



TECHNISCHE
UNIVERSITÄT
WIEN
Vienna University of Technology

DIPL O M A R B E I T

Simulation of Random Dopant Fluctuations with a Quantum-Corrected Drift-Diffusion Model

Ausgeführt am

Institut für Mikroelektronik
der Technischen Universität Wien

unter der Anleitung von
Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Klaus-Tibor Grasser
und
Dipl.-Ing. Markus Bina

durch

Benedikt Schwarz

Storkgasse 12/19
1050 Wien

Matrikelnummer 0725036
Studienkennzahl 066 439

Datum

Unterschrift

Abstract

The general purpose semiconductor device simulator MINIMOS-NT is extended to a “atomistic” quantum-corrected drift-diffusion simulator to study parameter fluctuations due to random discrete dopants. It has been confirmed that discrete dopants cannot straightforwardly be included in classical drift-diffusion simulators, because of unphysically large and grid dependent charge localization. This unphysical behavior can be eliminated by splitting the Coulomb potential into a long-range and a short-range part, explicitly including the long-range part only. The issue can also be solved by first-order quantum-correction to the classical drift-diffusion model via the density gradient model. Unfortunately, the density gradient model leads to worsened numerical robustness especially when discrete dopants are included. Thus several advanced discretization schemes of the quantum-correction equation are implemented, but their numerical benefit could not be confirmed. With the focus on sub-nanometer MOS devices the density gradient model has the advantage of additionally including basic quantum mechanical effects such as confinement due to energy quantization. With the density gradient model we are able to fit a CV-curve to the solution of the Schrödinger Poisson solver VSP2 using Cauchy boundary conditions for the quantum-correction potential at the oxide interface. In a simulation study focused on a 22 nm NMOS transistor 100 macroscopically identical samples are simulated showing random discrete dopant induced threshold voltage lowering and fluctuation.

Contents

1	Introduction	7
I	Quantum Corrected Drift-Diffusion Model	9
2	Modelling of Charge Transport in Sub-Nanometer Devices	10
2.1	Quasi-Electrostatic	11
2.2	Drift-Diffusion Model	12
2.3	Quantum Correction	14
2.4	The Complete Quantum Corrected Drift-Diffusion Equation System	15
2.5	Quasi-Fermi Level Formalism	16
3	Numerical Implementation	18
3.1	Finite Volume Discretization	19
3.2	Control Function and Numerical Flux	20
3.3	Assembly Module	21
3.4	Interface Conditions	21
3.5	Boundary Conditions	22
3.6	Discretization of Poisson's Equation	23
3.7	Discretization of the Continuity Equation	24
3.8	Discretization of the Density Gradient Equation	25
4	Results	29
4.1	Carrier Confinement and CV-Curve Matching	29
4.2	Dependence on Grid Spacing	35
II	Random Dopant Fluctuations	36
5	Splitting into Generation and Simulation Model	37
5.1	<i>kd</i> -tree	38
6	Generation of Random Discrete Dopants	40
6.1	Voronoi-Region Model	40
6.2	Segment Model	42
6.3	Read-from-File Model	43

6.4	Point-in-Segment Check	43
7	Modeling Discrete Dopants	45
7.1	Nearest-Grid-Point Model	46
7.2	Long-Range Model	48
7.3	The Density Gradient Model	52
8	Results	54
8.1	Silicon Resistor with Constant Doping	54
8.2	Ionized Impurity Scattering	57
8.3	NMOS Transistor with Random Dopant Fluctuations	58
9	Summary	64
A	Minimos-NT Model Defaults	66
A.1	Density Gradient Model	66
A.2	Random Dopant Generation Model	67
A.3	Random Dopant Simulation Models	68
B	Random Discrete Dopant Input File	69

List of Figures

3.1	Voronoi region of an inner and an interface grid point	19
4.1	CV-curve obtained by QDD with Dirichlet boundary conditions on the Si/SiO ₂ interface.	30
4.2	CV-curve obtained by QDD with zero-outflux Neumann boundary conditions on the Si/SiO ₂ interface.	31
4.3	CV-curve obtained by QDD with non-zero-outflux Neumann boundary conditions on the Si/SiO ₂ interface.	32
4.4	CV-curve obtained by QDD with Cauchy boundary conditions on the Si/SiO ₂ interface.	33
4.5	Dependency on the Oxide thickness of the CV-curve obtained by QDD with Cauchy boundary conditions on the Si/SiO ₂ interface.	34
4.6	Grid dependency of the CV-curve obtained by QDD with Cauchy boundary conditions on the Si/SiO ₂ interface.	35
5.1	Example of a 2D <i>kd</i> -tree out of 8 points.	38
6.1	Continuous doping of a 22 nm NMOS.	41
6.2	Random discrete doping of a 22 nm NMOS.	41
6.3	A 16 × 8 × 8 nm silicon resistor with random discrete dopants.	42
7.1	The electrostatic potential of a discrete donor assigned to a single grid point using the classical drift-diffusion model.	46
7.2	The electron concentration around a discrete donor assigned to a single grid point using the classical drift-diffusion model.	47
7.3	Long-range and short-range of a dopant located at the origin.	48
7.4	The electrostatic potential of a discrete donor using the long-rang model and the classical drift-diffusion model.	50
7.5	The electron concentration of a discrete donor and two discrete donors using the long-rang model in a drift diffusion simulation.	51
7.6	The electrostatic potential of a discrete donor assigned to a single grid point using the density gradient model.	52
7.7	The electrostatic potential, electron quantum-correction potential and electron concentration of a discrete donor assigned to a single grid point using the density gradient model.	53
8.1	Increase of resistance due to charge trapping.	55

8.2	Effect of the quantum-correction potential contact boundary conditions . . .	56
8.3	Adjusting Ionized impurity scattering parameters to fit the DG random dopant simulation to the continuous DG simulation.	57
8.4	Electron current density in the channel of a 22 nm NMOS with random discrete dopants.	58
8.5	Electrostatic potential in a 22 nm NMOS with random discrete dopants. . .	59
8.6	Electron concentration in a 22 nm NMOS with random discrete dopants. . .	60
8.7	I_D - V_G characteristics of a 22 nm NMOS with random discrete dopants assigned to the nearest grid point using the density gradient model.	61
8.8	I_D - V_G characteristics of a 22 nm NMOS with random discrete dopants using the long-range model and the drift-diffusion model.	62
8.9	I_D - V_G characteristics of a 22 nm NMOS with random discrete dopants using the long-range model and the density gradient model.	63

Acknowledgement

First of all, I would like to thank Prof. Tibor Grasser for his technical advice and support during the work on this thesis guiding me past all difficulties. Without his motivation perhaps I would have been yielded to despair.

Many thanks also go to Markus Bina, Franz Schanovsky for correction reading. I would also like to thank Markus Karner and Christian Kernstock for the fast implementation of desired features in the GTS-Framework.

Last but not least I thank my girlfriend Simone Schuler for her support especially during the last weeks. I am really sorry about my delayed reaction while the matter has preoccupied me completely.

Notation

Symbol	Meaning
∂_t	time derivative
$\vec{\nabla}$	Nabla operator
$\vec{\nabla} \cdot$	divergence
$\vec{\nabla} \times$	rotation
$\exp(x), e^x$	exponential function
$\mathcal{B}(x) = x/(e^x - 1)$	Bernoulli function
\vec{E}	electric field intensity
\vec{D}	electric flux density
ρ	electric charge density
\vec{H}	magnetic field intensity
\vec{B}	magnetic flux density
ψ	electrostatic potential
$\hat{\epsilon}$	dielectric permittivity (anisotropic)
ϵ	dielectric permittivity (isotropic)
$\hat{\mu}$	magnetic permeability (anisotropic)
μ	magnetic permeability (isotropic)
q	elementary charge ($1.602176487(40) \cdot 10^{-19}\text{C}$)
n, p	electron/hole concentration
C	net impurity concentration
c_0	the speed of light in vacuum ($2.99792458 \cdot 10^8\text{m/s}$)
\vec{J}	current density
\vec{J}_n, \vec{J}_p	electron/hole current density
R	net recombination rate
\vec{v}_n, \vec{v}_p	average electron/hole velocity
μ_n, μ_p	electron/hole mobility
D_n, D_p	electron/hole diffusion coefficient
k_B	Boltzmann constant ($1.3806503 \cdot 10^{-23}\text{m}^2\text{kg/s}^2\text{K}$)
T_L	lattice temperature
$\mathcal{E}_C, \mathcal{E}_V$	conductance/valence band-edge energy
N_C, N_V	conductance/valence effective density-of-states
γ_n, γ_p	electron/hole quantum potential
m_n^*, m_p^*	electron/hole effective mass
λ_n, λ_p	electron/hole quantum-correction fitting factor

Chapter 1

Introduction

With the aggressive scaling of MOS devices to the sub-nanometer regime, statistical variations have become one of the main challenges in microelectronic industry and cannot be neglected any more [1, 2]. Currently the start of 22 nm tri-gate technology in mass production in the second half of 2011 has been reported by Intel [3]. The random position and number of dopants in the channel is reported to be the main source of variability [4], inducing threshold voltage lowering and fluctuation [5]. Thus it is no longer sufficient to describe doping with a concentration in sub-nanometer devices.

Treating dopants as random and discrete often referred to as “atomistic” requires a 3D simulation with fine grain discretization, simulating a statistical representative number of macroscopically identical samples [5]. Thus a highly efficient simulation framework is needed. The drift-diffusion model [6] meets these requirement and has been the most widely used model during the past 40 years. However, with the high-grade device scaling the limits of the drift-diffusion model have been exceeded and more advanced models have to be used for appropriate device description. To a certain extent quantum mechanical effects, such as tunneling and confinement due to energy level quantization can be modeled by phenomenological extensions to the classical drift-diffusion model. Popular examples are the Tsu-Esaki tunneling model [7, 8] and the VanDort confinement model [9]. The inclusion of discrete dopants into the classical drift-diffusion model leads to unphysically large charge trapping by discrete dopants. In [10] the role of long-range and short-range of the Coulomb potential has been discussed and it has been pointed out that only the long-range part has to be included in a drift-diffusion simulation for an appropriate description of screening. All these models have limitations and thus a quantum mechanical treatment has to be included in the simulation framework to further improve simulation accuracy.

Apart from physical accuracy, numerical stability and computational efficiency are of central importance. There are several suitable methods for quantum mechanical treatment, but most do not meet the requirements of an efficient general purpose device simulation [11]. Unfortunately quantum mechanical non-locality results in with a high computational effort. Computational efficiency becomes increasingly important when treating three dimensional problems and reaches its ultimate extent when statistical fluctuations

are included. An alternative approach is to include a quantum-correction into the classical drift-diffusion model. One of this correction models is the density gradient model. It has the ability to appropriately describe screening of discrete dopants [1], carrier confinement and tunneling [12]. In [2] the long-range model and the density gradient model have been used together, leading to improved numerical stability of the density gradient model when discrete dopants are included. In the course of this thesis both the long-range model and the density gradient model have been implemented in the general purpose device simulator MINIMOS-NT [13] to provide a flexible simulation framework for statistical parameter fluctuations and advanced reliability issues.

In the first part of this thesis the quantum-corrected drift-diffusion model is described, starting with the basic semiconductor equations. After introducing quantum mechanical corrections based on the density gradient method, we turn towards numerical discretization of the quantum-corrected drift-diffusion model using the finite volume method. In Chapter 4 the simulation results are compared to the solution obtained by a Schrödinger-Poisson solver, to study the accuracy of the newly implemented quantum-correction term.

In the second part we turn towards modeling of random discrete dopants, starting with the splitting into generation and simulation models to provide high flexibility. After treating the generation of random discrete dopants, the appropriate inclusion of discrete dopants into the classical and quantum-corrected drift-diffusion simulation is described. The influence of random discrete dopants in sub-nanometer semiconductor devices is investigated in Chapter 8.

Part I

**Quantum Corrected Drift-Diffusion
Model**

Chapter 2

Modelling of Charge Transport in Sub-Nanometer Devices

In current and future microelectronic devices quantum effects cannot be neglected and thus must be treated in semiconductor device simulation. With the focus on metal oxide semiconductor field effect transistors (MOS-FET), especially confinement and tunneling are relevant effects. When considering random discrete dopants, also appropriate screening has to be included. The most physically accurate way is to model all these effects using a full quantum mechanical treatment [14]. One of these full quantum mechanical methods is the self-consistent solution of Schrödinger's and Poisson's equation. Unfortunately, solving Schrödinger's equation results in a direct coupling of all grid points as a consequence the Jacobian loses its sparse structure and the computational costs rise dramatically. Also the nearest neighbor couplings are lost, which is a requirement for the use of unstructured grids to keep the number of grid points in a realistic range. Although the non-equilibrium Green's function (NEGF) formalism [15] is often used for simulation of nanowire transistors because it accurately describes quantum mechanical transport in sub-nanometer transistors, its high computational expense is considerable. Because of the high computational cost, full quantum mechanical treatment is not suitable for a general purpose device simulation in three dimensions so far. Including statistical variations, the computational effort rises even more, resulting in the need of a highly efficient quantum mechanical treatment of sub-nanometer devices.

Several quantum-correction models have been discussed in [11], including the quantum-corrected drift-diffusion (QDD) model [12, 14]. The QDD model, also known as density gradient (DG) model, includes confinement, tunneling [12], screening of discrete dopants [1] and naturally fits into drift-diffusion device simulators. In the target simulator MINIMOS-NT [13], tunneling is covered by semi-classical models including trap-assisted tunneling. Confinement and appropriate screening of discrete dopants are treated with the newly implemented density gradient model. QDD introduces a first-order quantum-correction term to the classical drift-diffusion model and thus can be understood as a generalization of drift-diffusion to the level of quantum mechanics. With its simplicity, QDD is appropriate for efficient quantum-corrected semiconductor device simulation in three dimensions including statistic fluctuations.

In this Chapter the quantum-corrected drift-diffusion model is derived, starting at the classical point of view.

2.1 Quasi-Electrostatic

In semiconductor device simulation, field problems are assumed to be quasi-electrostatic. The decoupling of the electric and magnetic field simplifies Maxwell's equations significantly. In general, a field problem is called quasi-electrostatic if the electric field is dominant to the magnetic field. A detailed description is given in this section, starting with Maxwell's equations [16]

$$\begin{aligned}\vec{\nabla} \times \vec{E} &= -\partial_t \vec{B}, \\ \vec{\nabla} \cdot \vec{B} &= 0, \\ \vec{\nabla} \times \vec{H} &= \vec{J} + \partial_t \vec{D}, \\ \vec{\nabla} \cdot \vec{D} &= \rho,\end{aligned}\tag{2.1}$$

where \vec{E} and \vec{H} are the electric and magnetic field intensities, \vec{D} and \vec{B} are the electric and magnetic flux densities, \vec{J} is the current density and ρ is the space charge density. To get a mathematical condition for the quasi-electrostatic regime, (2.1)₁ is rewritten in scaled form

$$\tilde{\nabla} \times \tilde{E} = - \left[\frac{LB_0}{TE_0} \right] \tilde{\partial}_t \tilde{B},\tag{2.2}$$

using

$$\tilde{E} = \vec{E}/E_0, \quad \tilde{B} = \vec{B}/B_0, \quad \tilde{\nabla} = L\vec{\nabla}, \quad \tilde{\partial}_t = T\partial_t,\tag{2.3}$$

where E_0 , B_0 , L and T are the characteristic magnitudes of the electric field intensity, the magnetic flux density and the characteristic length and time scale. The condition for the quasi-electrostatic regime is now

$$\frac{LB_0}{TE_0} \ll 1,\tag{2.4}$$

in which case (2.2) reduces to

$$\vec{\nabla} \times \vec{E} = 0.\tag{2.5}$$

Because every curl-free vector-field can be described by a gradient-field, in respect to (2.5), a electrostatic scalar potential ψ is introduced by

$$\vec{E} = -\vec{\nabla}\psi.\tag{2.6}$$

With (2.5), (2.6) and the material property $\vec{D} = \hat{\epsilon}\vec{E}$, Poisson's equation can be formed

$$\vec{\nabla} \cdot (\hat{\epsilon}\vec{\nabla}\psi) = -\rho,\tag{2.7}$$

where $\hat{\epsilon}$ is the dielectric permittivity tensor. In semiconductor devices the space charge density ρ consists of electron and hole densities n and p and the concentration of ionized

impurities and dopants combined in C . Poisson's equation for semiconductor analysis can be written as

$$\vec{\nabla} \cdot (\hat{\epsilon} \vec{\nabla} \psi) = q(n - p - C), \quad (2.8)$$

with the elementary charge q .

A more suitable quasi-electrostatic state definition for device simulation can be found by assuming rapidly changing fields, so that $E_0 \approx cB_0$. Thus equation (2.4) can be rewritten as

$$\frac{L}{cT} = \frac{L}{\lambda} \ll 1. \quad (2.9)$$

A quasi-electrostatic situation can be assumed if the characteristic length scale is much smaller than the characteristic wavelength. To give an example, silicon has a relative dielectric permittivity $\epsilon_r = 11.68$ and a relative magnetic permeability $\mu_r = 1$, thus $c_{Si} = 1/\sqrt{\epsilon\mu} = c_0/\sqrt{\epsilon_r\mu_r} = c_0/\sqrt{11.68}$. The operating frequency $f = 100\text{GHz}$ corresponds to a wavelength of $\lambda = 877\mu\text{m}$, which is quite larger than the typical dimensions of semiconductor devices.

Applying the divergence on equation (2.1)₃ and inserting into (2.1)₄ results in the continuity equation

$$\vec{\nabla} \cdot \vec{J} = -\partial_t \varrho, \quad (2.10)$$

which is generally valid.

Because both electrons and holes contribute to an electric current in semiconductor device simulation, the continuity equation is split into one continuity equation for holes and one for electrons, introducing R as the net recombination rate

$$\begin{aligned} \vec{\nabla} \cdot \vec{J}_n - q\partial_t n &= qR, \\ \vec{\nabla} \cdot \vec{J}_p + q\partial_t p &= -qR. \end{aligned} \quad (2.11)$$

The net recombination rate describes the amount of electron-hole-pairs recombined minus those generated per time interval. In equilibrium the net recombination rate is always zero.

2.2 Drift-Diffusion Model

A simple and widely used model describing charge transport in semiconductors is the drift-diffusion model [6]. As its name indicates, it consists of two parts, the drift term and the diffusion term. The drift term is related to charge transport due to an electric field. A similar situation occurs in conductors, where the current density is directly proportional to the electric field intensity over the conductivity tensor $\hat{\gamma}$ with $\vec{J} = \hat{\gamma} \vec{E}$. The diffusion term describes the thermal motion of particles to reduce a non-equilibrium particle distribution. In conductors imbalances are immediately compensated by surrounding electrons, because both the number of free electrons and the mobility is high. Thus diffusion is neglected in classical electrodynamics.

To relate the average carrier velocities \vec{v}_n and \vec{v}_p to the electric field intensity, the carrier mobilities μ_n and μ_p are introduced by

$$\vec{v}_n = -\mu_n \vec{E} \quad \text{and} \quad \vec{v}_p = \mu_p \vec{E}. \quad (2.12)$$

In the absence of magnetic fields, the drift terms are

$$\vec{J}_n^{\text{Drift}} = -qn\vec{v}_n = qn\mu_n\vec{E} \quad \text{and} \quad \vec{J}_p^{\text{Drift}} = qp\vec{v}_p = qp\mu_p\vec{E}. \quad (2.13)$$

Due to the thermal particle motion, imbalances yield a flow of these particles to reduce any non-equilibrium distribution. The flow of charged particles results in a current flow and is modeled by the diffusion terms

$$\vec{J}_n^{\text{Diffusion}} = -q\vec{F}_n = qD_n\vec{\nabla}n \quad \text{and} \quad \vec{J}_p^{\text{Diffusion}} = q\vec{F}_p = -qD_p\vec{\nabla}p, \quad (2.14)$$

where F_n and F_p describe the particle flux due to a gradient of the particle concentration and D_n and D_p are the diffusion coefficients given by the Einstein relation

$$D_{n,p} = \frac{k_B T}{q} \mu_{n,p} = V_{\text{th}} \mu_{n,p}, \quad (2.15)$$

where k_B is the Boltzmann constant and V_{th} is the thermal voltage. At room temperature the thermal voltage V_{th} is approximately 26mV. The Einstein relation is only valid if the carrier temperature matches the lattice temperature, that is in equilibrium. Using the electrostatic potential defined by (2.6) the complete drift-diffusion current equations read

$$\begin{aligned} \vec{J}_n &= \vec{J}_n^{\text{Drift}} + \vec{J}_n^{\text{Diffusion}} = qD_n\vec{\nabla}n - qn\mu_n\vec{\nabla}\psi \\ \vec{J}_p &= \vec{J}_p^{\text{Drift}} + \vec{J}_p^{\text{Diffusion}} = -qD_p\vec{\nabla}p - qp\mu_p\vec{\nabla}\psi. \end{aligned} \quad (2.16)$$

The slightly different version implemented in the target simulator MINIMOS-NT [13] which also accounts for inhomogeneous materials reads

$$\begin{aligned} \vec{J}_n^{\text{DD}} &= q\mu_n n \left[\vec{\nabla} \left(\frac{\mathcal{E}_C}{q} - \psi \right) + \frac{k_B T_L}{q} \frac{N_{C,0}}{n} \vec{\nabla} \left(\frac{n}{N_{C,0}} \right) \right], \\ \vec{J}_p^{\text{DD}} &= q\mu_p p \left[\vec{\nabla} \left(\frac{\mathcal{E}_V}{q} - \psi \right) - \frac{k_B T_L}{q} \frac{N_{V,0}}{p} \vec{\nabla} \left(\frac{p}{N_{V,0}} \right) \right]. \end{aligned} \quad (2.17)$$

The implemented version considers position dependent band-edge energies \mathcal{E}_C and \mathcal{E}_V , and effective masses. The effective masses are taken into account by the position dependent effective density-of-states $N_{C,0}$ and $N_{V,0}$ evaluated at a fixed temperature T_0 .

2.3 Quantum Correction

To provide an approximate treatment to confinement and screening of discrete dopants an additional density gradient term is inserted in the classical drift-diffusion model. In particular, the first-order quantum-correction introduces an additional dependency of the equation of state on the density gradient and therefore a quantum mechanically characteristic non-locality. The density gradient model was derived from Wigner's equation by M. G. Ancona *et al.* [12] and A. Wettstein *et al.* [11] applying several assumptions. These assumptions are that many-body correlation effects and quantum effects associated with Fermi-Dirac statistic are negligible, the effective-mass theory applies and the electron gas has infinite extent. Under typical conditions all but the last are reasonable. The QDD model can approximately describe confinement, tunneling [12] and screening effects [1] in sub-nanometer devices. As we are interested in confinement and screening only, the quantum-corrected drift-diffusion model is only applied in semiconductor segments. Tunneling is still included in the target simulator via conventional models.

The quantum drift-diffusion current equation for electrons derived in Ref. [11] using slightly different variables reads

$$\vec{J}_n^{\text{QDD}} = qD_n \vec{\nabla} n + qn\mu_n \vec{\nabla} \left(\frac{\mathcal{E}_C}{q} - \psi - \gamma_n \right), \quad (2.18)$$

where the quantum-correction potential γ_n is introduced. The quantum-correction term includes a density depended part in the current relation to model first-order quantum effects. The quantum-correction potential γ_n is given by

$$\gamma_n = \frac{\hbar^2 \beta}{12m_n^*} \left[\nabla^2 \psi - \frac{\beta}{2} (\vec{\nabla} \psi)^2 \right], \quad (2.19)$$

where $\beta = 1/k_B T_L$ and m_n^* is the effective electron mass.

