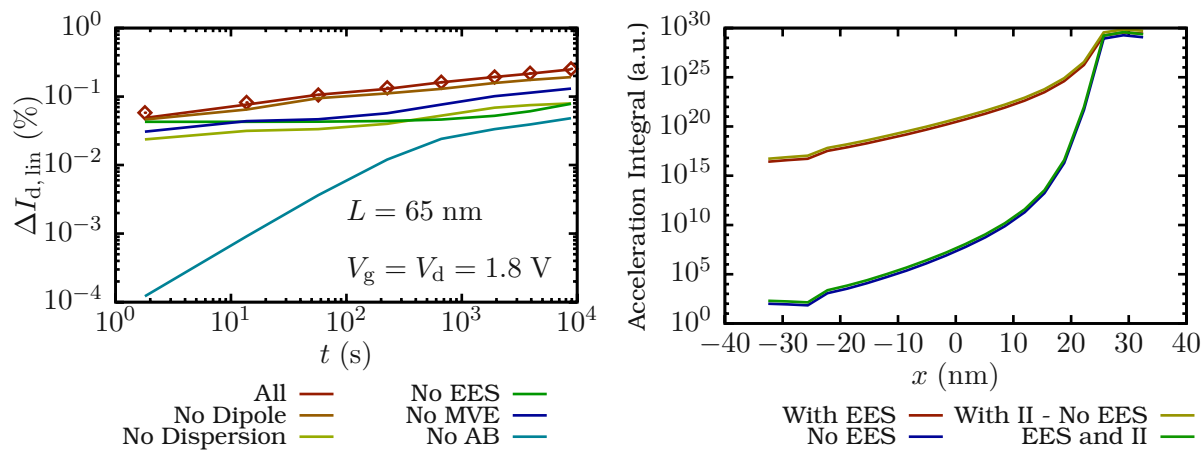


Figure 7.9: Left: The relative $\Delta I_{d,lin}$ degradation for the 65 nm n-channel MOSFET stressed under worst case conditions (lines) compared to the measurement data (diamonds). The influence of the various ingredients of the model is shown. The degradation predicted by the model is strongly influenced by electron-electron scattering and single hot carriers (anti-bonding processes). **Right:** Acceleration integrals along the channel (from source to drain at the interface) with and without considering electron-electron scattering and impact ionization.

change the N_{it} profiles, but causes a uniform shift of ΔI_d towards lower values (cf. Figure 7.9). The same line of reasoning applies to the 150 nm nMOSFET, for which experimental data is properly represented by the simulation results. Thus, our model is capable of predicting the $\Delta I_{d,lin}$ degradation for different devices (with various channel lengths) stressed using various combinations of operating voltages with the same set of model parameters. In Figure 7.11 the fits, using a single parameter set, for all stress conditions and devices are shown. All the ingredients of the model are necessary to represent the data. However, electron-electron scattering is most important for shorter channel devices as can be seen in Figure 7.11 for both stress conditions. A significant contribution from the majority charge carriers is, despite previous developments [46, 174], not necessary in this particular case to represent the experimental data for the long and short channel device for all stress conditions. This is due to the strong unipolar character of the MOSFETs under investigation, where an insignificant amount of majority carriers in the channel is typical. For the devices investigated here, the hole contribution HCD is negligibly small.

7.7.3 Importance of Electron-Electron Scattering

Electron-electron scattering is a two particle process by which one particle gains kinetic energy from another particle (cf. Figure 7.13), provided that the kinetic energy of both particles exceeds a certain threshold. Thereby it is, especially in short channels, possible to populate the high energy tails of the distribution function. The elevation of the high energy tail in turn then increases the probability of a single electron to break, upon interaction, a Si-H bond at the semiconductor-oxide interface. This has been exhaustively shown by [46, 163, 189] that solving the BTE without any electron-electron scattering leads to an underestimation of the hot-carrier degradation. In Figure 7.12 the electron and hole acceleration integrals and the carrier distribution functions are shown, if only phonon (acoustic and optical) and impurity

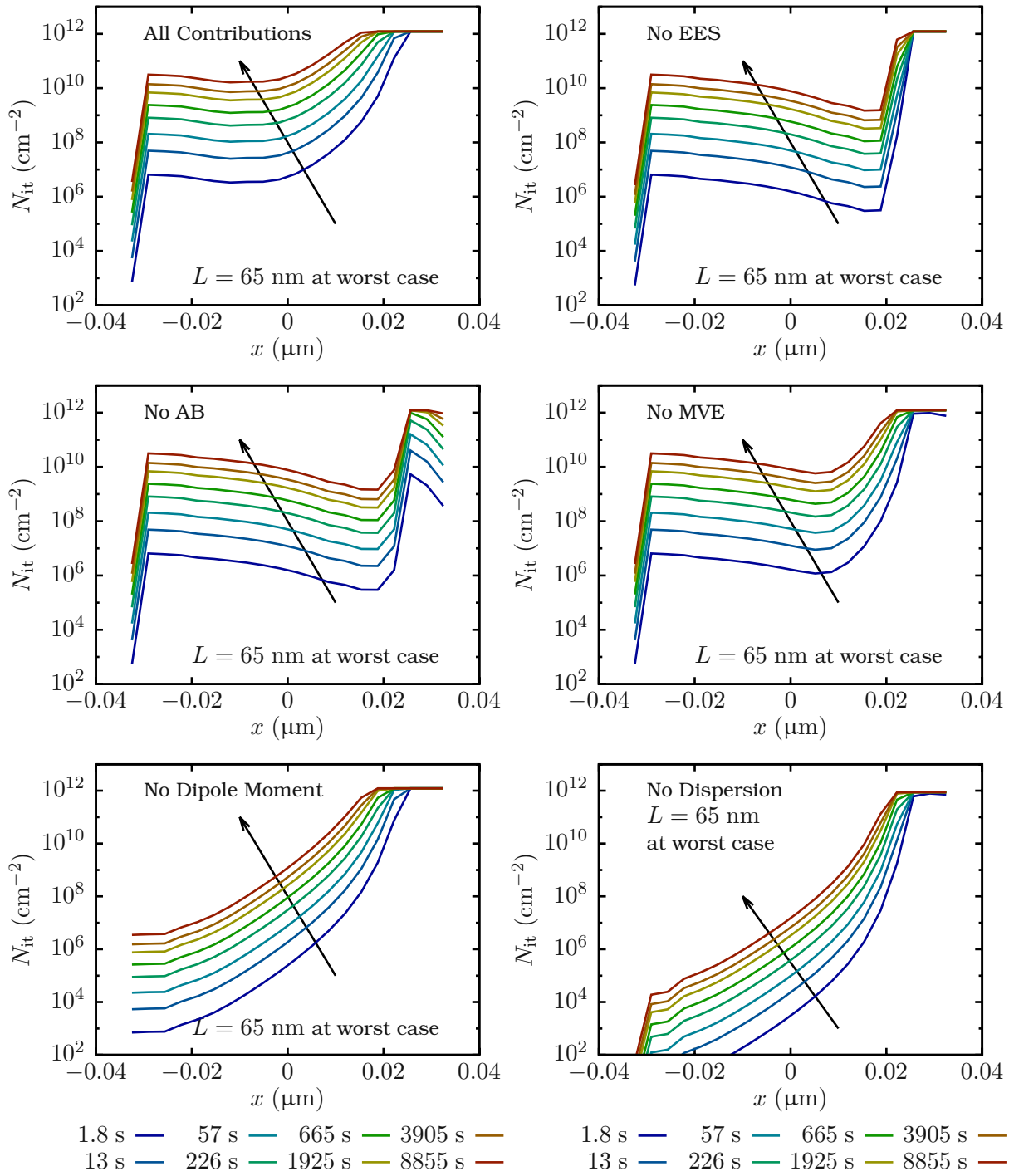


Figure 7.10: N_{it} profiles predicted by the model for the 65 nm n-channel MOSFET stressed under worst case conditions. The source side is located at negative x . Noteworthy is the strong influence of the dipole moment and the activation energy dispersion, which strongly affect N_{it} close to the source side, but shows a smaller influence on the drain current degradation as compared to electron-electron scattering. The arrow indicates the growing N_{it} with increasing stress time.

