Band Structure Effects and Quantum Transport

Diss. ETH No.

Band Structure Effects and Quantum Transport

A dissertation submitted to
ETH ZURICH

for the degree of
Dr. sc. ETH Zürich

presented by
ANIELLO ESPOSITO
Dipl. Phys. ETH
born August 06, 1979
citizen of Switzerland

accepted on the recommendation of
Prof. Dr. Wolfgang Fichtner, examiner
Prof. Dr. Giorgio Baccarani, co-examiner

2010

Acknowledgment

I wish to thank Prof. Wolfgang Fichtner for the opportunity to work in a well-equipped and comfortable environment at the Integrated Systems Laboratory during several years. For many helpful discussions and advices as well as for carefully reading my scientific contributions, including this thesis, a special thank goes to Prof. Andreas Schenk. I am also grateful to Prof. Giorgio Baccarani for having accepted to act as co-examiner during my PhD defense and for several helpful comments.

I thank the members of my group and companions since the beginning of my occupation at the IIS, Martin Frey and Simon Brugger, for countless helpful advices and discussions as well as for the moral support. I am particularly indebted to Mathieu Luisier for sharing his profound knowledge of the topic and for several insightful cooperations. The helpfulness of the scientific staff notably contributed to create a friendly atmosphere. My gratitude goes to Denis Dolgos, Philipp Kreuter, Jan Kupec, Hektor Meier, Vincent Peikert, Beat Sahli, Luca Sponton, Sebastian Steiger, Ratko Veprek, and Kilian Vollenweider.

Finally, I thank the technical staff of the IIS, particularly Christine Haller for the handling of administrative issues and Christoph Wicki for the indispensable computer support. Thanks to the staff of the supercomputing center CSCS in Manno TI for the technical support.

This work has been carried out in the context of the projects EU-IST-216171 (NANOSIL) and SNF 20021-109393 (NEQUATTRO) with the additional support of Fujitsu.

Abstract

Topic of this thesis is the development and extension of three-dimensional quantum ballistic transport simulators for the modeling of nanowire and planar field effect transistors (FETs) at the nanometer scale, as well as the investigation of band structure effects by various atomistic methods in order to improve the effective mass approximation (EMA) being at the base of the present simulators.
The accurate description of strongly confined nanostructures, with typical dimensions below 5 nm, requires approaches which go beyond the EMA. The empirical tight-binding method (ETB) is widely appreciated since the related computational burden is notably smaller compared to other approaches such as the empirical pseudopotential method (EPM). On the other hand, the EPM yields a more physical picture of the charge density and the related computational burden is still small compared to fully ab-initio approaches. Detailed comparisons between the EMA, ETB, and EPM are carried out for various nanostructures. For this purpose, and for a future use within a quantum transport simulator, a fully scalable band structure calculator for bulk, nanowires, and wells based on the EPM is provided. A popular nonparabolicity (NP) model for the improvement of the EMA is reconsidered, i.e. adapted, for the case of nanowires and quantum wells. The parametrization of the present NP models is accomplished by means of the ETB.
The scattering matrix algorithm (SMA) for the description of quantum transport within the Landauer-Büttiker framework as well as related numerical issues are thoroughly reviewed. For the description of nanowires with arbitrary cross sectional shapes, the finite element method (FEM) is implemented within the SMA. The parallelization of time consuming routines is accomplished by means of OpenMP and the MPI for the use on large scale compute clusters. Extensive calculations in order to investigate the impact of band structure effects on transfer characteristics of planar and nanowire FETs of various shapes are carried out. For this purpose, the NP models are implemented by means of a spectral method. Furthermore, detailed comparisons with a full-band tight-binding quantum transport simulator are carried out in order to validate the NP models.
Finally, inelastic scattering processes and NP are taken simultaneously into account for the simulation of nanowire FETs.

Zusammenfassung

In dieser Dissertation werden dreidimensionale ballistische Transportsimulatoren auf quantenmechanischer Basis entwickelt und erweitert. Sie dienen der Modellierung von Nanowire- und planaren Feld-Effekt-Transistoren (FETs) im Längenbereich von wenigen Nanometern, sowie der Untersuchung von Bandstruktur-Effekten mit verschiedenen Methoden. Letztere verbessern die gängige Effektivmassen-Approximation (EMA), die auch den vorliegenden Simulatoren zu Grunde liegt.
Die genaue Beschreibung der elektronischen Struktur von Bauelementen mit typischen Abmessungen im Bereich von 5 nm erfordert Methoden, die über die EMA hinausgehen. Hier ist die empirische Tight-Binding-Methode (ETB) sehr populär, da der mit ihr verbundene Rechenaufwand erheblich kleiner ist als bei anderen Verfahren, wie z.B. der empirischen Pseudo-Potential-Methode (EPM). Allerdings liefert die EPM eine physikalisch genauere Beschreibung der Ladungsträgerdichte, und der Rechenaufwand bleibt im Vergleich zu ab-initio-Methoden immer noch gering. Detaillierte Vergleiche zwischen EMA, ETB und EPM werden anhand verschiedener Nanostrukturen durchgeführt. Dazu wurde ein vollständig parallelisiertes Bandstruktur-Programm für Bulk, Wires und Wells basierend auf der EPM implementiert, das zukünftig in Transportsimulatoren eingesetzt werden kann. Ein weit verbreitetes Nichtparabolizitätsmodell (NP) zur Verbesserung der EMA wird auf Nanowires und Wells angepasst, wobei die Parametrisierung mithilfe der ETB-Methode erfolgt.
Der Streumatrix-Algorithmus (SMA) zur Beschreibung des Quantentransports im Rahmen des Landauer-Büttiker-Formalismus, sowie damit verbundene numerische Probleme werden tiefer analysiert. Für die Modellierung von Nanowires mit beliebigen Querschnittsflächen wird die Finite-Elemente-Methode im SMA benutzt. Die Parallelisierung von zeitaufwändigen Routinen wird mittels OpenMP und MPI für die Verwendung auf Grossrechnern bewerkstelligt. Umfangreiche Berechnungen von Transfer-Kennlinien verschiedener Nanowire- und planarer FETs verdeutlichen den Einfluss der Bandstruktur-Effekte. Dazu wurden die entwickelten NP-Modelle unter Benutzung einer spektralen Methode implementiert. Detaillierte Vergleiche mit einem atomistischen Tight-Binding-Transportsimulator ermöglichen die Validierung der NP-Modelle.
Schliesslich werden gleichzeitig NP-Effekte und inelastische Streuprozesse bei der Simulation von Nanowire FETs berücksichtigt.

Contents

Chapter 1

Introduction

Integrated circuits (ICs) are present in almost all electronic equipment used in the domestic or industrial field. Compared to the ICs back in the 1960's comprising around 50 components, nowadays high-end-processors such as graphic processing units (GPUs) contain up to 2.1×10^9 transitors on a surface of less than 6cm^2. This astonishing integration density is the result of long lasted research on the field of miniaturization giving rise to indespensable methods such as photo-lithography. A concurrent trend to improve the performance, beside miniaturization, consists in an increment of the microprocessor area or an optimization of the available resources by altering the circuit design. However, the stringent limits for the chip area often exclude these two alternatives thus highlighting the importance of miniaturization.

In virtue of the ITRS 2005 [1], sub-10-nm channel-length FETs will be manufactured in the year 2016 for the HP22 technology node while 5-nm channel lengths will be required for the HP14 node in year 2020. Silicon-on-insulator (SOI) transistors with comparable dimensions [2] are hard to fabricate and are known to suffer from severe short-channel effects due to a reduced electrostatic control by the gate contact. Gate leakage currents caused by the reduced thickness of the oxide layer further aggravate the circumstances. Nanoelectronic research is partially concerned with the development of novel device architectures to overcome these shortcomings and thus to exploit the limit of CMOS technology. Promising materials and device components such as graphene or carbon nanotubes represent also hot topics in this field beside emerging approaches such as spintronics.

A superior electrostatic control can be straightforwardly achieved by adding further gate contacts to a device such as in the case of triple-gate or gate-all-around nanowire field effect transistors (FETs). Nanowires are attractive not only from an electrostatic point of view [3], but also due to their capability to act both as FET and wire connector. Realizations of nanowires consisting of various materials such as Si[4], GaAs[5], or Ge[6] with different cross sectional shapes and channel orientations have been reported in the literature as well as nanowire FETs with a triangular [7], rectangular [8], or cylindrical [9] nanowire as channel.

Technology computer aided design (TCAD) can be used to support the fabrication of such novel devices by means of preliminary simulations to investigate their features and performance limits. The use of quantum mechanics is mandatory for the description of these devices since the typical dimensions are of the order of the de Broglie wave length. The effective mass approximation (EMA) is a widely used approach within nowadays quantum transport simulators since the implementation is rather simple and computation times are kept within reasonable limits. Many-body effects due to Coulomb interaction are typically accounted for by the Hartree approximation. The finite difference (FD) method on a tensorial grid is a popular choice for the

discretization of the Schrödinger and Poisson equation. Particularily in the case of nanowires with complex surfaces, the finite element method (FEM) provides more flexibility than the FD. However, the EMA is intended for situations, where the variation of an external perturbation is small compared to the lattice periodicity of the considered material. This prerequisite is clearly violated for nanowires with small (typically ≤ 5 nm) diameters. In this regime, fully atomistic quantum transport simulators are more accurate and corresponding results notably differ from the ones obtained by the EMA. The latter shortcomings of the EMA are referred to as band structure effects. Atomistic approaches to quantum transport range from empirical tight-binding (ETB) methods to fully ab-initio appraches. However, the considerable computational burden caused by these methods is the main reason why the EMA remains attractive. Nonparabolicity (NP) models gained a lot of attention since they aim at improving the EMA while keeping a comparable simulation efficiency. In this work, TCAD tools are developed (extended) in order to include suitable NP models for a more accurate description of aggressively scaled nanowire and planar FETs. The focus in this work is laid on band structure effects instead of scattering phenomenons without the intention to underrate the latter effect.

This work is organized as follows. In chapter 2 some important simplifications for the treatment of the inhomogeneous electron gas are briefly reviewed such as the Born Oppenheimer and the Hartree approximation for the treatment of the electronic system. The Hartree approximation is used within the quantum transport frameworks employed in this work in order to partially account for many-body effects.

Chapter 3 is concerned with the band structure calculation of bulk materials and nanostructures by means of different approaches. The ETB and the empirical pseudopotential method (EPM) are described and comparisons between the ETB and EPM are carried out for the case of nanowires and quantum wells. The EMA is revised and results are compared to atomistic data. A widely used NP model being able to capture several band structure effects is considered and appropriately modified for the case of nanowires and quantum wells. This modification is crucial for the implementation of NP whithin the present quantum transport framework.

Topic of chapter 4 is the elaboration of algorithmic, numerical, and implementational details concerning the present quantum transport simulator based on the EMA and the Landauer-Büttiker theory. The FEM for the solution of the Poisson and Schrödinger equation within the scattering matrix algorithm (SMA) is thouroughly described. In a corresponding appendix, some notes on the SMA for arbitrarily oriented nanowire FETs are given. A spectral method is employed in order to include NP within the SMA. The chapter ends with an investigation of the impact of band structute effects on transfer characteristics of nanowire and planar FETs.

The parametrization of the present NP models is accomplished by means of band structures derived from tight-binding models as described in chapter 5. Transfer characteristcs of nanowire FETs obtained by means of the NP model are compared to results from a TB-NEGF simulator. Finally, the combined effects of NP and scattering are investigated in chapter 6 after a short survey of NEGF for quantum transport.

Beside the notable improvements provided for the SIMNAD software such as OpenMP parallelization of time-consuming routines and the inclusion state-of-the-art libraries for linear and eigenvalue problems, the TCAD tools provided in this work comprise a bulk, well, and wire EPM band structure calculator as well as a fully MPI parallelized quantum transport simulator based on the FEM.

Chapter 2

Inhomogeneous Electron Gas

2.1 Introduction

The behavior of charge carriers in semiconductor devices reaching the nanometer scale is best described by means of quantum mechanics. Particularly in strongly confined systems, the Schrödinger equation has to account for many-particle effects as the Coulomb repulsion between electrons becomes non-negligible. A gas consisting of N electrons of mass m_e interacting by means of Coulomb repulsion and being subjected to a background potential $V(\vec{r})$ is referred to as the *inhomogeneous electron gas* (IEG). The Hamiltonian of the IEG reads

$$\mathbf{H}^{\mathrm{el}} = \sum_{j=1}^{N} \left[-\frac{\hbar^2}{2m_e} \Delta_j + V(\vec{r}_j) \right] + \frac{e^2}{8\pi\epsilon_0} \sum_{\substack{i,j \\ i \neq j}}^{1\ldots N} \frac{1}{|\vec{r}_i - \vec{r}_j|}, \tag{2.1}$$

where the coordinate $\vec{r}_j = (x_j, y_j, z_j)$ describes the j-th electron and $\Delta_j = (\partial^2/\partial x_j^2) + (\partial^2/\partial y_j^2) + (\partial^2/\partial z_j^2)$. If $V(\vec{r}_j)$ does not depend on $\vec{r}_j$ the system is referred to as the *homogeneous* electron gas. The Hamiltonian (2.1) can be modified to include the spin degree of freedom , and therefore, the spin-orbit coupling as well as any further relativistic correction which is not considered in this work. Three approaches for the solution of the IEG problem are given by the Hartree method, Hartree-Fock method, and density functional theory [10, 11], respectively. Generally, the electron gas is an integral part of a system consisting of many atoms whose positively charged cores generate the background potential $V(\vec{r})$. In this case, the ability to formulate an Hamiltonian for the electronic system alone, and thus treating the ionic cores as if they were frozen, is provided by the Born-Oppenheimer approximation (BOA). This simplification is of crucial importance not only for electronic systems in semiconductors and shall be discussed thoroughly.

First, the full quantum many-body problem describing electrons and atomic cores is introduced. Subsequently, the BOA is described and a simplified solution for the resulting equations is illustrated. Finally, the Hartree approximation for solving the IEG problem [1] will be addressed.

[1] The Hartree approximation is widely used within nowadays quantum transport simulators including the framework described in this work.