The inclusion of the quantum-correction term leads to a fourth-order differential equation, which can be split into two second-order differential equations to meet the requirements of finite volume discretization on unstructured grids [11].

Because $\vec{\nabla} \psi$ can be undefined at abrupt steps of ψ , ψ is replaced by $(\psi + \gamma_n)$. In the first-order approximation this is valid because γ_n is of $\mathcal{O}(\hbar^2)$.

The quantum potential equation then reads

$$\gamma_n = \frac{\hbar^2 \beta}{12\lambda_n m_0} \left[\nabla^2 (\psi + \gamma_n) - \frac{\beta}{2} (\vec{\nabla} \psi + \vec{\nabla} \gamma_n)^2 \right], \quad (2.20)$$

introducing the fitting parameter λ_n .

Due to the isotropic mass approximation it is unclear which mass has to be used in (2.20). This problem is treated with the introduction of the fitting parameter λ_n and using the free electron mass m_0 . In [11] the density of states mass is used instead of the effective mass as in [12, 17] and [18]. This only changes the value of the fitting factor. In later

implementations of A. Asenov *et al.* [1] an anisotropic effective mass is used.

A. Wettstein [11] has also introduced a simplified model neglecting the square term to obtain a more stable version of the quantum-correction. The simplified model reads

$$\gamma_n^{simplified} = \frac{\hbar^2 \beta}{12\lambda_n m_0} \nabla^2 (\psi + \gamma_n). \quad (2.21)$$

In most literature the quantum-corrected drift-diffusion model is denoted as density gradient model. This becomes clear when expressing the quantum-correction potential in terms of the carrier density n . The microscopic derivation of the density gradient model is only valid near thermal equilibrium, thus the quasi-Fermi potential is neglected in the course of this thesis. Assuming thermal equilibrium the carrier density reads

$$n = N_C \exp\left(\frac{\mathcal{E}_C - q(\psi + \gamma_n)}{k_B T_L}\right). \quad (2.22)$$

Replacing $\mathcal{E}_C/q - \psi - \gamma_n$ in (2.19) using (2.22) leads to the well known formula

$$\gamma_n = -\frac{\hbar^2}{12m_n^*} \left[\nabla^2 \log n + \frac{1}{2} (\vec{\nabla} \log n)^2 \right] = -\frac{\hbar^2}{6m_n^*} \frac{\vec{\nabla}^2 \sqrt{n}}{\sqrt{n}}, \quad (2.23)$$

often referred to as Bohm potential [19].

2.4 The Complete Quantum Corrected Drift-Diffusion Equation System

In respect to previous implementations the complete system of equations for the implementation of the quantum-corrected drift-diffusion model putting together (2.8), (2.11), (2.17), (2.18) and (2.20) reads

$$\begin{aligned} \vec{\nabla} \cdot (\hat{\varepsilon} \vec{\nabla} \psi) &= q(n - p - C), \\ \vec{\nabla} \cdot \vec{J}_n &= +q(\partial_t n + qR), \\ \vec{\nabla} \cdot \vec{J}_p &= -q(\partial_t p + qR), \\ \vec{J}_n &= q\mu_n n \left[\vec{\nabla} \left(\frac{\mathcal{E}_C}{q} - \psi - \gamma_n \right) + \frac{k_B T_L}{q} \frac{N_{C,0}}{n} \vec{\nabla} \left(\frac{n}{N_{C,0}} \right) \right], \\ \vec{J}_p &= q\mu_p p \left[\vec{\nabla} \left(\frac{\mathcal{E}_V}{q} - \psi - \gamma_p \right) - \frac{k_B T_L}{q} \frac{N_{V,0}}{p} \vec{\nabla} \left(\frac{p}{N_{V,0}} \right) \right], \\ \gamma_n &= \frac{\hbar^2 \beta}{12\lambda_n m_0} \left[\nabla^2 \left(\frac{\mathcal{E}_C}{q} - \psi - \gamma_n \right) - \frac{\beta}{2} \left(\frac{\vec{\nabla} \mathcal{E}_C}{q} - \vec{\nabla} \psi - \vec{\nabla} \gamma_n \right)^2 \right], \\ \gamma_p &= \frac{\hbar^2 \beta}{12\lambda_p m_0} \left[\nabla^2 \left(\frac{\mathcal{E}_V}{q} - \psi - \gamma_p \right) - \frac{\beta}{2} \left(\frac{\vec{\nabla} \mathcal{E}_V}{q} - \vec{\nabla} \psi - \vec{\nabla} \gamma_p \right)^2 \right]. \end{aligned} \quad (2.24)$$

This is the full equation system used as a basis in the following sections. In the literature various variable transformations are applied to obtain a numerical advantage. In [20] and [1] the variable $s = \sqrt{n}$ and $S_n = \sqrt{n/n_i}$ are used instead of n . In [17] the variables ψ , ψ_n , ψ_p , u_n and u_p are used by introducing $n = \sqrt{n} \exp(u_n)$ and $p = \sqrt{p} \exp(u_p)$, where ψ_n , ψ_p are the quasi-Fermi potentials for electrons and holes.

2.5 Quasi-Fermi Level Formalism

In the literature the quasi-Fermi levels are often used as unknowns to obtain a numerical advantage. In this Section a brief summary of carrier concentrations and quasi-Fermi levels is given.

Physically there is only one Fermi-Level for the whole device. Away from thermal equilibrium two so called quasi-Fermi levels are introduced. These quasi-Fermi levels refer to a subsystem such as the conduction and the valence band.

In semiconductor devices the carrier concentrations depend on the density of states and the distribution function [6]. The density of state functions $\rho_C(\mathcal{E})$ and $\rho_V(\mathcal{E})$ describe the amount of states available for the corresponding carrier type per energy. The distribution functions $f_n(\mathcal{E})$ and $f_p(\mathcal{E})$ give the occupation probabilities at each energy. Multiplying and integrating them over the whole band leads to the carrier concentrations

$$n = \int_{\mathcal{E}_C}^{\infty} \rho_C(\mathcal{E}) f_n(\mathcal{E}) d\mathcal{E} \quad \text{and} \quad p = \int_{-\infty}^{\mathcal{E}_V} \rho_V(\mathcal{E}) f_p(\mathcal{E}) d\mathcal{E}. \quad (2.25)$$

The distribution functions are the well known Fermi functions

$$f_n(\mathcal{E}) = \frac{1}{1 + \exp\left(\frac{\mathcal{E} - \mathcal{E}_{Fn}}{k_B T_L}\right)} \quad \text{and} \quad f_p(\mathcal{E}) = \frac{1}{1 + \exp\left(\frac{\mathcal{E}_{Fp} - \mathcal{E}}{k_B T_L}\right)}, \quad (2.26)$$

where \mathcal{E}_{Fn} and \mathcal{E}_{Fp} are the quasi-Fermi energies for electrons and holes.

For the calculations of the carrier concentrations, the Fermi integrals have to be used. In order to keep it simple the distribution functions are often reduced to the Boltzmann distribution functions

$$f_n(\mathcal{E}) = \exp\left(\frac{\mathcal{E}_{Fn} - \mathcal{E}}{k_B T_L}\right) \quad \text{and} \quad f_p(\mathcal{E}) = \exp\left(\frac{\mathcal{E} - \mathcal{E}_{Fp}}{k_B T_L}\right). \quad (2.27)$$

The conditions for assuming Boltzmann statistics are

$$\frac{\mathcal{E}_{Fn} - \mathcal{E}_C}{k_B T_L} \ll -1 \quad \text{and} \quad \frac{\mathcal{E}_V - \mathcal{E}_{Fp}}{k_B T_L} \ll -1. \quad (2.28)$$

So the Fermi energy has to be sufficiently lower than the conduction band edge energy and the Fermi energy of the holes has to be sufficiently higher than the valence band edge energy. In the target simulator MINIMOS-NT, Boltzmann statistics are assumed although the above conditions are not always fulfilled.

By defining the quasi-Fermi potentials ψ_n and ψ_p with

$$\psi_n = \frac{\mathcal{E}_{Fn} - \mathcal{E}_C}{q} \quad \text{and} \quad \psi_p = \frac{\mathcal{E}_{Fp} - \mathcal{E}_C}{q}, \quad (2.29)$$

assuming Boltzmann statistics and using

$$\mathcal{E}_C = \mathcal{E}_{C,0} - q(\psi + \gamma_n) \quad \text{and} \quad \mathcal{E}_V = \mathcal{E}_{V,0} - q(\psi + \gamma_p), \quad (2.30)$$

the carrier concentrations can be written as

$$\begin{aligned} n &= N_C \exp\left(\frac{\mathcal{E}_{Fn} - \mathcal{E}_C}{k_B T_L}\right) = N_C \exp\left(-\frac{q(\psi_n - \psi - \gamma_n) + \mathcal{E}_C}{k_B T_L}\right), \\ p &= N_V \exp\left(\frac{\mathcal{E}_V - \mathcal{E}_{Fp}}{k_B T_L}\right) = N_V \exp\left(\frac{q(\psi_p - \psi - \gamma_p) + \mathcal{E}_V}{k_B T_L}\right), \end{aligned} \quad (2.31)$$

where N_C and N_V are the effective density of states of the conduction and valence band.

It is interesting to note that in contrast to the classical treatment $np = n_i^2$ does not hold in thermal equilibrium due to the quantum-correction potentials γ_n and γ_p .

Using the quasi-Fermi potentials, the current relation (2.16) can be rewritten as

$$\vec{J}_n = -qn\mu_n \vec{\nabla}\psi_n \quad \text{and} \quad \vec{J}_p = -qp\mu_p \vec{\nabla}\psi_p. \quad (2.32)$$

As indicate by (2.32) the current densities depend on the gradient of the quasi-Fermi potentials only. In unipolar devices one majority carrier type is dominant.

In situations with vanishing current densities, such as during the calculations of a CV-curve of a MOS capacitor, both quasi-Fermi potential can be set constant. In this situation (2.31) are used to calculate the carrier concentrations in Poisson's equation (2.8) and the full quantum-corrected drift-diffusion equation system (2.24) reduces to

$$\begin{aligned} \vec{\nabla} \cdot (\hat{\varepsilon} \vec{\nabla} \psi) &= q \left[N_C \exp\left(\frac{q(\psi + \gamma_n) - \mathcal{E}_C}{k_B T_L}\right) - N_V \exp\left(\frac{\mathcal{E}_V - q(\psi + \gamma_p)}{k_B T_L}\right) - C \right], \\ \gamma_n &= \frac{\hbar^2 \beta}{12\lambda_n m_0} \left[\nabla^2 \left(\frac{\mathcal{E}_C}{q} - \psi - \gamma_n \right) - \frac{\beta}{2} \left(\frac{\vec{\nabla} \mathcal{E}_C}{q} - \vec{\nabla} \psi - \vec{\nabla} \gamma_n \right)^2 \right], \\ \gamma_p &= \frac{\hbar^2 \beta}{12\lambda_p m_0} \left[\nabla^2 \left(\frac{\mathcal{E}_V}{q} - \psi - \gamma_p \right) - \frac{\beta}{2} \left(\frac{\vec{\nabla} \mathcal{E}_V}{q} - \vec{\nabla} \psi - \vec{\nabla} \gamma_p \right)^2 \right]. \end{aligned} \quad (2.33)$$

Chapter 3

Numerical Implementation

In this thesis the general-purpose device and circuit simulator MINIMOS-NT [13] is used and extended with a density gradient model and a random dopant model to obtain a quantum-corrected drift-diffusion simulation of random dopant effects with appropriate screening. MINIMOS-NT provides a steady-state, transient and small-signal analysis of two- and three-dimensional devices with additional mixed-mode capabilities. It is the successor of Minimos 6, a specialized simulator for MOS structures. With its straightforward expandable material database and the state-of-the-art set of physical models MINIMOS-NT is a flexible simulation tool for modern semiconductor devices, like sub-nanometer MOS-, SOI-, and heterostructure devices.

In MINIMOS-NT devices consist of segments of arbitrary materials and corresponding physical models. Materials are treated in an abstract way using material classes, such as semiconductors, insulators and conductors. On this material classes miscellaneous models are available. With its flexible equation assembly module [21] it is possible to solve the quantum-corrected drift-diffusion model on one segment and the classical drift-diffusion model on another.

In this Chapter the implementation of the drift-diffusion model is briefly outlined and extended by the previously presented density gradient model to obtain the full quantum-corrected drift-diffusion model. We first start at finite volume discretization principles followed by a discretization of Poisson's and the continuity equations. We finish with three discretization variations of quantum-correction or density gradient equations now available in MINIMOS-NT.

3.1 Finite Volume Discretization

In Chapter 2 the well known and widely used drift-diffusion model was revised and expanded to the quantum-corrected drift-diffusion model [12, 14]. To numerically solve the corresponding equation system (2.24) or (2.33), a discretization scheme has to be found for each equation. In MINIMOS-NT the finite volume method, also known as box integration method, is used [22]. The finite volume method can be used to discretize conservation equations on unstructured grids. Unstructured grids are necessary to keep the number of grid points within a realistic range. In comparison to finite differences and the finite element method, the finite volume method has relatively high computational costs per grid point, but stability and robustness is superior, especially if large gradients occur. As the finite volume method is conservative and there is no requirement on continuity, the approximation of large gradients on a few grid points is possible [23].

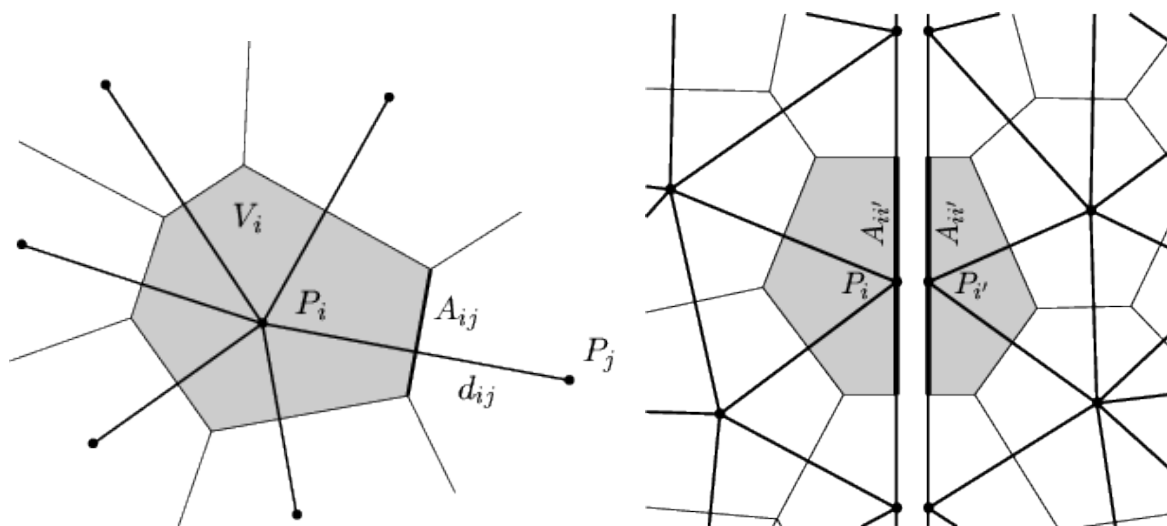


Figure 3.1: Left: Voronoi region \mathcal{V}_i of an inner grid point P_i , where V_i is the Voronoi volume, $A_{i,j}$ and $d_{i,j}$ are the connection area and distance to an adjacent grid point P_j . Right: Voronoi region of a interface grid point, split into two grid points P_i and P_i' for each segment, connected over the area $A_{i,i'}$. Taken from [24].

The simulation domain is partitioned into a finite set of finite volume subdomains, called control volumes. Accomplishing a suitable partition into subdomains is the main task of grid generators, where usually a Voronoi tessellation is used. Each subdomain, or Voronoi region in this case, consists of all points that are closer to the corresponding grid point than to any other. On each of these subdomains the differential equations are written in integral form to obtain the corresponding control function. Thus the solution of the control function is the weak solution of the partial differential equations. The quantities describe the average of the continuous solution within each subdomain and are not the local value at the grid points themselves. Using the finite volume method on unstructured grids only little grid information is needed by the simulator, as illustrated in Figure 3.1.

The necessary grid information can be split into three parts [24].

- **Grid point information**

For each grid point P_i its index i , location r_i and volume V_i of its subdomain \mathcal{V}_i are needed.

- **Subdomain connection information**

For each adjacent subdomains \mathcal{V}_j the corresponding grid point indices i and j , the distance between the grid points $d_{ij} = \|x_i - x_j\|$ and the connection area A_{ij} are needed.

- **Interface connection and boundary information**

On segment interfaces one grid point for each segment and one directly on the boundary are defined to offer complex interface modeling like interface charges. Thus, additional connection information for interface points is needed. It consists of the indices i and i' corresponding to the grid points P_i and $P_{i'}$ and the connection area of their subdomain $A_{ii'}$, see Figure 3.1. On boundary points P_i , where no adjacent subdomain is available, the corresponding boundary area A_{ii} is stored. The boundary models are defined on the interface and accessed using the boundary points.

Using this nearest neighbor information structure the partial differential equations must have a maximum order of two to be solved and have to be split if higher orders occur.

3.2 Control Function and Numerical Flux

From a general point of view, we can start with the conservation equation

$$\partial_t u + \vec{\nabla} \cdot \vec{f} = g, \quad (3.1)$$

where u is a conservative quantity, \vec{f} is the flux and g is the generation rate related to u . To obtain a discrete version Gauss' integral theorem

$$\int_{\mathcal{V}} \vec{\nabla} \cdot \vec{f} \, dV = \int_{\partial\mathcal{V}} \vec{n} \cdot \vec{f} \, dA \quad (3.2)$$

has to be applied, where \vec{f} is an arbitrary flux quantity, \vec{n} is the outgoing normal vector and $\partial\mathcal{V}$ denotes the boundary of the volume \mathcal{V} . The integral form of equation (3.1) on the Voronoi region \mathcal{V}_i reads

$$\int_{\mathcal{V}_i} \partial_t u \, dV + \int_{\partial\mathcal{V}_i} \vec{n} \cdot \vec{f} \, dA = \int_{\mathcal{V}_i} g \, dV. \quad (3.3)$$

Approximating the integrals we obtain the control function of u for an inner Voronoi region

$$F_u^S = (\partial_t u - g) V_i + \sum_j f_{ij} A_{ij} = 0, \quad (3.4)$$

where $f_{ij}A_{ij}$ denotes the sum of the outgoing flux to the adjacent Voronoi region \mathcal{V}_j over the connection area A_{ij} as indicated in figure 3.1. This form of equation (3.1) can be straightforwardly implemented using the described nearest neighbor grid information and a discrete version of the time derivative.

The main difficulty is finding a good approximation for the numerical flux f_{ij} . Where a linear approximation is sufficient for Poisson's equation, a linear approximation leads to instabilities for the continuity equation. Thus the well known Scharfetter-Gummel method is used. It gives a non-linear approximation of the numerical flux using an exponential ansatz. The numerical flux has to fulfill two conditions [23].

- **Lipschitz continuity**

It guarantees, that a small difference in the argument of f_{ij} leads to a small difference of f_{ij} .

- **Consistency condition**

The numerical flux f_{ij} has to be an appropriate approximation of the physical flux f .

3.3 Assembly Module

All control functions, interface and boundary models are implemented using the advanced equation assembly module of MINIMOS-NT [21]. With all control functions and their derivatives to every dependent quantity the segment right hand side vector \mathbf{b}_s and the segment Jacobian matrix \mathbf{A}_s are built up. Interface and boundary models are implemented with matrices \mathbf{A}_b , \mathbf{T}_b and vector \mathbf{b}_b . All together this leads to the full matrix system

$$\begin{aligned}\mathbf{A} &= \mathbf{A}_b + \mathbf{T}_b \cdot \mathbf{A}_s, \\ \mathbf{b} &= \mathbf{b}_b + \mathbf{T}_b \cdot \mathbf{b}_s, \\ \mathbf{A} \cdot \mathbf{u} &= \mathbf{b}.\end{aligned}\tag{3.5}$$

This matrix system is the basis for Newton's iteration method to solve for the quantity array \mathbf{u} . For each Newton iteration the matrix system is solved by direct or iterative methods. A detailed description is given in [21] or [22].

3.4 Interface Conditions

Interface points are split into one point for each segment and one directly on the interface. This offers complex modeling possibilities. After building up the discrete equation system within the segments, the control functions of the interface points remain incomplete, because the flux through the interface is not included. With the interface models these control function are completed. In the simplest case the incomplete control functions $F_{u_i}^S$ and $F_{u_{i'}}^S$ are connected over the interface flux $F_{u_{i,i'}}$ by

$$\begin{aligned}F_{u_i} &= F_{u_i}^S + F_{u_{i,i'}} = 0, \\ F_{u_{i'}} &= F_{u_{i'}}^S - F_{u_{i,i'}} = 0.\end{aligned}\tag{3.6}$$

When equation $F_{u_{i'}}$ is added to F_{u_i} the control function F_{u_i} is valid for both Voronoi regions \mathcal{V}_i and $\mathcal{V}_{i'}$ together. As a result the second equation is obsolete, thus must be deleted and can be used for an other interface equation for the corresponding grid points. For example $F_{u_{i'}} = u_i - u_{i'} = 0$, if the quantity u has to be continuous.

In MINIMOS-NT $F_{u_{i'}}$ is added to F_{u_i} by an additional non-diagonal entry in the transformation matrix \mathbf{T}_b . The second equation is deleted by clearing the corresponding diagonal entry in \mathbf{T}_b and entering the substitutive equation into \mathbf{b}_b and their derivatives into \mathbf{A}_b . It has to be noted that if a homogeneous segment is split, the result must not change. Thus appropriate interface conditions have to be defined. A special type of interface is the contact interface. Contacts only have one control function for the whole segment. All fluxes are summed up together in one equation.

3.5 Boundary Conditions

The implementation of boundary conditions is analogous to the implementation of interface conditions, but without adjacent points from other segments. Thus there is only one control function. Boundary conditions can be divided into three different types, Dirichlet, Neumann and Cauchy conditions.

Dirichlet Conditions

In the case of Dirichlet boundary conditions the quantity value u_i is set to a boundary value u_B , with the substitutive equation

$$F_{u_i} = u_i - u_B = 0. \quad (3.7)$$

The original control function is cleared by clearing the corresponding diagonal entry and the substitutive equation is entered into \mathbf{b}_b and its derivatives into \mathbf{A}_b .

To give an example, Dirichlet conditions are used for the electrostatic potential ψ at contacts.

Neumann Conditions

At boundaries with Neumann condition the normal component of the outgoing flux is specified by a boundary value F_B . This boundary flux is simply added to the control function of the boundary point

$$F_{u_i} = F_{u_i}^S + f_B A_i = 0, \quad (3.8)$$

by inserting $f_B A_i$ into \mathbf{b}_b , where A_i is the boundary area.

Neumann boundary conditions with zero outgoing flux are implicitly used if no boundary model is applied and are used at the simulation domain boundary of semiconductors for the current density and the electric flux density.

Cauchy Conditions

The Cauchy boundary condition can be seen as the generalization of Dirichlet and Neumann boundary conditions. In detail it is a linear combination of both, described by

$$f_B + \alpha u_i = \beta. \quad (3.9)$$

Cauchy boundary conditions are implemented similar to Neumann conditions, like

$$F_{u_i} = F_{u_i}^S + (\beta_i - \alpha_i u_i) A_i = 0, \quad (3.10)$$

where the outgoing flux additionally depends on the quantity itself. Thus also the Jacobian has to be modified by inserting $-\alpha A_i$ into \mathbf{A}_b .