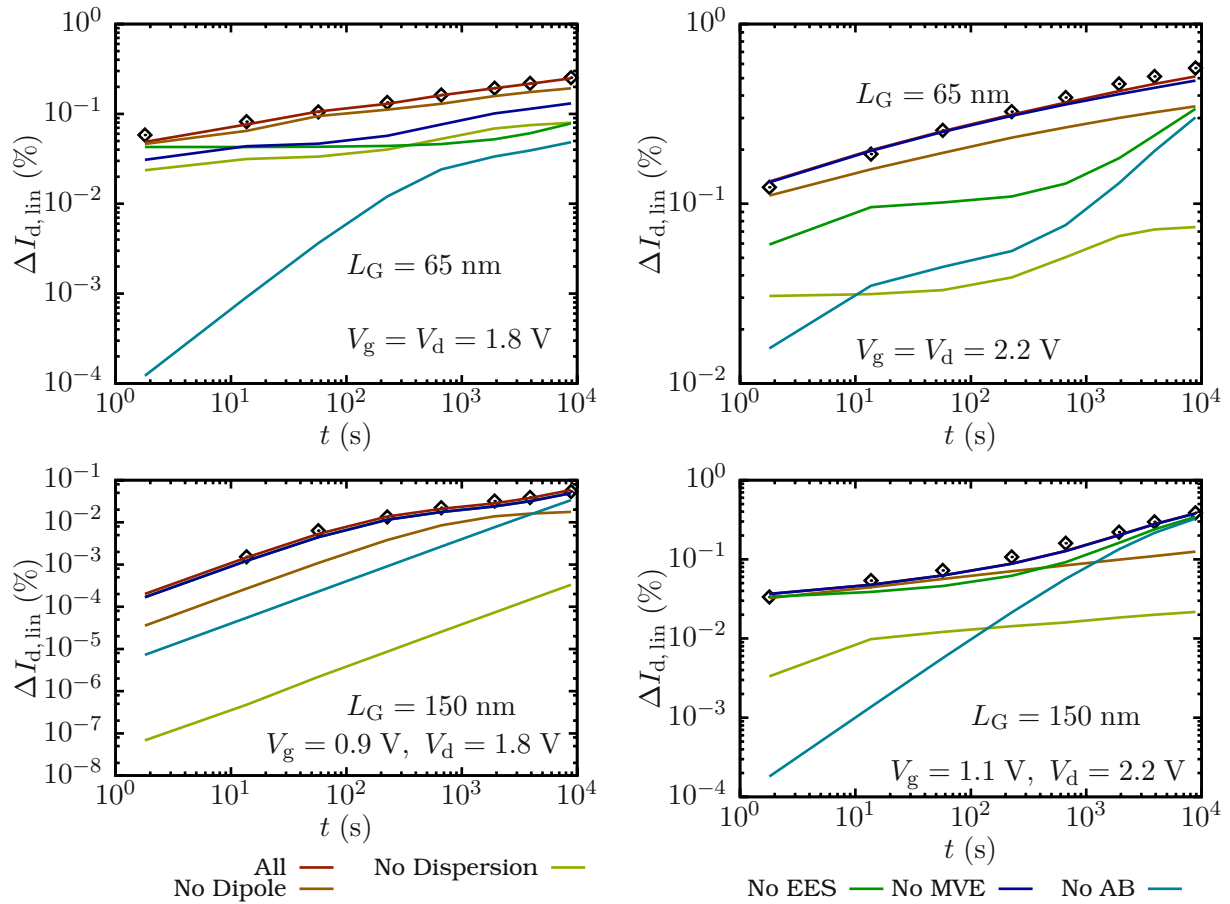


Figure 7.11: Relative drain current degradation as predicted by our model for both the 65 nm and the 150 nm (gate length) n-channel device for all stress conditions compared to measurement data (diamonds). Both short and long channel devices show a significant dependence on activation energy dispersion and hot carriers (anti-bonding processes). However, electron-electron scattering mostly affects only the short channel device.

scattering are considered in comparison to the acceleration integrals obtained by additionally considering electron-electron scattering. It has been shown that only the effect of impact ionization can, in certain cases, be safely neglected [163]. In Figure 7.14 a few selected electron distribution functions at the interface are shown for both devices stressed under worst case conditions. As we demonstrated before, HCD is very sensitive particularly to the high energy tails of the distribution function, and therefore EES is of great importance in the context of hot-carrier degradation modeling.

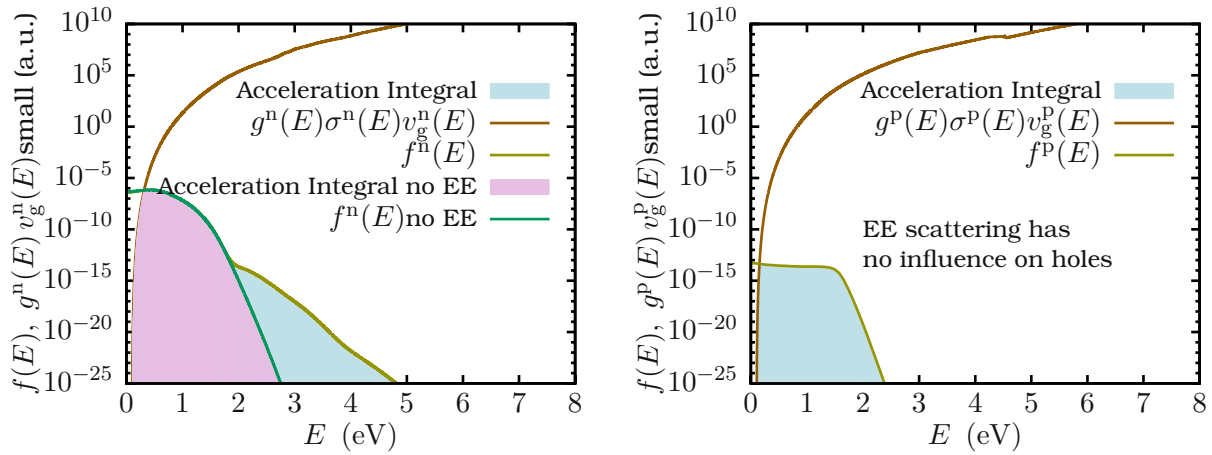


Figure 7.12: The influence of electron-electron scattering on the acceleration integral. The distribution functions have been obtained using ViennaSHE and the electron-electron scattering operator therein [61]. The graphical comparison of the acceleration integrals with and without electron-electron scattering shows the importance of EES for hot-carrier degradation modelling.

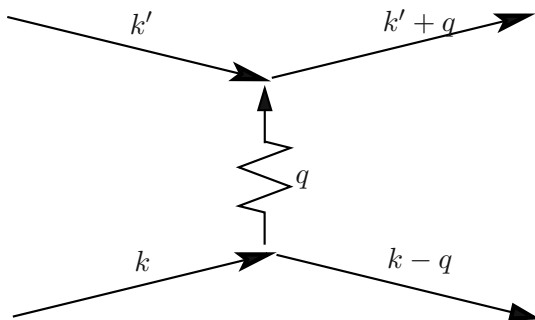


Figure 7.13: Electron-electron scattering by exchange of a virtual phonon. Both electrons with wave vector k' and k exchange a virtual phonon. Through this exchange, the wavevector k' is increased by q and the wavevector k is decreased by q . Thus the electron associated with k' has a higher kinetic energy than the electron associated with k .

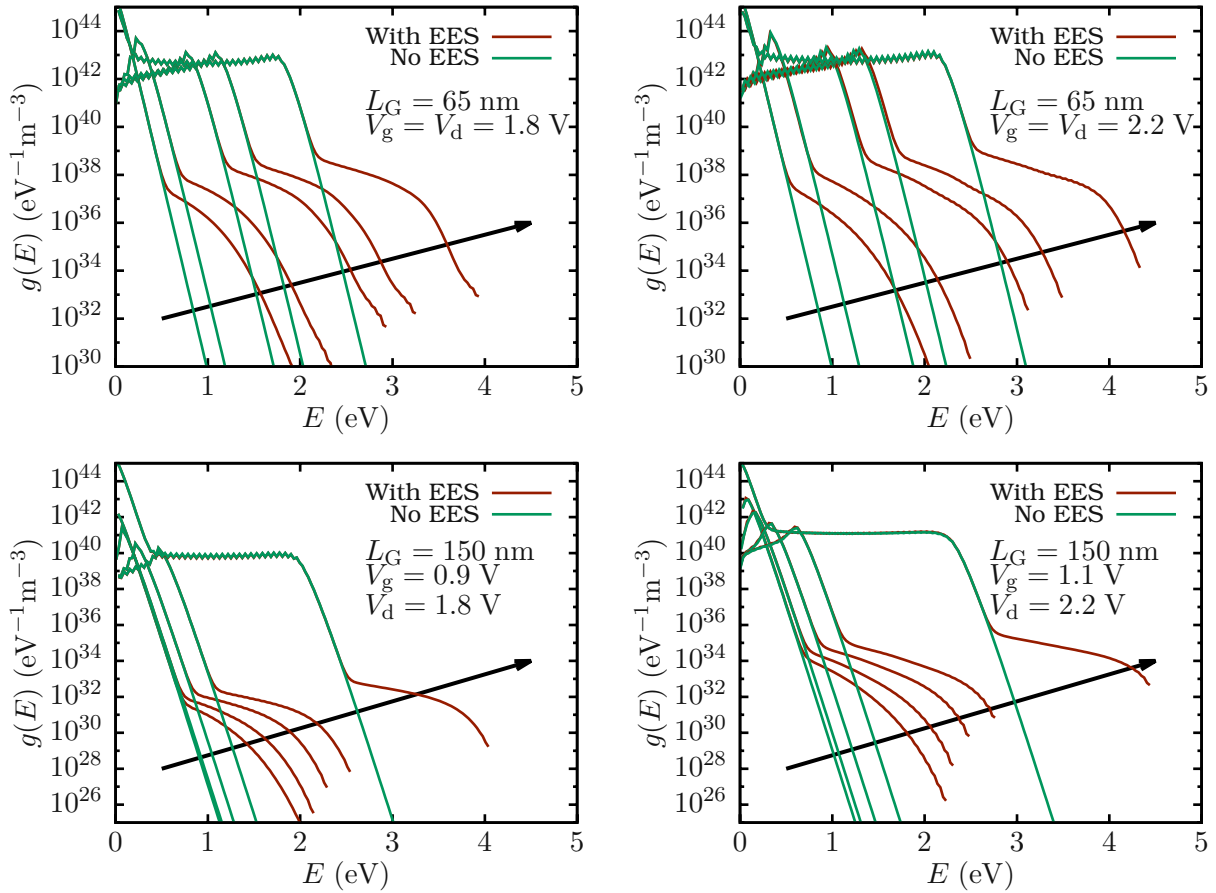


Figure 7.14: Selected distribution functions with and without electron-electron scattering at the interface along the channel for the 65 nm and the 150 nm n-channel device (the source to drain direction is indicated by the arrow). The significance of EES for both channel lengths and all stress conditions is evident. Due to EES the high energy tails of the distribution functions are elevated. Also notable are the phonon-cascades visible at smaller energies.