2.2 The Born-Oppenheimer Approximation

2.2.1 Full Many-Body Problem

A system consisting of M atomic cores [2] of mass M_j located at the positions $\{\vec{R}_j\}_{j=1,2,\ldots,M}$ and N electrons located at $\{\vec{r}_i\}_{i=1,2,\ldots,N}$ is now considered. For simplicity, the abbreviations

$$\begin{aligned} \mathbf{X} &= (\vec{R}_1, \vec{R}_2, \ldots, \vec{R}_M) \\ \mathbf{x} &= (\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N) \end{aligned}$$

for cores and electrons, respectively, are employed. The full Hamiltonian reads

$$\mathbf{H}(\mathbf{x}, \mathbf{X}) = \mathbf{H}^{\mathrm{el}}(\mathbf{x}, \mathbf{X}) + \mathbf{T}^{\mathrm{ion}}(\mathbf{X}) \tag{2.2}$$

with

$$\mathbf{H}^{\mathrm{el}}(\mathbf{x}, \mathbf{X}) = \mathbf{T}^{\mathrm{el}}(\mathbf{x}) + \mathbf{V}^{\mathrm{el-ion}}(\mathbf{x}, \mathbf{X}) + \mathbf{V}^{\mathrm{el-el}}(\mathbf{x}) + \mathbf{V}^{\mathrm{ion-ion}}(\mathbf{X}). \tag{2.3}$$

The operators

$$\begin{aligned} \mathbf{T}^{\mathrm{ion}}(\mathbf{X}) &= \sum_{j=1}^{M} -\frac{\hbar^2}{2M_j} \frac{\partial^2}{\partial \vec{R}_j^2} \\ \mathbf{T}^{\mathrm{el}}(\mathbf{X}) &= \sum_{i=1}^{N} -\frac{\hbar^2}{2m_e} \frac{\partial^2}{\partial \vec{r}_i^2} \end{aligned}$$

denote the kinetic energy of the electrons and cores respectively. The term

$$\mathbf{V}^{\mathrm{ion-ion}}(\mathbf{X}) = \frac{e^2}{8\pi\epsilon_0} \sum_{\substack{i,j \\ i \neq j}}^{1\ldots M} \frac{Z_i Z_j}{|\vec{R}_i - \vec{R}_j|} \tag{2.4}$$

is the potential energy of the ions with Z_i being the number of protons of the i-th ion and

$$\mathbf{V}^{\mathrm{el-el}}(\mathbf{x}) = \frac{e^2}{8\pi\epsilon_0} \sum_{\substack{i,j \\ i \neq j}}^{1\ldots N} \frac{1}{|\vec{r}_i - \vec{r}_j|} \tag{2.5}$$

is the potential energy of the electrons. Finally, the interaction between the electrons and the ions is described by

$$\mathbf{V}^{\mathrm{el-ion}}(\mathbf{x}, \mathbf{X}) = \sum_{i=1}^{N} \sum_{j=1}^{M} \Phi_j^{\mathrm{ion}}(\vec{r}_i), \tag{2.6}$$

where

$$\Phi_j^{\mathrm{ion}}(\vec{r}) = -\frac{e^2}{4\pi\epsilon_0} \frac{Z_j}{|\vec{r} - \vec{R}_j|} \tag{2.7}$$

is the Coulomb potential of the j-th ion. The Schrödinger equation in the product Hilbert space $\mathrm{H} = \mathcal{H}^{\mathrm{el}} \otimes \mathcal{H}^{\mathrm{ion}}$ reads

$$\begin{aligned} \mathbf{H}(\mathbf{x}, \mathbf{X})\Psi(\mathbf{x}, \mathbf{X}) &= \left[\mathbf{H}^{\mathrm{el}}(\mathbf{x}, \mathbf{X}) + \mathbf{T}^{\mathrm{ion}}(\mathbf{X})\right] \Psi(\mathbf{x}, \mathbf{X}) \\ &= E\Psi(\mathbf{x}, \mathbf{X}), \end{aligned} \tag{2.8}$$

where $\mathcal{H}^{\mathrm{el}}$ is the Hilbert space relative to the configuration space of the electrons $\mathbf{x}$ and $\mathcal{H}^{\mathrm{ion}}$ relative to $\mathbf{X}$.

[2] The terms core and ion will be used interchangeably.

2.2.2 Approximation

The wave function $\Psi(\mathbf{x}, \mathbf{X})$ from Eq. (2.8) can be written as

$$\Psi(\mathbf{x}, \mathbf{X}) = \sum_{\nu} \phi_\nu(\mathbf{x}, \mathbf{X})\chi_\nu(\mathbf{X}), \tag{2.9}$$

with $\{\phi_\nu(\mathbf{x}, \mathbf{X})\}_\nu$ being the eigensolutions of the electronic problem

$$\mathbf{H}^{\mathrm{el}}(\mathbf{x}, \mathbf{X})\phi_\nu(\mathbf{x}, \mathbf{X}) = E_\nu^{\mathrm{el}}(\mathbf{X})\phi_\nu(\mathbf{x}, \mathbf{X}) \tag{2.10}$$

for a fixed $\mathbf{X}$. Equation (2.10) corresponds to the IEG problem (2.1) with an appropriate $V(\vec{r})$. The ansatz (2.9) is used in Eq. (2.8). Based on the assumption that the electron mass is much smaller than any of the core masses, the mixing between different ν in Eq. (2.9) due to $\mathbf{T}^{\mathrm{ion}}$ can be neglected [10, 11] and Eq. (2.8) simplifies to

$$\left[\mathbf{T}^{\mathrm{ion}}(\mathbf{X}) + E_\mu^{\mathrm{el}}(\mathbf{X})\right] \chi_\mu(\mathbf{X}) = E\chi_\mu(\mathbf{X}) \tag{2.11}$$

for the different ν. For instance, an eigensolution of Eq. (2.8) is given by $\Psi(\mathbf{x}, \mathbf{X}) = \phi_\nu(\mathbf{x}, \mathbf{X})\chi_\nu(\mathbf{X})$. The $E_\nu^{\mathrm{el}}(\mathbf{X})$ acts as a potential energy in the vibrational problem (2.11) including the effect of $V^{\mathrm{ion-ion}}(\mathbf{X})$. The subsequent solution of the electronic (2.10) and vibrational problem (2.11) is actually referred to as the BOA or *adiabatic* approximation. For a discussion on non-adiabatic phenomena, i.e. effects which can not be reproduced within the BOA, such as the *electron-phonon* interaction, see Ref. [10].

2.2.3 Solution

In general, where no analytic dependence of E_μ^{el} on $\mathbf{X}$ can be found, a solution of Eqs. (2.11) and (2.10) can be accomplished by numerically solving the electronic problem for a finite set of $\{\mathbf{X}\}$ and use the discretized E_μ^{el} for the solution of the ionic problem. Except for very simple cases such as the H_2 molecule it is obvious that the sampling of $E_\mu^{\mathrm{el}}(\mathbf{X})$ represents a non-trivial task. A more simplified scheme is based on a classical description of the atomic cores. Given an initial configuration $\mathbf{X}^{\mathrm{init}}$, the forces

$$-F(\mathbf{X}_0) \equiv \left.\frac{\partial E_\nu^{\mathrm{el}}(\mathbf{X})}{\partial \mathbf{X}}\right|_{\mathbf{X}=\mathbf{X}_0} \tag{2.12}$$

are used to propagate the cores during a small time step. The iterative application of this procedure, as summarized in Fig. (2.1), leads to a quantum mechanical description of molecular dynamics (MD), where the motion of the cores is usually described by the ground state energy $E_g^{\mathrm{el}}(\mathbf{X})$ of the electronic system. This represents an improvement compared to the purely classical motion of electrons and atomic cores. At some point the cores are supposed to rest at the positions $\mathbf{X}_0$ defined by

$$F(\mathbf{X}_0) = 0. \tag{2.13}$$

The calculation of the atomic forces without knowing the dependence of E_ν^{el} on $\mathbf{X}$ can be accomplished by the Hellmann-Feynman theorem

$$-F(\mathbf{X}) = \frac{\partial E_\nu^{\mathrm{el}}(\mathbf{X})}{\partial \mathbf{X}} = \langle \phi_\nu, \mathbf{X}| \frac{\partial H^{\mathrm{el}}}{\partial \mathbf{X}} |\phi_\nu, \mathbf{X}\rangle, \tag{2.14}$$

where $\langle \mathbf{x}|\phi_\nu, \mathbf{X}\rangle = \phi_\nu(\mathbf{x}, \mathbf{X})$. Of course, modern MD simulators are much more involved than the procedure scheched above, especially when a finite temperature is taken into account, and

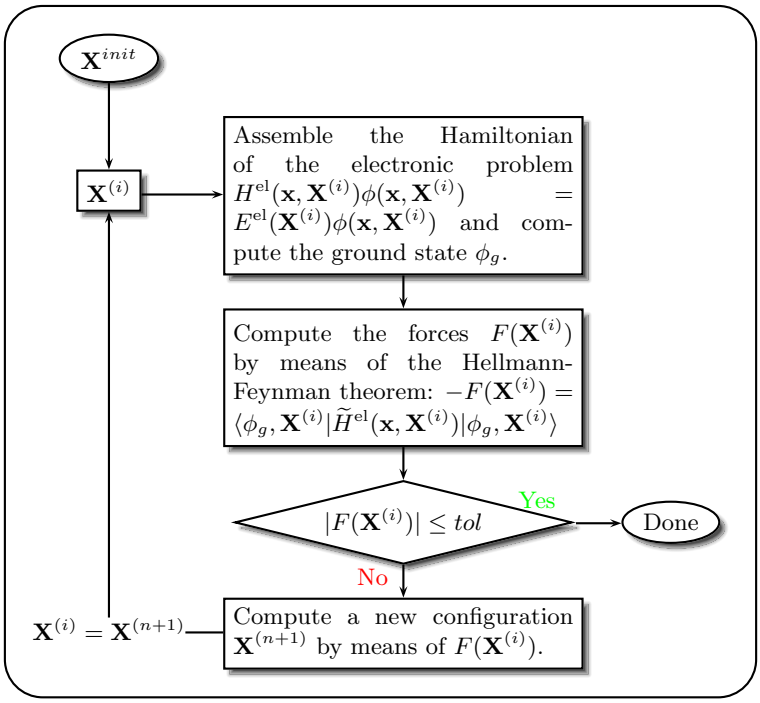

Figure 2.1: Minimization procedure to find the rest positions $\mathbf{X}_0$ of the atomic cores defined in Eq. (2.13). The Hellmann-Feynman theorem (2.14) is used to compute of the forces on the atomic cores. The abbreviation $\widetilde{H}^{\mathrm{el}}(\mathbf{x}, \mathbf{X}) = (\partial H^{\mathrm{el}})/(\partial \mathbf{X})$ is used.

the electronic energy is rather complicated having many local minima. For more details see Refs. [12, 13]. Energy minimizations allow to investigate important physical quantities such as the crystal structure of solids for instance.
The atomic cores being trapped at some positions, i.e. the rest positions found by a minimization procedure for instance, is the picture that will be retained within the ballistic quantum transport framework described in chapter 4. Obviously, the atomic cores are never perfectly at rest but rather oscillate around the positions $\mathbf{X}_0$. Further details on these lattice vibrations can be found in Ref. [10]. Furthermore, the impact of an applied voltage bias due to electric contacts, such as in semiconductor devices, on the positions of the atomic cores is ignored.

2.3 The Hartree Approximation

A direct solution of the electronic problem

$$\mathbf{H}^{\mathrm{el}}(\mathbf{x})\phi(\mathbf{x}) = E^{\mathrm{el}}\phi(\mathbf{x}) \tag{2.15}$$

can be accomplished by expanding the solution ϕ in an orthonormal basis $\{\psi_\mu\}_\mu$ of $\mathcal{H}^{\mathrm{el}}$

$$\phi(\mathbf{x}) = \sum_\mu c_\mu \psi_\mu(\mathbf{x}). \tag{2.16}$$

The $\mathbf{X}$-dependence in $\mathbf{H}^{\mathrm{el}}$,E^{el}, and ϕ is omitted for simplicity. Using the expansion (2.16) in Eq. (2.15), multiplying on the left by ψ_ν^*, and integrating over $\mathbf{x}$ yields an algebraic eigenvalue problem. This approach, also referred to as *exact diagonalization*, is rather impracticable especially when a large number of electrons is considered as the resulting matrices can not be diagonalized within reasonable time limits even on large scale computers. As mentioned in Sec. (2.1), the DFT and the Hartree Fock method are two widely used approaches to tackle the electronic problem. A simpler approach is given by the *Hartree approximation*. Within this approximation, the ansatz for the total wave function reads

$$\phi(\mathbf{x}) = \phi(\vec{r}_1, \vec{r}_1, \ldots, \vec{r}_N) = \psi_1(\vec{r}_1)\psi_2(\vec{r}_2)\ldots\psi_N(\vec{r}_N). \tag{2.17}$$

The single particle states ψ_i are obtained by minimizing the functional

$$\begin{aligned}
&\langle\phi|\mathbf{H}^{\mathrm{el}}|\phi\rangle \\
&= \langle\phi| \sum_{i=1}^{N} \left[-\frac{\hbar^2}{2m_e}\Delta_{\vec{r}_i} + \sum_{k=1}^{M} \Phi_k^{\mathrm{ion}}(\vec{r}) + \frac{e^2}{8\pi\epsilon_0} \sum_{\substack{j=1 \\ j\neq i}}^{N} \frac{1}{|\vec{r}_i - \vec{r}_j|} \right] |\phi\rangle \\
&= \sum_{i=1}^{N} \int \psi_i^*(\vec{r}_i) \left[-\frac{\hbar^2}{2m_e}\Delta_{\vec{r}_i} + \sum_{k=1}^{M} \Phi_k^{\mathrm{ion}}(\vec{r}) \right] \psi_i(\vec{r}_i) d\vec{r}_i \\
&+ \frac{e^2}{8\pi\epsilon_0} \sum_{\substack{i,j \\ i\neq j}}^{1\ldots N} \int\int \frac{|\psi_i(\vec{r}_i)|^2 |\psi_j(\vec{r}_j)|^2}{|\vec{r}_i - \vec{r}_j|} d\vec{r}_i d\vec{r}_j
\end{aligned} \tag{2.18}$$

under the constraint that the single electron wave functions are normalized

$$\frac{\delta}{\delta\psi_i^*} \left[\langle\phi|H^{\mathrm{el}}|\phi\rangle - \sum_{j=1}^{N} \epsilon_j \int |\psi_j(\vec{r}_j)|^2 d\vec{r}_j \right] = 0 \qquad i = 1\ldots N, \tag{2.19}$$

with $\{\epsilon_j\}_j$ being Lagrange multipliers. The resulting equations, i.e. the *Hartree* equations, read

$$\left[-\frac{\hbar^2}{2m_e}\Delta + V_i(\vec{r})\right]\psi_i(\vec{r}) = \epsilon_i\psi_i(\vec{r}) \tag{2.20}$$

with

$$V_i(\vec{r}) = \underbrace{\frac{e^2}{4\pi\epsilon_0}\sum_{\substack{j=1\\ j\neq i}}^{N}\int\frac{|\psi_j(\vec{r'})|^2}{|\vec{r}-\vec{r'}|}d\vec{r'}}_{V_H^{(i)}(\vec{r})} + V_c(\vec{r}). \tag{2.21}$$

The term $V_H^{(i)}(\vec{r})$ is referred to as the *Hartree potential* and $V_c(\vec{r}) = \sum_{k=1}^{M}\Phi_k^{\text{ion}}(\vec{r})$ is the crystal potential. In a generalized sense, the $\psi_i(\vec{r})$ may be understood as single-electron states [10]. The Hartree potential $V_H^{(i)}(\vec{r})$ can then be interpreted as the sum of the Coulomb potentials of all electrons except the i-th. Obviously, as the Hartree potential depends on the single-electron states, the Hartree equations have to solved in a self-consistent way. Assuming that the number of electrons N in the system is macroscopically large, as in the case of a solid for instance, and that the wave functions of the electrons are widely spread over the considered region, no significant difference can be expected if the sum over j in Eq. (2.21) is extended to include the term $j = i$. In this way the i-dependence of the Hartree potential is eliminated, where the resulting error is of order of magnitude $1/N$. The conditions (2.19) collapse to a single Schrödinger equation [3] yielding the eigensystem $\{\epsilon^{(\nu)}, \psi^{(\nu)}\}_\nu$. It remains to specify a representative configuration of the ψ_i which enter the product ansatz (2.17). The existence of such a configuration implies that external perturbations do not notably change the state of the electron gas. From an energy point of view then, a reasonable choice would be to use the ground state configuration of the N-particle system. At zero temperature the first N states, i.e. the energetically lowest, are considered to construct this state, where $\epsilon^{(N)}$ is referred to as the Fermi energy. The configuration-dependence of the Hartree potential is strongly affected by the kind of one-particle states involved. Plane-wave-like states of ideal crystals for instance are expected to yield a weaker configuration-dependence than localized states of real semiconductors including perturbations [10].
The Hartree potential is finally given by

$$V_H(\vec{r}) = \frac{e}{4\pi\epsilon_0}\int \underbrace{e\sum_{\nu=1}^{N}|\psi_\nu(\vec{r}')|^2}_{\rho(\vec{r}')}\frac{1}{|\vec{r}-\vec{r}'|}d\vec{r}' \tag{2.22}$$

with $\rho(\vec{r}')$ being the electron density. The Hartree equation is then solved by the iterative procedure described in Fig. (2.2). Instead of using Eq. (2.22) which implies an integration in three dimensions, the Hartree potential can be calculated by means of Poisson's equation

$$-\Delta V_H(\vec{r}) = \frac{n(\vec{r})}{\epsilon_0} \tag{2.23}$$

In this way the presence of electric contacts in semiconductor devices can be accounted for by suitable boundary conditions for Eq. (2.23). In particular, a given voltage on the gate contact is

[3] This implies that the different ψ_i in Eq. (2.17) are orthonormal. Furthermore, by ensuring that the ψ_i are different in Eq. (2.17), the exclusion principle is partially accounted for.