For instance a Cauchy boundary condition is implemented for the quantum-correction potential at the semiconductor-oxide interface to obtain advanced fitting possibilities for a CV-curve compared with Dirichlet or Neumann boundary conditions. A detailed discussion is given in Section 4.1.

3.6 Discretization of Poisson's Equation

With respect to equation (3.1), a discrete version of Poisson's equation (2.8) can be obtained by setting

$$\partial_t u = 0, \quad \vec{f} = \vec{D} \quad \text{and} \quad g = -q(n - p - C). \quad (3.11)$$

The discrete version then reads

$$F_{\psi}^S = q(n_i - p_i - C_i) V_i + \sum_j D_{ij} A_{ij} = 0, \quad (3.12)$$

where the index i denotes the quantity value on the Voronoi region \mathcal{V}_i . The electric flux density D_{ij} is discretized as

$$D_{ij} = \epsilon_{ij} E_{ij} = \frac{\epsilon_i + \epsilon_j}{2} \frac{\psi_j - \psi_i}{d_{ij}}, \quad (3.13)$$

although $\epsilon_{ij} = (\epsilon_i + \epsilon_j)/2$ is not necessarily a good approximation of ϵ_{ij} [25].

3.7 Discretization of the Continuity Equation

The discretization of the continuity equation is shown for electrons. Using (3.1) we set

$$\partial_t u = -qn, \quad \vec{f} = \vec{J}_n \quad \text{and} \quad g = qR. \quad (3.14)$$

This leads to the spatially discrete version of the continuity equation

$$F_n^S = -q(\partial_t n + R)V_i + \sum_j J_{n,ij} A_{ij} = 0. \quad (3.15)$$

While a linear standard difference scheme can be used for the numerical flux in the discrete Poisson equation, this is not suitable for the discrete continuity equation. The linear treatment leads to instabilities if the spatial potential variation exceed $2k_B T_L/q$ between two grid points. In the Scharfetter-Gummel scheme [26] the differential equation is solved between two grid points P_i and P_j assuming $J_{n,ij}$, $J_{p,ij}$, E_{ij} , ϵ_{ij} , μ_n and μ_p to be constant along the edge between P_i and P_j . This leads to a robust numerical flux approximation even on coarse grids for both diffusive and convective problems. The current relation for electrons on the edge reads

$$\frac{\partial n(x)}{\partial x} - n(x) \frac{\Delta_{ij}}{d_{ij}} = \frac{J_n}{q} \mu_n V_{th}, \quad (3.16)$$

with $\Delta_{ij} = (\psi_j - \psi_i)/V_{th}$. This first-order linear non-homogeneous differential equation can be solved straightforwardly by using the exponential ansatz including the particular solution

$$n(x) = a \exp(bx) + c \quad (3.17)$$

and using the boundary conditions $n(x_i) = n_i$ and $n(x_j) = n_j$. The current relation for electrons within the Scharfetter-Gummel discretization reads

$$J_{n,ij} = \frac{q\mu_n V_{th}}{d_{ij}} \cdot (n_j \mathcal{B}(\Delta_{ij}) - n_i \mathcal{B}(-\Delta_{ij})), \quad (3.18)$$

where $\mathcal{B}(x) = x/(e^x - 1)$ is the Bernoulli function.

3.8 Discretization of the Density Gradient Equation

For the density gradient equation three different discretization schemes are implemented. Wettstein has introduced various discretization in [11], but most suffer from convergence problems. The two preferred schemes with the best numerical stability are the so called simplified and full scheme. In this work both have been implemented. In the simplified scheme the second-order term in (2.20) is neglected, thus the introduced general procedure described in Section 3.2 can be applied without any variable transformation. For the full scheme the quantum potential is rewritten in terms of carrier concentration and an exponential fitting is applied. The third scheme is similar to the high-resolution method developed by S. Odanaka [17, 27] using a simplified source term.

Simplified Scheme

In the simplified scheme [11] the second-order terms in the equations (2.24)_{4,5} are neglected. Thus, we can start again by comparing the simplified density gradient equations

$$\gamma_n^{\text{simplified}} = \frac{b_n}{q} \vec{\nabla} \cdot \vec{\nabla} u_n \quad \text{and} \quad \gamma_p^{\text{simplified}} = \frac{b_p}{q} \vec{\nabla} \cdot \vec{\nabla} u_p \quad (3.19)$$

with the conservative equation (3.1), using

$$u_n = \frac{q(\psi + \gamma_n) - \mathcal{E}_C}{k_B T_L}, \quad u_p = \frac{q(\psi + \gamma_p) - \mathcal{E}_V}{k_B T_L}, \quad (3.20)$$

$$b_n = \frac{\hbar^2}{12\lambda_n m_0} \quad \text{and} \quad b_p = \frac{\hbar^2}{12\lambda_p m_0}. \quad (3.21)$$

With the focus on electrons, setting

$$\partial_t u = 0, \quad \vec{f} = -b_n \vec{\nabla} u_n, \quad g = -\gamma_n \quad (3.22)$$

and using the linear standard difference numerical flux approximation

$$f_{ij} = b_n \vec{\nabla} u_{n,ij} = b_n \frac{u_{n,j} - u_{n,i}}{d_{ij}}, \quad (3.23)$$

this leads to the discretization

$$F_{\gamma_n}^S = \gamma_{n,i} V_i - b_n \sum_j \frac{A_{ij}}{d_{ij}} (u_{n,j} - u_{n,i}) = 0. \quad (3.24)$$

The density gradient discretization for holes is derived analogously and reads

$$F_{\gamma_p}^S = \gamma_{p,i} V_i - b_p \sum_j \frac{A_{ij}}{d_{ij}} (u_{p,j} - u_{p,i}) = 0. \quad (3.25)$$

Full Scheme

Due to the second order term in the full scheme [11], a discretization of equations (2.24)_{4,5} cannot be done analogously to the procedure described in Section 3.2. Thus, the Bohm potential formulation (2.23) with an exponential fitting is used instead. The Bohm potentials for electrons and holes read

$$\gamma_n = -\frac{2b_n}{q} \frac{\vec{\nabla}^2 \sqrt{n}}{\sqrt{n}} \quad \text{and} \quad \gamma_p = \frac{2b_p}{q} \frac{\vec{\nabla}^2 \sqrt{p}}{\sqrt{p}}. \quad (3.26)$$

By setting

$$\partial_t u = 0, \quad \vec{f} = \frac{2b_n}{q} \vec{\nabla} \sqrt{n}, \quad g = -\sqrt{n} \gamma_n \quad (3.27)$$

and the numerical flux approximation

$$f_{ij} = \frac{2b_n}{q} \vec{\nabla} \sqrt{n_{ij}} = \frac{2b_n}{q} \vec{\nabla} (\sqrt{n_j} - \sqrt{n_i}), \quad (3.28)$$

we obtain

$$F_{\gamma_n}^S = \gamma_{n,i} V_i + \frac{2b_n}{q} \sum_j \frac{A_{ij}}{d_{ij}} \left[\frac{\sqrt{n_j}}{\sqrt{n_i}} - 1 \right] = 0. \quad (3.29)$$

This discretization suffers from convergence problems due to the exponential behavior of the electron concentration. Using (2.31)₁ and assuming thermal equilibrium, the electron concentration is substituted with

$$\sqrt{n_i} = \sqrt{N_C} \cdot \exp\left(\frac{\mathcal{E}_{C,i} - q(\psi_i + \gamma_{n,i})}{2k_B T_{L,i}}\right) = \sqrt{N_C} \cdot \exp\left(-\frac{u_{n,i}}{2}\right). \quad (3.30)$$

This leads to the full scheme for electrons and analogous for holes

$$\begin{aligned} F_{\gamma_n}^S &= \gamma_{n,i} V_i + \frac{2b_n}{q} \sum_j \frac{A_{ij}}{d_{ij}} \left[\exp\left(\frac{u_{n,i} - u_{n,j}}{2}\right) - 1 \right] = 0 \\ F_{\gamma_p}^S &= \gamma_{p,i} V_i + \frac{2b_p}{q} \sum_j \frac{A_{ij}}{d_{ij}} \left[\exp\left(\frac{u_{p,i} - u_{p,j}}{2}\right) - 1 \right] = 0. \end{aligned} \quad (3.31)$$

An alternative derivation can be done, by integrating (3.27)₂ along the edge between the adjacent grid points P_i and P_j .

In [28] the discretization investigations done in [11, 14] have been further improved. It was reported that the exponential function leads to numerical overflow for $u_i > u_j$ during the first steps of Newton iteration. Thus, the discretization

$$F_{\gamma_n}^S = \gamma_{n,i} V_i + \frac{2b_n}{q} \sum_j \frac{A_{ij}}{d_{ij}} \left[u_{n,j} - u_{n,i} - \frac{1}{4}(u_{n,i} - u_{n,i})^2 \right] = 0 \quad \text{for} \quad u_{n,i} > u_{n,j} \quad (3.32)$$

is used.

The numerical advantage of (3.32) could not be confirmed in simulations done within this thesis. If convergence problems occur, the more robust simple scheme can be used instead.

High Resolution Scheme

The high resolution scheme reported by Odanaka [17] can be considered an extension of the non-linear scheme described by Ancona *et al.* [20, 29]. Tang *et al.* [30] reports that the non-linear scheme leads to high sensitivity to boundary conditions at the silicon-oxide interface but is superior in smooth regions. In [27] a slope-limiter is introduced to switch between a high-order and a low-order numerical flux. This makes it possible to use a linear scheme near the interface and the non-linear scheme otherwise. In the course of this thesis the high resolution scheme with slope-limiter is implemented using a simplified source term approximation.

To derive the high resolution scheme we employ a similar procedure as used for the full scheme, but with a different numerical flux approximation. In the full scheme the exponential behavior is considered after the numerical flux approximation. In the high resolution scheme the exponential behavior (3.30) is used to obtain a high-order and a low-order numerical flux approximation. The flux for electrons reads

$$f = \frac{2b_n}{q} \vec{\nabla} \sqrt{n} = -\frac{b_n}{q} \sqrt{N_C} \exp\left(-\frac{u_n}{2}\right) \vec{\nabla} u_n. \quad (3.33)$$

For the low-order flux the linear approximation

$$f_{ij}^{\text{low}} = -\frac{b_n}{q} \sqrt{n_i} \exp\left(\frac{u_{n,i} - u_{n,j}}{4}\right) \frac{u_{n,j} - u_{n,i}}{d_{ij}} \quad (3.34)$$

is used. To obtain the high-order flux equation, (3.33) is rewritten and integrated over the edge between the grid points P_i and P_j

$$\int_{x_i}^{x_j} \vec{\nabla} u \, dx = -\frac{q}{b_n} \frac{1}{\sqrt{N_C}} \int_{x_i}^{x_j} f_{ij}^{\text{high}} \exp\left(\frac{u_n}{2}\right) \, dx. \quad (3.35)$$

Solving the integrals assuming $\nabla u = \partial u / \partial x = (u_j - u_i) / d_{ij}$ and f_{ij}^{high} are constant along the edge leads to

$$u_{n,j} - u_{n,i} = -2 \frac{q}{b_n} \frac{1}{\sqrt{N_C}} f_{ij}^{\text{high}} \left[\exp\left(\frac{u_{n,j}}{2}\right) - \exp\left(\frac{u_{n,i}}{2}\right) \right] \frac{d_{ij}}{u_{n,j} - u_{n,i}} \quad (3.36)$$

which results in the high-order numerical flux approximation

$$f_{ij}^{\text{high}} = -\frac{b_n}{q} \sqrt{n_i} \mathcal{B}\left(\frac{u_{n,j} - u_{n,i}}{2}\right) \frac{u_{n,j} - u_{n,i}}{d_{ij}}. \quad (3.37)$$

Introducing the spatially dependent slope-limiter θ_i the combination of both low-order and high-order numerical flux approximations reads

$$\begin{aligned} f_{ij} &= (1 - \theta_i) f_{ij}^{\text{low}} + \theta_i f_{ij}^{\text{high}} \\ &= -\frac{b_n}{q} \sqrt{n_i} \exp\left(\frac{u_{n,i} - u_{n,j}}{4} (1 - \theta_i)\right) \mathcal{B}\left(\frac{u_{n,j} - u_{n,i}}{2} \theta_i\right) \frac{u_{n,j} - u_{n,i}}{d_{ij}} \end{aligned} \quad (3.38)$$

3.8. DISCRETIZATION OF THE DENSITY GRADIENT EQUATION 28

The control functions for the high-resolution discretization scheme with slope-limiter for both electrons and holes read

$$\begin{aligned}
 F_{\gamma_n}^S &= \gamma_{n,i} V_i + \frac{b_n}{q} \sum_j \frac{A_{ij}}{d_{ij}} \exp\left(\frac{u_{n,i} - u_{n,j}}{4}(1 - \theta_i)\right) \mathcal{B}\left(\frac{u_{n,j} - u_{n,i}}{2}\theta_i\right) (u_j - u_i) = 0, \\
 F_{\gamma_p}^S &= \gamma_{p,i} V_i + \frac{b_p}{q} \sum_j \frac{A_{ij}}{d_{ij}} \exp\left(\frac{u_{p,i} - u_{p,j}}{4}(1 - \theta_i)\right) \mathcal{B}\left(\frac{u_{p,j} - u_{p,i}}{2}\theta_i\right) (u_j - u_i) = 0
 \end{aligned}
 \tag{3.39}$$

A numerical benefit of this advanced discretization scheme could not be confirmed.

Chapter 4

Results

In this Chapter the effect of the quantum-correction to the classical drift-diffusion model is investigated. In particular we are interested in the confinement within sub-nanometer MOS devices. In the first Section various boundary conditions are discussed to match the quantum-corrected drift-diffusion model with the Vienna-Schrödinger-Poisson solver VSP2 [31] using a one dimensional MOS capacitor. After studying the dependence on the grid spacing we consider a PMOS transistor. Investigations to full device behavior and appropriate screening of discrete dopants are covered in in the second part of this thesis, which focuses on random discrete dopant effects.

4.1 Carrier Confinement and CV-Curve Matching

In this Section the ability of the quantum-corrected drift-diffusion model to describe confinement effects within a MOS device is investigated. For this purpose a one dimensional MOS capacitor is used and the results are compared with the Vienna-Schrödinger-Poisson solver VSP2. All simulations are done with the simplified discretization scheme (3.24) and (3.24). While a simple Dirichlet boundary condition for the quantum-correction potential can be used to obtain appropriate modeling of carrier confinement at a specific voltage, it cannot fit the whole CV-curve in Figure 4.1. Thus four Si/SiO₂ boundary conditions are explored. For each example the CV-curve for three different dopings and the electron concentration for one doping and various gate voltages is shown. The implicit zero out-flux Neumann boundary condition in Figure 4.2 fails in accumulation and inversion and can be improved by a Neumann boundary condition with a small non-zero out-flux to appropriately describe depletion and threshold voltage shift in Figure 4.3 and the fourth example the Cauchy boundary condition is a linear combination of Dirichlet and Neumann boundary conditions. With the Cauchy boundary condition it is possible to fit the whole CV-curve, but its fitting parameters are valid for one specific doping only. Thus the parameters have to be adjusted where different channel doping is used. In the PMOS example in Figure 4.5 (top) the same situation occurs. While the same fitting parameters from the NMOS simulation are used, again they only fit the curve with doping $N_{A,D} = 3 \times 10^{17} \text{cm}^{-2}$. In Figure 4.5 (bottom) the CV-curve for various oxide thicknesses is given. The density gradient model with Cauchy boundary condition fits all curves without adjusting the parameters.

Dirichlet Boundary Condition

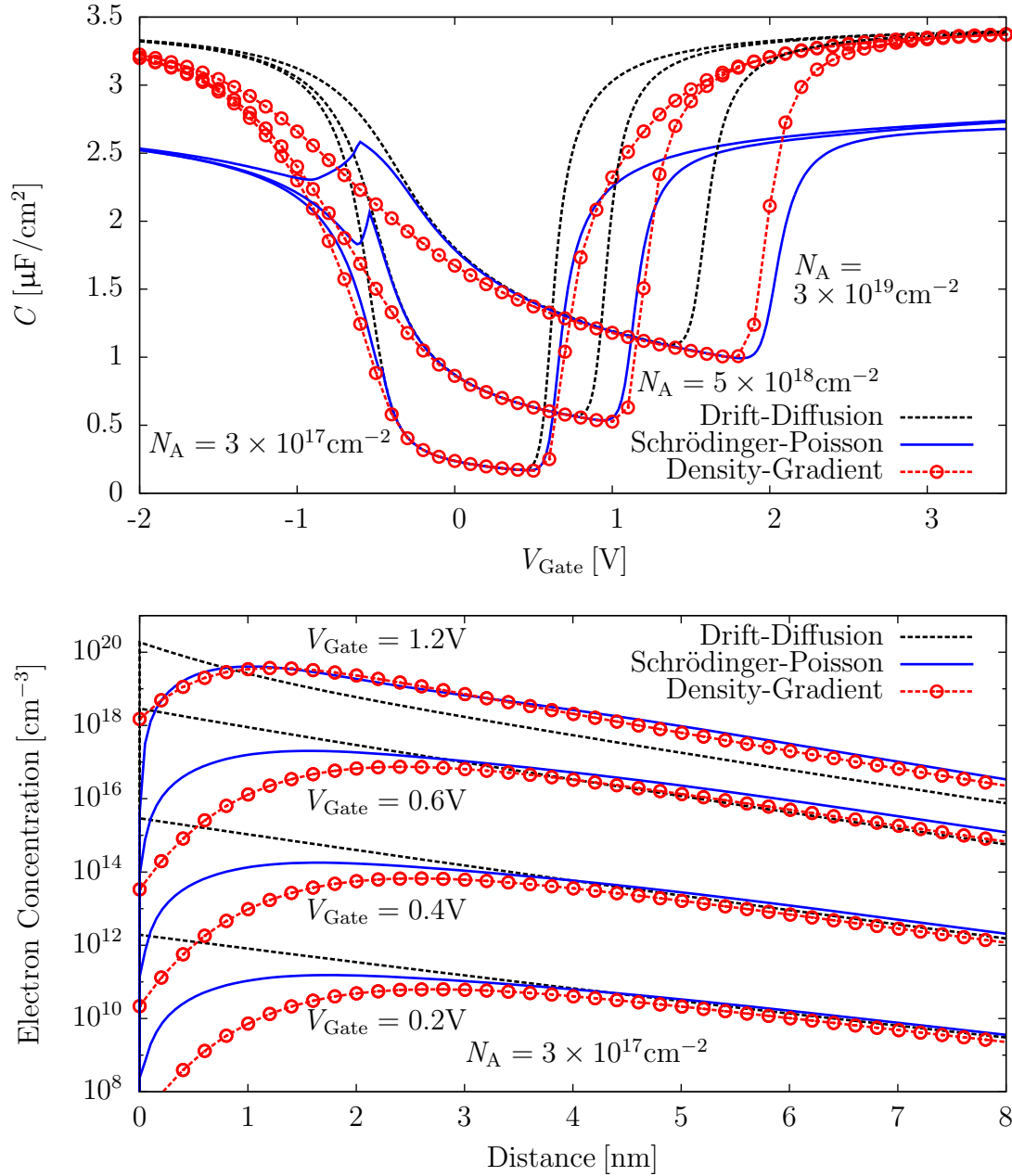


Figure 4.1: Top: CV-curve of a silicon 1D NMOS capacitor with a metal gate, a SiO_2 oxide with thickness $t_{\text{ox}} = 1 \text{ nm}$ and an acceptor substrate doping of $N_A = 3 \times 10^{17} \text{ cm}^{-2}$, $5 \times 10^{18} \text{ cm}^{-2}$ and $3 \times 10^{19} \text{ cm}^{-2}$. Bottom: Electron concentration in the substrate for $V_{\text{Gate}} = 0.2 \text{ V}$, 0.4 V , 0.6 V , 1.2 V and an acceptor doping of $N_A = 3 \times 10^{17} \text{ cm}^{-2}$. For the quantum-correction potential a Dirichlet boundary condition is used on the Si/SiO₂ interface. The simulation parameters for the quantum-correction potential are $\lambda_n = 0.22$, $\lambda_p = 0.2$, $\gamma_n^b = -0.3$ and $\gamma_p^b = 0.3$ (cf. Appendix A). With Dirichlet boundary conditions the carrier distribution can be described in a good approximation for a specific voltage, but not for the whole range. Thus the CV-curve cannot be reproduced using Dirichlet boundary conditions on the interface.

Implicit Zero-Outflux Neumann Boundary Condition

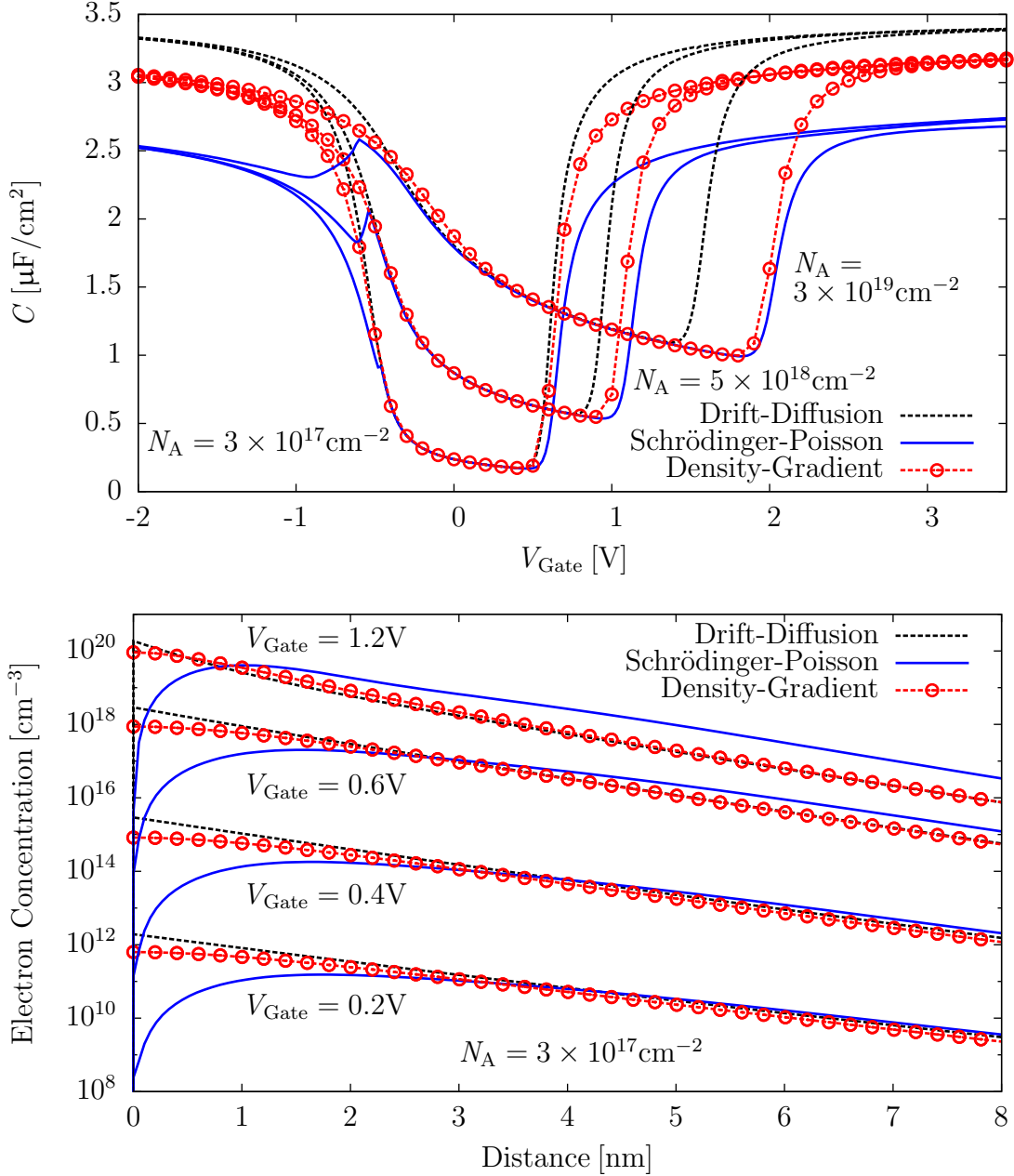


Figure 4.2: Top: CV-curve of a silicon 1D NMOS capacitor with a metal gate, a SiO_2 oxide with thickness $t_{\text{ox}} = 1 \text{ nm}$ and an acceptor channel doping of $N_A = 3 \times 10^{17} \text{ cm}^{-2}$, $5 \times 10^{18} \text{ cm}^{-2}$ and $3 \times 10^{19} \text{ cm}^{-2}$. Bottom: Electron concentration in the channel for $V_{\text{Gate}} = 0.2 \text{ V}$, 0.4 V , 0.6 V , 1.2 V and an acceptor doping of $N_A = 3 \times 10^{17} \text{ cm}^{-2}$. For the quantum-correction potential of the DG model the implicit zero-outflux Neumann boundary condition is used at the Si/SiO₂ interface. The simulation parameters for the quantum-correction potential are $\lambda_n = 0.22$ and $\lambda_p = 0.2$ (cf. Appendix A). Fitting the whole CV-curve and the electron distribution within the channel is not possible.