It's my experience that in modelling tradeoffs 'pain' is conserved.

(Tibor Grasser)

8 Conclusions

Scaling devices down to the nano-meter regime increases the probability of functional failure due to single point defects or parametric fluctuations, such as fluctuations in the doping. The effect of single point defects and parametric fluctuations on the figures of merit, such as the threshold voltage, of MOSFETs are intertwined. In the first part of this thesis it has been shown that, in scaled devices, the characteristic discrete threshold voltage shifts caused by charged single defects (step-heights) is highly dependent on the exact positions of the dopants and defects in the device. Thus the effect of potential fluctuations caused by the random discrete doping (RDD) has been assessed first, since in sub-100nm node devices the discreteness of the doping cannot be ignored anymore. Afterwards, the electrostatic interaction of charged single point defects and random discrete dopants has been investigated, neglecting the charge reaction kinetics of the single point defects. In these studies it has been shown that an electrostatic and drift-diffusion-based picture to determine the characteristic step-heights caused by single defects is insufficient. Subsequently it has been shown that effects other than pure electrostatic influences, such as fluctuations in the charge carrier mobility, need to be taken into account. In this respect a drift-diffusion-based transport model, even with quantum mechanical corrections, is insufficient to reproduce the step-heights caused by single charged defects. When taking the trapping kinetics into account, two main modes of defect creation have been discussed, namely the bias temperature instability (BTI) and hot-carrier degradation (HCD). For BTI, the modelling of the tremendously electric field and temperature dependent defect creation as well as the charge trapping kinetics have been discussed. The most accurate model for the recoverable component of BTI today is the four-state non-radiative multiphonon model, where it was shown that this model can not only reproduce classic BTI degradation experiments, but also can explain trap assisted tunneling after bias temperature stress to high accuracy. Additionally, it has been shown that the electric field dependence of BTI, which is modelled using a non-radiative multiphonon theory, can lead to problems of trap self-interaction in self-consistent simulations. This self-interaction is independent of the employed transport model and has been shown to artificially broaden the distribution of the trapping time constants. However, it is possible to assess the amount of self-interacting traps by comparing first-order trapping kinetics to the trapping kinetics predicted by a self-consistent simulation. Since trap self-interaction is closely related to the field dependence of the characteristic trapping time constants, trap self-interaction is not an issue when investigating HCD. In the latest models, developed in the course of this thesis, standard SRH charge trapping kinetics are assumed for hot-carrier induced traps which show a weak dependence of the trapping time constants on the electric field. Contrary to the trapping kinetics of hot-carrier induced defects, the creation of the same defects is highly dependent on the electric field. However, it has been shown that defect creation caused by hot carriers is highly sensitive to the kinetic energy of the charge carriers and cannot be modelled by taking the electric field into account alone. Thus, to inves-

tigate HCD one needs exact information on the energy distribution of charge carriers. This in turn requires an efficient method to solve the Boltzmann Transport equation (BTE) for at least 2D devices. As laid out in this thesis, the usual Monte Carlo approach is insufficient for HCD, due to the inherent numerical noise in the solution. In this respect the deterministic approach to expand the BTE in spherical harmonics (SHE) has proven to be the method of choice for investigations into hot-carrier degradation. Finally, employing the SHE-based simulator ViennaSHE it was exhaustively shown that the developed hot-carrier model can predict HCD in devices of different channel lengths for various stress conditions.

9 Outlook

It has been shown to be insufficient to determine the characteristic step-heights caused by single defects in electrostatic and drift-diffusion-based simulations. Consequently, variability models for RDD and first order quantum correction models have been integrated for future investigations into the SHE-based simulator ViennaSHE. The advantage of directly solving the BTE is that changes in the mobility and charge carrier energy in the presence of RDD can be directly investigated. This is necessary since in the following [98] one needs to reproduce the low-field mobility-doping relation found by Caughy and Thomas [191], when using RDD in a large resistor with a, at least one nanometer, fine mesh. To calibrate the RDD model in ViennaSHE one should undertake a statistical significant number of simulations with RDD to find the same resistance in the limit of an infinite large slab of silicon with RDD as one would find using continuous doping under low-field conditions. In all simulations with RDD fine meshes are required to correctly resolve the granularity of the doping and leads to a high computational demand. Thus, for this investigation the simulator ViennaSHE needs to be parallelized using MPI [192] and an algebraic multi-grid [193] solver such as Prometheus [194] is required.

Bibliography

- [1] M. Masuduzzaman, A. Islam, and M. Alam, "Exploring the Capability of Multifrequency Charge Pumping in Resolving Location and Energy Levels of Traps Within Dielectric," *IEEE Transactions on Electron Devices*, vol. 55, no. 12, pp. 3421–3431, 2008.
- [2] W. Gös and T. Grasser, "Charging and Discharging of Oxide Defects in Reliability Issues," in *IEEE International Integrated Reliability Workshop Final Report*, pp. 27–32, 2007.
- [3] A. Asenov, "Random Dopant Induced Threshold Voltage Lowering and Fluctuations in sub-0.1 μm MOSFET's: A 3-D "atomistic" Simulation Study," *IEEE Transactions on Electron Devices*, vol. 45, no. 12, pp. 2505–2513, 1998.
- [4] M. Bina, O. Triebel, B. Schwarz, M. Karner, B. Kaczer, and T. Grasser, "Simulation of Reliability on Nanoscale Devices," pp. 109–112, 2012.
- [5] A. Wettstein, O. Penzin, and E. Lyumkis, "Integration of the Density Gradient Model Into a General Purpose Device Simulator," *VLSI Design*, vol. 15, no. 4, pp. 751–759, 2002.
- [6] N. Sano, K. Matsuzawa, M. Mukai, and N. Nakayama, "Role of Long-Range and Short-Range Coulomb Potentials in Threshold Characteristics Under Discrete Dopants in sub-0.1 μm Si-MOSFETs," in *IEEE International Electron Devices Meeting (IEDM)*, pp. 275–278, IEEE, 2000.
- [7] C. Jungemann, C. Nguyen, B. Neinhüs, S. Decker, and B. Meinerzhagen, "Improved Modified Local Density Approximation for Modeling of size Quantization in NMOSFETs," in *Proceedings of the International Conference on Modeling and Simulation of Microsystems*, 2001.
- [8] D. Hisamoto, W. Lee, J. Kedzierski, H. Takeuchi, K. Asano, C. Kuo, E. Anderson, T. King, J. Bokor, and C. Hu, "FinFET—a self-aligned double-gate MOSFET scalable to 20 nm," *IEEE Transactions on Electron Devices*, vol. 47, no. 12, pp. 2320–2325, 2000.
- [9] A. Wettstein, *Quantum Effects in MOS Devices*. PhD thesis, ETH Zürich, 2000.
- [10] M. Born and R. Oppenheimer, "Zur Quantentheorie der Moleküle," *Annalen der Physik*, vol. 389, pp. 457–484, 1927.
- [11] J. M. Ziman, *Electrons and Phonons: The Theory of Transport Phenomena in Solids*. Oxford University Press, 2001.
- [12] O. Madelung, *Introduction to Solid-State Theory*. Springer Series in Solid-State Sciences, Springer, 1996.

- [13] H. Kosina, *Simulation des Ladungstransportes in elektronischen Bauelementen mit Hilfe der Monte-Carlo-Methode*. PhD thesis, TU Wien, 1992.
- [14] C. Kittel, *Introduction to Solid State Physics, 7th Edition*. Wiley India Pvt. Limited, 2007.
- [15] F. Bloch, "Über die Quantenmechanik der Elektronen in Kristallgittern," *Zeitschrift für Physik*, vol. 52, no. 7-8, pp. 555-600, 1929.
- [16] L. Brillouin and A. Sommerfeld, *Wave Propagation and Group Velocity*, vol. 960. Academic Press New York, 1960.
- [17] L. Brillouin, *Wave Propagation in Periodic Structures: Electric Filters and Crystal Lattices*. Dover Publications, 2003.
- [18] O. Madelung, U. Rössler, and M. Schulz, eds., vol. 41A1b of *Landolt-Börnstein - Group III Condensed Matter*, ch. Silicon (Si), band structure.
- [19] J. R. Chelikowsky and M. L. Cohen, "Nonlocal Pseudopotential Calculations for the Electronic Structure of eleven Diamond and Zinc-Blende Semiconductors," *Physical Review B*, vol. 14, pp. 556-582, 1976.
- [20] C. Herring and E. Vogt, "Transport and Deformation-Potential Theory for Many-Valley Semiconductors with Anisotropic Scattering," *Physical Review*, vol. 101, pp. 944-661, 1955.
- [21] E. O. Kane, "Band Structure of Indium Antimonide," *Journal of Physics and Chemistry of Solids*, vol. 1, no. 4, pp. 249 - 261, 1957.
- [22] M. Lundstrom, *Fundamentals of Carrier Transport*. Cambridge University Press, 2009.
- [23] C. Jacoboni and P. Lugli, *The Monte Carlo Method for Semiconductor Device Simulation*. Computational Microelectronics, Springer, 1989.
- [24] C. Jacoboni and L. Reggiani, "The Monte Carlo Method for the Solution of Charge Transport in Semiconductors with Applications to Covalent Materials," *Reviews of Modern Physics*, vol. 55, pp. 645-705, 1983.
- [25] C. Jungemann and B. Meinerzhagen, *Hierarchical Device Simulation: The Monte-Carlo Perspective*. Computational Microelectronics, Springer, 2003.
- [26] H. Brooks and C. Herring, "Scattering by Ionized Impurities in Semiconductors," *Physical Review*, vol. 83, no. 4, p. 879, 1951.
- [27] E. Cartier, M. V. Fischetti, E. A. Eklund, and F. R. McFeely, "Impact ionization in Silicon," *Applied Physics Letters*, vol. 62, no. 25, pp. 3339-3341, 1993.
- [28] P. Lugli and D. Ferry, "Effect of Electron-Electron and Electron-Plasmon Interactions on Hot Carrier Transport in Semiconductors," *Physica B+C*, vol. 129, no. 1, pp. 532-536, 1985.
- [29] T. Grasser, "Stochastic Charge Trapping in Oxides From Random Telegraph Noise to Bias Temperature Instabilities," *Microelectronics Reliability*, vol. 52, no. 1, pp. 39-70, 2012.