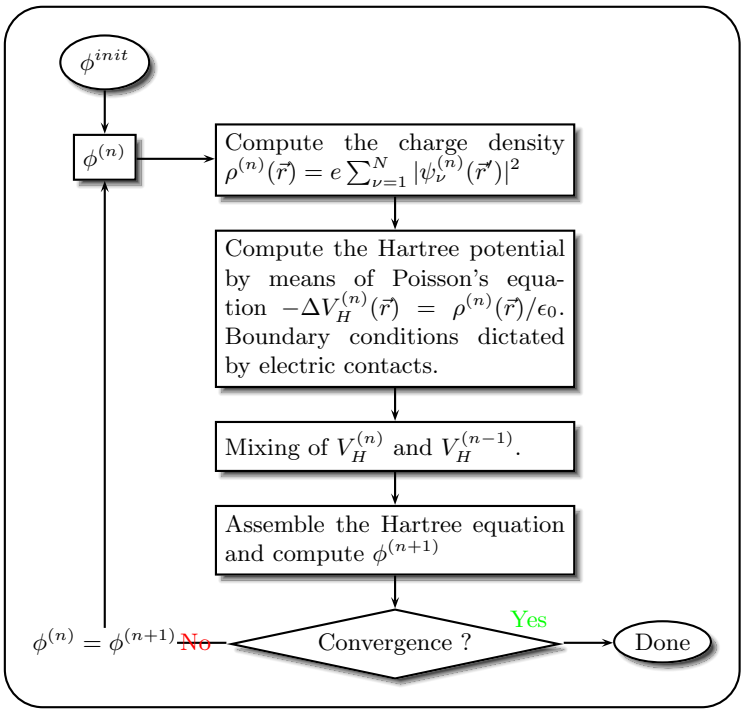

Figure 2.2: Iteration scheme for the self-conistent solution of the Hartree equations. The iteration cycle usually requires advanced damping schemes which are thouroughly discussed in chapter 4.

described by a Dirichlet condition

$$V_H(\vec{r}) = V_g, \quad \vec{r} \in \text{gate} \tag{2.24}$$

while the source and drain contacts require more specific treatment as explained in chapter 4. The calculation of the electron density $\rho(\vec{r})$ at finite temperatures is addressed in Sec. 3.8.

2.4 Summary

The stationary Schrödinger equation describing a system of electrons and atomic cores is presented at the beginning of this chapter. An important approximation due to Born and Oppenheimer notably simplifies this problem which is almost intractable in its full form especially in the presence of many particles. The problem reduces to the solution of a Schrödinger equation for the electronic system depending on the configuration of the atomic cores. Subsequently, the electronic spectrum can be used to compute the properties of the cores. However, the motion of the cores is not further addressed. The Hartree approximation for the solution of the electronic problem is described which will be employed within the transport framework described in chapter 4. This approximation requires an iteration procedure to reach self-consistency between the potential and the density. The presence of electric contacts can be straightforwardly included within this approximation.

Chapter 3

Band Structure Effects

3.1 Introduction

Several simplifications were outlined in the previous chapter leading to a simplified description of electrons moving in the presence of a fixed ionic background. The electronic problem

$$\underbrace{\left[-\frac{\hbar^2}{2m_e}\Delta + U(\vec{r})\right]}_{\mathbf{H}} \quad \psi(\vec{r}) = \epsilon\psi(\vec{r}) \tag{3.1}$$

with $U(\vec{r})$ being the sum of the potential $V_c(\vec{r})$ generated by the ionic cores and a mean field potential describing the interactions between the electrons, is the starting point for further simplifications needed for the simulation of semiconductor devices. The case of an infinite crystal is considered first, followed by a description of the empirical tight-binding (ETB) and the empirical pseudopotential method (EPM). These two methods are used to compute the energies of carriers in nanostructures. Subsequently, the effective mass approximation (EMA) is introduced and applied to specific examples. Comparisons to tight-binding and pseudopotential methods are carried out in order to investigate the shortcomings of the effective mass approximation [1]. Nonparabolicity models for quantized structures are introduced, followed by some details on the calculation of charge densities in thermodynamic equilibrium.

3.2 Infinite Crystal (Bulk)

The description of a crystal structure starts with the choice of a Bravais lattice [14]

$$\mathbf{\Gamma} \equiv \left\{\mathbf{A}\vec{n} \mid \vec{n} \in \mathbb{Z}^3\right\} \tag{3.2}$$

where $\mathbf{A} \equiv (\vec{a}_1|\vec{a}_2|\vec{a}_3)$ consists of three linearly independent column vectors $\vec{a}_1$, $\vec{a}_2$, and $\vec{a}_3$. A unit cell

$$\Omega_c \equiv \{\lambda_1\vec{a}_1 + \lambda_2\vec{a}_2 + \lambda_3\vec{a}_3 \mid \lambda_i \in [0,1), i = 1,2,3\} \tag{3.3}$$

of the Bravais lattice (3.2) can be defined which, translated through all vectors in $\mathbf{\Gamma}$, fills the whole space without overlaps. The unit cell contains a basis of atoms which is translated together with the cell to generate the entire crystal. Of course, other unit cells can be defined with an

[1] These shortcomings are referred to as band structure effects.

appropriate atomic basis to describe the same crystal structure. The definition (3.2) implies that the crystal is infinitely large. This idealization can be used for actual crystals which, of course, have a finite dimension but are large enough such that the majority of atoms do not feel the presence of the surface [2]. Given a Bravais lattice (3.2), the corresponding *reciprocal lattice* is defined by

$$\widetilde{\mathbf{\Gamma}} \equiv \left\{ \mathbf{B}\vec{n} \mid \vec{n} \in \mathbb{Z}^3 \right\}, \tag{3.4}$$

where $\mathbf{B} = 2\pi \mathbf{A}^{-T} = (\vec{b}_1|\vec{b}_2|\vec{b}_3)$. A function $f(\vec{r})$ which is periodic with respect to the Bravais lattice (3.2) can be expanded in terms of plane waves

$$f(\vec{r}) = \sum_{\vec{G} \in \widetilde{\mathbf{\Gamma}}} \widehat{f}(\vec{G}) e^{i\vec{G}\vec{r}} \tag{3.5}$$

with

$$\widehat{f}(\vec{G}) = \frac{1}{|\Omega_c|} \int_{\Omega_c} f(\vec{r}) e^{-i\vec{G}\vec{r}}. \tag{3.6}$$

When a material crystallizes in a given structure, the potential generated by the atomic cores

$$V_c(\vec{r}) = \sum_{j=1}^{\infty} \Phi_j^{\text{ion}}(\vec{r}) \tag{3.7}$$

is periodic with respect to a Bravais lattice (3.2) and the potential $U(\vec{r})$ is assumed to have the same periodicity of the crystal potential $V_c(\vec{r})$.
The eigenstates ψ of the Hamiltonian (3.1) can then be chosen to have the form of a plane wave times a function with the periodicity of the Bravais lattice, i.e.

$$\psi(\vec{r}) \to \psi(\vec{k}, \vec{r}) = e^{i\vec{k}\vec{r}} u(\vec{k}, \vec{r}), \tag{3.8}$$

where

$$u(\vec{k}, \vec{r} + \vec{R}) = u(\vec{k}, \vec{r}) \;\; \forall \vec{R} \in \mathbf{\Gamma}. \tag{3.9}$$

Accordingly, the wave vector $\vec{k}$ enters the energy $\epsilon \to \epsilon(\vec{k})$. This statement is known as Bloch theorem and the property

$$\psi(\vec{k}, \vec{r} + \vec{R}) = e^{i\vec{k}\vec{R}} \psi(\vec{k}, \vec{r} + \vec{R}) \tag{3.10}$$

is referred to as Bloch condition[3]. For a proof see Ref. [14]. The energy and the wave function have the following important properties

$$\psi(\vec{k} + \vec{G}, \vec{r}) = \psi(\vec{k}, \vec{r}) \tag{3.11}$$

$$\epsilon(\vec{k} + \vec{G}) = \epsilon(\vec{k}) \tag{3.12}$$

for all $\vec{G} \in \widetilde{\mathbf{\Gamma}}$. Using the ansatz (3.8) for the Schrödinger problem (3.1) yields

$$\left[-\frac{\hbar^2}{2m_e} (\vec{\nabla} + i\vec{k})^2 + U(\vec{r}) \right] \psi(\vec{k}, \vec{r}) = \epsilon(\vec{k}) \psi(\vec{k}, \vec{r}) \tag{3.13}$$

which is solved on Ω_c with periodic boundary conditions (3.9). The wave vector $\vec{k}$ is varied over a primitive unit cell of the reciprocal lattice, i.e.

$$\widetilde{\Omega}_c \equiv \{ \lambda_1 \vec{b}_1 + \lambda_2 \vec{b}_2 + \lambda_3 \vec{b}_3 \mid \lambda_i \in [0, 1), i = 1, 2, 3 \} \tag{3.14}$$

[2] The term *bulk* is often used to describe such crystals.

[3] Using wave functions of the form (3.8) within the DFT or Hartree aproximation guarantees that the potential $U(\vec{r})$ possesses the periodicity of the Bravais lattice.

for instance. Solving Eq. (3.8) for a fixed $\vec{k}$ yields the eigensystem $\{\epsilon_n(\vec{k}), \psi_n(\vec{k},\vec{r})\}$ which is labeled[4] by the band index n. Usually, the *Born-von Karman* boundary conditions

$$\psi_n(\vec{k},\vec{r}+\vec{v}) \stackrel{!}{=} \psi_n(\vec{k},\vec{r}) \;\; \forall \;\; \vec{v} \in \mathbf{\Gamma}_{\vec{N}}, \tag{3.15}$$

where $\vec{N} = (N_1, N_2, N_3)$ and

$$\mathbf{\Gamma}_{\vec{N}} \equiv \{n_1 N_1 \vec{a}_1 + n_2 N_2 \vec{a}_2 + n_3 N_3 \vec{a}_3 \mid n_i \in \mathbb{Z} \;\; i = 1,2,3\}\,. \tag{3.16}$$

The lattice $\mathbf{\Gamma}_{\vec{N}}$ is a subset of $\mathbf{\Gamma}$ and can be characterized by a unit cell

$$\Omega_{\vec{N}} \equiv \{\lambda_1 N_1 \vec{a}_1 + \lambda_2 N_2 \vec{a}_2 + \lambda_3 N_3 \vec{a}_3 \mid \lambda_i \in [0,1), i = 1,2,3\} \tag{3.17}$$

being a multiple of Ω_c. In this context it is useful to define the reduced lattice

$$\mathbf{\Gamma}^{\mathrm{red}}_{\vec{N}} \equiv \{n_1 \vec{a}_1 + n_2 \vec{a}_2 + n_3 \vec{a}_3 \mid n_i = 0, \ldots, N_i - 1 \;\; i = 1,2,3\}\,. \tag{3.18}$$

The condition (3.15) restricts the allowed wave vectors $\vec{k}$ in the reciprocal space to the set

$$\widetilde{\mathbf{\Gamma}}^{\mathrm{red}}_{\vec{N}} \equiv \left\{ \frac{n_1 \vec{b}_1}{N_1} + \frac{n_2 \vec{b}_2}{N_2} + \frac{n_3 \vec{b}_3}{N_3} \mid n_i = 0, \ldots, N_i - 1 \;\; i = 1,2,3 \right\} \tag{3.19}$$

which is a subset of

$$\widetilde{\mathbf{\Gamma}}_{\vec{N}} \equiv \left\{ \frac{n_1 \vec{b}_1}{N_1} + \frac{n_2 \vec{b}_2}{N_2} + \frac{n_3 \vec{b}_3}{N_3} \mid n_i \in \mathbb{Z} \;\; i = 1,2,3 \right\}. \tag{3.20}$$

The cell $\Omega_{\vec{N}}$ is used to normalize the Bloch states

$$\int_{\Omega_{\vec{N}}} |\psi_n(\vec{k},\vec{r})|^2 d\vec{r} = 1. \tag{3.21}$$

The number N of allowed wave vectors $\vec{k}$ is equal to the number of cells Ω_c in the reduced lattice $\mathbf{\Gamma}^{\mathrm{red}}_{\vec{N}}$, i.e. $N = N_1 N_2 N_3$.

3.3 Methods

The calculation of the potential $U(\vec{r})$ from Eq. (3.1) often requires an iterative procedure to achieve self-consistency. This is the case when DFT, Hartree, or the Hartree-Fock approximation are used for instance to simplify the original many-particle problem (2.15). Other methods implicitly assume the form (3.1) without being related to a simplification of the original many-particle problem. The latter methods rely for instance upon experimental data to construct $U(\vec{r})$ and do not require self-consistency. Consequently, these empirical approaches are particularily interesting from a computational point of view. Two examples are given by the ETB and EPM described in appendix A. The band structure of bulk silicon is computed by means of the ETB and EPM and a comparison of characteristic quantities such as the band gap and effective masses along different directions [10, 15] is given in Tab. 3.1. The band structures are plotted in

Table 3.1: Effective masses and band gap of band structures computed by means of the parameterizations of Chelikowski et al. (Ch) and Wang et al. (Wa) given in Eqs. (A.17) and (A.18) respectively and by means of the tight-binding (TB) formalism described in appendix A.0.1. Two different cutoffs $co \in \{3.5, 4.5\}[1/a_B]$ are used for the Wa case while $co = 8\pi/a_{Si}$ is used for the Ch case. Conduction band masses and minimum are extracted at $0.85[\Gamma X]$ in the case of the pseudopotentials. The experimental data is taken from Ref. [16] and references therein.