Nonzero-Outflux Neumann Boundary Condition

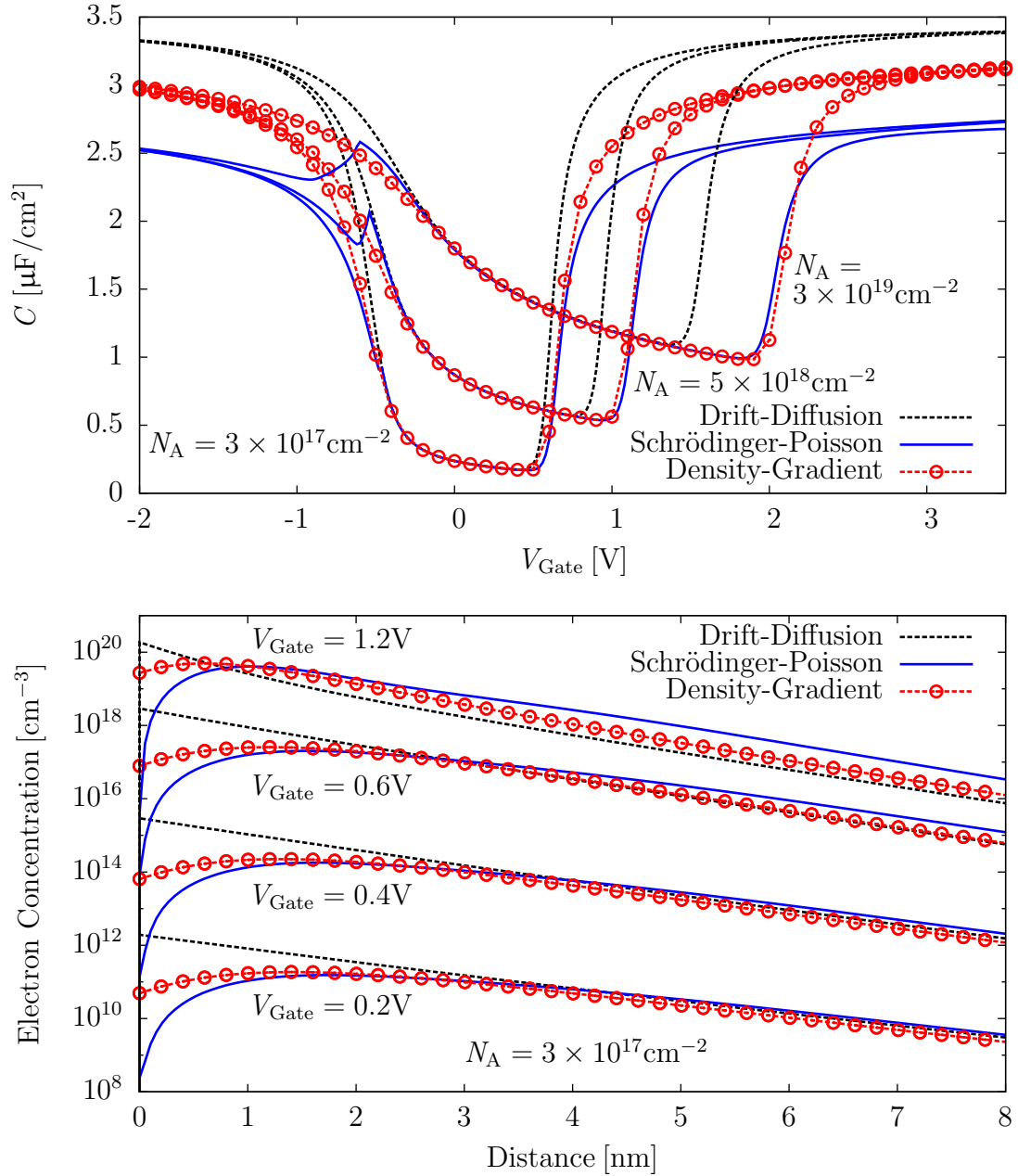


Figure 4.3: Top: CV-curve of a silicon 1D NMOS capacitor with a metal gate, a SiO_2 oxide with thickness $t_{\text{ox}} = 1 \text{ nm}$ and an acceptor channel doping of $N_A = 3 \times 10^{17} \text{ cm}^{-2}$, $5 \times 10^{18} \text{ cm}^{-2}$ and $3 \times 10^{19} \text{ cm}^{-2}$. Bottom: Electron concentration in the channel for $V_{\text{Gate}} = 0.2 \text{ V}$, 0.4 V , 0.6 V , 1.2 V and an acceptor doping of $N_A = 3 \times 10^{17} \text{ cm}^{-2}$. For the quantum-correction potential of the DG model a non-zero out-flux Neumann boundary condition is used on the Si/SiO₂ interface. The simulation parameters for the quantum-correction potential are $\lambda_n = 0.22$, $\lambda_p = 0.2$ and $F_n^b = F_p^b = 5 \times 10^{-3}$ (cf. Appendix A). An appropriate prediction of the threshold voltage shift is possible but the boundary model fails in accumulation and strong inversion.

Cauchy Boundary Condition

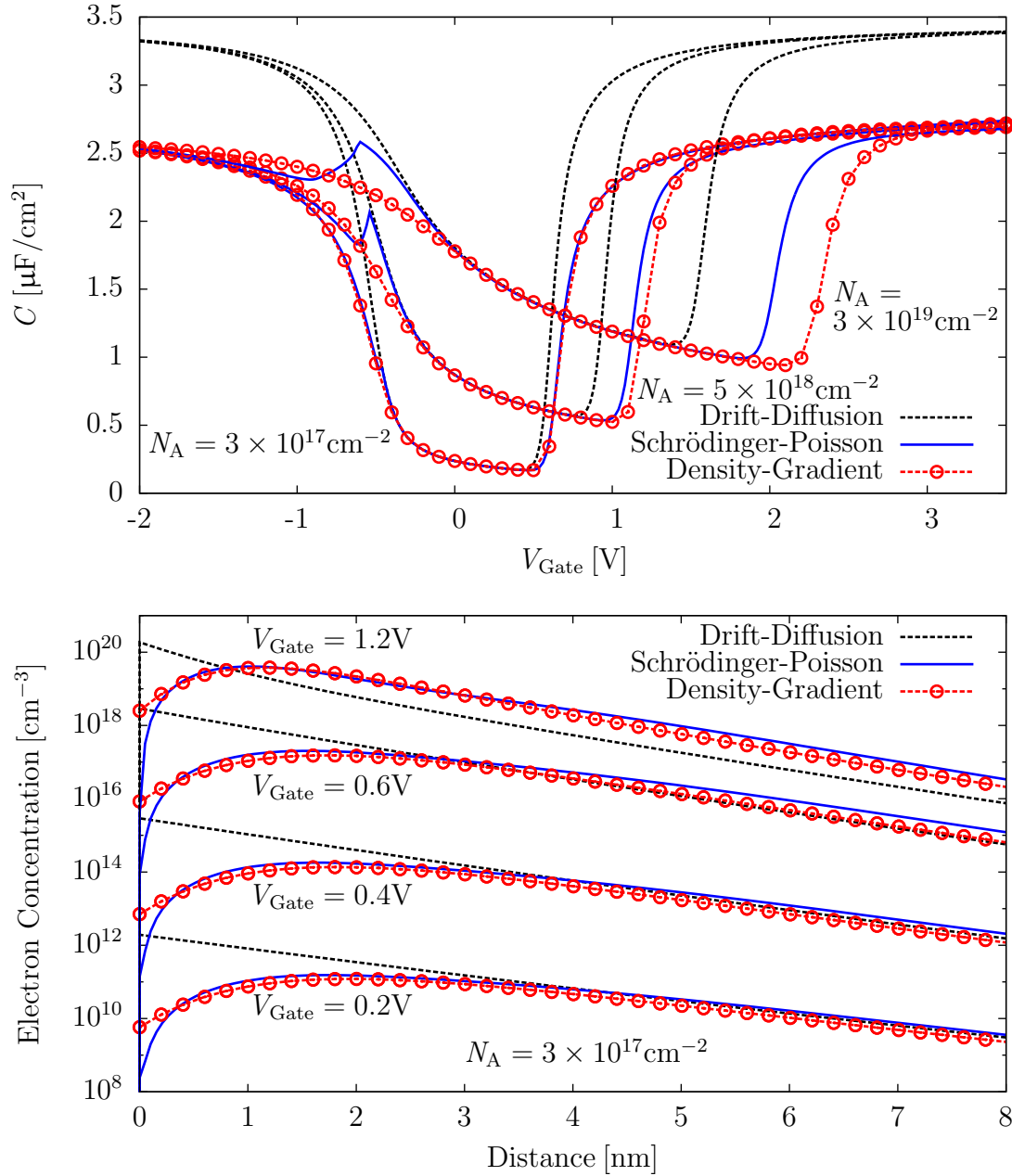


Figure 4.4: Top: CV-curve of a silicon 1D NMOS capacitor with a metal gate, a SiO₂ oxide with thickness $t_{\text{ox}} = 1$ nm and an acceptor channel doping of $N_A = 3 \times 10^{17} \text{ cm}^{-2}$, $5 \times 10^{18} \text{ cm}^{-2}$ and $3 \times 10^{19} \text{ cm}^{-2}$. Bottom: Electron concentration in the channel for $V_{\text{Gate}} = 0.2 \text{ V}$, 0.4 V , 0.6 V , 1.2 V and an acceptor doping of $N_A = 3 \times 10^{17} \text{ cm}^{-2}$. For the quantum-correction potential of the DG model a Cauchy boundary condition is used at the Si/SiO₂ interface. The simulation parameters for the quantum-correction potential are $\lambda_n = 0.22$, $\lambda_p = 0.2$, $\alpha_n = -3.7 \times 10^{-4}$, $\beta_n = -7.5 \times 10^{-5}$, $\alpha_p = -4.3 \times 10^{-4}$ and $\beta_p = -8.3 \times 10^{-5}$ (cf. Appendix A). Using a Cauchy boundary condition it is possible to fit the whole CV-curve at one specific doping. For different doping concentration the parameters have to be adjusted.

Dependency on Oxide Thickness and PMOS

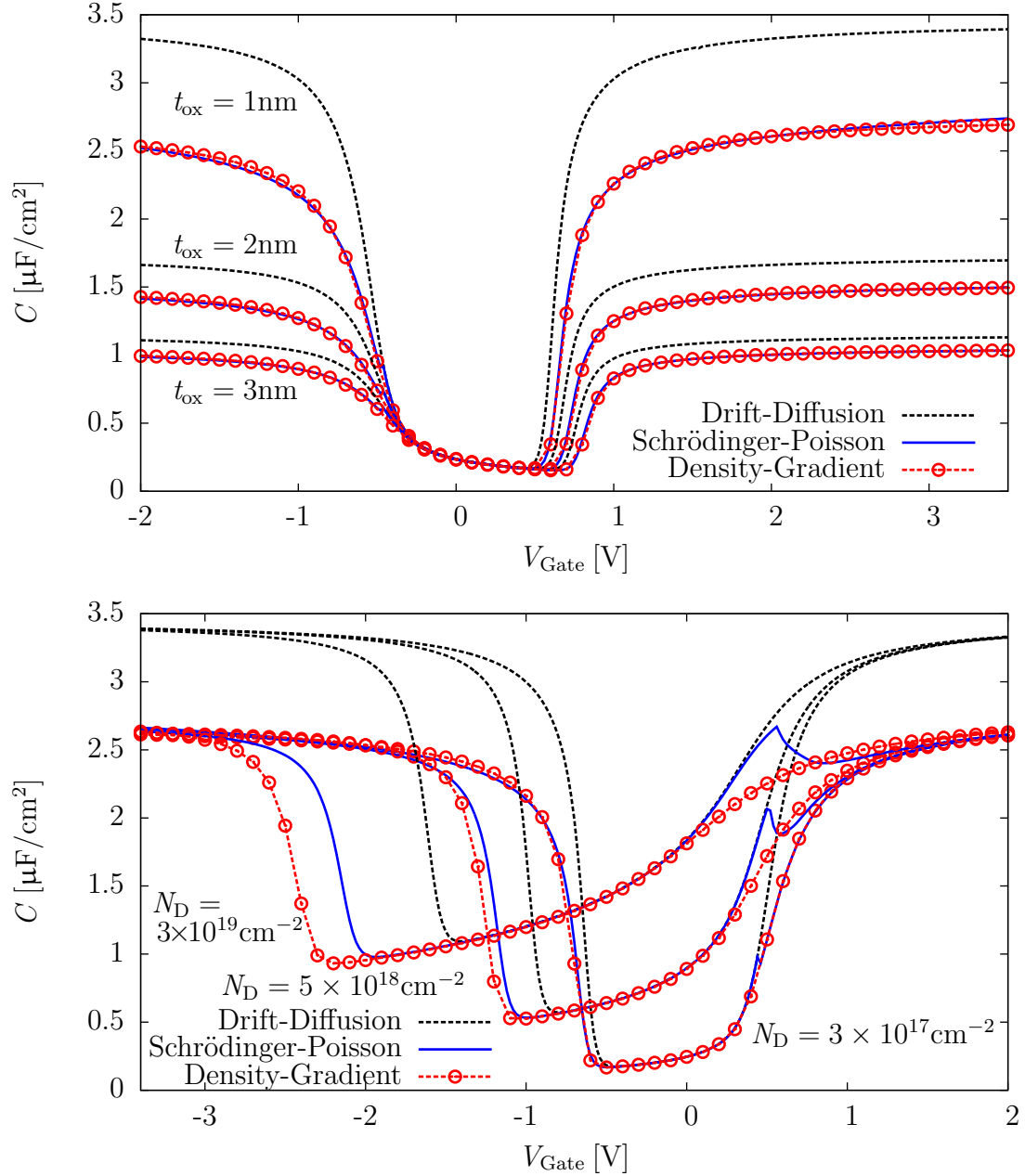


Figure 4.5: Top: CV-curve of a silicon 1D PMOS capacitor with a metal gate, a SiO₂ oxide with thickness $t_{\text{ox}} = 1$ nm and a channel doping of $N_D = 3 \times 10^{17} \text{ cm}^{-2}$, $5 \times 10^{18} \text{ cm}^{-2}$ and $3 \times 10^{19} \text{ cm}^{-2}$. Bottom: CV-curve of a silicon 1D NMOS capacitor with a metal gate, an channel doping of $N_A = 3 \times 10^{17} \text{ cm}^{-2}$ and SiO₂ oxide with thickness $t_{\text{ox}} = 1$ nm, 2 nm and 3 nm. For the quantum-correction potential of the DG model a Cauchy boundary condition is used at the Si/SiO₂ interface. The simulation parameters for the quantum-correction potential are $\lambda_n = 0.22$, $\lambda_p = 0.2$, $\alpha_n = -3.7 \times 10^{-4}$, $\beta_n = -7.5 \times 10^{-5}$, $\alpha_p = -4.3 \times 10^{-4}$ and $\beta_p = -8.3 \times 10^{-5}$ (cf. Appendix A). In the PMOS example it is again possible to fit a CV-curve for one specific doping with Cauchy boundary conditions using the same parameters as in the NMOS example. For a change in the oxide thickness the results remain valid without readjustment of the fitting parameters.

4.2 Dependence on Grid Spacing

Figure 4.6 shows the electron concentration for different grid spacings. While a grid spacing $d_g < 0.4$ nm leads to an appropriate description, the maximum cannot be predicted right for coarser grids and the quantum-correction become negligible. For the simulation the simplified model was used, because all other schemes do not converge for all grid spacings. For the solution on a fine grid no difference between the various discretization schemes has been experienced. For $d_g = 1$ nm, 1.5 nm and 2 nm also the result obtained from the high-resolution scheme is included in the figure. The numerical advantage of the high-resolution scheme or the full scheme pointed out in [28] could not be confirmed using Cauchy boundary conditions.

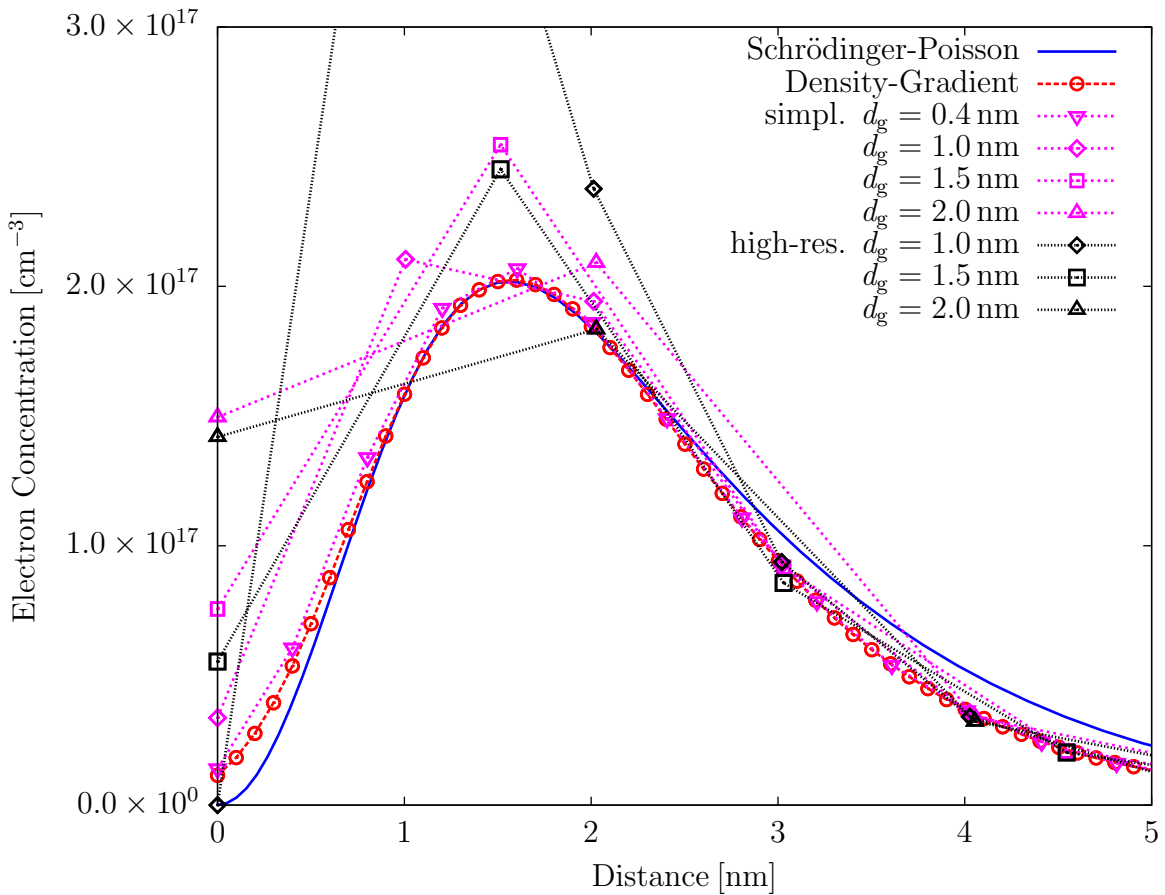


Figure 4.6: Electron concentration using various grid spacings for a Si NMOS with metal gate, SiO₂ of thickness $t_{\text{ox}} = 1$ nm and a donor doping $N_D = 3 \times 10^{17} \text{ cm}^{-2}$. On the Si/SiO₂ interface a Cauchy boundary condition is used. The simulation parameters are $\lambda_n = 0.286$, $\alpha_e = -3.9 \times 10^{-4}$ and $\beta_e = -5.4 \times 10^{-5}$ (cf. Appendix A). Only the simplified discretization scheme converges for all grid spacings. No numerical advantage of the high-resolution or full scheme has been found.

Part II

Random Dopant Fluctuations

Chapter 5

Splitting into Generation and Simulation Model

The random dopant model is split into generation and simulation models. For both purposes several models are available. In the generation model the dopants positions are randomly chosen using the continuous doping concentration for the local mean values of the random process. In the simulation model the random dopant doping concentration is calculated using the dopant positions. This random dopant doping concentration is then included in the Poisson equation instead of the continuous doping concentration. For all other models, such as the ionized impurity scattering model, the original continuous doping concentration is used.

By choosing the seed of the random number generator in the input-deck various devices can be generated. Because of the flexible equation assembly of MINIMOS-NT, it is possible to use random dopants on one segment and continuous doping on another. The models are implemented using a *kd*-tree for an efficient three dimensional nearest neighbor and points in range search.

In this Chapter the implementation of the random discrete dopant models is described. We start with the general procedure followed by the description of the *kd*-tree. In Section 6 the generation of random discrete dopants is outlined. The simulation models are described in Section 7.

For every impurity of each segment where a random dopant model is enabled, the doping concentration is randomized using the following procedure.

- **Building the *kd*-tree and information arrays:** The grid points are stored in a *kd*-tree to provide efficient search algorithms. The continuous and random dopant doping concentrations, Voronoi region volumes and the Voronoi region bounds are stored separately in arrays. This data is necessary to generate and model random discrete dopants.
- **Generation model:** The generation of random discrete dopants with one of the three generation models. The *Voronoi-region* model, the *segment* model or the *read-from-file* model.

- **Simulation model:** The new random dopant doping concentration is calculated using the generated dopants. The available models are the nearest grid point model and the long-range model.

5.1 kd-tree

A *kd*-tree is the generalization of a binary tree to k dimensions [32]. In one dimension a record is defined by one key. By choosing one partition key the file containing all records is partitioned into one subfile with records that have a lower key and one with all records that have a higher key. Each subfile is again partitioned in the same way. The partition keys are stored in the nodes of the tree with the partition keys of the subfiles as child nodes. The tree ends at the leafs, where the subfile consists of one record only and thus has no child nodes.