- [30] A. Piazza, C. Korman, and A. Jaradeh, "A Physics-based Semiconductor Noise Model Suitable for Efficient Numerical Implementation," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 18, no. 12, pp. 1730-1740, 1999.
- [31] W. Shockley and W. T. Read, "Statistics of the Recombinations of Holes and Electrons," *Physical Review*, vol. 87, pp. 835-842, 1952.
- [32] R. Stratton, "Diffusion of Hot and Cold Electrons in Semiconductor Barriers," *Physical Review*, vol. 126, pp. 2002-2014, 1962.
- [33] K. Blotekjaer, "High-Frequency Conductivity, Carrier Waves, and Acoustic Amplification in Drifted Semiconductor Plasmas," *Ericsson Technics*, vol. 2, pp. 126-183, 1966.
- [34] M. Gritsch, *Numerical Modeling of Silicon-on-Insulator MOSFETs*. PhD thesis, TU Wien, 2002.
- [35] S. Selberherr, *Analysis and Simulation of Semiconductor Devices*. Springer-Verlag, 1984.
- [36] C. Jungemann and B. Meinerzhagen, "On the High Frequency Limit of the Impedance Field Method for Si," in *International Conference on Noise in Physical Systems and 1/f Fluctuations*, vol. 780, pp. 799-802, 2005.
- [37] J. Ruch, "Electron Dynamics in Short Channel Field-Effect Transistors," *IEEE Transactions on Electron Devices*, vol. 19, no. 5, pp. 652-654, 1972.
- [38] T. Grasser and S. Selberherr, "Limitations of hydrodynamic and energy-transport models," in *Proceedings of the International Society for Optical Engineering*, vol. 1, pp. 584-591, 2002.
- [39] T. Grasser, H. Kosina, M. Gritsch, and S. Selberherr, "Using Six Moments of Boltzmann's Transport Equation for Device Simulation," *Journal of Applied Physics*, vol. 90, no. 5, pp. 2389-2396, 2001.
- [40] J. Hammersley and D. Handscomb, *Monte Carlo Methods*. Methuen's monographs on applied probability and statistics, Methuen, 1964.
- [41] H. Kosina, M. Nedjalkov, and S. Selberherr, "Theory of the Monte Carlo Method for Semiconductor Device Simulation," *IEEE Transactions on Electron Devices*, vol. 47, no. 10, pp. 1898-1908, 2000.
- [42] P. A. Childs and C. C. C. Leung, "A one-dimensional Solution of the Boltzmann Transport Equation including Electron-Electron Interactions," *Journal of Applied Physics*, vol. 79, no. 1, pp. 222-227, 1996.
- [43] A. Zaka, P. Palestri, Q. Rafhay, R. Clerc, M. Iellina, D. Rideau, C. Tavernier, G. Pananakakis, H. Jaouen, and L. Selmi, "An Efficient Nonlocal Hot Electron Model Accounting for Electron-Electron Scattering," *IEEE Transactions on Electron Devices*, vol. 59, pp. 983-993, April 2012.
- [44] B. Meinerzhagen, A. Pham, S.-M. Hong, and C. Jungemann, "Solving Boltzmann Transport Equation without Monte-Carlo Algorithms - New Methods for Industrial TCAD Applications," in *International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pp. 293-296, 2010.

- [45] C. Jungemann, S. Keith, and B. Meinerzhagen, "Full-band Monte Carlo Device Simulation of a Si/SiGe-HBT with a Realistic Ge Profile," *IEICE Transactions on Electronics*, vol. 83, no. 8, pp. 1228–1234, 2000.
- [46] M. Bina, K. Rupp, S. Tyaginov, O. Triebel, and T. Grasser, "Modeling of Hot Carrier Degradation using a Spherical Harmonics Expansion of the bipolar Boltzmann Transport Equation," in *IEEE International Electron Devices Meeting (IEDM)*, pp. 30.5.1–30.5.4, 2012.
- [47] S. Hong, A. Pham, and C. Jungemann, *Deterministic Solvers for the Boltzmann Transport Equation*. Springer, 2011.
- [48] J. Seonghoon, S. Hong, and C. Jungemann, "An Efficient Approach to Include Full-Band Effects in Deterministic Boltzmann Equation Solver Based on High-Order Spherical Harmonics Expansion," *IEEE Transactions on Electron Devices*, vol. 58, no. 5, pp. 1287–1294, 2011.
- [49] D. Ruić and C. Jungemann, "A self-consistent Solution of the Poisson, Schrödinger and Boltzmann Equations by a Full Newton-Raphson Approach for Nanoscale Semiconductor Devices," in *International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, 2013.
- [50] R. Oberhuber, G. Zandler, and P. Vogl, "Subband Structure and Mobility of Two-Dimensional Holes in Strained Si/SiGe MOSFETs," *Physical Review B*, vol. 58, no. 15, p. 9941, 1998.
- [51] O. Baumgartner, Z. Stanojevic, K. Schnass, M. Karner, and H. Kosina, "VSP-A quantum-electronic Simulation Framework," *Journal of Computational Electronics*, vol. 12, no. 4, pp. 701–721, 2013.
- [52] A. Gehring, *Simulation of Tunneling in Semiconductor Devices*. PhD thesis, TU Wien, 2003.
- [53] M. P. Anantram, M. S. Lundstrom, and D. E. Nikonov, "Modeling of Nanoscale Devices," *Proceedings of the IEEE*, vol. 96, no. 9, pp. 1511–1550, 2008.
- [54] S. Datta, "The Non-Equilibrium Green's Function (NEGF) Formalism: An Elementary Introduction," in *IEEE International Electron Devices Meeting (IEDM)*, pp. 703–706, 2002.
- [55] L. Kadanoff and G. Baym, *Quantum Statistical Mechanics: Green's Function Methods in Equilibrium and Nonequilibrium Problems*. Frontiers in Physics, W.A. Benjamin, 1962.
- [56] O. Baumgartner, *Simulation of Quantum Transport Using the Non-Equilibrium Green's Functions Formalism*. PhD thesis, TU Wien, 2006.
- [57] K. Rupp, T. Grasser, and A. Jüngel, "On the Feasibility of Spherical Harmonics Expansions of the Boltzmann Transport Equation for Three-Dimensional Device Geometries," in *IEEE International Electron Devices Meeting (IEDM)*, 2011.
- [58] M. Vecchi and M. Rudan, "Modeling Electron and Hole Transport with Full-band Structure Effects by means of the Spherical-Harmonics Expansion of the BTE," *IEEE Transactions on Electron Devices*, vol. 45, no. 1, pp. 230–238, 1998.

- [59] S.-M. Hong, G. Matz, and C. Jungemann, "A Deterministic Boltzmann Equation Solver Based on a Higher Order Spherical Harmonics Expansion with Full-Band Effects," *IEEE Transactions on Electron Devices*, vol. 57, no. 10, pp. 2390–2397, 2010.
- [60] C. Jungemann, A. Pham, S. Hong, and B. Meinerzhagen, "Deterministic Simulation of 3D and quasi-2D Electron and Hole Systems in SiGe Devices," in *Proceedings of the European Solid-State Device Research Conference (ESSDERC)*, pp. 318–321, 2012.
- [61] K. Rupp, P. W. Lagger, T. Grasser, and A. Jüngel, "Inclusion of Carrier-Carrier-Scattering Into Arbitrary-Order Spherical Harmonics Expansions of the Boltzmann Transport Equation," in *International Workshop on Computational Electronics*, pp. 1–4, 2012.
- [62] K. Rupp, C. Jungemann, M. Bina, A. Jungel, and T. Grasser, "Bipolar Spherical Harmonics Expansions of the Boltzmann Transport Equation," in *International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pp. 19–22, 2012.
- [63] C.-K. Lin, N. Goldsman, Z. Han, I. Mayergoyz, S. Yu, M. Stettler, and S. Singh, "Frequency Domain Analysis of the Distribution Function by Small Signal Solution of the Boltzmann and Poisson Equations," in *International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pp. 39–42, 1999.
- [64] N. Goldsman, C. Lin, Z. Han, and C. Huang, "Advances in the Spherical Harmonic-Boltzmann-Wigner Approach to Device Simulation," *Superlattices and Microstructures*, vol. 27, pp. 159–175, 2000.
- [65] "ViennaSHE Device Simulator." <http://viennashe.sourceforge.net/>.
- [66] S.-M. Hong and C. Jungemann, "A Fully Coupled Scheme for a Boltzmann-Poisson Equation Solver Based on a Spherical Harmonics Expansion," *Journal of Computational Electronics*, vol. 8, pp. 225–241, 2009.
- [67] K. Rupp, *Deterministic Numerical Solution of the Boltzmann Transport Equation*. PhD thesis, TU Wien, 2011.
- [68] C. Jungemann, A. Pham, B. Meinerzhagen, C. Ringhofer, and M. Bollhofer, "Stable Discretization of the Boltzmann Equation based on Spherical Harmonics, Box Integration, and a Maximum Entropy Dissipation Principle," *Journal of Applied Physics*, vol. 100, no. 2, pp. 024502–024502, 2006.
- [69] A. Gnudi, D. Ventura, G. Bacarani, and F. Odeh, "Two-dimensional MOSFET Simulation by means of a Multidimensional Spherical Harmonics Expansion of the Boltzmann Transport Equation," *Solid-State Electronics*, vol. 36, no. 4, pp. 575 – 581, 1993.
- [70] C. Ringhofer, "Numerical Methods for the Semiconductor Boltzmann Equation Based on Spherical Harmonics Expansions and Entropy Discretizations," *Transport Theory and Statistical Physics*, vol. 31(4-6), pp. 431–452, 2002.
- [71] K. Rupp, A. Jüngel, and T. Grasser, "Matrix Compression for Spherical Harmonics Expansions of the Boltzmann Transport Equation for Semiconductors," *Journal of Computational Physics*, vol. 229, no. 23, pp. 8750–8765, 2010.