Quantity	Ch	Wa (3.5)	Wa (4.5)	TB	experiment
$m_{lh}^{[100]}$	0.167	0.150	0.151	0.214	0.15
$m_{lh}^{[111]}$	0.097	0.092	0.093	0.144	0.11
$m_{hh}^{[100]}$	0.272	0.279	0.276	0.276	0.34
$m_{hh}^{[111]}$	0.671	0.682	0.662	0.734	0.69
m_l	0.911	0.911	0.913	0.891	0.916
m_t	0.195	0.200	0.200	0.201	0.19
E_g [eV]	1.056	1.165	1.173	1.131	1.124

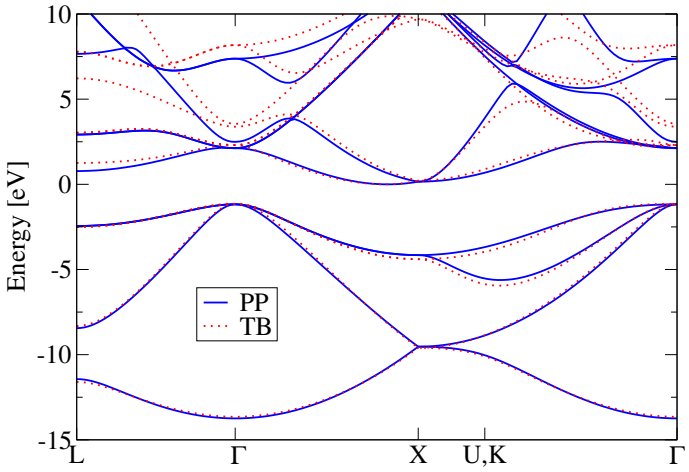

Figure 3.1: Band structure of bulk silicon obtained by means of the empirical pseudopotential (PP) method using the parametrization (A.17) and the present tight-binding (TB) method. A comparison of characteristic quantities is given in Tab. 3.1.

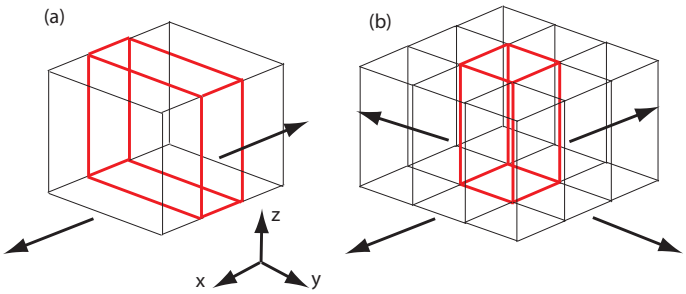

Figure 3.2: Schematic representation of nanostructures confined in two (a) and one (b) direction. The red polygons denote the unit cells of the structures which contain a basis of atoms. The cells are translated along the arrows to obtain the entire structure.

Fig. 3.1. Conduction band effective masses are in good agreement with experimental data while the agreement for the valence band masses is less obvious in some cases [5]. An increment of the cutoff *co* from $3.5/a_B$ to $4.5/a_B$ has a minor impact on the results obtained by the pseudopotential of Wang et al.. This is of crucial importance for the simulation of nanostructures as shown in Sec. 3.4.
Using $co = 3.5/a_B$ for the potential of Wang et al. (see appendix A) yields a Hamilton matrix of size 181×181. The tight-binding method has the advantage that the matrices involved in the band structure calculations are rather small, i.e. 20×20. On the other hand, within the tight-binding formalism, the functional form of the basis set is undermined, since only the empirically adjusted Hamilton matrix elements are used. Without explicit basis functions it becomes difficult to describe the impact of the atomic geometry on the matrix elements or to compare the real space representation of the Bloch states with results from other methods such as ab-initio calculations. Furthermore, the small basis set suggests limited variational flexibility compared to the continuous form factors used within the pseudopotential method.

3.4 Nanostructures

3.4.1 Overview

The critical dimensions of nowadays semiconductor devices have reached a regime where the bulk picture described in Sec. 3.2 does not apply any longer. These semiconductor structures can be generally divided in three classes - structures periodic in two dimensions (2D quantum wells), those periodic in one dimension (1D quantum wires)[6], and finally structures that are periodic in zero dimensions (0D quantum dots). With respect to the 3D periodic bulk case, these structures exhibit localization effects in one, two, and three dimensions, respectively. In the case of wires and wells, a unit cell containing a basis of atoms similar to the bulk case can be employed to describe the reduced periodicity as illustrated in Fig. 3.2. The use of the EPM to

[4]The solutions of Eq. (3.8) are usually referred to as *Bloch states*, *Bloch functions*, or *Bloch waves* while the term *band structure* is used for $\epsilon_n(\vec{k})$. Note that not all methods for band structure calculations require a reduction of the problem (3.1) to Eq. (3.8).

[5]Minor deviations from the results given in Ref. [16] are observed which is probably due to different extraction methods.

[6]The terms quantum wire and nanowire will be used interchangeably.

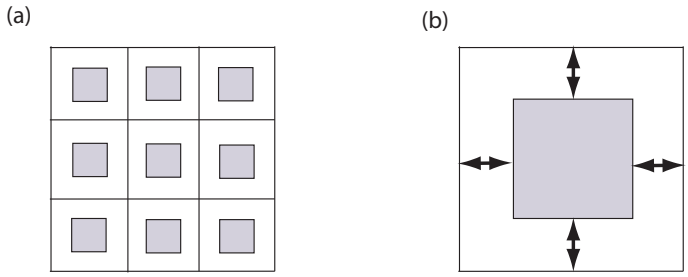

Figure 3.3: (a) The unit cell of a nanowire is periodically repeated to fill the whole space. This generates an artificial array of nanowires giving rise to a potential which is periodic in three dimensions. (b) Each cell contains a wire (dark area) which is surrounded by vacuum (arrows).

compute energies and states of wires and wells is described in the following [7]. The treatment of nanostructures within the semi-empirical $sp^3d^5s^*$ tight-binding framework is explained elsewhere [17]. Quantum dots are not considered in this work.

3.4.2 Wire

The unit cell given in Fig. 3.2.a is specified by

$$\Omega_c = [0, a] \times [0, L] \times [0, L]. \tag{3.22}$$

While the length a is fixed by the atomic structure of the wire, the lateral extension L can can be set to any value provided that the entire atomic basis is contained in the cell [8]. In order to employ Bloch's theorem for the calculation of the electronic states of the nanowire, a periodicity in all three dimensions has to be recovered. This is achieved by constructing an artificial array of nanowires as shown in Fig. 3.3 generating a potential $U(\vec{r})$ that is periodic with respect to Ω_c. The corresponding Brillouin zone in the reciprocal space reads

$$\widetilde{\Omega}_c = [-\pi/a, \pi/a] \times [-\pi/L, \pi/L] \times [-\pi/L, \pi/L]. \tag{3.23}$$

The presence of the artificial wire replicas is eliminated by enlarging the vacuum layer around the atomic basis, i.e. $L \to \infty$. Consequently, the Brillouin zone (3.23) becomes quasi one-dimensional and it is sufficient to consider $\vec{k} = (k_x, 0, 0)$ with $k_x \in [-\pi/a, \pi/a]$. For computational reasons, the vacuum layer is increased until the energies on the trajectory $k_x \in [0, 2\pi/a]$ become independent of the vacuum width. The Brillouin zone retains a certain width and the energies and the Bloch states are assumed to be independent of k_y and k_z, i.e. $\epsilon(\vec{k}) \simeq \epsilon(k_x, 0, 0)$ and $\psi(\vec{k}) \simeq \psi(k_x, 0, 0)$. The atomic structure of a typical silicon nanowire grown along the $\langle 100 \rangle$ direction is shown in Fig. 3.4.a. Two possible wire cross sections with surfaces oriented along the $\langle 100 \rangle$ and $\langle 110 \rangle$ directions are given in Figs.3.4.b and 3.4.c respectively, both having a unit cell of length $a = a_{Si}$. The dangling bonds at the surface are terminated by hydrogen atoms in this particular case. A continuous form factor for hydrogen is provided by Wang et al.

$$\widehat{V}_H^{Wa}(q) = \begin{cases} v_{Si}(a_1 + a_2 q + a_3 q^2 + a_4 q^3) & q \leq 2/a_B \\ v_{Si}(a_5/q + a_6/q^2 + a_7/q^3 + a_8/q^4) & q > 2/a_B \end{cases} \tag{3.24}$$

[7] The approach adopted in fully ab-initio methods is similar.

[8] A rectangular unit cell with a square cross section is not always the optimal choice. However, it simplifies the formalism and does not restrict the shape of the nanowire cross section, i.e. the atomic basis.

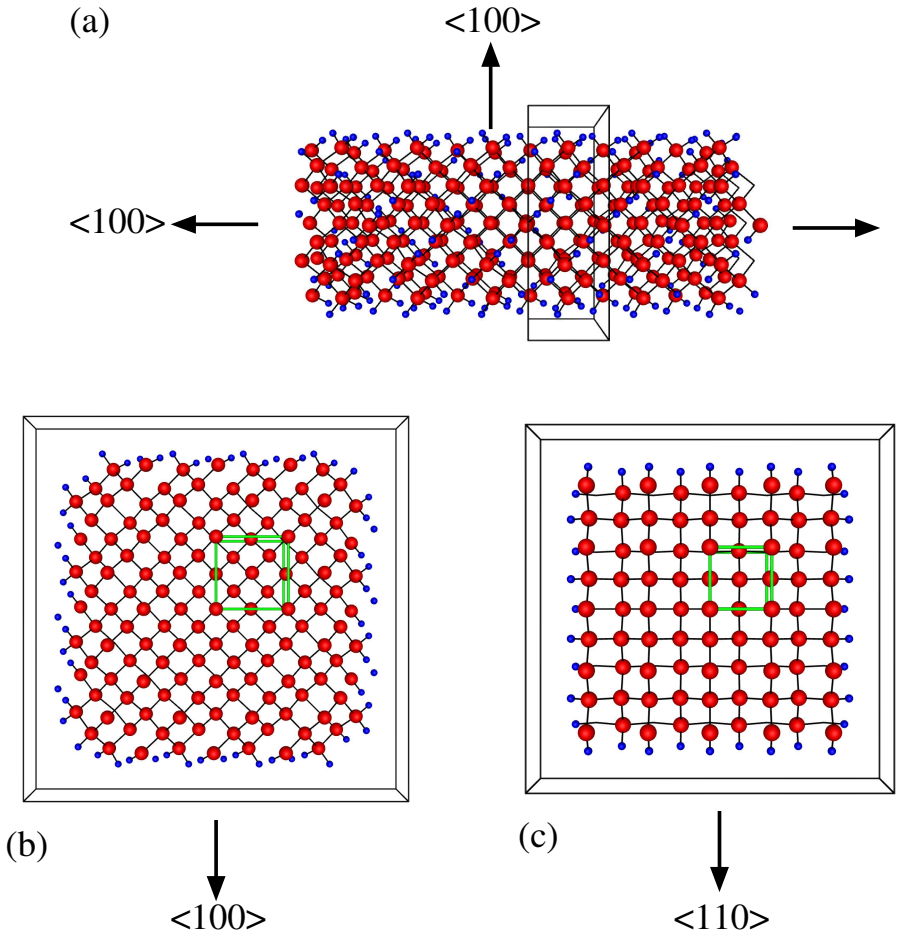

Figure 3.4: (a) Silicon nanowire grown along the ⟨100⟩ direction. The cross sectional shape of this wire can be square for instance with surfaces oriented along the ⟨100⟩ and ⟨110⟩ directions as shown in (b) and (c) respectively. The green blocks are used to construct the skeleton of the wires (red spheres). Bond lengths and orientations within the skeleton are identical to bulk silicon. The dangling bonds are terminated by hydrogen (blue spheres) and the surface is deformed as explained in Ref. [16]. This deformation affects the silicon skeleton in the vicinity of the surfaces. Both unit cells have a length of $a = a_{Si}$ while the lateral extensions are $L = 2.89$nm and $L = 2.28$nm for (b) and (c) respectively including a vacuum layer of 0.25nm.

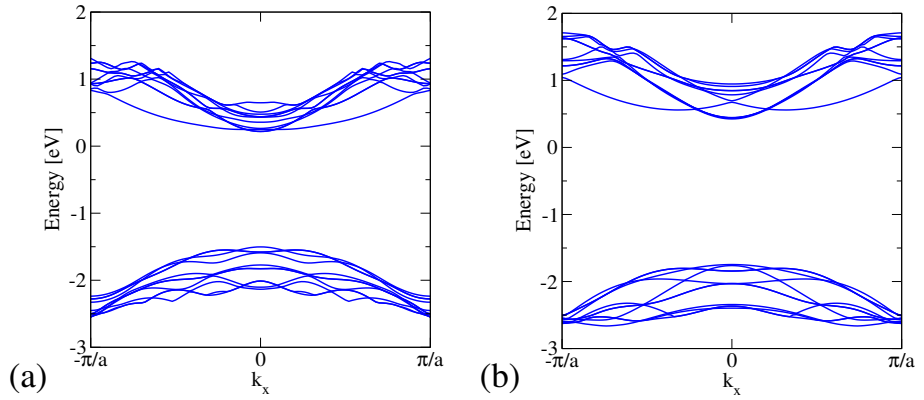

Figure 3.5: Band structures of silicon nanowires computed by means of the empirical pseudopotential method explained in Sec 3.4.2. The dispersions of the wires from Figs. 3.4.b and 3.4.c are given in (a) and (b), respectively. The energy scales are gauged such that the corresponding bulk conduction band minimum is equal to zero.

where $a_1 = -0.1416\text{Ry}$, $a_2 = 9.802 \times 10^{-3} a_B^{-1}\text{Ry}$, $a_3 = 6.231 \times 10^{-2} a_B^{-2}\text{Ry}$, $a_4 = -1.895 \times 10^{-2} a_B^{-3}\text{Ry}$, $a_5 = 2.898 \times 10^{-2} a_B\text{Ry}$, $a_6 = -0.3877 a_B^2\text{Ry}$, $a_7 = 0.9692 a_B^3\text{Ry}$, and $a_8 = -1.022 a_B^4\text{Ry}$. Together with the form factor (A.18), the Fourier transform (A.12) of a wire unit cell (slab) consisting of N_H hydrogen and N_{Si} silicon atoms is given by

$$\widehat{U}(\vec{G}) = \frac{1}{|\Omega_c|}\left[\widehat{V}_H^{Wa}(|\vec{G}|)\sum_{i=1}^{N_H} e^{-i\vec{G}\vec{b}_i} + \widehat{V}_{Si}^{Wa}(|\vec{G}|)\sum_{i=N_H+1}^{N_H+N_{Si}} e^{-i\vec{G}\vec{b}_i}\right]. \tag{3.25}$$

The potential (3.25) is used in the secular problem (A.14) to compute the band structures of the nanowires from Figs. 3.4.b and 3.4.c. A cutoff $co = 3.5/a_B$ is used for the calculations. These band structures are plotted in Fig. 3.5. Contour plots of the potentials as well as the energetically highest valence band state (HVS) and lowest conduction band state (LCS) are given in Fig. 3.6. Using $co = 3.5/a_B$, yields a secular problem of size 22135^2 for the $\langle 100\rangle$ wire. As in the bulk case, the tight-binding formalism involves Hamilton matrices of much more moderate sizes, i.e. 1280^2 for a comparable nanowire of the same orientation [9]. Furthermore, within the tight-binding formalism there is no need to account for the surrounding vacuum explicitly.