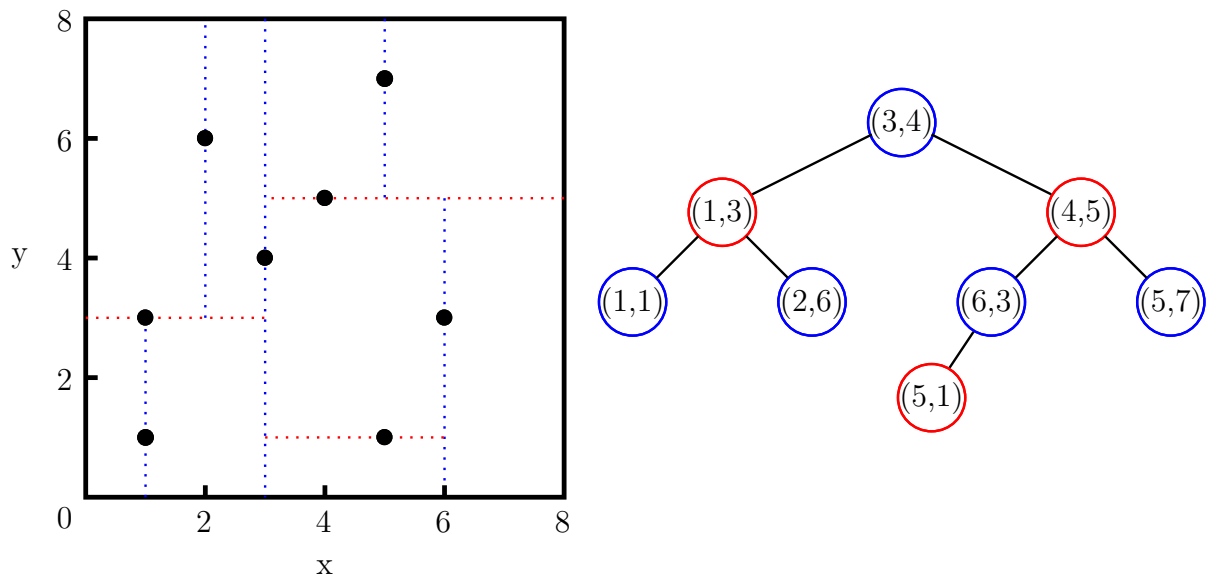


Figure 5.1: Example of a 2D *kd*-tree (right) out of 8 points (left). Starting with the point at $(3, 4)$ as a lucky choice the domain is partitioned into two subdomains, one for lower $x \leq 3$ and one for $x > 3$ using the points as the root point. All records of the left subfile are in the left subdomain and all records of the right subfile are in the right subdomain. The right subdomain is then partitioned with the point at $(4, 5)$ into one subdomain for $y \leq 5$ and one for $y > 5$ adding the point as the right child of the root point. These subdomains are again partitioned using the x -coordinates of the points.

In k dimensions each record is defined by k keys. The subfiles are cyclically partitioned in the k dimensions using the first key for the first partition, the second for the second and so on. For the $k + 1$ th partition the first key is used again.

Figure 5.1 shows an example of a two dimensional *kd*-tree (right) constructed from 8 points (left). By proper selection of the partition keys the tree can be balanced so that the subfile size difference of each partition is low. In this case the depth of the tree and the computational costs for searching can be reduced.

A *kd*-tree offers an efficient nearest neighbor search with $\mathcal{O}(\log(n))$ effort and an efficient record within bounds search. The record within bounds search is used in the *point-in-segment-check* described in the next section. The nearest neighbor search is used to find the nearest grid point and thus to check whether a random discrete dopant is inside a particular Voronoi region.

Chapter 6

Generation of Random Discrete Dopants

In this Chapter the three implemented generation models are described. The *Voronoi-region* model is an efficient procedure to generate random discrete dopants on segments with non-constant doping. The *segment* model considers the lattice spacing and is designed for small device segments. To offer the ability of importing random dopant data from a file the *read-from-file* model is included. Because a *point-in-segment-check* is necessary for all models the check is described separately in Subsection 6.4. By setting the seed of the random number generator in the input-deck of the simulator various reproducible device samples can be generated using one main device with continuous doping representing the mean values.

6.1 Voronoi-Region Model

The *Voronoi-region* model is an efficient method to generate random discrete dopants with non-constant doping. As the doping concentration varies from zero to more than 10^{20} cm^{-3} within a device a simple rejection technique is highly inefficient.

The *Voronoi-region* model can be described with three steps.

Step 1: The number of dopants for each Voronoi region is calculated using a Poisson distribution.

Step 2: For each Voronoi region a cuboid is defined by the minimum and maximum coordinates including the whole Voronoi region. This cuboid is used to generate the dopants for the corresponding Voronoi region by choosing random x , y and z coordinates.

Step 3: With the nearest neighbor search the nearest grid point is determined. If this grid point corresponds to the particular Voronoi region and the dopant is inside the segment the dopant is saved in the random dopant list with the coordinates and the corresponding grid point index. Otherwise it is rejected.

To test whether the dopant is inside of segment or not is done by the *point-in-segment-check* described in Subsection 6.4.

Figure 6.2 shows random discrete dopants in a 22 nm NMOS generated with the *Voronoi-region* model from the device with continuous doping in Figure 6.1.

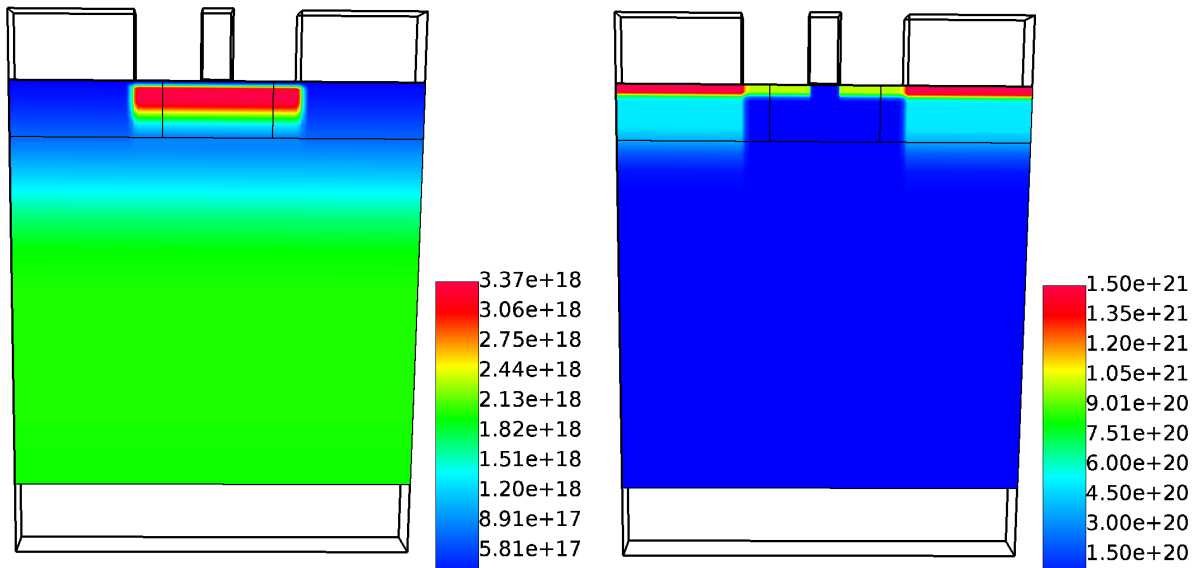


Figure 6.1: Continuous acceptor doping (left) and donor doping (right) of a NMOS with 22 nm gate length.

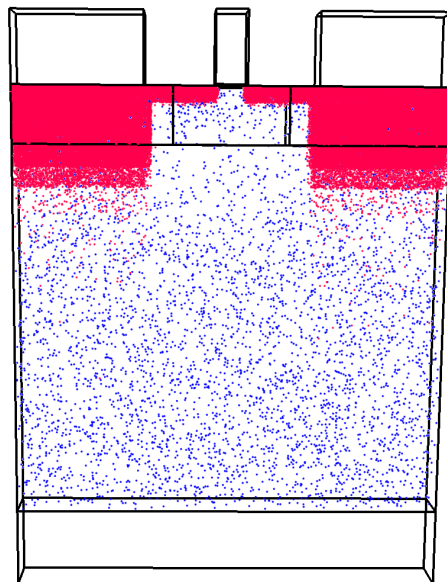


Figure 6.2: Random discrete doping of a NMOS with 22 nm gate length obtained by the *Voronoi-region* model. The continuous doping profile in Figure 6.1 is used for the mean value of Poisson's distribution. The donors are visualized by red dots, the acceptors with blue dots.

6.2 Segment Model

The *segment* model is especially designed to distinguish between effects due to total dopant number fluctuations and the random position of the individual dopants. Figure 6.3 shows random discrete dopants in a silicon resistor generated with the *segment* model.

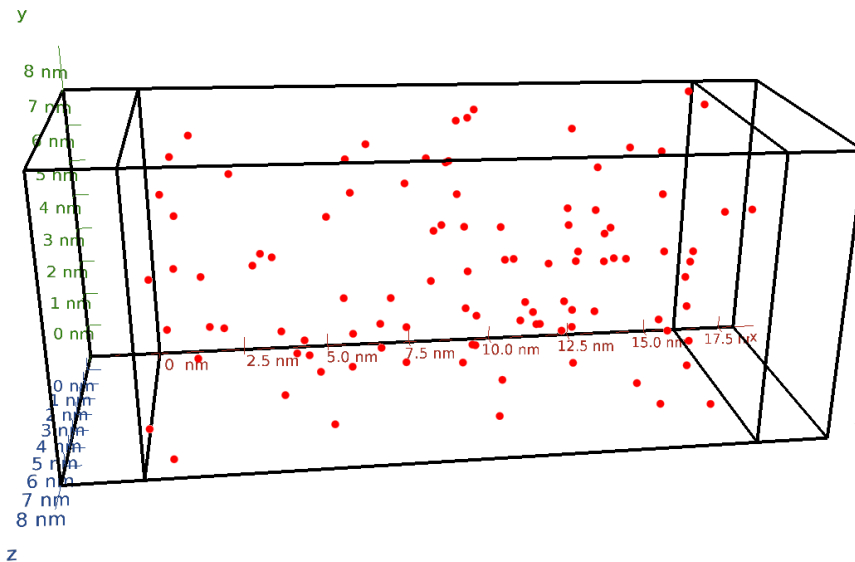


Figure 6.3: A $16 \times 8 \times 8$ nm silicon resistor with random discrete dopants obtained by the *segment* model. The donors are visualized by red dots.

The algorithm of the *segment* model can be described with the following steps.

- Step 1:** A cuboid is defined with the minimum and maximum coordinates of the corresponding segment so that the whole segment is inside.
- Step 2:** The number of dopants within the cuboid can be calculated using Poisson's statistics to consider total dopant fluctuations or set to the total dopant number of the reference doping to include random position only. For each of these dopants three random numbers for the x , y and z coordinates are chosen.
- Step 3:** With the nearest neighbor search the nearest grid point is determined to obtain the point index and the local doping concentration.
- Step 4:** With an additional random number it is decided whether the dopant is rejected or not. This is necessary to meet the mean values represented by the local doping concentration.
- Step 5:** If the dopant is not rejected and is inside the segment the dopant is stored in the random dopant list with the coordinates and the corresponding grid point index.

If the local doping concentration is far below the maximum doping concentration most dopants are rejected. Thus the rejection technique is a highly inefficient method when large doping concentration differences occur within one segment.

6.3 Read-from-File Model

One way to obtain random discrete dopants is to read their positions from a file. It offers the possibility of importing discrete dopant positions from a previously done simulation or setting the position manually. In [33] deterministic single-atom doping via single-ion implantation have been investigated. To model deterministic single-atom doping the dopant position can be set manually in the input file using the *read-from-file* model.

An example input file is given in Appendix B. The dopant positions are given in a list of three dimensional points analogous to the grid points. The dopant material number is given as an attribute defined on this list and describes the dopant type, for example donors and acceptors. The segment number is an optional attribute and identifies the particular segment where the dopants should be added. If the segment number attribute is not specified in the input file the *point-in-segment-check* is used instead. After the determination of the nearest grid point using the nearest neighbor search of the *kd-tree* the dopants are added to the random dopant list.

6.4 Point-in-Segment Check

For non-rectangular segments it has to be checked whether a dopant lies inside or outside the segment. In this Section a *point-in-segment-check* is described.

We start at the point and follow a ray in an arbitrary direction counting all intersections with the segment surface. If the number of intersections is odd the point lies inside the segment and if it is even the point lies outside.

In MINIMOS-NT a segment surface is built up with triangles. For the ray direction the positive z -direction is used. In this situation the problem can be split into two conditions with reduced dimensionality.

Conditions

- The point has to be inside the projection of the surface triangle on the xy -plane.
- The normal vector of the surface triangle plane tending to the point must have a negative z -component.

All triangles that meet both requirements are counted and if and only if the value is odd the point is considered inside the segment.

The first condition is similar to the three dimensional problem reduced to two dimensions. We again start at the point and follow a ray in positive y direction and count the intersected lines. The point lies inside the triangle if and only if the value is odd. This test is generally applicable to two dimensional polygons and thus also valid for a simple triangle.

Pre-Test

Because it is not efficient to test all surface triangles pre-tests are included. The pre-test selects all triangles where the xy -coordinates of the point lies between the minimum and maximum xy -coordinates of the particular triangle. This is done with a four dimensional kd -tree using the minimum and maximum xy -coordinates of each triangle. With the *find-points-in-region* function of the kd -tree these triangles can be found. We search for all triangles with minimum coordinate in between of the minimum device coordinate and the point coordinate and the the maximum coordinate in between of the point coordinate and the maximum device coordinate.

To further reduce the computational effort the normal vector test is only done if the z -coordinate of the point lies in between the minimum and maximum z -coordinate of the triangle. If all z -coordinates are larger than the point the triangle counts or if all are lower it does not count.

Chapter 7

Modeling Discrete Dopants

Discrete dopants cannot be straightforwardly included in classical drift-diffusion simulation [34]. The solution obtained by Poisson's equation represents the long-range part of the Coulomb potential connected to the amount of band-bending [10, 35, 36]. The short-range part is including via scattering. In the "atomistic" situation, where the grid spacing is less than the mean dopant distance, the ionized discrete dopants are represented by a delta function like charge density when the dopant is simply added to the nearest grid point. Then both long-range and short-range parts are included and the particular solution depends on the grid spacing. Because the carrier concentrations resulting from Boltzmann or Fermi statistics follow the coulomb potential, the mobile carriers become sharply localized at the discrete dopants. Quantum mechanically, this strong localization is not possible because of the high ground state due to confinement and at room temperature most carriers are relieved from any bound states of the coulomb potential. Thus, when using the classical drift-diffusion model, dopants cannot be simply added to nearest grid points.

To include the long-range part only, the coulomb potential is explicitly split into a long-range and short-range part based on screening considerations, with the drawback of finding the right cut-off point. The other solution to this problem is a quantum mechanical correction to the classical drift-diffusion model [1]. One of these quantum-correction model is the density gradient model described in the first part of this thesis.

In this Chapter the modeling of discrete dopants is investigated. We start with the nearest grid point model, showing that it lacks appropriate discrete dopant screening when the classical drift-diffusion model is used. After continuing with the long-range model we turn towards screening of discrete dopants using the density gradient model.

7.1 Nearest-Grid-Point Model

Using finite volume discretization, each quantity of a grid point describes the average of the quantity in the corresponding Voronoi region and not the local value on the grid point itself. Thus all discrete dopants within a Voronoi region are assigned to its grid point by spreading the charges over the Voronoi region. In the generation models described in Chapter 6 the discrete dopants are stored in the random dopant list. The nearest grid point model walks through this list and adds each dopant to the nearest grid point. The doping concentration of the each grid point then reads

$$N_{D,A} = \frac{n_{D,A}}{V}, \quad (7.1)$$

where $n_{D,A}$ is the number of donors or acceptors in the corresponding voronoi region with the volume V .

Screening of Discrete Dopants using the classical Drift-Diffusion Model

Using the classical drift-diffusion model an unphysical situation occurs when a discrete dopant is assigned to the nearest grid point only. The carriers become localized at the dopants with a grid dependent amount of localization.

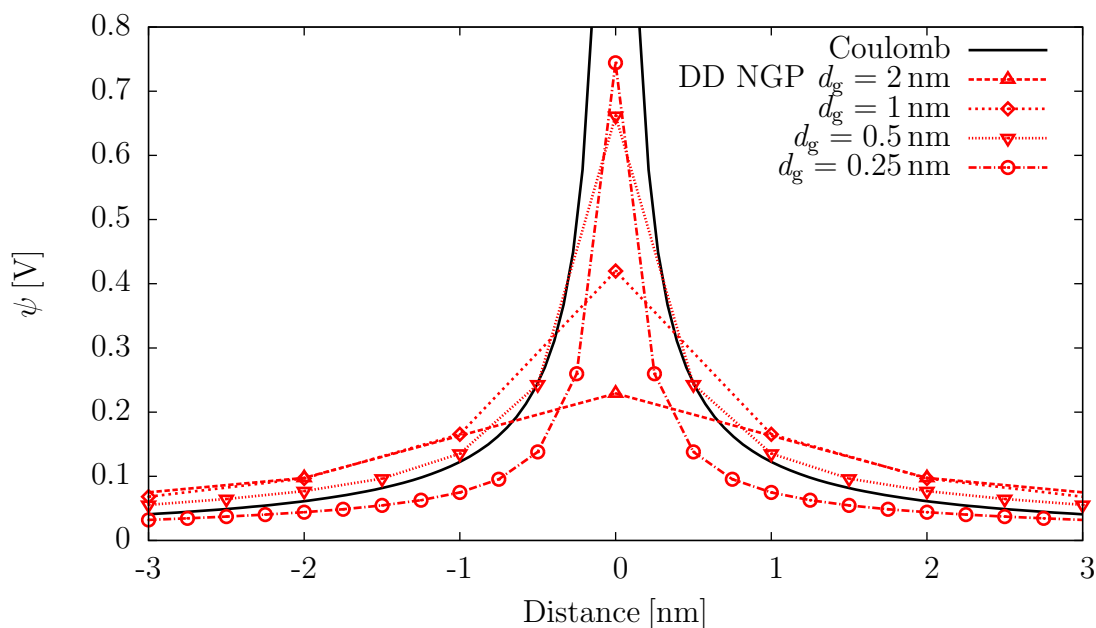


Figure 7.1: The electrostatic potential of a discrete donor assigned to a grid point in the middle of a $16 \times 8 \times 8$ nm silicon resistor using the classical drift-diffusion model. The high grid dependency is due to the quantity values representing the average over the Voronoi region. This leads to unphysically strong carrier localization due to Boltzmann or Fermi statistics.

To investigate the origin of this unphysical situation a single donor is assigned to a grid point in the middle of a $16 \times 8 \times 8$ nm silicon resistor. Figure 7.1 shows the electrostatic

potential obtained by the classical drift-diffusion model. At a first sight, the potential seems to be highly grid dependent. It has to be noted, that the quantity values describe the average within the Voronoi region and not the local values on the grid point itself. This leads to a unphysically high and grid dependent carrier localization around the discrete dopants due to Boltzmann or Fermi statistics.

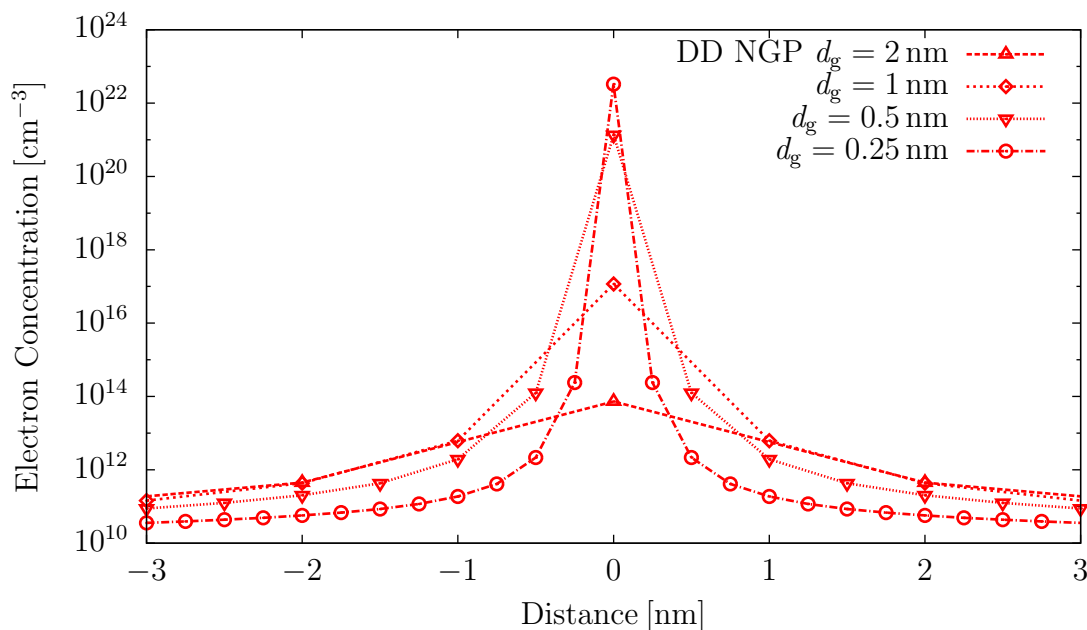


Figure 7.2: The electron concentration around a discrete donor assigned to a grid point in the middle of a $16 \times 8 \times 8$ nm silicon resistor using the classical drift-diffusion model. The electrons are strongly localized and the solution is highly grid dependent.

A closer analysis reveals that the problem is not the potential itself, but the fact that the carrier concentrations exponentially follow the potential. Figure 7.2 shows the electron concentration around the discrete dopant corresponding to the electrostatic potential in Figure 7.1. Even though the average potential values are correct, the corresponding carrier concentrations do not correspond to the average values due to the exponential dependency. Thus the solution is naturally grid dependent and does not represent the physical behavior.

From a physical point of view, the assignment of the dopants to the nearest grid point using the classical drift-diffusion model leads to strong carrier localization. In reality such a sharp localization of electrons or holes is not possible because of energy quantization due to confinement in the Coulomb potential. This quantization leads to a high ground state level and at room temperature the electrons or holes are relieved from the dopants. The dopants remain ionized and the mobile electrons or holes contribute to a current flow in the presence of an electric field. A naive classical treatment cannot reflect this behavior and is thus not sufficient, when discrete dopants are assigned to the nearest grid point.

7.2 Long-Range Model

In the previous Section it was pointed out that the assignment of the discrete dopants to the nearest grid point suffers from unphysically strong carrier localization and high grid dependency when the classical drift-diffusion model is used. In [10] this situation is explained to be due to the inclusion of the short-range part of the Coulomb potential, when a dopant is represented by a delta-function like dopant concentration. In the classical drift-diffusion model naturally only the long-range part is included via Poisson's equation representing the amount of band-bending. In the presence of discrete dopants, the charge density has to be explicitly split into a short-range and a long-range part. Figure (7.3) shows the long-range and short-range functions of a dopant located at the origin.