- [72] S.-M. Hong and C. Jungemann, "Inclusion of the Pauli Principle in a Deterministic Boltzmann Equation Solver Based on a Spherical Harmonics Expansion," *Journal of Computational Electronics*, vol. 9, no. 3-4, pp. 153-159, 2010.
- [73] K. Levenberg, "A Method for the Solution of Certain Problems in Least Squares," *Quarterly Applied Mathematics*, vol. 2, pp. 164-168, 1944.
- [74] D. W. Marquardt, "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," *Journal of the Society for Industrial & Applied Mathematics*, vol. 11, no. 2, pp. 431-441, 1963.
- [75] C. Lin, N. Goldsman, I. D. Mayergoyz, and C. Chang, "A Transient Solution of the Boltzmann Equation Exposes Energy Overshoot in Semiconductor Devices," *Journal of Applied Physics*, vol. 86, no. 1, pp. 468-475, 1999.
- [76] G. Dimarco, L. Pareschi, and V. Rispoli, "Implicit-Explicit Runge-Kutta schemes for the Boltzmann-Poisson System for Semiconductors," *ArXiv e-prints*, May 2013.
- [77] J. Butcher, *Numerical Methods for Ordinary Differential Equations*. Wiley, 2008.
- [78] J. Strikwerda, *Finite Difference Schemes and Partial Differential Equations*. Society for Industrial and Applied Mathematics, 2007.
- [79] W. Shockley, G. L. Pearson, and J. R. Haynes, "Hole Injection in Germanium - Quantitative Studies and Filamentary Transistors," *Bell System Technical Journal*, vol. 28, pp. 344-366, 1949.
- [80] G. Paasch and H. Übensee, "A Modified Local Density Approximation. Electron Density in Inversion Layers," *Physica Status Solidi (B)*, vol. 113, no. 1, pp. 165-178, 1982.
- [81] M. G. Ancona and G. J. Iafrate, "Quantum Correction to the Equation of State of an Electron Gas in a Semiconductor," *Physical Review B*, vol. 39, pp. 9536-9540, 1989.
- [82] A. Wettstein, A. Schenk, and W. Fichtner, "Quantum Device-Simulation with the Density-Gradient Model on Unstructured Grids," vol. 48, no. 2, pp. 279-284, 2001.
- [83] A. Wettstein, A. Schenk, and W. Fichtner, "Quantum Device-Simulation with the Density-Gradient Model on Unstructured Grids," vol. 15, no. 4, p. 751-759, 2002.
- [84] H. Tsuchiya, B. Fischer, and K. Hess, "A Full-Band Monte Carlo Model for Silicon Nanoscale Devices with a Quantum Mechanical Correction of the Potential," in *IEEE International Electron Devices Meeting (IEDM)*, pp. 283-286, Dec 2000.
- [85] S. Odanaka, "Multidimensional Discretization of the Stationary Quantum Drift-Diffusion Model for Ultrasmall MOSFET Structures," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 23, no. 6, pp. 837-842, 2004.
- [86] M. G. Ancona and B. A. Biegel, "Nonlinear Discretization Scheme for the Density-Gradient Equations," in *International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pp. 196-199, 2000.
- [87] S. Odanaka, "A High-Resolution Method for Quantum Confinement Transport Simulations in MOSFETs," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 26, no. 1, pp. 80-85, 2007.

- [88] M. G. Ancona, "Finite-Difference Schemes for the Density-Gradient Equations," *Journal of Computational Electronics*, vol. 1, no. 3, pp. 435–443, 2002.
- [89] T.-W. Tang, X. Wang, and Y. Li, "Discretization Scheme for the Density-Gradient Equation and Effect of Boundary Conditions," *Journal of Computational Electronics*, vol. 1, no. 3, pp. 389–393, 2002.
- [90] B. Schwarz, *Simulation of Random Dopant Fluctuations with a Quantum-Corrected Drift-Diffusion Model*. PhD thesis, TU Wien, 2012.
- [91] T. Höhr, A. Schenk, A. Wettstein, and W. Fichtner, "On Density-Gradient Modeling of Tunneling through Insulators," *IEICE Transactions on Electronics*, vol. 86, no. 3, pp. 379–384, 2003.
- [92] N. Sano and M. Tomizawa, "Random Dopant Model for Three-Dimensional Drift-Diffusion Simulations in Metal-Oxide-Semiconductor Field-Effect-Transistors," *Applied Physics Letters*, vol. 79, no. 14, pp. 2267–2269, 2001.
- [93] A. Asenov, G. Slavcheva, A. R. Brown, J. H. Davies, and S. Saini, "Increase in the Random Dopant Induced Threshold Fluctuations and Lowering in sub-100 nm MOSFETs due to Quantum Effects: a 3-D Density-Gradient Simulation Study," *IEEE Transactions on Electron Devices*, vol. 48, no. 4, pp. 722–729, 2001.
- [94] A. R. Brown, N. M. Idris, J. R. Watling, and A. Asenov, "Impact of Metal Gate Granularity on Threshold Voltage Variability: A full-scale three-dimensional Statistical Simulation Study," *IEEE Electron Device Letters*, vol. 31, no. 11, pp. 1199–1201, 2010.
- [95] T. Grasser, *Bias Temperature Instability for Devices and Circuits*. Springer London, Limited, 2013.
- [96] H.-S. Wong and Y. Taur, "Three-dimensional 'atomistic' Simulation of Discrete Random Dopant Distribution Effects in sub-0.1 μm MOSFET's," in *IEEE International Electron Devices Meeting (IEDM)*, pp. 705–708, 1993.
- [97] E. Conwell and V. F. Weisskopf, "Theory of Impurity Scattering in Semiconductors," *Physical Review*, vol. 77, no. 3, p. 388, 1950.
- [98] C. Alexander, A. R. Brown, J. R. Watling, and A. Asenov, "Impact of Scattering in 'atomistic' Device Simulations," *Solid-state Electronics*, vol. 49, no. 5, pp. 733–739, 2005.
- [99] A. Asenov and S. Saini, "Suppression of random dopant-induced threshold voltage fluctuations in sub-0.1- μm mosfet's with epitaxial and δ -doped channels," *IEEE Transactions on Electron Devices*, vol. 46, no. 8, pp. 1718–1724, 1999.
- [100] B. Kaczer, S. Mahato, V. de Almeida Camargo, M. Toledano-Luque, P. Roussel, T. Grasser, F. Catthoor, P. Dobrovolny, P. Zuber, G. Wirth, and G. Groeseneken, "Atomistic approach to variability of bias-temperature instability in circuit simulations," in *IEEE International Reliability Physics Symposium (IRPS)*, pp. XT.3.1–XT.3.5, 2011.
- [101] B. Kaczer, P. Roussel, T. Grasser, and G. Groeseneken, "Statistics of Multiple Trapped Charges in the Gate Oxide of Deeply Scaled MOSFET Devices-Application to NBTI," *IEEE Electron Device Letters*, vol. 31, no. 5, pp. 411–413, 2010.