3.4.3 Well

The calculation of electronic states in semiconductor quantum wells is similar to the procedure employed for nanowires. The unit cell shown in Fig. 3.2.b

$$\Omega_c = \{(\lambda_1 a_1^x + \lambda_2 a_1^y, \lambda_1 a_2^x + \lambda_2 a_2^y, \lambda_3 L) \mid \lambda_i \in [0,1], i = 1,2,3\} \tag{3.26}$$

is specified by means of two vectors $\mathbf{a}_1 = (a_1^x, a_1^y)$ and $\mathbf{a}_2 = (a_2^x, a_2^y)$ in the xy-plane and a height L. The atomic structure of the quantum well fixes the vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ while L is variable. Figure 3.7 shows the artificial array of quantum well unit cells needed to generate the periodic potential $U(\vec{r})$. A possible unit cell in the reciprocal space is given by

[9] More details on how the pseudopotential and tight-binding frameworks used in this work are compared can be found in Sec. 3.6.3. This is not a straightforward task, as hydrogen is not explicitly parametrized within the present tight-binding framework.

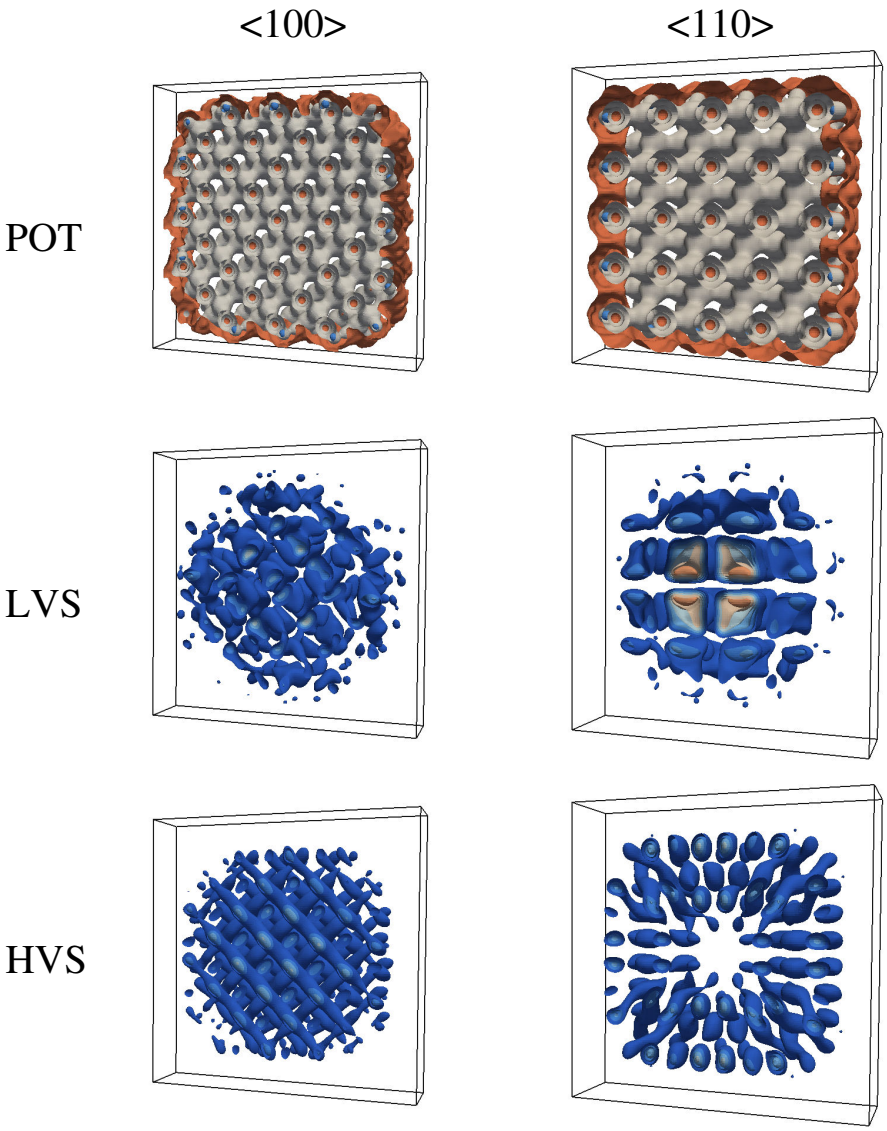

Figure 3.6: Contour plots of the potential (POT) and the absolute square of the HVS and LCS (see text) of the wires from Fig. 3.4.b and 3.4.c. The contours are computed for equally spaced values, ten for the states and five for the potentials, between the maximum and minimum of the corresponding quantity. Blue represents low values, red high.

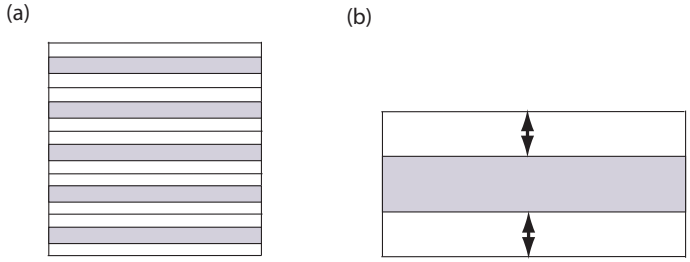

Figure 3.7: Same as in Fig. 3.3 for the case of quantum wells. (a) Array of quantum well unit cells. (b) Each unit cell contains a basis of the quantum well sandwiched between vacuum layers.

$$\widetilde{\Omega}_c = \{(\lambda_1 b_1^x + \lambda_2 b_1^y, \lambda_1 b_2^x + \lambda_2 b_2^y, \lambda_3 2\pi/L) \mid \lambda_i \in [-0.5, 0.5], i = 1, 2, 3\}, \tag{3.27}$$

where $b_1^x = 2\pi a_2^y/d$, $b_1^y = -2\pi a_1^y/d$, $b_2^x = -2\pi a_2^x/d$, $b_2^y = 2\pi a_1^x/d$, and $d = a_1^x a_2^y - a_2^x a_1^y$. These reciprocal unit cells become quasi two-dimensional as the height L of the well, i.e. the vacuum layer shown in Fig. 3.7.b, is increased to remove the artificial well replicas. The energies and Bloch states are computed for $k_z = 0$, assuming that these quantities are approximately independent of k_z even though a finite height L is retained for computational reasons.
Two typical silicon quantum well structures with surfaces perpendicular to the $\langle 100 \rangle$ and $\langle 110 \rangle$ direction are shown in Fig. 3.8. The unit vectors in the $xy-$plane are $\{\mathbf{a}_1 = (-a_{Si}, a_{Si})/2, \mathbf{a}_2 = (a_{Si}, a_{Si})/2\}$ and $\{\mathbf{a}_1 = (a_{Si}, 0), \mathbf{a}_2 = (0, a_{Si}/\sqrt{2})\}$ for the $\langle 100 \rangle$ and $\langle 110 \rangle$ case respectively.
The pseudopotential is constructed according to Eq. (3.25) using the same parameterizations employed for the wire case in Sec. 3.4.2, i.e. Eqs. (A.18) and (3.24), and the potential is used within the secular problem (A.14). The two-dimensional band structures of the quantum wells shown in Fig. 3.8 are given in Fig. 3.9 while one-dimensional cuts are plotted in Fig. 3.10. In both cases, $co = 3.5/a_B$ has been employed. The size of the secular problem for the $\langle 100 \rangle$ case is 2029^2 while the tight-binding framework generates a 170^2 matrix for a well of comparable height.

3.4.4 Computational Remarks

As mentioned in Secs. 3.4.2 and 3.4.3, the Hamilton matrices assembled within the EPM are notably larger than the ones used in the tight-binding framework. While the eigenvalue problems generated in the latter approach can be satisfactorily solved by means of a standard library. i.e. the LAPACK [18] for instance, more advanced diagonalization algorithms have been developed to efficiently simulate nanostructures described by pseudopotentials [19]. In this work, ScaLAPACK [20] is used to compute the pseudopotential band structures of nanowires while quantum wells are treated by means of LAPACK. Parallelization in the latter case is achieved by distributing the different $\vec{k}$'s among the the compute nodes. The diagonalization routines can be accellerated by providing the exact indices of the desired eigenvalues. Generally, only a few eigenvalues around the gap are required. Provided that the nanostructure is a semiconductor, the number of valence bands is just half the number of electrons in the unit cell.

3.5 Effective Mass Approximation

The calculation of electronic states in semiconductor nanostructures as outlined in Sec. 3.4 is related to a notable computational burden. A widely used simplification is given by the effective

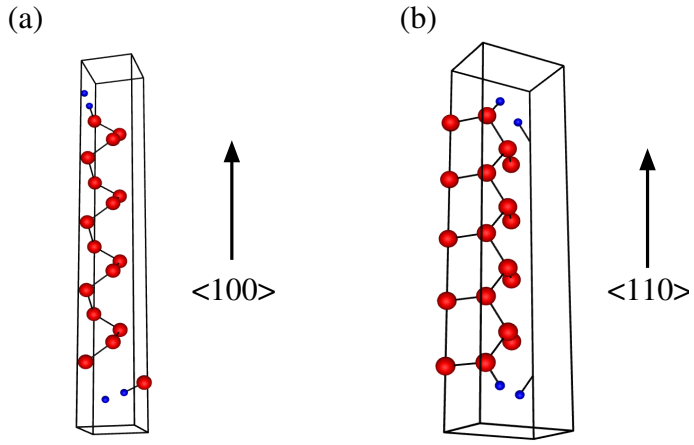

Figure 3.8: Unit cells of silicon quantum wells with surfaces perpendicular to the $\langle 100 \rangle$ (a) and $\langle 110 \rangle$ direction (b). The green polygons are used to construct the skeleton of the wells (red spheres) which is carved out of bulk silicon as in the wire case. Hydrogen (blue spheres) passivation of the surface is explained in Ref. [16]. The height of the unit cells is $L = 2.89$ nm and $L = 2.29$ nm for (a) and (b) respectively including a vacuum layer of 0.25 nm.

mass approximation (EMA) which has been originally introduced to describe an infinite crystal subjected to an external perturbation $U^{ext}(\vec{r})$

$$\left[-\frac{\hbar^2}{2m_e}\Delta + U(\vec{r}) + U^{ext}(\vec{r}) \right] \Psi = E\Psi, \tag{3.28}$$

where $U(\vec{r})$ is the periodic crystal potential [10].The problem (3.28) is solved on the enlarged unit cell $\Omega_{\vec{N}}$ from Eq. (3.17) and Born-von Karman boundary conditions (3.15) are used for $\Psi(\vec{r})$. In the following, $\vec{N}$ is macroscopically large and omitted in the notation. The wave function $\Psi(\vec{r})$ can be expanded in terms of Bloch states of the potential $U(\vec{r})$

$$\Psi(\vec{r}) = \sum_{n,\vec{k}\in\tilde{\mathbf{\Gamma}}^{\mathrm{red}}} \langle n, \vec{k}|\Psi\rangle \psi_n(\vec{k},\vec{r}) \tag{3.29}$$

with $\langle n, \vec{k}|\Psi\rangle = \int_\Omega \psi_n^*(\vec{k},\vec{r})\Psi(\vec{r})d\vec{r}$. Using the ansatz (3.29) for the Schrödinger problem (3.28) yields the secular equation

$$\sum_{n',\vec{k}'\in\tilde{\mathbf{\Gamma}}^{\mathrm{red}}} \left[\epsilon_n(\vec{k})\delta_{n,n'}\delta_{\vec{k},\vec{k}'} + \langle n, \vec{k}|U^{ext}|n', \vec{k}'\rangle \right] \langle n', \vec{k}'|\Psi\rangle = E\langle n, \vec{k}|\Psi\rangle. \tag{3.30}$$

The external potential $U^{ext}(\vec{r})$ is assumed to be periodic with respect to the cell Ω and can be expressed by means of plane waves

$$U^{ext}(\vec{r}) = \sum_{\vec{k}\in\tilde{\mathbf{\Gamma}}^{\mathrm{inf}}} \langle \vec{k}|U^{ext}\rangle e^{i\vec{k}\vec{r}}, \tag{3.31}$$

[10]For a discussion on the mean field potentials describing the interactions between the electrons, see Ref. [10]. In the following, the periodic potential $U(\vec{r})$ is assumed to be unaffected by $U^{ext}(\vec{r})$.

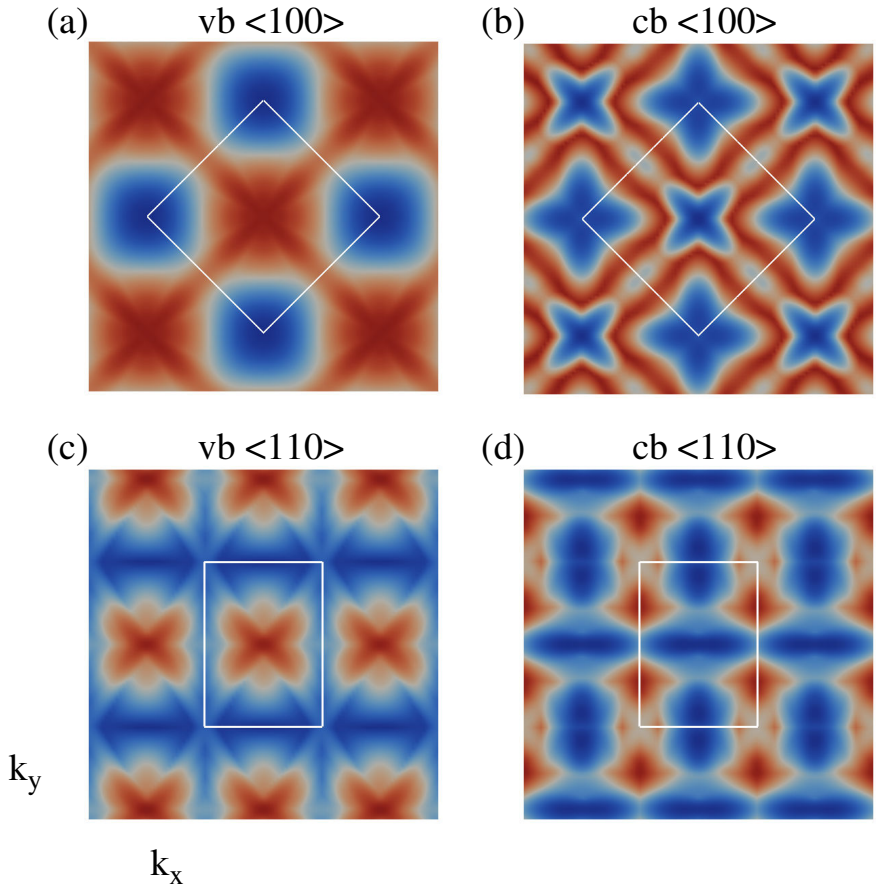

Figure 3.9: Two-dimensional band structures of silicon quantum wells computed by means of the EPM explained in Sec 3.4.3. The valence bands (vb) of the wells shown in Figs. 3.8.a and 3.8.b are plotted in (a) and (c), respectively and the conduction bands are given in (b) and (d), respectively. Plotted is the range $(k_x, k_y) \in [-3\pi/a_{Si}, 3\pi/a_{Si}]^2$. The white lines denote the unit cells (3.27) which in this case coincide with the first Brillouin zones. Blue represents low energies, red high energies.