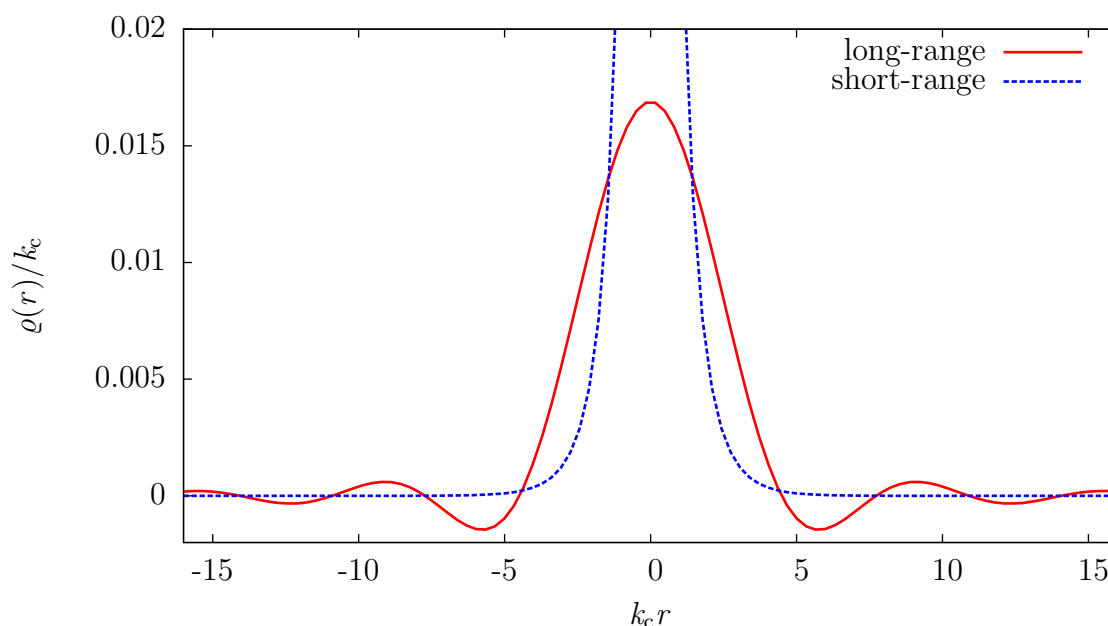


Figure 7.3: Long-range and short-range of a dopant located at the origin.

This splitting is done using the Fourier transform of the delta function and reads

$$\varrho_A(r) = -q\delta(r) = -\frac{q}{V} \sum_{k < k_c} \exp(ik \cdot r) - \frac{q}{V} \sum_{k > k_c} \exp(ik \cdot r) = \varrho_A^{\text{long}}(r) + \varrho_A^{\text{short}}(r), \quad (7.2)$$

with the long-range function

$$\varrho_A^{\text{long}}(r) = \frac{qk_c^3}{2\pi^2} \frac{\sin(k_c r) - (k_c r) \cos(k_c r)}{(k_c r)^3} \quad (7.3)$$

and the short-range function

$$\varrho_A^{\text{short}}(r) = \frac{qk_c^3}{4\pi} \frac{\exp(-k_c r)}{(k_c r)}. \quad (7.4)$$

The cut-off k_c is related to the screening length of the Conwell-Weisskopf model and is given by the half of the mean distance between the dopants [35]

$$\frac{1}{k_c} = \frac{N_A^{-1/3}}{\kappa}, \quad (7.5)$$

where N_A is the acceptor doping concentration and κ is a fitting parameter with a default value $\kappa = 2$. It has to be noted that for the long-range model a reference doping concentration is necessary to calculate the cut-off parameter k_c . This can lead to inconsistencies, when the long-range model is used in combination with the *read-from-file* model without defining a corresponding doping concentration in the input structure file.

Implementation

In semiconductor device simulation the grid spacing can vary over a decade or more within a single device. The implemented model is valid for both fine and coarse grids. Because of convergence problems due to negative values of the long-range function only the first positive part around the dopant is included. Figure 7.3 shows the long-range and short-range functions given by (7.3) and (7.4). In the following description we will denote the region around the dopant in which the long-range function is included as the long-range region. This long-range region has a radius of $r_{LR} = 4.4/k_c$. This radius is chosen to exclude all negative parts of the long-range function and includes the first positive part around the dopant. The doping concentration $N_{A,i}$ of the grid point P_i is obtained by integrating the long-range function over the whole Voronoi region divided by the volume of the Voronoi region. For a very fine grid $N_{A,i}$ can be calculated straightforwardly using the following approximation

$$N_{(D,A),i} = \varrho_{D,A}^{\text{long}}(\|r_i - r_{\text{dopant}}\|), \quad (7.6)$$

where $\|r_i - r_{\text{dopant}}\|$ denotes the distance between the grid point P_i and the dopant. For a coarse grid this approximation is not sufficient and thus the long-range function is calculated on a sub-grid within each Voronoi region averaging over all sub-grid points. In the extreme situation, when the long-range region includes only one or no grid-point, the long-range model converges to the nearest grid point model, if the whole long-range region lies inside the Voronoi region of the nearest grid point.

The implementation of the long-range model, which is valid for both fine and coarse grids can be described in a few steps.

Step 1: Calculate the cut-off parameter k_c from the local reference doping concentration of the nearest grid point.

Step 2: Find all points within the long-range region using the *kd*-tree.

Step 3: If one or no dopant is inside the long-range region, use the nearest point model.

Step 4: Iterate over all grid points in the long-range region.

If the Voronoi region is relatively large, calculate and average the long-range function on a sub-grid inside the Voronoi region and the long-range region to get a temporary doping concentration. If not, use the grid point only.

Step 5: Sum up the temporary doping concentration values of all grid point in the long-range region multiplied by the Voronoi region volume to obtain a normalization factor to one.

Step 6: Add the normalized temporary doping concentration values to the doping concentration.

Screening of Discrete Dopants

Figure 7.4 shows the electrostatic potential of a discrete donor located at the origin obtained by the classical drift-diffusion model and the long-range model. In contrast to the nearest point model, the solution is independent from the grid spacing. The electrostatic potential does not follow the Coulomb potential, because the long-range model only accounts for the amount of band-bending due to discrete dopants.

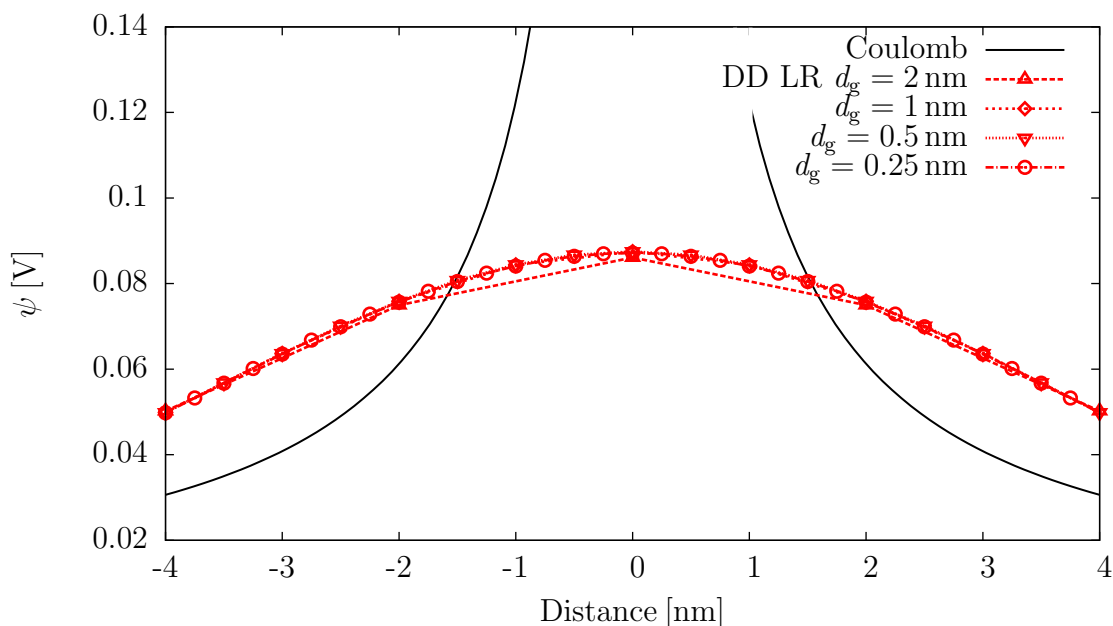


Figure 7.4: The electrostatic potential of a discrete donor assigned to a grid point in the middle of a $16 \times 8 \times 8$ nm silicon resistor using the classical drift-diffusion and the long-range model. For the long-range cut-off a reference doping $N_D = 10^{20} \text{ cm}^{-3}$ is used. The solution and thus the amount of localization is not dependent on the grid spacing d_g .

In Figure 7.5 (top) the electron concentration around the discrete dopant without unphysical charge trapping is shown. The explicit splitting of the Coulomb potential into a long-range and short-range part and the inclusion of the long-range part only leads to a low and grid independent amount of charge trapping. Thus the long-range model solves the previously described problems treating discrete dopants with the classical drift-diffusion model. But another problem appears, the proper choice of the cut-off parameter. Figure 7.5 (bottom) shows the electrostatic potential resulting from two donors located within a distance of $d = 4$ nm using three different fitting parameters κ . A larger κ leads to higher charge trapping. The reduction of the number of mobile charged carriers leads to a reduction of the current through a device using discrete dopants. Thus the macroscopic device behaviour depends on the cut-off parameter.

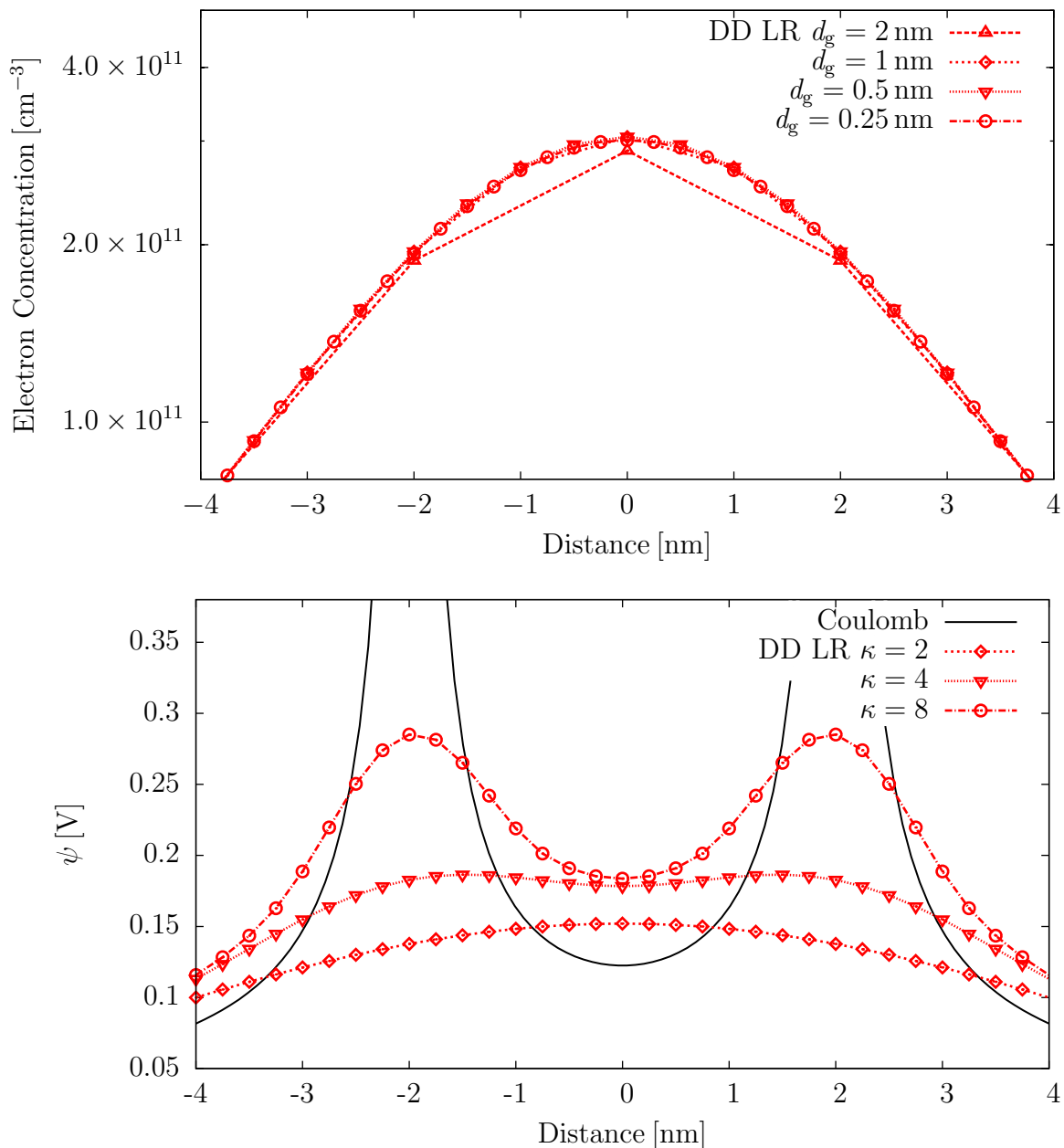


Figure 7.5: Top: The electron concentration of a discrete donor assigned to a grid point in the middle of a $16 \times 8 \times 8 \text{ nm}$ silicon resistor using the classical drift-diffusion and the long-range model. The long-range model prevents the classical drift-diffusion model from unphysical high charge localization and leads to a grid independent solution. Bottom: The electrostatic potential of two donors located at -2 nm and 2 nm for three different long-range fitting parameters. For the long-range cut-off a reference doping $N_D = 10^{20} \text{ cm}^{-3}$ is used. The electrons become more and more localized by the donors with the increase of the cut-off fitting parameter.

7.3 The Density Gradient Model

In [1] it has been pointed out that the density gradient model has the ability to correctly describe screening effects of discrete dopants. In the density gradient model an additional term is included in the classical drift-diffusion model to obtain a first-order quantum-correction, as described in the first part of this thesis.

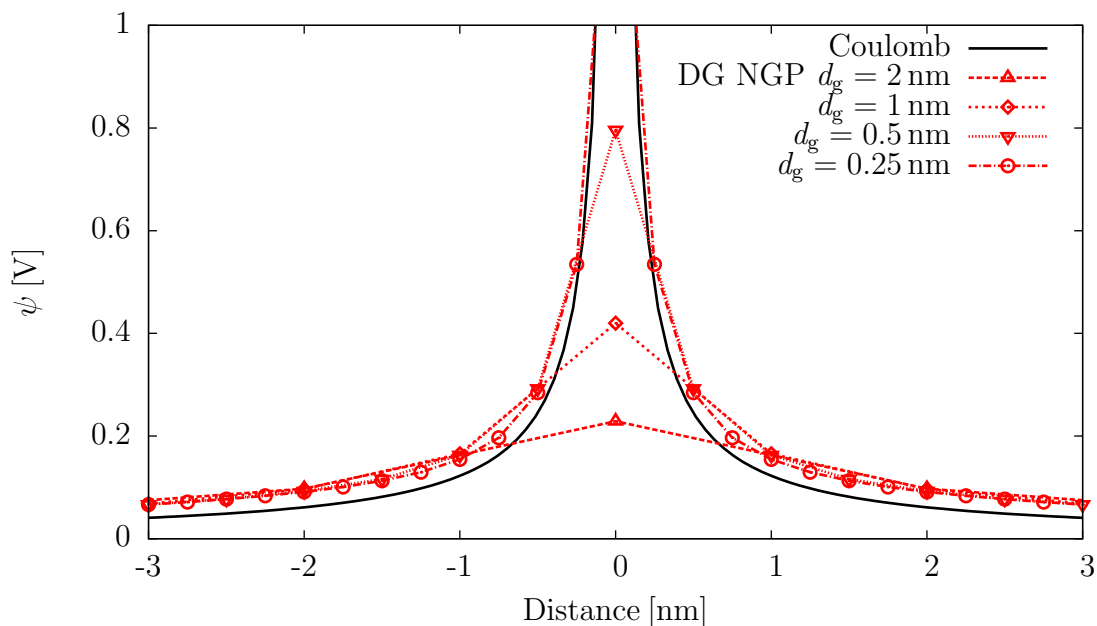


Figure 7.6: The electrostatic potential of a discrete donor assigned to a grid point in the middle of a $16 \times 8 \times 8$ nm silicon resistor using the quantum-corrected drift-diffusion model (density gradient). The electrostatic potential is grid dependent, but the dependency is compensated via the electron quantum-correction potential

The electrostatic potential obtained by the density gradient model and the nearest grid point model is shown in Figure 7.6. Again the electrostatic potential is grid dependent. However in contrast to the classical drift-diffusion model this does not lead to unphysical strong and grid dependent charge localization. In the density gradient model the carrier concentration exponentially depends both on the electrostatic potential and the quantum-correction potential. This situation is shown in Figure 7.7. The sum of both potentials remain grid independent and thus also the electron concentration for grid spacings d_g smaller than 1 nm. Also the unphysical sharp carrier localization compared to the classical drift-diffusion model is reduced. The remaining small amount of localization still has an effect on the macroscopic device behavior and is discussed in detail in the next Chapter.

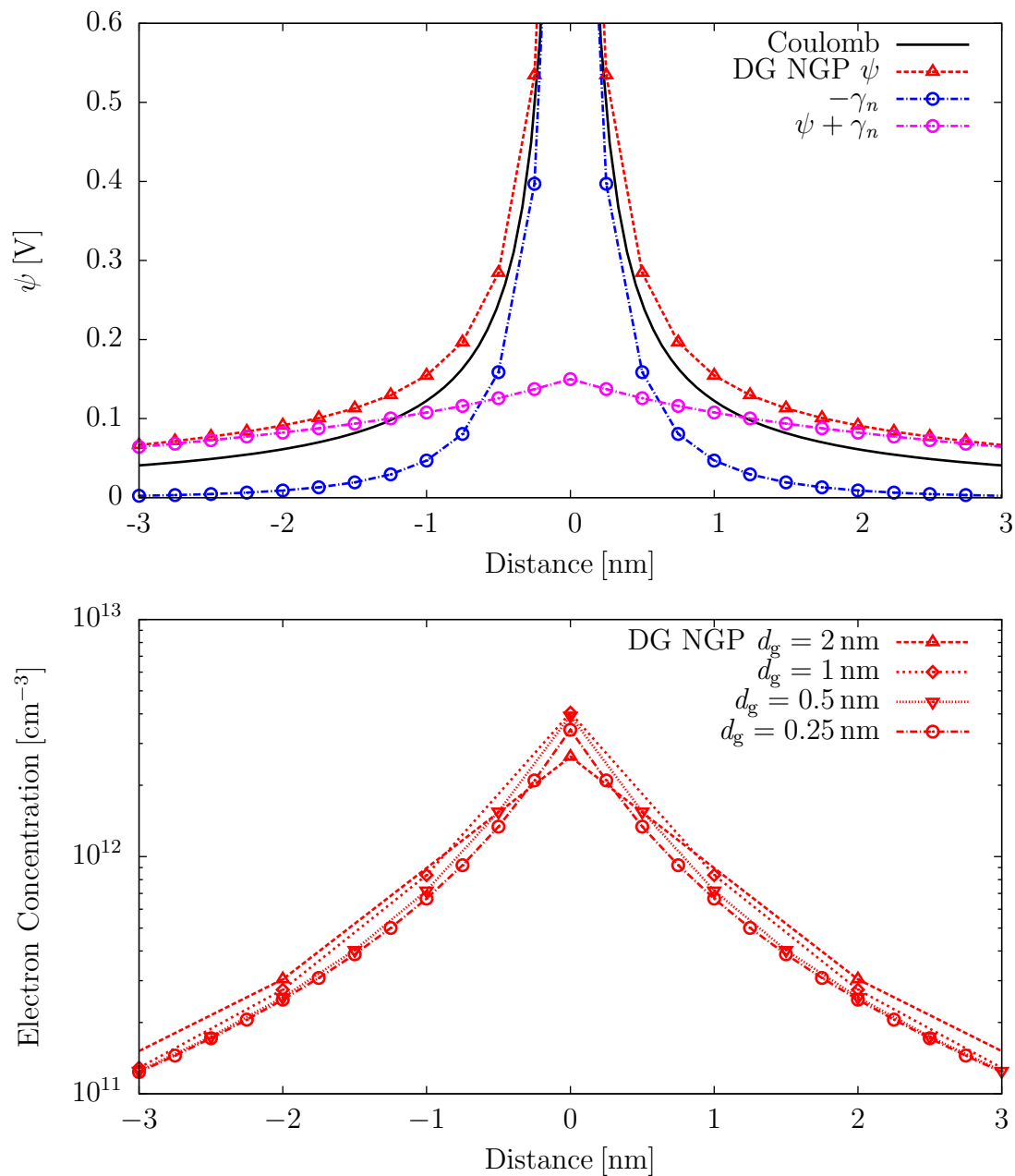


Figure 7.7: Top: The electrostatic potential ψ , the electron quantum-correction potential γ_n and the sum of both $\psi + \gamma_n$ obtained by a discrete donor assigned to a grid point in the middle of a $16 \times 8 \times 8$ nm silicon resistor using the density gradient model. Bottom: The electrostatic potential follows the Coulomb potential near the dopant and is corrected with the electron quantum-correction potential to a grid independent complete potential $\psi + \gamma_n$. The corresponding electron concentration exponentially depends on $\psi + \gamma_n$ and is thus also grid independent for grid spacings $d_g \leq 1$ nm, with a small amount of localization.

Chapter 8

Results

In this Chapter the effects of random dopant fluctuations are investigated. We first start with the influence of charge localization by discrete dopants on the device behavior. Using a simple silicon resistor with constant doping the correlation of carrier localization and the increase of resistance is pointed out. This increase of the resistance can be compensated via ionized impurity mobility and velocity saturation models. After that we turn towards the goal of this thesis, the simulation of random dopant fluctuations in a sub-nanometer MOS device.

8.1 Silicon Resistor with Constant Doping

The strong and grid dependent localization of the charged carriers around the dopants leads to a reduction of mobile carriers when the dopant is assigned to the nearest grid point and the classical drift-diffusion model is used. As a result the resistance increases and depends on the grid spacing. In Chapter 7 it was pointed out that this strong and highly grid dependent charge localization can be corrected by splitting the Coulomb potential into a long-range part and a short-range part and explicitly considering the long-range part only. A second solution is the use of the density gradient model presented in the first part of this thesis. In the density gradient model a quantum-correction term is added to the classical drift-diffusion model. With the electron and hole quantum-correction potentials the grid dependent amount of localization is compensated and the carrier concentrations become grid independent and far less localized when compared to the classical drift-diffusion model.

In this Section the correlation of charge trapping and increase of resistance is investigated. Figure 8.1 shows both the charge localization due to a single donor (top) and the IV-curve of a $16 \times 8 \times 8$ nm silicon resistor using the presented random dopant simulation models (bottom). For the random dopant generation the segment model is used with a constant total dopant number only choosing the random positions of the dopants. Also the positions of the individual dopants are independent from the grid spacing when the segment model is used. Thus effects due to total dopant number fluctuation and position fluctuation vanish when comparing the various random dopant simulation models.