- [102] T. Grasser, B. Kaczer, W. Göös, H. Reisinger, T. Aichinger, P. Hehenberger, P.-J. Wagner, F. Schanovsky, J. Franco, M. Toledano-Luque, and M. Nelhiebel, "The Paradigm Shift in Understanding the Bias Temperature Instability: From Reaction-Diffusion to Switching Oxide Traps," in *IEEE Transactions on Electron Devices*, vol. 58, pp. 3652–3666, 2011.
- [103] M. F. Bukhori, S. Roy, and A. Asenov, "Simulation of statistical aspects of charge trapping and related degradation in bulk mosfets in the presence of random discrete dopants," *IEEE Transactions on Electron Devices*, vol. 57, no. 4, pp. 795–803, 2010.
- [104] A. Mauri, N. Castellani, C. Compagnoni, A. Ghetti, P. Cappelletti, A. Spinelli, and A. Lacaita, "Impact of atomistic doping and 3D electrostatics on the variability of RTN time constants in flash memories," in *IEEE International Electron Devices Meeting (IEDM)*, 2011.
- [105] T. Grasser, H. Reisinger, P.-J. Wagner, and B. Kaczer, "Time-Dependent Defect Spectroscopy for Characterization of Border Traps in Metal-Oxide-Semiconductor Transistors," *Physical Review B*, vol. 82, no. 24, p. 245318, 2010.
- [106] S. Markov, S. M. Amoroso, L. Gerrer, F. Adamu-Lema, and A. Asenov, "Statistical interactions of multiple oxide traps under bti stress of nanoscale mosfets," *IEEE Electron Device Letters*, vol. 34, no. 5, pp. 686–688, 2013.
- [107] A. Asenov, A. R. Brown, J. H. Davies, and S. Saini, "Hierarchical Approach to 'atomistic' 3-D MOSFET Simulation," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 18, no. 11, pp. 1558–1565, 1999.
- [108] C. Alexander, G. Roy, and A. Asenov, "Random-Dopant-Induced Drain Current Variation in Nano-MOSFETs: A Three-Dimensional Self-Consistent Monte Carlo Simulation Study Using "Ab Initio" Ionized Impurity Scattering," *IEEE Transactions on Electron Devices*, vol. 55, no. 11, pp. 3251–3258, 2008.
- [109] A. Cathignol, B. Cheng, D. Chanemougame, A. Brown, K. Rochereau, G. Ghibaudo, and A. Asenov, "Quantitative Evaluation of Statistical Variability Sources in a 45-nm Technological Node LP N-MOSFET," *IEEE Electron Device Letters*, vol. 29, no. 6, pp. 609–611, 2008.
- [110] C. Millar, D. Reid, G. Roy, S. Roy, and A. Asenov, "Accurate Statistical Description of Random Dopant-Induced Threshold Voltage Variability," *IEEE Electron Device Letters*, vol. 29, no. 8, pp. 946–948, 2008.
- [111] T. Grasser, P.-J. Wagner, P. Hehenberger, W. Göös, and B. Kaczer, "A Rigorous Study of Measurement Techniques for Negative Bias Temperature Instability," *IEEE Transactions on Device and Materials Reliability*, vol. 8, no. 3, pp. 526–535, 2008.
- [112] F. Schanovsky, *Atomistic Modeling in the Context of the Bias Temperature Instability*. PhD thesis, TU Wien, 2013.
- [113] W. Göös, *Hole Trapping and the Negative Bias Temperature Instability*. PhD thesis, TU Wien, 2011.
- [114] D. K. Schroder, "Negative Bias Temperature Instability: What do we understand?," *Micron Electronics Reliability*, vol. 47, no. 6, pp. 841–852, 2007.

- [115] T. Grasser, W. Göss, V. Sverdlov, and B. Kaczer, "The Universality of NBTI Relaxation and its Implications for Modeling and Characterization," in *IEEE International Reliability Physics Symposium Proceedings*, pp. 268–280, 2007.
- [116] H. Reisinger, U. Brunner, W. Heinrigs, W. Gustin, and C. Schlunder, "A Comparison of Fast Methods for Measuring NBTI Degradation," *IEEE Transactions on Device and Materials Reliability*, vol. 7, no. 4, pp. 531–539, 2007.
- [117] H. Reisinger, O. Blank, W. Heinrigs, A. Muhlhoff, W. Gustin, and C. Schlunder, "Analysis of NBTI Degradation- and Recovery-Behavior based on Ultra fast VT-Measurements," in *IEEE International Reliability Physics Symposium Proceedings*, pp. 448–453, 2006.
- [118] T. Grasser, B. Kaczer, W. Göss, H. Reisinger, T. Aichinger, P. Hehenberger, P.-J. Wagner, F. Schanovsky, J. Franco, M. Toledano-Luque, and M. Nelhiebel, "The Paradigm Shift in Understanding the Bias Temperature Instability: From Reaction-Diffusion to Switching Oxide Traps," *IEEE Transactions on Electron Devices*, vol. 58, no. 11, pp. 3652–3666, 2011.
- [119] T. Aichinger, *On the Role of Hydrogen in Silicon Device Degradation and Metalization Processing*. PhD thesis, TU Wien, 2010.
- [120] G. Pobegen, *Degradation of electrical parameters of power semiconductor devices - process influences and modeling*. PhD thesis, TU Wien, 2013.
- [121] T. Aichinger, M. Nelhiebel, and T. Grasser, "On the Temperature Dependence of NBTI Recovery," *Microelectronics Reliability*, vol. 48, pp. 1178–1184, 2008.
- [122] B. Kaczer, T. Grasser, P. Roussel, J. Martin-Martinez, R. O'Connor, B. O'Sullivan, and G. Groeseneken, "Ubiquitous Relaxation in BTI Stressing - New Evaluation and Insights," in *Conference Proceedings of International Reliability Physics Symposium (IRPS)*, pp. 20–27, 2008.
- [123] T. Grasser, P.-J. Wagner, P. Hehenberger, W. Göss, and B. Kaczer, "A Rigorous Study of Measurement Techniques for Negative Bias Temperature Instability," *IEEE Transactions on Device and Materials Reliability*, vol. 8, no. 3, pp. 526–535, 2008.
- [124] A. T. Krishnan, V. Reddy, S. Chakravarthi, J. Rodriguez, S. John, and S. Krishnan, "NBTI Impact on Transistor and Circuit: Models, Mechanisms and Scaling Effects," in *IEEE International Electron Devices Meeting (IEDM)*, pp. 14–5, 2003.
- [125] S. Mahapatra, K. Ahmed, D. Varghese, A. Islam, G. Gupta, L. Madhav, D. Saha, and M. Alam, "On the Physical Mechanism of NBTI in Silicon Oxynitride p-MOSFETs: Can Differences in Insulator Processing Conditions Resolve the Interface Trap Generation versus Hole Trapping Controversy?," in *IEEE International Reliability Physics Symposium (IRPS)*, pp. 1–9, 2007.
- [126] V. Huard, M. Denais, and C. Parthasarathy, "NBTI Degradation: From Physical Mechanisms to Modelling," *Microelectronics Reliability*, vol. 46, no. 1, pp. 1–23, 2006.
- [127] C. Sah, "Effects of Surface Recombination and Channel on p-n Junction and Transistor Characteristics," *IEEE Transactions on Electron Devices*, vol. 9, pp. 94–108, 1962.

- [128] A. Neugroschel, C.-T. Sah, K. Han, M. Carroll, T. Nishida, J. Kavalieros, and Y. Lu, "Direct-Current Measurements of Oxide and Interface Traps on Oxidized Silicon," *IEEE Transactions on Electron Devices*, vol. 42, no. 9, pp. 1657–1662, 1995.
- [129] M. Waltl, P.-J. Wagner, H. Reisinger, K. Rott, and T. Grasser, "Advanced Data Analysis Algorithms for the Time-Dependent Defect Spectroscopy of NBTI," in *IEEE International Integrated Reliability Workshop Final Report*, pp. 74–79, 2012.
- [130] M. Alam, "A Critical Examination of the Mechanics of Dynamic NBTI for PMOSFETs," in *IEEE International Electron Devices Meeting (IEDM)*, pp. 14.4.1–14.4.4, 2003.
- [131] S. Chakravarthi, A. Krishnan, V. Reddy, C. Machala, and S. Krishnan, "A Comprehensive Framework for Predictive Modeling of Negative Bias Temperature Instability," in *IEEE International Reliability Physics Symposium Proceedings (IRPS)*, pp. 273–282, 2004.
- [132] S. Mahapatra, M. Alam, P. B. Kumar, T. Dalei, D. Varghese, and D. Saha, "Negative Bias Temperature Instability in CMOS Devices," *Microelectronic Engineering*, vol. 80, no. 0, pp. 114 – 121, 2005. 14th biennial Conference on Insulating Films on Semiconductors {INFOS2005}.
- [133] T. Grasser, T. Aichinger, G. Pobegen, H. Reisinger, P.-J. Wagner, J. Franco, M. Nelhiebel, and B. Kaczer, "The 'Permanent' Component of NBTI: Composition and Annealing," in *Conference Proceedings of International Reliability Physics Symposium (IRPS)*, 2011.
- [134] T. Grasser, B. Kaczer, W. Göss, T. Aichinger, P. Hehenberger, and M. Nelhiebel, "A Two-Stage Model for Negative Bias Temperature Instability," in *IEEE International Reliability Physics Symposium Proceedings*, pp. 33–44, 2009.
- [135] P. McWhorter and P. Winokur, "Simple Technique for Separating the Effects of Interface Traps and Trapped-Oxide Charge in Metal-Oxide-Semiconductor Transistors," *Applied Physics Letters*, vol. 48, no. 2, pp. 133–135, 1986.
- [136] W. Shockley and W. Read, "Statistics of the Recombinations of Holes and Electrons," *Physical Review*, vol. 87, no. 5, pp. 835–842, 1952.
- [137] V. Huard, M. Denais, and C. Parthasarathy, "NBTI Degradation: From Physical Mechanisms to Modelling," *Microelectronics Reliability*, vol. 46, no. 1, pp. 1 – 23, 2006.
- [138] G. Wentzel, "Eine Verallgemeinerung der Quantenbedingungen für die Zwecke der Wellenmechanik," *Zeitschrift für Physik*, vol. 38, no. 6-7, pp. 518–529, 1926.
- [139] H. Kramers, "Wellenmechanik und Halbzahlige Quantisierung," *Zeitschrift für Physik*, vol. 39, no. 10-11, pp. 828–840, 1926.
- [140] L. Brillouin, "La mécanique ondulatoire de Schrödinger; Une Méthode Générale de résolution successive," *Comptes Rendus (Paris)*, vol. 138, pp. 24–26, 1926.
- [141] R. Pässler, "Description of Nonradiative Multiphonon Transitions in the Static Coupling Scheme," *Czechoslovak Journal of Physics B*, vol. 24, no. 3, pp. 322–339, 1974.
- [142] M. Kirton and M. Uren, "Noise in Solid-State Microstructures: A new Perspective on Individual Defects, Interface States and Low-Frequency (1/f) Noise," *Advances in Physics*, vol. 38, no. 4, pp. 367–468, 1989.