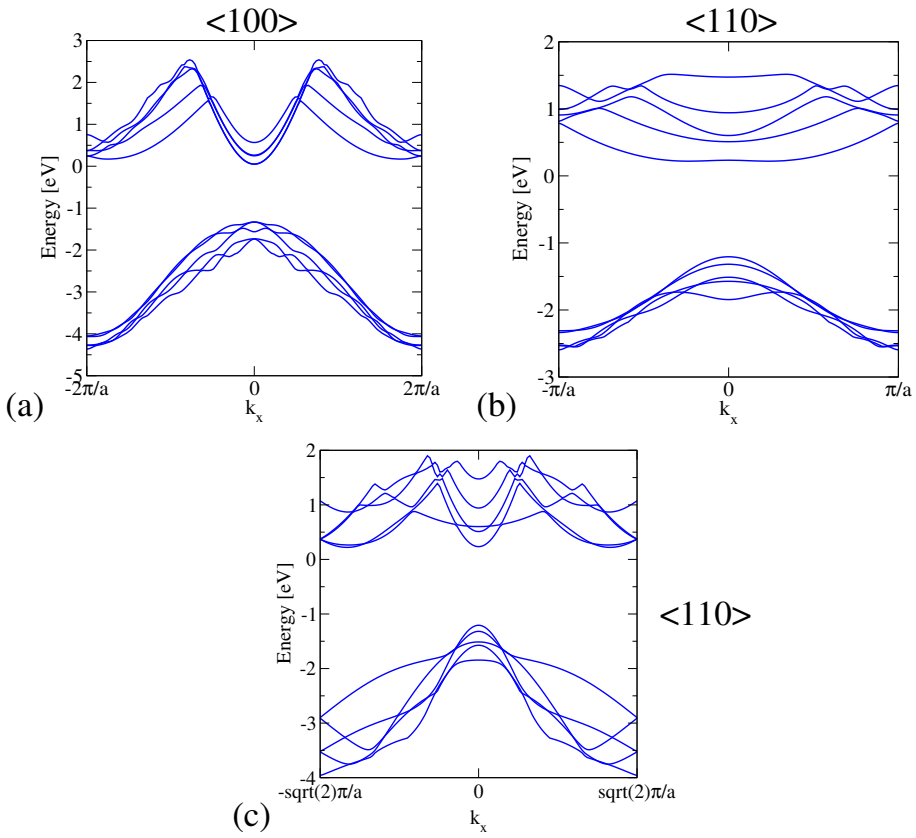

Figure 3.10: Band structure cuts of the silicon quantum wells shown in Figs. 3.8.a and 3.8.b are plotted in (a) and (b) respectively. In both cases, the reciprocal space is cutted along the cartesian k_x-axis. A cut along the k_y-axis for the $\langle 110 \rangle$ case is plotted in (c). The present EPM has been employed for the calculations. Energies are gauged as in Fig. 3.5.

where $\widetilde{\mathbf{\Gamma}}^{\inf} \equiv \left\{\mathbf{B}_{\vec{N}}\vec{n} \mid \vec{n} \in \mathbb{Z}\right\}$ and $\langle\vec{k}|U^{ext}\rangle = (1/\Omega)\int_\Omega \exp(-i\vec{k}\vec{r})U^{ext}(\vec{r})d\vec{r}$. The EMA is based on several prerequisites. The first is that the external potential has to be smooth compared to the periodic potential $U(\vec{r})$. Quantitatively, this means that the main contributions to the sum (3.31) come from wave vectors $\vec{k}$ inside the first Brillouin zone and the remaining coefficients $\langle\vec{k}|U^{ext}\rangle$ for $\vec{k}$'s outside the Brillouin zone are much smaller than the corresponding Fourier coefficients of the crystal potential $U(\vec{r})$. The $|\langle\vec{k}|U^{ext}\rangle|$ are roughly related to the change of the external potential over the unit cell of the crystal which can be estimated as $a|\vec{\nabla}U^{ext}|$. On the other hand, the Fourier coefficients of the crystal potential can be estimated by a typical band energy, i.e. the band gap, which yields the approximate condition $a|\vec{\nabla}U^{ext}| \ll E_g$ for smoothness.
The second prerequisite is that the wave function $\Psi(\vec{r})$ can be approximated by a single energy band $n = b$

$$\Psi(\vec{r}) = \sum_{\vec{k}\in\widetilde{\mathbf{\Gamma}}^{\mathrm{red}}} \langle b,\vec{k}|\Psi\rangle\psi_b(\vec{k},\vec{r}) \equiv \sum_{\vec{k}\in\widetilde{\mathbf{\Gamma}}^{\mathrm{red}}} \widehat{F}_b(\vec{k})\psi_b(\vec{k},\vec{r}) \tag{3.32}$$

This single-band picture is consistent with the requirement of smoothness for the external potential $U^{ext}(\vec{r})$. A consequence of smoothness is namely that the mixing $\langle n,\vec{k}|U^{ext}|n',\vec{k}'\rangle \simeq \delta_{n,n'}\delta_{\vec{k},\vec{k}'}\langle n,\vec{k}|U^{ext}|n,\vec{k}\rangle$ is suppressed in Eq. (3.30). The band b is assumed to have an non-degenerated extremum at $\vec{k}_0 = 0$ which can be expanded in second order in $\vec{k}$

$$\epsilon_b(\vec{k}) \simeq E_{edge} + \frac{\hbar^2}{2m_e}\vec{k}^T\mathbf{M}\vec{k} + O(\vec{k}^2), \tag{3.33}$$

where $\mathbf{M}$ is referred to as the effective mass tensor. The function $\widehat{F}_b(\vec{k})$ is essentially non-zero for wave vectors $\vec{k}$ in the first Brillouin zone similar to Fourier coefficients of the external potential. Finally, it is assumed that for these vectors $\vec{k}$, the Bloch factor $u_b(\vec{k},\vec{r})$ can be approximated by

$$u_b(\vec{k},\vec{r}) \simeq u_b(\vec{k}_0,\vec{r}). \tag{3.34}$$

Using the approximation (3.34) in Eq. (3.32) yields

$$\begin{aligned}\Psi(\vec{r}) &\simeq u_b(\vec{k}_0,\vec{r}) \sum_{\vec{k}\in\widetilde{\mathbf{\Gamma}}^{\mathrm{red}}} \widehat{F}_b(\vec{k})e^{i\vec{k}\vec{r}} \\ &\equiv u_b(\vec{k}_0,\vec{r})F_b(\vec{r}).\end{aligned} \tag{3.35}$$

The *envelope function* $F_b(\vec{r})$ is determined [10] by means of the effective mass equation

$$\left[E_{edge} - \frac{\hbar^2}{2m_e}\vec{\nabla}^T\mathbf{M}\vec{\nabla} + U(\vec{r})\right]F_b(\vec{r}) = EF_b(\vec{r}). \tag{3.36}$$

The extension to bands with multiple extrema located at $\vec{k}^{(i)}, i = 1,\ldots,\nu$ is straightforward [21]. Each extremum is related to an envelope function $F_b^{(i)}(\vec{r})$ and a corresponding effective mass equation

$$\left[E_{edge} - \frac{\hbar^2}{2m_e}\vec{\nabla}^T\mathbf{M}^{(i)}\vec{\nabla} + U^{ext}(\vec{r})\right]F_b^{(i)}(\vec{r}) = E^{(i)}F_b^{(i)}(\vec{r}), \tag{3.37}$$

where E_{edge} is the same for all extrema [11]. The total wave function reads

$$\Psi(\vec{r}) = \sum_{i=1}^{\nu}\psi_b(\vec{k}^{(i)},\vec{r})F_b^{(i)}(\vec{r}). \tag{3.38}$$

[11] For a general U^{ext}, the energies $E^{(i)}$ will have different values. On the other hand, if U^{ext} is spherically symmetric or has the symmetry of the crystal point group, the solutions to different extrema i will have the same energy [21]. When open boundary conditions are employed, the energy spectrum is continuous for a general potential and a common energy can be found for each extremum.

For the calculation of charge densities, the $|\Psi|^2$ is important. As long as the region of interest extends over several unit cells of the crystal, it is a good approximation to average the oscillating Bloch factors within a cell. Furthermore, the mixing between different extrema in $|\Psi|^2$ is omitted, yielding

$$|\Psi|^2 \simeq \sum_{i=1}^{\nu} |F_b^{(i)}(\vec{r})|^2. \tag{3.39}$$

As outlined above, the EMA and, hence Eq. (3.37), describes a single material subjected to an external perturbation. In order to describe nanostructures consisting of several materials, some modifications are necessary. A possible approach would be to employ a position-dependent effective mass tensor and band edge in Eq. (3.37) which account for the local material, together with suitable boundary conditions at the material interfaces [22]. More advanced envelope-function theories are based on transfer matrices to describe the connection rules at the material interfaces [23]. However, as the regular calculation of these matrices represents an extremely tedious task, the use of the standard approach [22] is encouraged. The effective mass equation becomes

$$\left[-\frac{\hbar^2}{2m_e} \vec{\nabla}^T \mathbf{M}^{(i)}(\vec{r}) \vec{\nabla} + \underbrace{E_{edge}(\vec{r}) + U^{ext}(\vec{r})}_{U^{eff}(\vec{r})} \right] F_b^{(i)}(\vec{r}) = E^{(i)} F_b^{(i)}(\vec{r}), \tag{3.40}$$

where $E_{edge}(\vec{r})$ is piecewise constant and determined by means of the local affinity $\chi(\vec{r})$ and band gap $E_g(\vec{r})$. Transitions from a material to the surrounding vacuum are approximated by hard wall conditions, i.e. zero Dirichlet conditions. A further simplification is given by the single-material approximation [24, 25, 26]. Within this approach, the effective mass tensor is position-independent and determined by the material which hosts the majority of the envelope function. This approximation is particularly suited for large band edge offsets and moderate confinements.

The case of degenerate band extrema, i.e. beyond the single-band picture, has not been discussed so far and will not be considered in this work. The $k \cdot p$ method for instance, accounts for band degeneracies within an envelope function framework similar to the EMA and is thoroughly described in Refs. [21, 10, 27].

The effective mass tensors related to the six minima of the silicon conduction band (see Fig. A.2) are diagonal, i.e. $\mathbf{M} = diag(1/m_x, 1/m_y, 1/m_z)$. Examples for the masses are given in Tab. 3.1. The entries of the six mass tensors are obtained by permuting the set $\{1/m_l, 1/m_t, 1/m_t\}$. Within the effective mass approximation, a silicon nanowire grown along the $\langle 100 \rangle$ direction, i.e. the x direction for instance, can be described by a potential [12] $U^{eff}(y,z)$ being independent of x. Therefore, the general form of the corresponding envelope functions reads

$$F(\vec{r}) = e^{ik_x x} \psi_n(y,z), \tag{3.41}$$

where $\psi_n(y,z)$ is determined by means of the Schrödinger equation

$$\left[-\frac{\hbar^2}{2m_e} \left(\frac{\partial}{\partial y} \frac{1}{m_y} \frac{\partial}{\partial y} + \frac{\partial}{\partial z} \frac{1}{m_z} \frac{\partial}{\partial z} \right) + U^{eff}(y,z) \right] \psi(y,z) = \epsilon \psi(y,z) \tag{3.42}$$

and $E = \epsilon_n + \hbar^2 k_x^2/(2m_e m_x)$ is the total energy. The special form of the kinetic operator in Eq. (3.42), i.e. effective mass sandwiched between the derivatives, is particularly important when the masses are position-dependent as this form ensures hermiticity of the Hamiltonian.

[12] This form for the effective potential can be used for any material and growth direction.

Accordingly, for a silicon quantum well with $\langle 100 \rangle$ oriented surfaces, the solution of the envelope equation has the form

$$F(\vec{r}) = e^{i(k_x x + k_y y)} \psi_n(z), \tag{3.43}$$

with

$$\left[-\frac{\hbar^2}{2m_e} \frac{1}{m_z} \frac{\partial^2}{\partial z^2} + U^{eff}(z) \right] \psi(z) = \epsilon \psi(z). \tag{3.44}$$

and a total energy $E = \epsilon_n + (\hbar^2/2m_e)(k_x^2/m_x + k_y^2/m_y)$. In this case, the surfaces are perpendicular to the z-axis.

3.6 Limitations of the EMA

3.6.1 Overview

The prerequisites which led to the effective mass equation (3.37), in particular the smoothness of the external potential, become questionable in the case of very small nanostructures, where the spatial inhomogeneities are of the order of the lattice constant. A quantitative comparison to direct calculations [13] for specific systems shall give more insights regarding the suitability of the EMA and suggest some efficient improvements.