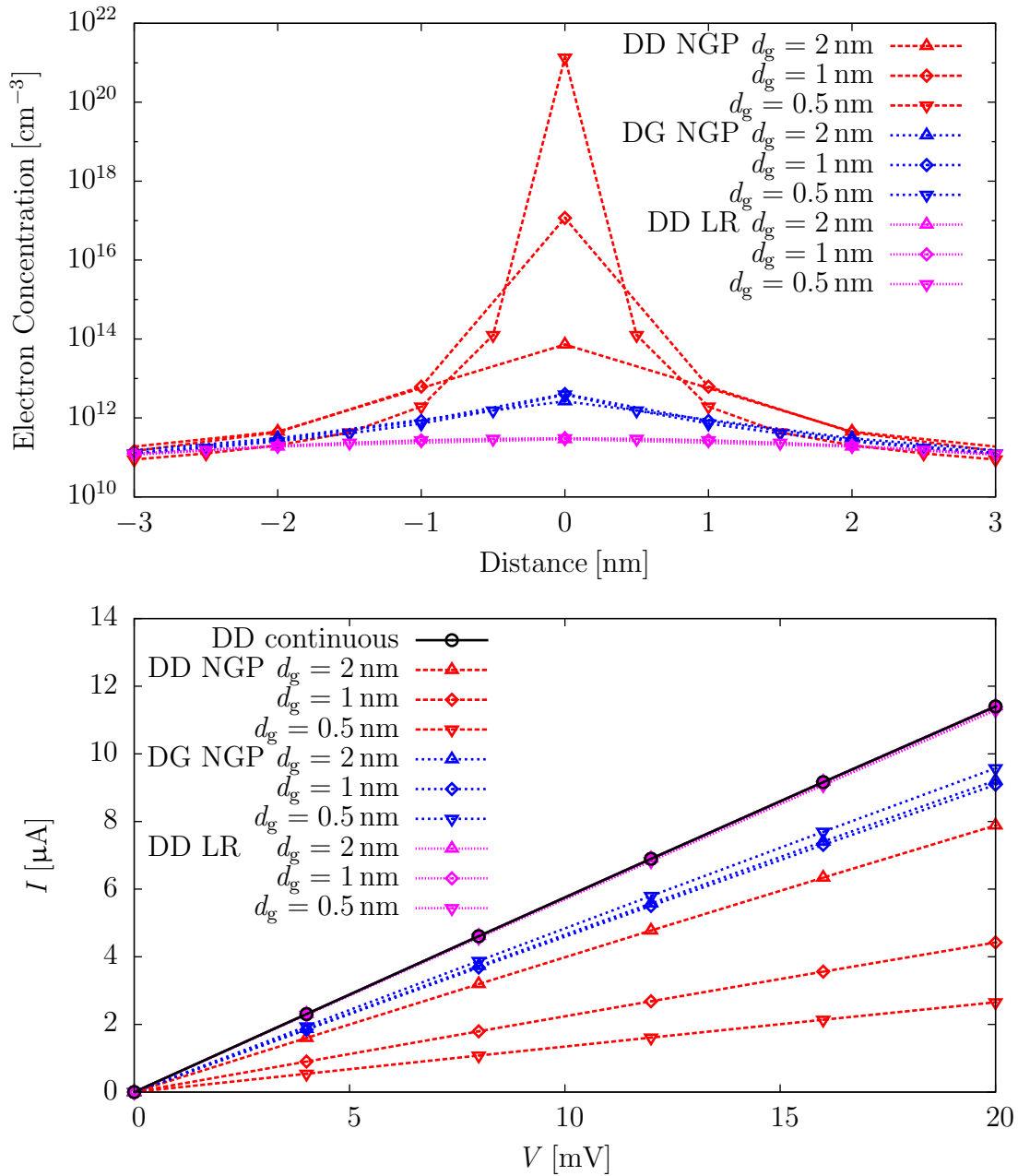


Figure 8.1: Top: The electron concentration obtained by a discrete donor assigned to a grid point in the middle of a $16 \times 8 \times 8$ nm silicon resistor. Bottom: The IV-curve of a $16 \times 8 \times 8$ nm silicon resistor with a donor doping $N_A = 10^{20} \text{ cm}^{-3}$. The generation of the random dopants is done with the segment model with random positions only. A strong localization of charged carriers around the donors leads to reduction of mobile carriers and thus to an increase in resistance. The long range model appropriately describes the resistance of this example. The density gradient model overestimates the resistance, but the solution is less grid dependent and can be corrected via recalibration of the mobility models.

The sharp and grid dependent localization of electrons around the donor correlates with the grid dependent increase of the resistance. With the explicit inclusion of the long-range part only the long-range model leads to a far lower and grid independent localization and thus also the resistance does not change significantly compared to the drift-diffusion simulation with continuous doping. The density gradient model leads to a almost grid independent but significant carrier localization. Thus the resistance obtained by the density gradient model is almost grid independent but higher compared to the continuous approach. This grid independent increase of resistance can be compensated via recalibration of the ionized impurity scattering parameters and is described in Section 8.2.

Quantum-Correction Potential Contact Boundary Condition

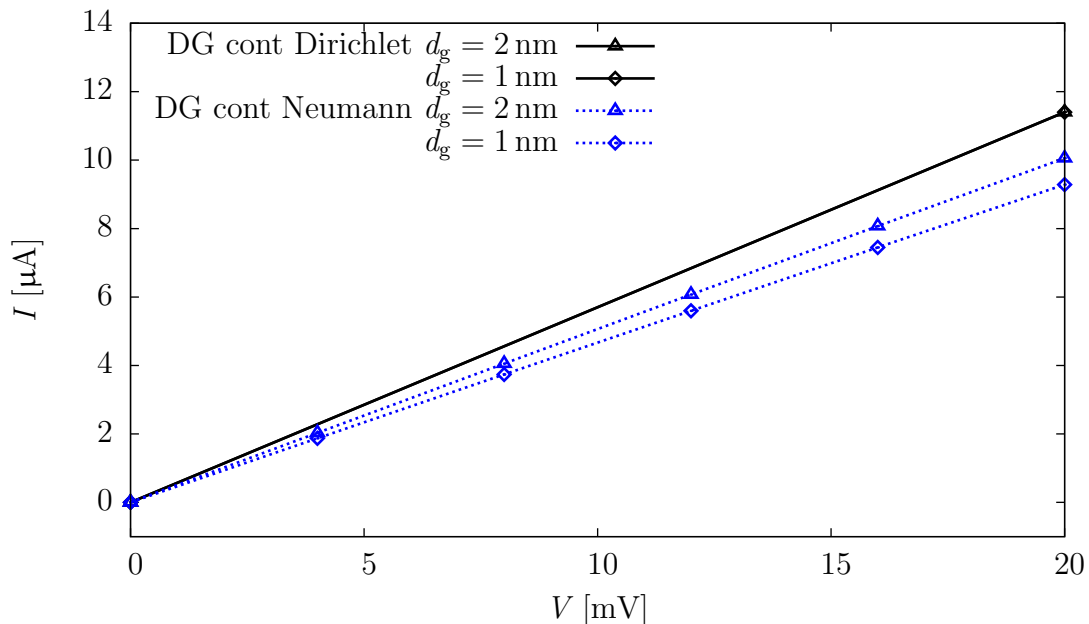


Figure 8.2: The IV-curve of a $16 \times 8 \times 8$ nm silicon resistor with a donor doping $N_A = 10^{20} \text{ cm}^{-3}$ for both Neumann and Dirichlet contact boundary conditions for the quantum-correction potentials. The zero-outflux Neumann boundary condition leads to a grid dependent current. With Dirichlet boundary conditions quantum effects near the contacts are neglected and the solutions only slightly depend on the grid spacing for both continuous and discrete doping.

In the previous example Dirichlet boundary conditions are used for the quantum-correction potentials at the semiconductor-contact interfaces. Figure 8.2 shows the IV-curves obtained by the density gradient model using implicit zero-outflux Neumann boundary conditions. These boundary conditions lead to a grid dependent resistance. Thus Dirichlet boundary conditions are used instead of setting the electron and hole quantum potentials to zero at the contact interface neglecting quantum effects near the contact interface.

8.2 Ionized Impurity Scattering

In the previous Section the correlation of carrier localization at the dopants and the increase of the resistance of a $16 \times 8 \times 8$ nm silicon resistor is investigated. The density gradient model leads to a grid independent but higher resistance. This increase of the resistance is due to the carrier localization at the dopants and thus the reduction of mobile carriers. Because the solution is grid independent it can be fitted to the simulation with continuous doping via recalibration of the ionized impurity scattering parameters. With the increase of the mobilities the current increases and the resistance decreases.

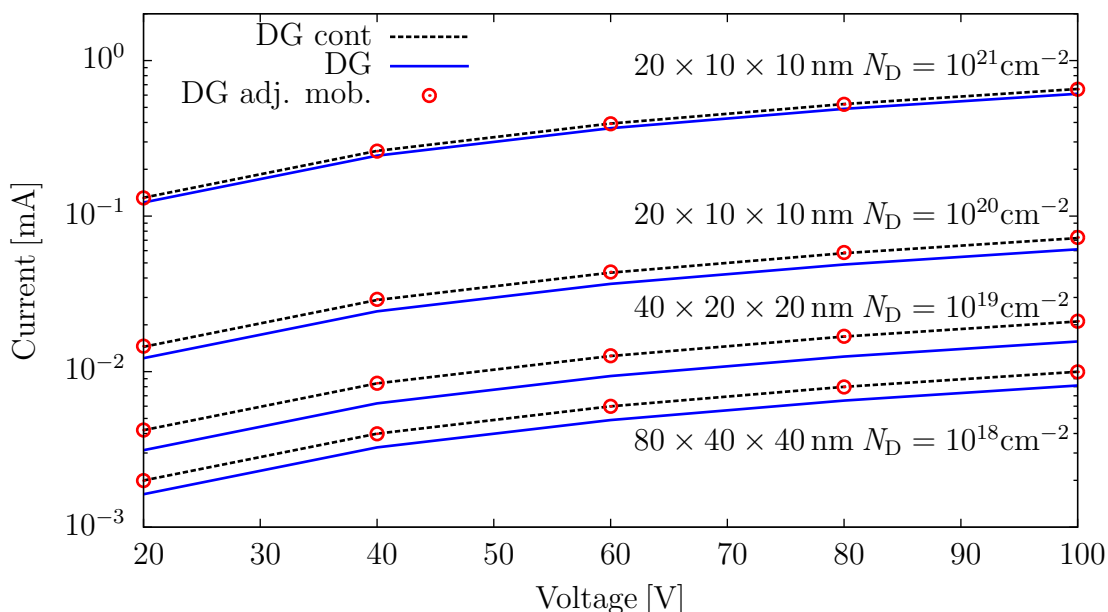


Figure 8.3: Various silicon resistor samples with different size and doping are simulated to adjust the ionized impurity scattering parameters. Larger resistors are used for the smaller dopings, as otherwise the number of dopants is too small and the statistical effects do not allow such fitting. The density gradient simulation with random discrete dopants (blue curve) can be fitted to the continuous simulations (black curve) by adjusting the ionized impurity scattering parameters (red circles).

To find the correct parameters for the mobility models, various devices with different doping are simulated. In particular we are interested in random dopant fluctuation effects in the channel of sub-nanometer MOS devices and we only consider the doping range $N_{A,D} = 10^{18} \text{ cm}^{-3}$ to 10^{21} cm^{-3} . Because the resistance of the various samples varies on a large scale, the IV-curves are plotted on a logarithmic scale in Figure 8.3 neglecting velocity saturation. The random dopant effects become more and more important, when the number of dopants decreases. Thus a larger resistor is used for the lower dopings. It has to be noted that even with these device dimensions the statistical variations due to the limited number of dopants cannot be neglected, but the results are far better than without adjusting the parameters. In the next Section we will observe that the statistical fluctuations are of order ten greater than the influence of the parameter adjustment.

8.3 NMOS Transistor with Random Dopant Fluctuations

Coming to the goal of this thesis, the influence of random discrete dopants and carrier confinement on device behavior of a sub-nanometer MOS transistor is investigated. In particular we are interested in threshold voltage lowering and fluctuation due to the total doping number and dopant position variations as reported in [5]. For the simulations a NMOS transistor with a channel length of 22 nm is used. The continuous doping profiles are depicted in Figure 6.1. From this reference device with continuous doping, macroscopically identical device samples with random discrete dopants are generated. The discrete dopants of one particular sample are shown in Figure 6.2.

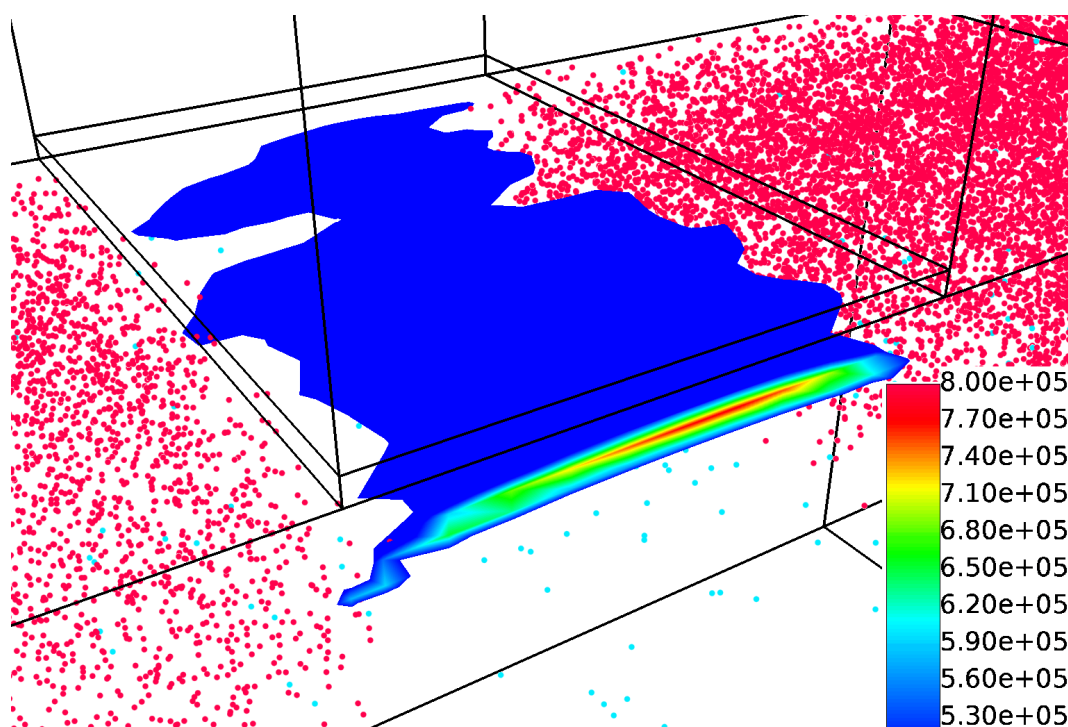


Figure 8.4: Electron current density in the channel of a 22 nm NMOS with random discrete dopants. The donors are visualized by red dots, the acceptors by light blue dots. The current follows the percolation paths around the acceptors.

The reduction of the threshold voltage is associated with percolation paths around the dopants due to an inhomogeneous potential distribution allowing an early turn [5]. Figure 8.4 shows the current density in the channel 2 nm under the oxide interface. The current flows around the discrete dopants through the individual percolation paths. The electrostatic potential and the electron concentration in the sample with random dopants are shown in Figure 8.5 and 8.6.

To further investigate the threshold voltage lowering and fluctuation 100 samples are simulated using the density gradient model, the long-range model with classical drift-diffusion and the long-range model with density gradient. In Figure 8.7 (top) the $I_{\text{Drain-}}$

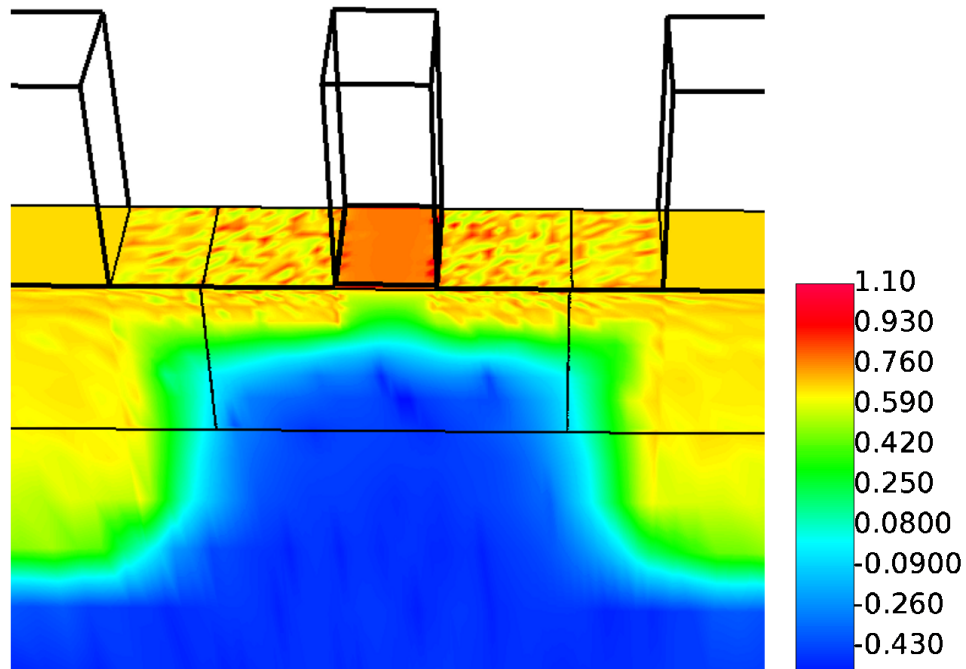


Figure 8.5: Electrostatic potential in a 22 nm NMOS with random discrete dopants.

V_{Gate} characteristics obtained by the density gradient model is shown. The continuous simulation represents the mean of the random dopant devices apart from the subthreshold slope, where the percolation paths due to the discrete dopants lead to a higher drain current. Figure 8.7 (bottom) shows the threshold voltage distribution of the random discrete dopant samples. The threshold voltages are calculated using the current condition $I_{\text{Drain}} = 1 \text{ nA}$ and is shifted from $V_{\text{th}} = 0.498 \text{ V}$ for the continuous doping to the mean threshold voltage $\langle V_{\text{th}} \rangle = 0.482 \text{ V}$ with standard deviation $\sigma_{V_{\text{th}}} = 37.57 \text{ mV}$. To first-order, the threshold voltage fluctuations are tested to be a normal distributed as reported in [5]. A KS-test is applied on the threshold voltage of the sample leading to a p-value of $p_{\text{val}} = 0.60029$

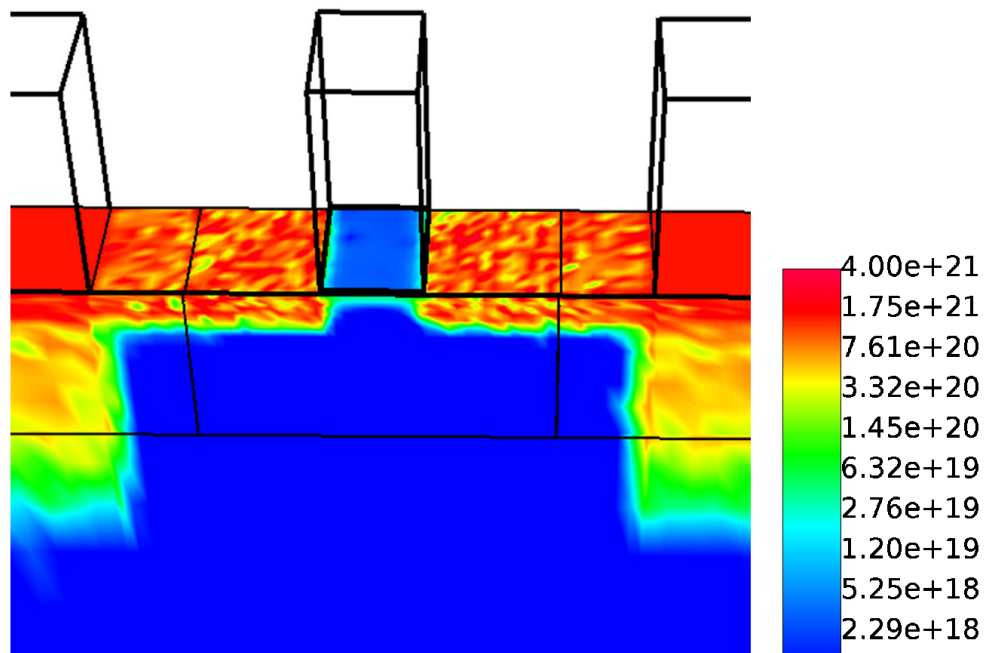


Figure 8.6: Electron concentration in a 22 nm NMOS with random discrete dopants.

The simulation results using the long-range model are depicted in Figure 8.8 and 8.9 for both classical drift-diffusion and density gradient. In the classical case threshold voltage is shifted from $V_{th} = 4.37$ V to $\langle V_{th} \rangle = 0.318$ V. The standard deviation of the assumed normal distribution is $\sigma_{V_{th}} = 28.62$ mV and the p-value evaluated with a KS-test is $p_{val} = 0.384$. In the quantum-corrected case the threshold voltage is shifted from $V_{th} = 0.498$ V to $\langle V_{th} \rangle = 0.323$ V, with a standard deviation $\sigma_{V_{th}} = 39.02$ mV. The p-value of the KS-test is $p_{val} = 0.749$. Treating random discrete dopants with the long-range model leads to a significant drain current increase no matter whether the drift-diffusion or density gradient model are used.

Simulation of Random Discrete Dopants with the Density Gradient Model

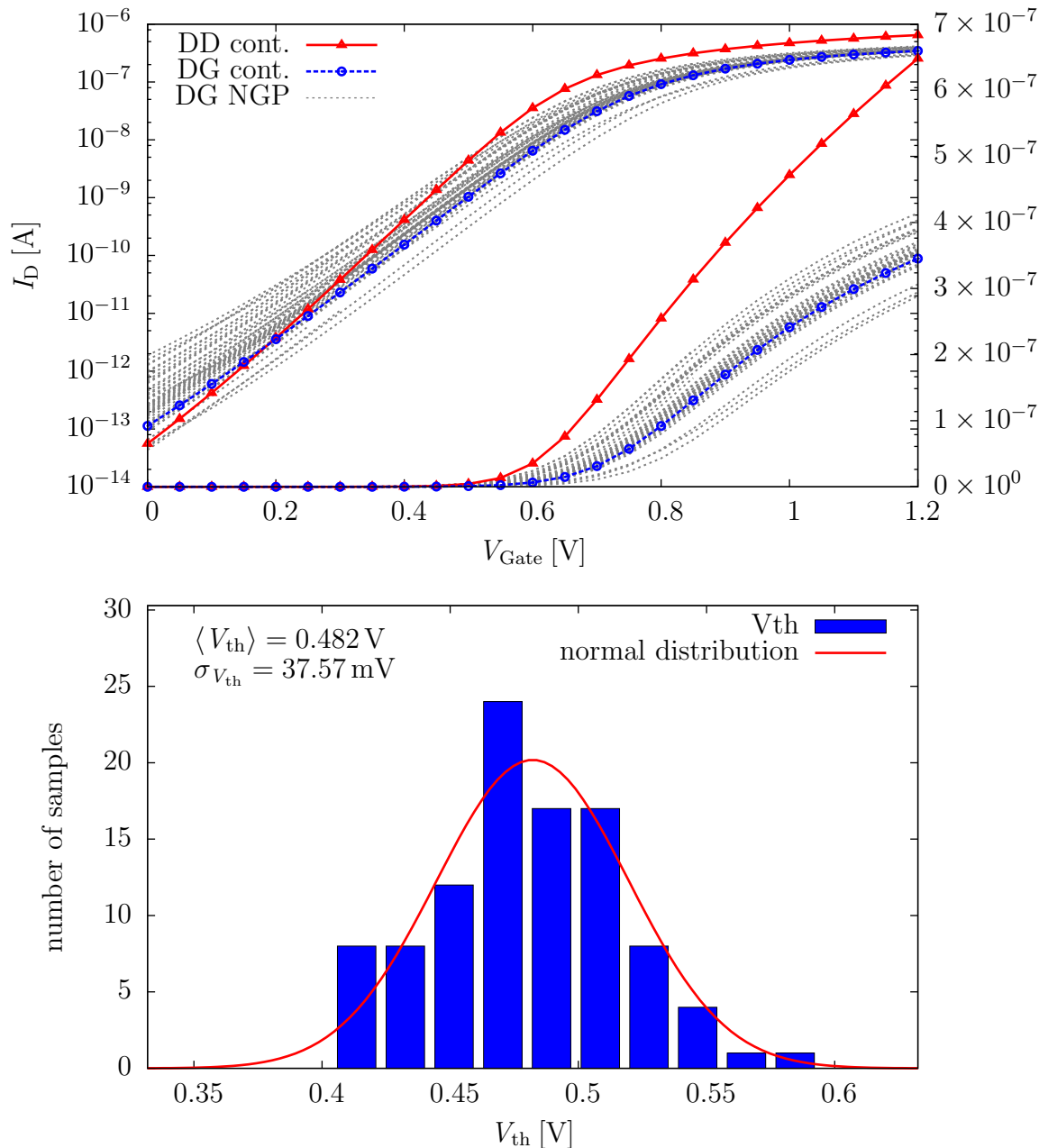


Figure 8.7: Top: I_D - V_G characteristics of a 22 nm NMOS with random discrete dopants assigned to the nearest grid point using the density gradient model. The drain voltage V_D is 1 mV. Bottom: Histogram of the threshold voltage with a mean value $\langle V_{th} \rangle = 0.482$ V and standard deviation $\sigma_{V_{th}} = 37.57$. The threshold voltage is defined via a current condition with $I_{D,th} = 1$ nA. Using the density gradient model, the random dopant curves only differ from continuous simulation in the subthreshold region.