- [143] O. Ibe, *Markov Processes for Stochastic Modeling*. Stochastic modeling, Elsevier Science, 2008.
- [144] K. Huang and A. Rhys, "Theory of Light Absorption and Non-Radiative Transitions in F-Centres," *Proceedings of the Royal Society of London. Series A*, vol. 204, pp. 406-423, 1950.
- [145] M. Lax, "The Franck-Condon Principle and Its Application to Crystals," *Journal of Chemical Physics*, vol. 20, no. 11, pp. 1752-1760, 1952.
- [146] F. Schanovsky, W. Gös, and T. Grasser, "Ab-Initio Calculation of the Vibrational Influence on Hole-Trapping," in *Proceedings of the International Workshop on Computational Electronics (IWCE)*, pp. 163-166, 2010.
- [147] A. D. McNaught and A. Wilkinson, *Compendium of Chemical Terminology, 2nd ed. (the "Gold Book")*. Blackwell Scientific Publications, Oxford, 1997.
- [148] F. Schanovsky, O. Baumgartner, W. Gös, and T. Grasser, "A Detailed Evaluation of Model Defects as Candidates for the Bias Temperature Instability," in *Proceedings of the 18th International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pp. 1-4, 2013.
- [149] M. Bina, T. Aichinger, G. Pobegen, W. Gös, and T. Grasser, "Modeling of DCIV Recombination Currents Using A Multistate Multiphonon Model," in *Final Report of IEEE International Integrated Reliability Workshop (IIRW)*, pp. 27-31, 2011.
- [150] T. Aichinger, P. M. Lenahan, T. Grasser, G. Pobegen, and M. Nelhiebel, "Evidence for Pb Center-Hydrogen Complexes after Subjecting PMOS Devices to NBTI Stress - a Combined DCIV/SDR Study," in *Conference Proceedings of International Reliability Physics Symposium (IRPS 2012)*, 2012.
- [151] T. Grasser, H. Reisinger, P.-J. Wagner, W. Goes, F. Schanovsky, and B. Kaczer, "The Time Dependent Defect Spectroscopy (TDDS) for the Characterization of the Bias Temperature Instability," in *Conference Proceedings of International Reliability Physics Symposium (IRPS)*, pp. 16-25, May 2010.
- [152] P. Hehenberger, W. Gös, O. Baumgartner, J. Franco, B. Kaczer, and T. Grasser, "Quantum-Mechanical Modeling of NBTI in High-k SiGe MOSFETs," in *International Conference on Simulation of Semiconductor Processes and Devices*, pp. 11-14, 2011.
- [153] T. Aichinger, M. Nelhiebel, S. Einspieler, and T. Grasser, "In Situ Poly Heater-A Reliable Tool for Performing Fast and Defined Temperature Switches on Chip," *IEEE Transactions on Device and Materials Reliability*, vol. 10, no. 1, pp. 3-8, 2010.
- [154] T. Grasser, T. Aichinger, H. Reisinger, J. Franco, P.-J. Wagner, M. Nelhiebel, C. Ortolland, and B. Kaczer, "On the 'Permanent' Component of NBTI," in *IEEE International Integrated Reliability Workshop Final Report (IRW)*, 2010.
- [155] J. Cai and C.-T. Sah, "Interfacial Electronic Traps in Surface Controlled Transistors," *IEEE Transactions on Electron Devices*, vol. 47, pp. 576 -583, mar 2000.

- [156] T. Grasser, H. Reisinger, W. Gös, T. Aichinger, P. Hehenberger, P.-J. Wagner, M. Nelhiebel, J. Franco, and B. Kaczer, "Switching Oxide Traps as the Missing Link Between Negative Bias Temperature Instability and Random Telegraph Noise," in *IEEE International Electron Devices Meeting (IEDM)*, 2009.
- [157] D. Lang and C. Henry, "Nonradiative Recombination at Deep Levels in GaAs and GaP by Lattice-Relaxation Multiphonon Emission," *Physical Review Letters*, vol. 35, no. 22, pp. 1525–1528, 1975.
- [158] F. Schanovsky, W. Gös, and T. Grasser, "Multiphonon Hole Trapping from First Principles," *Journal of Vacuum Science & Technology B*, vol. 29, no. 1, pp. 01A2011–01A2015, 2011.
- [159] M. Toledano-Luque, B. Kaczer, E. Simoen, R. Degraeve, J. Franco, P. J. Roussel, T. Grasser, and G. Groeseneken, "Correlation of Single Trapping and Detrapping Effects in Drain and Gate Currents of Nanoscaled nFETs and pFETs," in *Conference Proceedings of International Reliability Physics Symposium (IRPS)*, 2012.
- [160] O. Baumgartner, M. Bina, W. Gös, F. Schanovsky, M. Toledano-Luque, B. Kaczer, H. Kosina, and T. Grasser, "Direct Tunneling and Gate Current Fluctuations," in *Proceedings of the 18th International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pp. 17–20, 2013.
- [161] W. Gös, M. Toledano-Luque, O. Baumgartner, M. Bina, F. Schanovsky, B. Kaczer, and T. Grasser, "Understanding Correlated Drain and Gate Current Fluctuations," in *Proceedings of the 20th International Symposium on the Physical & Failure Analysis of Integrated Circuits*, pp. 51–56, 2013.
- [162] A. Bravaix, C. Guerin, V. Huard, D. Roy, J.-M. Roux, and E. Vincent, "Hot-Carrier Acceleration Factors for Low Power Management in DC-AC Stressed 40nm NMOS Node at High Temperature," in *IEEE International Reliability Physics Symposium*, pp. 531–548, 2009.
- [163] S. Tyaginov, M. Bina, F. Jacopo, D. Osintsev, Y. Wimmer, B. Kaczer, and T. Grasser, "Essential Ingredients for Modeling of Hot-Carrier Degradation in Ultra-Scaled MOSFETs," in *IEEE International Integrated Reliability Workshop Final Report*, 2013.
- [164] C. Hu, S. C. Tam, F.-C. Hsu, P.-K. Ko, T.-Y. Chan, and K. W. Terrill, "Hot-Electron-Induced MOSFET Degradation-Model, Monitor, and Improvement," *IEEE Journal of Solid-State Circuits*, vol. 20, no. 1, pp. 295–305, 1985.
- [165] C. A. Billman, P. M. Lenahan, and W. Weber, "Identification of the Microscopic Structure of New Hot Carrier Damage Centers in Short Channel Mosfets," *MRS Proceedings*, vol. 473, 1 1997.
- [166] J. T. Krick, P. M. Lenahan, and G. J. Dunn, "Direct Observation of Interfacial Point Defects Generated by Channel Hot Hole Injection in nâchannel Metal Oxide Silicon Field Effect Transistors," *Applied Physics Letters*, vol. 59, no. 26, pp. 3437–3439, 1991.
- [167] I. Starkov, H. Enichlmair, S. Tyaginov, and T. Grasser, "Analysis of the Threshold Voltage Turn-Around Effect in High-Voltage n-MOSFETs Due to Hot-Carrier Stress," in *Conference Proceedings of International Reliability Physics Symposium (IRPS)*, 2012.

- [168] P. Heremans, J. Witters, G. Groeseneken, and H. E. Maes, "Analysis of the Charge Pumping Technique and its Application for the Evaluation of MOSFET Degradation," *IEEE Transactions on Electron Devices*, vol. 36, no. 7, pp. 1318–1335, 1989.
- [169] J. F. Chen, S. Chen, K. Wu, and C. M. Liu, "Investigation of Hot-Carrier-Induced Degradation Mechanisms in p-Type High-Voltage Drain Extended Metal–Oxide–Semiconductor Transistors," *Japanese Journal of Applied Physics*, vol. 48, no. 4, p. 04C039, 2009.
- [170] S. Tyaginov and T. Grasser, "Modeling of Hot-Carrier Degradation: Physics and Controversial Issues," in *IEEE International Integrated Reliability Workshop Final Report*, pp. 206–215, 2012.
- [171] G. Pobegen, S. Tyaginov, M. Nelhiebel, and T. Grasser, "Observation of Normally Distributed Energies for Interface Trap Recovery After Hot-Carrier Degradation," *IEEE Electron Device Letters*, vol. 34, no. 8, pp. 939–941, 2013.
- [172] S. Rauch, F. Guarin, and G. La Rosa, "Impact of E-E scattering to the hot carrier degradation of deep submicron NMOSFETs," *IEEE Electron Device Letters*, vol. 19, no. 12, pp. 463–465, 1998.
- [173] "The International Technology Roadmap for Semiconductors (ITRS)."
- [174] S. Tyaginov, I. Starkov, O. Triebel, H. Enichlmair, C. Jungemann, J. Park, H. Ceric, and T. Grasser, "Secondary Generated Holes as a Crucial Component for Modeling of HC Degradation in High-voltage n-MOSFET," in *International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pp. 123–126, 2011.
- [175] R. H. Fowler and L. Nordheim, "Electron Emission in Intense Electric Fields," *Proceedings of the Royal Society of London. Series A*, vol. 119, no. 781, pp. 173–181, 1928.
- [176] D. J. DiMaria and J. W. Stasiak, "Trap Creation in Silicon Dioxide Produced by Hot Electrons," *Journal of Applied Physics*, vol. 65, no. 6, pp. 2342–2356, 1989.
- [177] D. J. DiMaria and J. H. Stathis, "Anode Hole Injection, Defect Generation, and Breakdown in Ultrathin Silicon Dioxide Films," *Journal of Applied Physics*, vol. 89, no. 9, pp. 5015–5024, 2001.
- [178] I. Starkov and H. Enichlmair, "Local oxide capacitance as a crucial parameter for characterization of hot-carrier degradation in long-channel n-MOSFETs," *Journal of Vacuum Science & Technology B*, vol. 31, no. 1, pp. 01A118–1–01A118–7, 2013.
- [179] F.-C. Hsu and S. Tam, "Relationship Between MOSFET Degradation and Hot-Electron-Induced Interface-State Generation," *IEEE Electron Device Letters*, vol. 5, no. 2, pp. 50–52, 1984.
- [180] T. Mizuno, A. Toriumi, M. Iwase, M. Takahashi, H. Niiyama, M. Fukmoto, and M. Yoshimi, "Hot-Carrier Effects in 0.1 μm Gate Length CMOS Devices," in *IEEE International Electron Devices Meeting (IEDM)*, pp. 695–698, 1992.
- [181] A. Bravaix and V. Huard, "Hot-Carrier Degradation Issues in Advanced CMOS Nodes," in *European Symposium on the Reliability of Electron Devices*, 2010.