3.6.2 Optical Lattice

An optical lattice is generated by superimposing two counterpropagating running-wave laser beams of wavelength λ in all three spatial directions, yielding a static potential $U(\vec{r})$ with periodicity $a = \lambda/2$

$$U(\vec{r}) = U_0 \sum_{i=1}^{3} \sin^2 \left(\frac{2\pi x_i}{\lambda} \right), \tag{3.45}$$

where U_0 is the amplitude of the laser beam. The potential (3.45) can be used to confine cold atoms. This system of trapped atoms resembles a crystal in the sense that the atoms move in a periodic potential. Several interesting effects such as the quantum phase transition from the superfluid to the Mott insulator phase in bosonic systems or ramping fermionic atoms across a Feshbach resonance can be investigated (For more details see Ref. [28] and references therein.) In this section, the optical lattice is used as a prototypic system to illustrate the shortcomings of the EMA. The computation of the bandstructure of the optical lattice (3.45) is relatively easy, as the corresponding Schrödinger equation separates in the three cartesian directions. It is therefore sufficient to consider the one-dimensional problem

$$\left[-\frac{\hbar^2}{2m_e} \frac{\partial^2}{\partial x^2} + U_0 \sin^2 \left(\frac{2\pi x}{\lambda} \right) \right] \psi = \epsilon \psi \tag{3.46}$$

which is solved by means of a Bloch ansatz and a secular equation of the form (A.14) in one dimension. The resulting eigenenergies $\epsilon_n(k)$ form the band structure of the optical lattice

$$E_{\vec{n}}(\vec{k}) = \epsilon_{n_1}(k_1) + \epsilon_{n_2}(k_2) + \epsilon_{n_3}(k_3), \tag{3.47}$$

where $\vec{n} = (n_1, n_2, n_3)$ and $\vec{k} = (k_1, k_2, k_3)$. A plot of $\epsilon_n(k)$ is given in Fig. 3.11. Increasing the amplitude U_0 results in a separation of the lowest band from all the others. The same holds

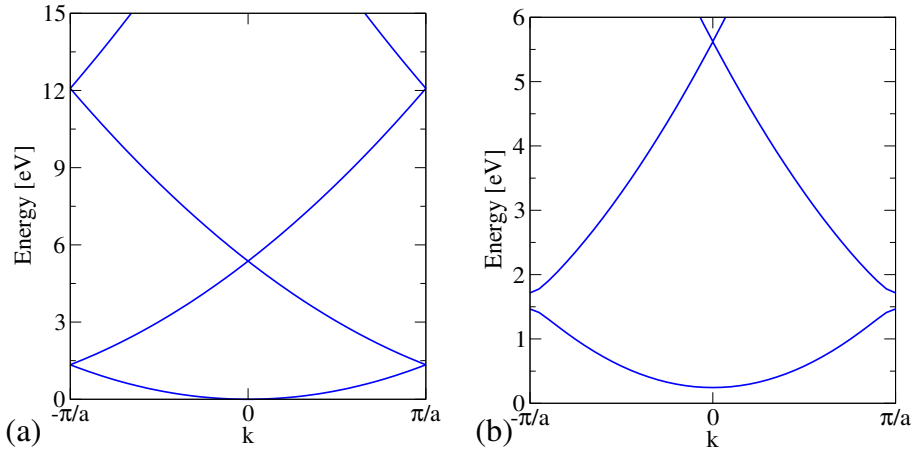

Figure 3.11: Band structure of the one-dimensional optical lattice (3.46) for $\{a = 10a_B, U_0 = 0\text{eV}\}$ (a) and $\{a = 10a_B, U_0 = 5\text{eV}\}$ (b).

for the three-dimensional band structure $E_{\vec{n}}(\vec{k})$. The effective mass equation is formulated for the lowest band having a minimum located at $\vec{k} = 0$ and an effective mass tensor of the form $\mathbf{M} = \text{diag}(1/m_0, 1/m_0, 1/m_0)$. The unit cell $\Omega_c = [-Na/2, Na/2]^3$ is employed to describe the optical lattice subjected to the perturbation

$$\widetilde{U}^{ext}(\vec{r}) = U_0^{ext}\left[1 - \chi_b(\vec{r})\right], \tag{3.48}$$

where χ_b is unity on $[-Nb/2, Nb/2]^3$ and zero elsewhere. The external potential is characterized by the width $d \equiv \sqrt{3}Nb$. In order to obtain a smooth confinement, the potential (3.48) undergoes a convolution with a Gaussian mollifier

$$U^{ext}(\vec{r}) \equiv \frac{1}{|\Omega_c|}\int_{\Omega_c} \widetilde{U}^{ext}(\vec{r}')F_\epsilon(\vec{r} - \vec{r}')d\vec{r}' \tag{3.49}$$

$$F_\epsilon(\vec{r}) \equiv \sum_{\vec{v}\in\Lambda(\Omega_c)} \frac{|\Omega_c|}{(2\pi)^{3/2}\epsilon^3}\exp\left[-\frac{1}{2}\frac{(\vec{r} - \vec{v})^2}{\epsilon^2}\right], \tag{3.50}$$

where $\Lambda(\Omega_c) \equiv \{Na\vec{n}|\vec{n} \in \mathbb{Z}^3\}$. In the following

$$\{N = 15, a = 10a_B, \epsilon = 0.2a_B, U_0^{ext} = 4.1eV, U_0 = 0.5eV\} \tag{3.51}$$

with $a_B = 0.052$nm. The resulting potentials $U(\vec{r})$ and $U^{eff}(\vec{r})$ are illustrated in Fig. 3.12. For the parameter set (3.51), the effective mass $m_0 = 1.006$ is close to unity and the band edge is $E_0 \equiv E_{\vec{0}} = 0.79$eV. The groundstate energies of the full Hamiltonian

$$\widehat{\mathbf{H}}^{\text{Full}} \equiv -\frac{\hbar^2}{2m_e}\Delta + U(\vec{r}) + U^{ext}(\vec{r}) \tag{3.52}$$

$$\tag{3.53}$$

[13]The terms *direct calculations* or *atomistic approach* refer to methods which explicitly account for the periodic background potential such as the tight-binding or pseudopotential framework. Comparisons to other envelope theories such as the kp method [27] are not carried out in this work.

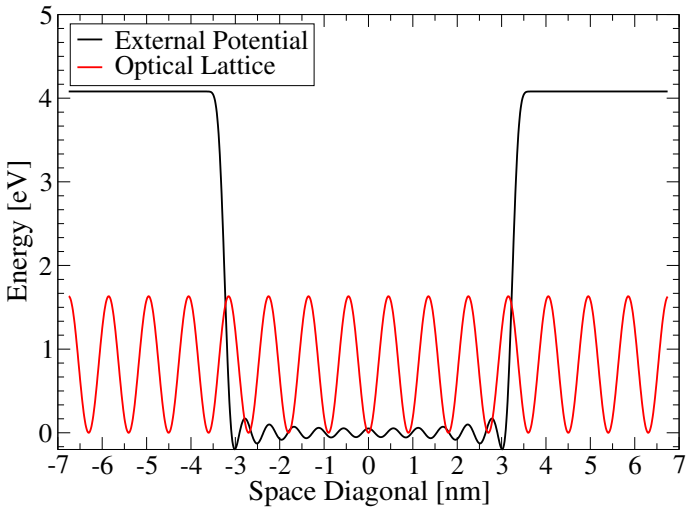

Figure 3.12: Potentials $U(\vec{r})$ and $U^{eff}(\vec{r})$ from Eqs. (3.45) and (3.49) for the parameter set (3.51). Plotted is a cut along the space diagonal of Ω_c.

and of the effective mass Hamiltonian

$$\widehat{\mathbf{H}}^{\mathrm{EMA}} \equiv -\frac{\hbar^2}{2m_e}\vec{\nabla}\mathbf{M}\vec{\nabla}^T + \underbrace{U^{ext}(\vec{r}) + E_0}_{U^{eff}(\vec{r})} \tag{3.54}$$

(3.55)

are computed for different widths d of the external potential and plotted in Fig. 3.13. For $d \lesssim$ 2nm, the EMA notably overestimates the groundstate energy compared to the full solution while in the remaining regime the agreement is quite satisfactory. The discrepancy for $d \lesssim 2$ reaches some tenths of an eV and can be related to one or more prerequisites for the EMA which are not fulfilled any longer such as the single-band picture or parabolicity[14]. Of course, these findings are related to the specific choice (3.51) and can not be generalized. These parameters roughly mimic silicon surrounded by an insulating material such as silicon dioxide. The qualitative behavior of the discrepancy however is assumed to be transferable to other systems as shall be seen in Secs. 3.6.3 and 3.6.4.

The groundstate energies of the Hamiltonians (3.52) and (3.54) are computed by means of the usual procedure, i.e. using a Bloch ansatz and solving the secular problem. In this particular case, the periodicity is broken in all three dimension and the secular problem has to be considered only at $\vec{k} = \vec{0}$ provided that Ω_c is sufficiently large. In order to obtain well converged groundstate energies, i.e. up to a meV, a notably larger plane wave basis set is required for the full Hamiltonian ($\sim$ 60000 plane waves) compared to the EMA case ($\sim$ 36000) [29]. This is due to strong oscillations of the optical lattice compared to the size of the unit cell or the width of

[14]The term *parabolicity* refers to the situation where a few Bloch states near the extremum of a single band are sufficient to describe the solution of the full problem. In this regime, the parabolic approximation is a good approximation to the band structure of the optical lattice. Conversely, *nonparabolicity* describes the case when Bloch states far from the minimum become relevant. In this case, the parabolic approximation provides too much dispersion around the extremum compared to the real band structure. More dispersion tendentiously yields higher energies. The changeover from the parabolic to nonparabolic regime is assumed to precede the involvement of further energy bands for the description of the full solution.

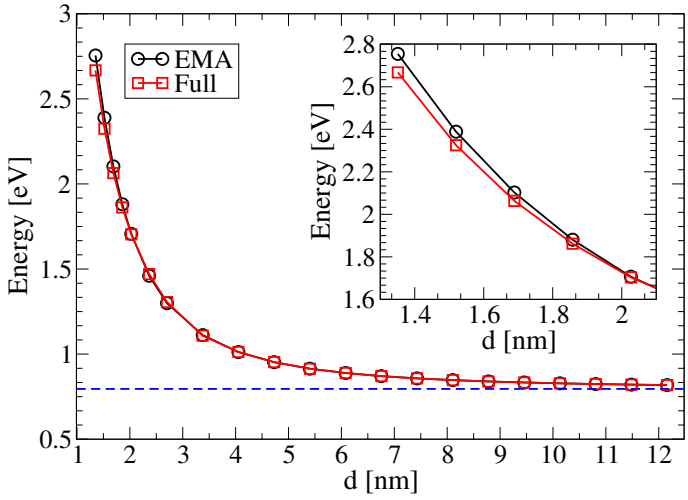

Figure 3.13: Lowest eigenvalues of the Hamiltonians (3.52) and (3.52) as a function of the width $d = \sqrt{3}Nb$. The dashed line denotes the edge $E_0 = 0.79$eV.

the external potential.

3.6.3 Silicon Nanowire

Circular and square shaped silicon nanowires grown along the $\langle 100 \rangle$ direction are now considered to investigate the suitability of the EMA. For each cross-sectional shape, a series of nanowires is constructed by varying the diameter. The present empirical tight-binding and pseudopotential method rely on different approaches to generate the wire structures and, more importantly, passivate the surface. Some details on the wire generation in the pseudopotential case are given in the caption of Fig. 3.4. Larger sized nanowires of these shapes can be obtained by increasing the number of unit blocks (green cells) before the surface is passivated. For the tight-binding case, a detailed description of the wire generation can be found in Ref. [17]. In particular, no hydrogen is used and the silicon skeleton is not distorted in the vicinity of the surface[15]. Thus, a direct comparison between the frameworks is not as straightforward as in the bulk case and it is mandatory to define an effective geometry parameter which can be generically assigned to a wire structure within both frameworks [16]. For the square case, a possibility [31] would be to add an atomic layer of vacuum to the silicon skeleton and measure the width. The approach [30] used in this work employs a cylinder of length a_{Si} made of silicon, whose diameter d_e is determined by the condition

$$\rho_{Si}\pi \left(\frac{d_e}{2}\right)^2 a_{Si} \stackrel{!}{=} N_{Si}, \tag{3.56}$$

[15] A tight-binding framework which explicitly employs hydrogen for the surface passivation is described in Ref. [30].

[16] The effective parameter can then be used to formulate the Dirichlet boundary conditions for a comparison with the EMA.

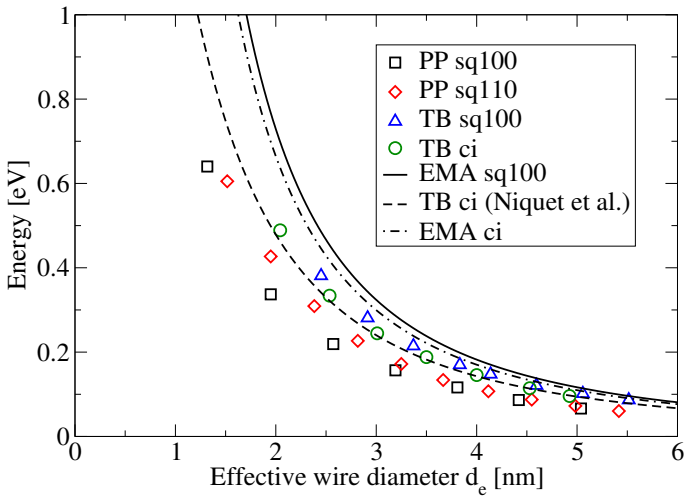

Figure 3.14: Conduction band edges of circular (ci) and square silicon nanowires with surfaces perpendicular to the ⟨100⟩ (sq100) and ⟨110⟩ (sq110) direction. Energies are computed by means of the present tight-binding (TB) and pseudopotential (PP) framework (see text). The EMA band edges $E^{sq100}(0,1,1)$ and $E^{ci}(0,1,1)$ for one of the Δ_4 valleys are denoted by the solid and dash-dot line respectively. Tight-binding results for the circular case by Niquet et al. [32] are plotted as well (dashed line).

i.e. the bulk silicon density $\rho_{Si} = 8/a_{Si}^3$ times the volume of the cylinder yields the number of silicon atoms N_{Si} in the wire unit cell. The diameter

$$d_e = a_{Si}\sqrt{\frac{N_{Si}}{2\pi}} \tag{3.57}$$

is the effective geometry parameter which can be assigned to wires with any cross-sectional shape. This procedure might seem counterintuitive for square nanowires. In this case, the wire can be thought to have an effective width

$$l_e = d_e\frac{\sqrt{\pi}}{2} = a_{Si}\sqrt{\frac{N_{Si}}{8}} \tag{3.58}$$

which yields the same cross-sectional area of a circle with diameter d_e.
The tight-binding framework is used to compute the band structures of circular and square nanowires with surfaces perpendicular to the ⟨100⟩ direction (cf. Fig. 3.4.b). The dispersions of square nanowires with ⟨110⟩ (cf. Fig. 3.4.c) and ⟨100⟩ oriented surfaces are computed by means of the present EPM [17]. For the latter approach, $co = 3.5/a_B$ and a vacuum width of 0.25nm has been used. A plot of the resulting conduction band minima, i.e. the edges, is given in Fig. 3.14 as a function of the effective diameter d_e. The energies tend to zero for $d_e \to \infty$ as they are shifted by the corresponding bulk edges.
The effective mass equation for a silicon nanowire grown along the ⟨100⟩ direction reduces to the two-dimensional Schrödinger problem (3.42) for the conduction band. The potential U^{eff} is

[17] Both, the tight-binding and pseudopotential approach will be referred to as the *atomistic* approaches in the following.

replaced by zero Dirichlet conditions at the effective boundary of the nanowire. For the square case with ⟨100⟩ oriented surfaces, the total energy reads

$$E^{sq100}(k_x, n, m) = \frac{\hbar^2}{2m_e}\left[\frac{k_x^2}{m_x} + \frac{\pi^2}{l_e^2}\left(\frac{n^2}{m_y} + \frac{m^2}{m_z}\right)\right] \tag{3.59}$$

with $n, m = 1, 2, \ldots$ and transverse eigenfucntions[18]

$$F^{sq100}_{k_x,n,m}(\vec{r}) = e^{ik_x x}\sin\left(\frac{n\pi y}{l_e}\right)\sin\left(\frac{m\pi z}{l_e}\right). \tag{3.60}$$