Simulation of Random Discrete Dopants with the Long-Range and the Drift-Diffusion Model

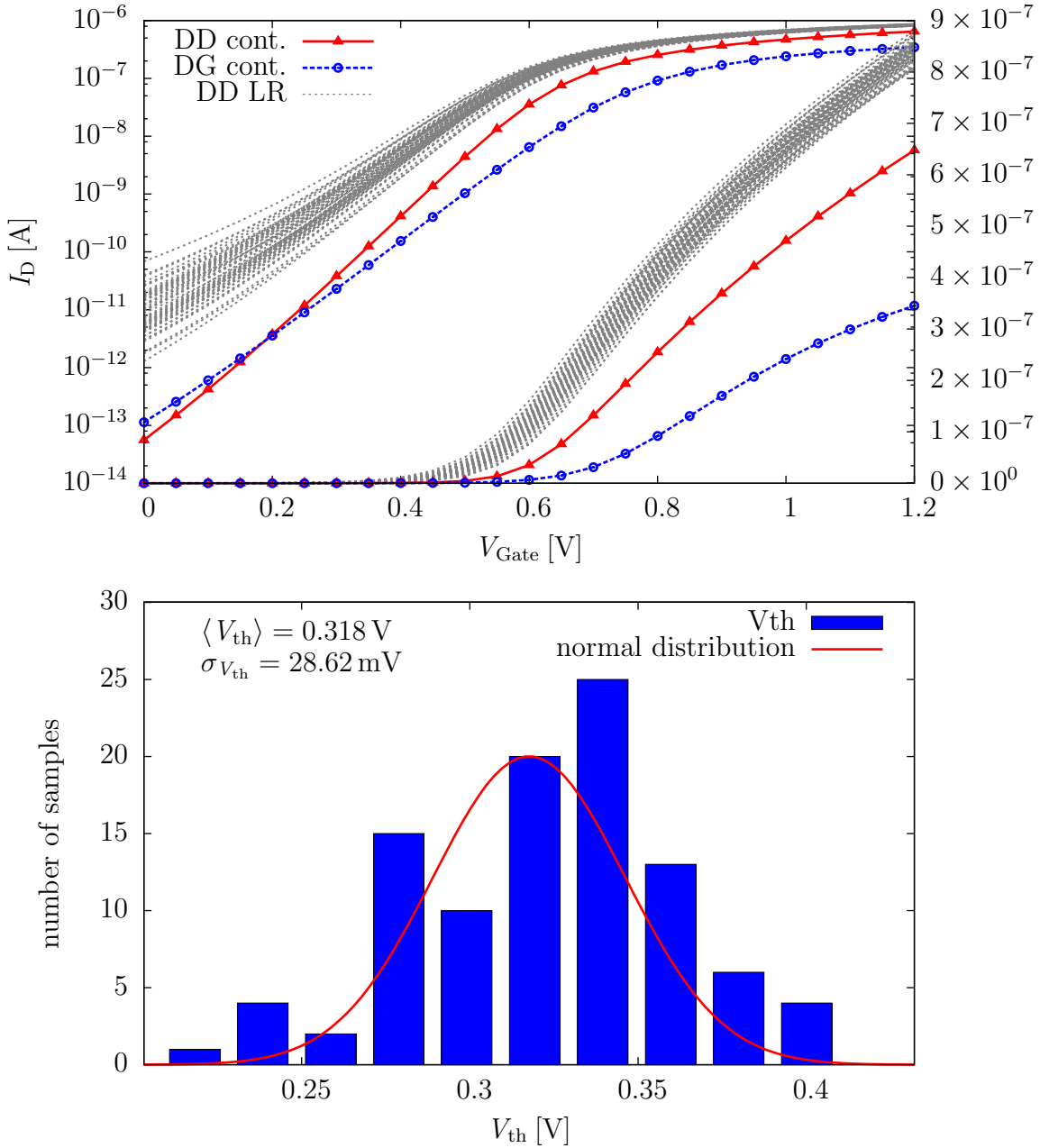


Figure 8.8: Top: I_D - V_G characteristics of a 22 nm NMOS with random discrete dopants using the long-range model and the drift-diffusion model. The drain voltage V_D is 1 mV. Bottom: Histogram of the threshold voltage with a mean value $\langle V_{th} \rangle = 0.318$ V and standard deviation $\sigma_{V_{th}} = 28.62$ mV. The threshold voltage is defined via a current condition with $I_{D,th} = 1$ nA. The long-range model leads to a significant current increase and thus to a threshold lowering from of $V_{th} - \langle V_{th} \rangle = 0.437$ V $-$ 0.318 V = 0.119 V.

Simulation of Random Discrete Dopants with the Long-Range and the Density Gradient Model

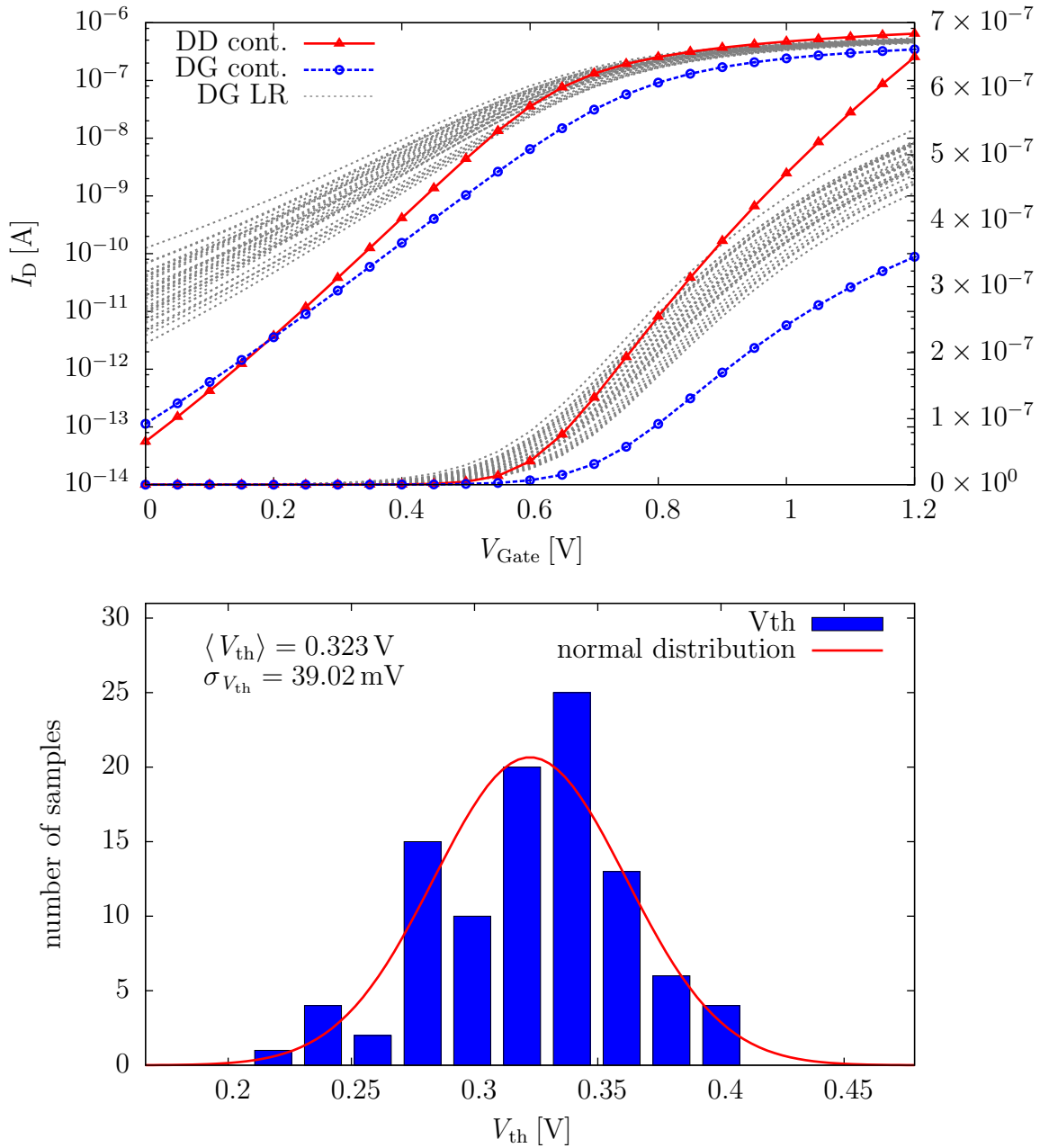


Figure 8.9: Top: I_D - V_G characteristics of a 22 nm NMOS with random discrete dopants using the long-range model and the density gradient model. The drain voltage V_D is 1 mV. Bottom: Histogram of the threshold voltage with a mean value $\langle V_{th} \rangle = 0.323$ V and standard deviation $\sigma_{V_{th}} = 39.02$ mV. The threshold voltage is defined via a current condition with $I_{D,th} = 1$ nA. Again the long-range model leads to a significant current increase compared to the continuous simulation and a reduction of the threshold voltage of $V_{th} - \langle V_{th} \rangle = 0.498$ V - 0.323 V = 0.176 V.

Chapter 9

Summary

Due to the rapid scaling of MOS transistors into the deep sub-nanometer regime, statistical variations and quantum effects cannot be neglected any more. In the course of this thesis the target simulator MINIMOS-NT was extended to an “atomistic” quantum-corrected drift-diffusion simulator. The simulator is now capable to further investigate reliability issues, such as charge trapping by random position defects in the oxide, connected with random dopant fluctuations. Instead of a single device with continuous doping, a statistical representative ensemble of macroscopically identical devices has to be simulated to extract the distribution of characteristic device parameters.

It has been confirmed that discrete dopants cannot straightforwardly be included in classical drift-diffusion simulation, because of grid dependent unphysically large charge localization. With the implementation of the long-range model [10] and the density gradient quantum-correction model [14] it has been confirmed that the amount of grid dependent unphysically large charge localization by discrete dopant can be reduced with both models. When using the density gradient model, the remaining small amount of charge trapping was compensated by adjusting the parameters of the ionized impurity scattering model.

In a 3D “atomistic” quantum-corrected simulation study centered around a 22 nm NMOS, the threshold voltage lowering and fluctuation has been shown as reported in [5]. The inhomogeneous electrostatic potential distribution due to random discrete dopants in the channel affects the drain current and thus the threshold voltage. The carriers flow around the discrete dopants through percolation path resulting in an dopant-position-dependent threshold voltage lowering.

Apart from random dopant fluctuation effects, the inclusion of confinement via the density gradient quantum-correction model partially compensates threshold voltage lowering. It has been shown that fitting the density gradient quantum-corrected drift-diffusion model to the Vienna Schrödinger Poisson solver VSP2 for the whole CV-curve of the 1D MOS capacitor example was possible using Cauchy boundary conditions on the semiconductor/oxide interface by applying the quantum-correction on the semiconductor only. The numerical benefit of advanced discretization schemes for the quantum-correction equation as reported in [17, 20, 27, 28] have not been confirmed. No convergence problems of

the density gradient model have been observed using the simplest model with continuous doping.

A few issues have been only investigated partially in this work due to the lack of time. The inclusion of discrete dopants in the implemented quantum-corrected drift-diffusion model leads to a decrease of numerical stability. Further investigations on stability issues can be done by exploring the numerical benefit due to the advanced source term approximation described in [17]. The assignment of discrete dopants to the nearest grid point discards the exact dopant position. This can be solved by spreading the dopant to several adjacent grid points as has been done in [34]. This very likely also has an numerical impact. Apart from numerical implementation issues, the physical matter of subthreshold slope difference between the drift-diffusion and density gradient simulations and the current magnification when using the long-range model have to be further investigated in future work.

Appendix A

Minimos-NT Model Defaults

MINIMOS-NT can be controlled in a very flexible way via the C++-like input-deck. In the input-deck file the individual models can be switched on and reconfigured on each segment. In the following sections the default input-deck structure of the newly implemented models is given.

A.1 Density Gradient Model

Segment Models

```
Device : ~DeviceDefaults
{
  Phys
  {
    Semiconductor
    {
      densityGradient = "Simple";
      DensityGradient
      {
        DGAll
        {
          Lambda_e = 0.22;
          Lambda_h = 0.2;
        }
        Simple : DGAll;
        Full : DGAll;
        Nonlinear : DGAll;
      }
    }
  }
}
```

Interface Models

```
Device : ~DeviceDefaults
{
  Phys
  {
    Semiconductor_Insulator
    {
      densityGradientInterface = "Cauchy";
      densityGradientInterface_gts = "options=None|Cauchy|Neumann|Dirichlet";

      DensityGradientInterface
      {
        Cauchy
        {
          Alpha_e = -3.7e-4;
          Beta_e = -7.5e-5;
          Alpha_h = -4.2e-4;
          Beta_h = 8.3e-5;
        }
        Neumann
        {
          Flux_e = 0.005;
          Flux_h = -0.005;
        }
        Dirichlet
        {
          Bnd_e = -2;
          Bnd_h = 2;
        }
      }
    }
  }
}
```

A.2 Random Dopant Generation Model

```
Device : ~DeviceDefaults
{
  Phys
  {
    Semiconductor
    {
      randomDopantGeneration = "None";
      randomDopantGeneration_gts = "options=None|ReadIn|Segment|Grid";
      RandomDopantGeneration
      {
        ReadIn
      }
    }
  }
}
```

```
    {
      InputListName = "RdopPoints";
      InputFileName = "RdopPoints.gts";
    }
  Segment
  {
    RandomSeed = 1;
    ConstDopantNumber = no;
  }
  Grid
  {
    RandomSeed = 1;
  }
}
}
```

A.3 Random Dopant Simulation Models

Device : ~DeviceDefaults

```
{
  Phys
  {
    Semiconductor
    {
      randomDopantSimulation = "None";
      RandomDopantSimulation
      {
        Nearest;

        Long_Range
        {
          fitting = 2; //cut-off fitting parameter
        }
      }
    }
  }
}
```

Appendix B

Random Discrete Dopant Input File

The *read-from-file* model offers the possibility of importing discrete dopant positions from a previously done simulation or setting the position manually via an input file. An example of the gts-file structure used is give here.

```
GTSDDevice("structure.ipds")
{
  Info
  {
    Version("1.0") // Fileformat version
    Dimension(3) // Device dimension
    Comment
    {
    }
  }
}

// ***** BEGIN OF Point-List *****
Pointlists(1)
{
  Pointlist("RdopPoints", "um", 3, 1)
  {
    Intern {
      0.008, 0.004, 0.004, // coordinates of the dopant
    } // End of Block 'Intern'
  } // End of Pointlist 'RdopPoints'
} // End of Pointlists

// ***** BEGIN OF Polyhedron-List *****
Polyhedronlists(0)
{
} // Polyhedronlists

// ***** BEGIN OF SEGMENT-List *****
Segments(0)
{
} // End of Segments
```

```
// ***** END OF SEGMENT-List *****

// ***** BEGIN OF ATTRIBUTE-List *****
Attributes(1)
{
  Attribute("DopingMaterialNumber", 1) {
    Ref(Pointlist, "RdopPoints")
    Type(double)
    Dimension(1)
    Datalocation(Point)
    Unit("-")
    Intern {
      1.00000, // 0.0 for acceptor, 1.0 for donors
    } // End of Block 'Intern'
  } // End of Attribute 'DopingMaterialNumber'
} // End of Attributes
// ***** END OF ATTRIBUTE-List *****
}
```


Bibliography

- [1] A. Asenov, A. Brown, G. Roy, B. Cheng, C. Alexander, C. Riddet, U. Kovac, A. Martinez, N. Seoane, and S. Roy, "Simulation of statistical variability in nano-CMOS transistors using drift-diffusion, Monte Carlo and non-equilibrium Green's function techniques," *J. of Comput. Electron.*, vol. 8, pp. 349–373, 2009.
- [2] S. Toriyama, D. Hagishima, K. Matsuzawa, and N. Sano, "Device simulation of random dopant effects in ultra-small MOSFETs based on advanced physical models," *SISPAD*, pp. 111–114, 2006.
- [3] I. Corporation, "Intel 22nm 3-d tri-gate transistor technology." <http://newsroom.intel.com/docs/DOC-2032>, Last accessed on 2011-6-7.
- [4] C. L. Alexander, G. Roy, and A. Asenov, "Random impurity scattering induced variability in conventional nano-scaled MOSFETs: Ab initio impurity scattering Monte Carlo simulation study," *IEDM*, pp. 1–4, 2006.
- [5] A. Asenov, "Random dopant induced threshold voltage lowering and fluctuations in sub-0.1 μm MOSFET's: A 3-d "atomistic" simulation study," *IEEE Trans. on Electron Devices*.
- [6] S. Selberherr, *Analysis and Simulation of Semiconductor Devices*. Springer-Verlag Wien-New York, 1984.
- [7] C. B. Duke, "Tunneling in solids," *Academic Press*, 1969.
- [8] R. Tsu and L. Esaki, "Tunneling in a finite superlattice," *Appl.Phys.Lett.*, vol. 22, no. 11, pp. 562–564, 1973.
- [9] M. J. van Dort, P. H. Woerlee, and A. J. Walker, "A simple model for quantisation effects in heavily-doped silicon MOSFETs at inversion conditions," *Solid-State Electron*, vol. 37, pp. 411–414, 1994.
- [10] 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- μm Si-MOSFETs," *IEDM Tech. Dig.*, pp. 275–278, 2000.
- [11] A. Wettstein, *Quantum Effects in MOS Devices*. Hartung-Gorre Verlag, Konstanz, 2000.

- [12] M. Ancona and G. Iafrate, “Quantum correction to the equation of state of an electron gas in semiconductor,” *Phys. Rev. B*, vol. 39, pp. 9536–9540, 1989.
- [13] Institut für Mikroelektronik, Technische Universität Wien, MINIMOS-NT USER’S GUIDE, 2002.
- [14] A. Wettstein, A. Schenk, and W. Fichtner, “Finite-difference schemes for the density gradient equations,” *IEEE Trans. Electron Devices*, vol. 48, pp. 279–284, 2001.
- [15] O. Baumgartner, M. Karner, and H. Kosina, “Modeling of high-k-metal-gate-stacks using the non-equilibrium Green’s function formalism,” *SISPAD*, pp. 353–356, 2008.
- [16] A. Prechtel, “Vorlesung über elektrodynamik.” Institut für Grundlagen und Theorie der Elektrotechnik, Technische Universität Wien, 2005.
- [17] S. Odanaka, “Multidimensional discretization of the stationary quantum drift–diffusion model in ultra-small MOSFET structures,” *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 23, pp. 837–842, 2004.
- [18] A. Asenov, G. Slavcheva, A. Brown, 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 Trans. on Electron Devices*, vol. 48, pp. 722–729, 2001.
- [19] D. Bohm, “A suggested interpretation of the quantum theory in terms of “hidden” variables,” *Phys. Rev.*, vol. 85, pp. 166–179, 1952.
- [20] M. G. Ancona and B. Biegel, “Nonlinear discretization scheme for density-gradient equations,” *Proc. SISPAD*, pp. 196–199, 2000.
- [21] S. Wagner, T. Grasser, C. Fischer, and S. Selberherr, “An advanced equation assembly module,” *Engineering with Computers*, vol. 21, pp. 151–163, 2005.
- [22] S. Wagner, *Small-Signal Device and Circuit Simulation*. PhD thesis, Institut für Mikroelektronik, Technische Universität Wien, 2005.
- [23] C.-D. Munz and T. Westermann, *Numerische Behandlung gewöhnlicher und partieller Differenzialgleichungen*. Springer-Verlag Berlin Heidelberg, 2009.
- [24] R. Klima, *Three-Dimensional Mesh Generation for Device and Process Simulation*. PhD thesis, Institut für Mikroelektronik, Technische Universität Wien, 2004.
- [25] J. Cervenka, *Small-Signal Device and Circuit Simulation*. PhD thesis, Institut für Mikroelektronik, Technische Universität Wien, 2005.
- [26] D. Scharfetter and H. Gummel, “Large-signal analysis of a silicon read diode oscillator,” *IEEE Trans. Electron Devices*, vol. 16, pp. 64–77, 1969.
- [27] S. Odanaka, “A high-resolution method for quantum confinement transport simulations in MOSFETs,” *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 26, pp. 80–85, 2007.

-
- [28] A. Wettstein, O. Penzin, and E. Lyumkis, "Integration of the density gradient model into a general purpose device simulator," *VLSI Design*, vol. 15, pp. 751–759, 2002.
- [29] M. G. Ancona, "Finite-difference schemes for the density gradient equations," *J. Comput. Electron.*, vol. 1, pp. 435–443, 2002.
- [30] X. W. T.-W. Tang and Y. Li, "Discretization scheme for the density-gradient equation and effect of boundary conditions," *J. Comput. Electron.*, vol. 1, pp. 389–, 2002.
- [31] M. K. et al., "A multi-purpose Schrödinger-Poisson solver for TCAD applications," *J. of Comput. Electron.*, vol. 6, pp. 179–182, 2007.
- [32] J. Friedman, J. Bentley, and R. Finkel, "An algorithm for finding best matches in logarithmic expected time," *ACM Trans. on Mathematical Software*, vol. 3, pp. 209–226, 1977.
- [33] T. S. M. Hori, Y. Ono, K. Taira, A. Komatsubara, T. Tanii, T. Endoh, and I. Ohdomari, "Reliable single atom doping and discrete dopant effects on transistor performance," *IEDM*, pp. 592–595, 2010.
- [34] G. Roy, A. Brown, A. Asenov, and S. Roy, "Quantum aspects of resolving discrete charges in atomistic device simulation," *J. of Comput. Electron.*, pp. 323–327, 2003.
- [35] N. Sano and M. Tomizawa, "Random dopant model for three-dimensional drift-diffusion simulations in metal-oxide-semiconductor field-effect-transistors," *App. Phys. Lett.*, vol. 79, pp. 2267–2269, 2001.
- [36] N. Sano, A. Hiroki, and K. Matsuzawa, "Device modeling and simulations toward sub-10 nm semiconductor devices," *IEEE Trans. on Nanotechnology*, vol. 1, pp. 63–71, 2002.