- [182] I. Starkov, S. Tyaginov, H. Enichlmair, J. Cervenka, C. Jungemann, S. Carniello, J. M. Park, H. Ceric, and T. Grasser, "Hot-Carrier Degradation Caused Interface State Profile-Simulation versus Experiment," *Journal of Vacuum Science & Technology B*, vol. 29, pp. 01AB09-1-01AB09-8, 2011.
- [183] W. McMahon and K. Hess, "A Multi-Carrier Model for Interface Trap Generation," *Journal of Computational Electronics*, vol. 1, no. 3, pp. 395-398, 2002.
- [184] W. McMahon, K. Matsuda, J. Lee, K. Hess, and J. Lyding, "The Effect of a Multiple Carrier Model of Interface Trap Generation on Lifetime Extraction for MOSFETs," *Proceedings of the International Conference on Modeling and Simulation of Microsystems*, vol. 1, p. 576, 2002.
- [185] C. Guerin, V. Huard, and A. Bravaix, "General Framework about Defect Creation at the Si SiO₂ Interface," *Journal of Applied Physics*, vol. 105, no. 11, pp. 114513-114513, 2009.
- [186] S. Tyaginov, I. Starkov, H. Enichlmair, J. Park, C. Jungemann, and T. Grasser, "Physics-Based Hot-Carrier Degradation Models," *ECS Transactions*, 2011.
- [187] W. McMahon, A. Haggaag, and K. Hess, "Reliability Scaling Issues for Nanoscale Devices," *IEEE Transactions Nanotechnology*, vol. 2, no. 1, pp. 33-38, 2003.
- [188] A. Stesmans and V. Afanasev, "Electrical Activity of Interfacial Paramagnetic Defects in Thermal (100) Si/SiO₂," *Physical Review B*, vol. 57, no. 16, p. 10030, 1998.
- [189] A. Zaka, P. Palestri, Q. Rafhay, R. Clerc, M. Iellina, D. Rideau, C. Tavernier, G. Pananakakis, H. Jaouen, and L. Selmi, "An Efficient Nonlocal Hot Electron Model Accounting for Electron-Electron Scattering," *IEEE Transactions on Electron Devices*, vol. 59, no. 4, pp. 983-993, 2012.
- [190] S. Tyaginov, I. Starkov, C. Jungemann, H. Enichlmair, J. Park, and T. Grasser, "Impact of the Carrier Distribution Function on Hot-Carrier Degradation Modeling," in *Proceedings of the European Solid-State Device Research Conference*, pp. 151-154, 2011.
- [191] D. Caughey and R. Thomas, "Carrier Mobilities in Silicon Empirically Related to Doping and Field," in *Proceedings of the IEEE*, vol. 52, pp. 2192-2193, 1967.
- [192] M. Snir, S. W. Otto, D. W. Walker, J. Dongarra, and S. Huss-Lederman, *MPI: The Complete Reference*. Cambridge, MA, USA: MIT Press, 1995.
- [193] U. Trottenberg, C. Oosterlee, and A. Schuller, *Multigrid*. Elsevier Science, 2000.
- [194] M. F. Adams, *Multigrid Equation Solvers for Large Scale Nonlinear Finite Element Simulations*. PhD thesis, University of California, Berkeley, 1998.

Own Publications

- [1] O. Baumgartner, M. Bina, W. Gös, F. Schanovsky, M. Toledano-Luque, B. Kaczer, H. Kosina, and T. Grasser, “Direct Tunneling and Gate Current Fluctuations,” in *Proceedings of the 18th International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, pp. 17–20, 2013.
- [2] M. Bina, T. Aichinger, G. Pobegen, W. Gös, and T. Grasser, “Modeling of DCIV Recombination Currents Using A Multistate Multiphonon Model,” in *Final Report of IEEE International Integrated Reliability Workshop (IIRW 2011)*, pp. 27–31, 2011.
- [3] M. Bina, K. Rupp, S. Tyaginov, O. Triebel, and T. Grasser, “Modeling of Hot Carrier Degradation Using a Spherical Harmonics Expansion of the Bipolar Boltzmann Transport Equation,” in *Proceedings of the 2012 IEEE International Electron Devices Meeting (IEDM)*, pp. 713–716, 2012.
- [4] M. Bina, O. Triebel, B. Schwarz, M. Karner, B. Kaczer, and T. Grasser, “Simulation of Reliability on Nanoscale Devices,” in *Proceedings of the 17th International Conference on Simulation of Semiconductor Processes and Devices*, pp. 109–112, 2012.
- [5] J. Franco, B. Kaczer, M. Toledano-Luque, P. J. Roussel, G. Groeseneken, B. Schwarz, M. Bina, M. Walzl, P.-J. Wagner, and T. Grasser, “Reduction of the BTI Time-Dependent Variability in Nanoscaled MOSFETs by Body Bias,” in *Conference Proceedings of International Reliability Physics Symposium (IRPS 2013)*, pp. 1–6, 2013.
- [6] W. Gös, M. Toledano-Luque, O. Baumgartner, M. Bina, F. Schanovsky, B. Kaczer, and T. Grasser, “Understanding Correlated Drain and Gate Current Fluctuations,” in *Proceedings of the 20th International Symposium on the Physical & Failure Analysis of Integrated Circuits*, pp. 51–56, 2013.
- [7] Y. Illarionov, S. Tyaginov, M. Bina, and T. Grasser, “A method to determine the lateral trap position in ultra-scaled MOSFETs,” in *Extended Abstracts of the 2013 International Conference on Solid State Devices and Materials(SSDM)*, pp. 728–729, 2013.
- [8] B. Kaczer, J. Franco, M. Toledano-Luque, P. J. Roussel, M. F. Bukhori, A. Asenov, B. Schwarz, M. Bina, T. Grasser, and G. Groeseneken, “The Relevance of Deeply-Scaled FET Threshold Voltage Shifts for Operation Lifetimes,” in *Conference Proceedings of International Reliability Physics Symposium (IRPS 2012)*, 2012.
- [9] K. Rupp, C. Jungemann, M. Bina, A. Jüngel, and T. Grasser, “Bipolar Spherical Harmonics Expansions of the Boltzmann Transport Equation,” in *Proceedings of the 17th International Conference on Simulation of Semiconductor Processes and Devices*, pp. 19–22, 2012.

- [10] M. Bina, "Extension of a Parallel Nonlinear Finite Element Simulation Software for Human Bone," Master's thesis, Institut für Leichtbau und Struktur-Biomechanik, 2013.
- [11] M. Bina, "Simulation of Interface States Generated During Stress in MOSFETs," Master's thesis, Institut für Mikroelektronik, 2010.

Curriculum Vitae

Personal Information

25.09.1985 Born in St. Pölten, Austria



Primary Education

- 1992-1996 Primary School, Volksschule der Erzdiözese St. Pölten Campus Krems-Mitterau
1996-2000 Secondary School, Bundesrealgymnasium Krems Ringstrasse

Further Education

- 2000-2005 High School, HTBLuVA St. Pölten - Department for Informatics and Organisation (EDVO)
Dez 2003 Medical examination for military service (not required to enroll)
May 2005 Passed the matura exams with honors
Jul 2005 Enrolled in electrical engineering at the TU Wien
Jul 2008 Received degree *Bakkalaureus der Technik* (BSc.)
Jun 2008 Enrolled in microelectronics (master program) at TU Wien
Jun 2008 Enrolled in biomedical engineering (master program) at TU Wien
Mar 2010 Received degree *Diplomingenieur* (MSc. equivalent)
Diploma thesis:
Simulation of Interface States Generated During Stress in MOSFETs
Entered PhD program at the Institute for Microelectronics, TU Wien
Jun 2013 Received degree *Diplomingenieur* (MSc. equivalent)
Diploma thesis:
Extension of a Parallel Nonlinear Finite Element Simulation Software for Human Bone

Work Experience

- 2000 Summer internship as account assistant
Ing. Fladenhofer Schweißerei und Spezialreperaturen
2003 Summer internship as programmer for web applications
at Strasser Informationstechnik GmbH
2004 Summer internship as programmer and system administrator
at the Niederösterreichische Versicherung AG
2007 Internship as programmer for test engineering
at Infineon Technologies AG, Villach
2005 - 2010 Freelancing programmer, mostly for Prof. Tibor Grasser
at the Institute for Microelectronics, TU Wien
2013 - 2014 Cooperation with Prof. Dieter Pahr as developer for parallel FE simulation software
at the Institute of Lightweight Design and Structural Biomechanics, TU Wien