An analytic expression for the total energy can be given in the circular case for $m_y = m_z = \overline{m}$:

$$E^{ci}(k_x, n, m) = \frac{\hbar^2}{2m_e}\left(\frac{k_x^2}{m_x} + \frac{4\mu_{n,m}^2}{\overline{m}d_e^2}\right) \tag{3.61}$$

with the transverse states expressed in radial coordinates (r, ϕ)

$$F^{ci}_{k_x,n,m}(\vec{r}) = e^{ik_x x}J_n\left(\frac{2\mu_{n,m}r}{d_e}\right)e^{in\phi}, \tag{3.62}$$

where $J_n(x)$ is the Bessel function of the first kind and order n and $\mu_{n,m}$ is the m-th zero of $J_n(x)$. Unfortunately, for the more interesting case $m_y \neq m_z$, an analytical solution is not straightforward [33] and therefore, the band edge has to be computed numerically (see Sec. 5.2). The square nanowire with ⟨110⟩ surfaces as depicted in Fig. 3.4.c requires a rotation of the two-dimensional effective mass tensor in Eq. (3.42) by $\pi/4$. Conversely, the form (3.42) can be retained but the Dirichlet boundary has to be rotated by $\pi/4$. Both considerations yield the same total energy E^{sq110} which is equal to E^{sq100}. The energetically lowest conduction bands in the square cases are given by $E^{sq100}(k_x, 1, 1)$. The band edge is fourfold degenerate and related to the four mass tensors (Δ_4) with $m_y \neq m_z$. The two remaining mass tensors with $m_y = m_z$ are referred to as the Δ_2 valleys [19]. The subband structure obtained by the atomistic approaches shows some similarities. For this purpose, a closer look at the dispersion given in Fig. 3.5.a for instance might be helpful. The conduction band minimum at $k_x = 0$ is related to the effective mass dispersion from the Δ_4 valleys. Two further local minima of the conduction band located at $k_x \simeq \pm 0.3\pi/a_{Si}$ due to zone folding are related to the Δ_2 valleys [20]. These two minima move towards the gamma point ($k_x = 0$) as the wire diameter decreases. Note that zone folding is not accounted for by the effective mass approximation, where all subbands are located at $k_x = 0$. Furthermore, in the atomistic approaches a splitting [31] of the four lowest conduction bands can be observed while the effective mass model inherently assumes a perfect degeneracy for the Δ_4 valleys (circular and square case). This effect as well as the movement of the local minima are not further discussed in this work. The observations made for the square wires can be directly transferred to the circular case. In particular, the presence of Δ_4 and Δ_2 valleys and the relations to the corresponding band structure. A plot of the band edges $E^{sq100}(0, 1, 1)$ and $E^{ci}(0, 1, 1)$ for one of the Δ_4 valleys are given in Fig. 3.14. The effective masses $m_l = 0.891$ and $m_t = 0.201$ are taken from the TB case in Tab. 3.1. The EMA overestimates the band edge of the atomistic approaches as already seen in the case of the optical lattice. Possible reasons are mentioned in Sec. 3.6.2. However, for large diameters the atomistic results approach the effective

[18] Normalization is not considered yet.

[19] As a mass tensor is uniquely related to a minimum (valley), the terms will be used interchangeably.

[20] A more detailed discussion on zone folding in the case of silicon nanowires and the relation to the effective mass subband structure is given in Refs. [32, 34].

Table 3.2: Analytical fits (3.63) to the atomistic conduction band edges shown in Fig 3.14.

Wire type	K_c [eVnm^2]	a_c [nm]	b_c[nm^2]
PP sq100	2.626	2.469	-0.886
PP sq110	1.561	-1.015	1.820
TB sq100	2.922	0.605	0.175
TB ci	2.5875	0.383	0.337

mass curve. The tight-binding results for square and circular wires are not supposed to collapse on a single line since even in the EMA the band edges of a square and circular nanowire with the same effective diameter are different as can be seen explicitely from Eqs. (3.59) and (3.61) for $m_y = m_z$. The same is assumed for $m_y \neq m_z$. Conversely, the band edges of the $\langle 100 \rangle$ and $\langle 110 \rangle$ nanowires are supposed to match, which can be observed for the pseudopotential data at large diameters. This is partially due to the surface modification in the $\langle 100 \rangle$ case which prevents the wire from being perfectly square for small diameters. However, a notable discrepancy can be observed between the tight-binding and pseudopotential data even at large diameters. In particular, the pseudopotential approach still predicts notable band structure effects for large diameters. Comparing with a fit provided by Ref. [32], which includes the present tight-binding parametrization but employs a different passivation model, yields a good agreement with the present tight-binding results for the circular case [21]. A comparison with fully ab-initio simulations might shed more light on the discrepancy between the tight-binding and pseudopotential results [22].

Analytic fits of the form [32]

$$\epsilon_c(d_e) = \frac{K_c}{d_e^2 + a_c d_e + b_c} \tag{3.63}$$

for the atomistic band edges $\epsilon_c(d_e)$ shown in Fig. 3.14 are reported in Tab. 3.2. The form (3.63) is more accurate than the widely used fit K_c/d_e^β when a large range of diameters is considered. It behaves like $1/d_e^2$ in large structures, thus being consistent with the effective mass approximation which becomes more accurate in this regime.

Beside the overestimation of the subband energies, a further important band structure effect is the distortion of the atomistic wire dispersions compared to the effective mass band structures [31]. In other words, the effective masses extracted from the minima of the atomistic conduction bands located at $k_x \in \{0, \sim \pm 0.3\pi/a_{Si}\}$ differ from the corresponding bulk values. These *conduction* masses are extracted from the present band structures by means of

$$\frac{1}{m_c} = \frac{m_e}{\hbar^2} \frac{\partial^2}{\partial k_x^2} E(k_x)\Big|_{k_x=0}, \tag{3.64}$$

where $E(k_x)$ is the conduction band, and plotted in Fig. 3.15. As for the band edges, the conduction masses increase for a decreasing diameter.

For a quantitative investigation of the suitability of the single-band picture in the case of silicon nanowires, see Ref. [35].

[21] In Ref. [32] the nanowires are carved out of bulk silicon in a cylindrical shape.

[22] Since the bulk band structures agree fairly well (cf. Tab. 3.1), the source of discrepancy most probably resides in the surface termination models.

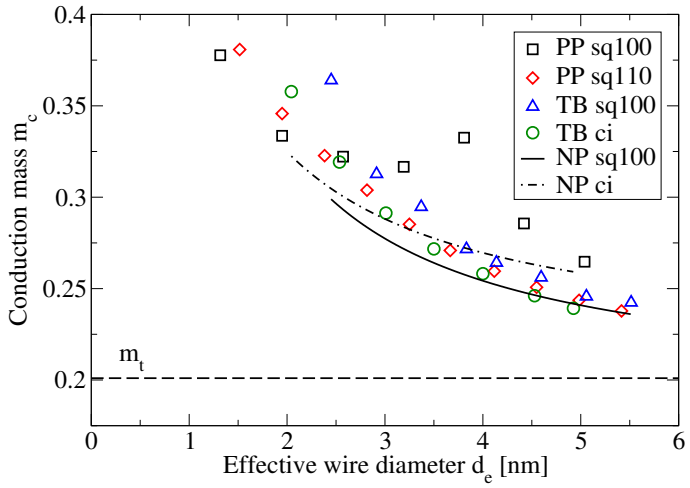

Figure 3.15: Conduction masses of the nanowires mentioned in the caption of Fig. 3.14. The masses are extracted by means of Eq. (3.64). The dashed line denotes the bulk value $m_t = 0.201$ (TB case in Tab. 3.1). For the TB sq100 and circular case, the conduction masses (3.80) computed by means of the nonparabolicity (NP) model described in Sec. 3.7 are denoted by the solid and dash-dot line respectively. See Sec. 5.2 for details on the extraction of the NP conduction mass.

3.6.4 Silicon Quantum Well

band structure effects for the case of the silicon quantum wells shown in Fig. 3.8 are investigated by means of the present pseudopotential method ($co = 3.5/a_B$). Similar to the wire case from Sec. 3.6.3, conduction band edges and conduction masses will be compared to the effective mass approach. To increase the height of the quantum wells, the silicon skeleton is enlarged by employing a higher number of unit blocks (green cells in Fig. 3.8) before the surface is passivated. The effective height h_e of the quantum well is determined by the condition

$$\rho_{Si} h_e |\mathbf{a}_1 \times \mathbf{a}_2| \stackrel{!}{=} N_{Si}, \tag{3.65}$$

where $\mathbf{a}_1$ and $\mathbf{a}_2$ span the projection of the unit cell in the (x, y)-plane as explained in Sec. 3.4.3 and $\mathbf{a}_1 \times \mathbf{z}_2$ is the cross product between $\mathbf{v}$ and $\mathbf{w}$. This definition is equivalent to the one used by Ref. [30]. Compared to the wire case, the conduction bands shown in Fig. 3.9 are more complicated [36]. For the quantum well with $\langle 100 \rangle$ oriented surfaces, the minima of the conduction band are located at $\mathbf{k} \equiv (k_x, k_y) = (\pm 0.85, \pm 0.85) 2\pi/a_{Si}$ and $\mathbf{k} = (0, 0)$, where the global minimum is at the latter point. In the $\langle 110 \rangle$ case, two global minima are in $\mathbf{k} = (0, \pm 0.85/\sqrt{2}) 2\pi/a_{Si}$ while two further local minima can be found in $\mathbf{k} = (\pm 0.15, 0) 2\pi/a_{Si}$. The position of the minima move towards zero as the well height decreases. This movement is not accounted for during the measurement of the edges and masses. Figure 3.16 shows the band edges of the quantum wells as a function of the effective height h_e. For large heights, the pseudopotential data agrees fairly well with the tight-binding results from Ref. [30].

Analytical fits of the form (3.63) are computed for the band edges plotted in Fig. 3.16 and reported in Tab. 3.3.

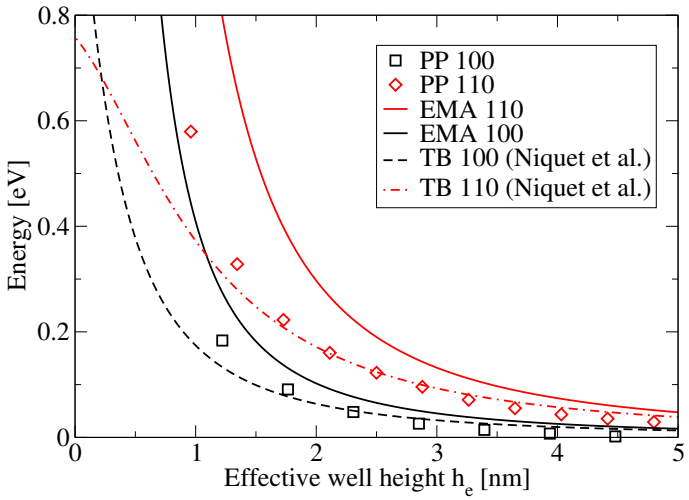

Figure 3.16: Conduction band edges of silicon quantum wells with ⟨100⟩ and ⟨110⟩ oriented surfaces. The corresponding energies computed by the EMA and tight-binding (TB) results by Niquet et al. [30] are plotted as well.

Table 3.3: Analytical fits (3.63) to the conduction band edges plotted in Fig 3.16 (pseudopotential data).

Well type	K_c [eVnm^2]	a_c [nm]	b_c[nm^2]
100	0.093	2.068	1.541
110	0.955	0.814	-0.051

The effective mass equation (3.44) yields the total energy

$$E^{(100)}(k_x,k_y,n)=\frac{\hbar^2}{2m_e}\left[\frac{k_x^2}{m_x}+\frac{k_y^2}{m_y}+\frac{\pi^2}{h_e^2}\frac{n^2}{m_z}\right] \tag{3.66}$$

and envelope functions

$$F^{(100)}_{k_x,k_y,n}(\vec{r})=e^{i(k_x x+k_y y)}\sin\left(\frac{n\pi z}{h_e}\right) \tag{3.67}$$

with $n=1,2,\ldots$ for the $\langle 100\rangle$ well. In the $\langle 110\rangle$ case, Eq. (3.41) has to be employed, whereas the effective mass tensor is rotated by $\pi/4$

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} \frac{1}{m_y} & 0 \\ 0 & \frac{1}{m_z} \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \tag{3.68}$$

and U^{eff} is replaced by Dirichlet conditions at $z\in\{0,h_e\}$. The total energy is

$$E^{(110)}(k_x,k_y,n)=\frac{\hbar^2}{2m_e}\left[\frac{k_x^2}{m_x}+\frac{2k_y^2}{m_y+m_z}+\frac{\pi^2}{n^2h_e^2}\frac{m_y+m_z}{2m_ym_z}\right] \tag{3.69}$$

with the envelope function

$$F^{(110)}_{k_x,k_y,n}(\vec{r})=e^{i(k_x x+k_y y+\gamma k_y z)}\sin\left(\frac{n\pi z}{h_e}\right), \tag{3.70}$$

where $\gamma=(m_y-m_z)/(m_y+m_z)$ and $n=1,2,\ldots$. The relations between the minima of the atomistic dispersions and the effective mass approximation can be straightforwardly derived by zone folding considerations as in the wire case. The energy $E^{(100)}(0,0,1)$ with $m_z=m_t$ corresponds to the band edge of the $\langle 100\rangle$ well, while $E^{(110)}(0,0,1)$ with $m_y=m_l$ and $m_z=m_t$ for instance yields the band edge of the $\langle 110\rangle$ well. The edges obtained by the EMA are plotted in Fig. 3.16, where $m_l=0.918$ and $m_t=0.191$ from Ref. [30] are employed.
As in the wire case, the conduction masses along the k_x-direction can be extracted at the global minima of the atomistic conduction bands. The masses extracted from the pseudopotential data are plotted in Fig. 3.17 as a function of the effective height h_e. For the $\langle 100\rangle$ case, the conduction masses converge very quickly to the corresponding bulk value m_t. The small peak in the $\langle 110\rangle$ curve at $h_e\simeq 3$ nm is most probably due to the movement of the conduction band minimum which is not accounted for during the measurement.
An investigation of silicon quantum wells within the $k\cdot p$ framework and the linear combination of bulk bands method can be found in Refs. [36, 25, 24]. In particular, a good agreement between the kp method and the effective mass approximation has been observed for the conduction band edges. Consideration on free-standing silicon quantum wells simulated by the empirical pseudopotential method are given in Ref. [37].

3.7 Nonparabolicity

In Secs. 3.6.2, 3.6.3, and 3.6.4 two important band structure effects, i.e. shortcomings of the EMA, have been highlighted, namely the overestimation of the subband energies and the underestimation of the conduction masses. More advanced approaches have been mentioned such as the $k\cdot p$ method or the linear combination of bulk bands. However, these methods still imply a large computational burden compared to the EMA and their implementation within the

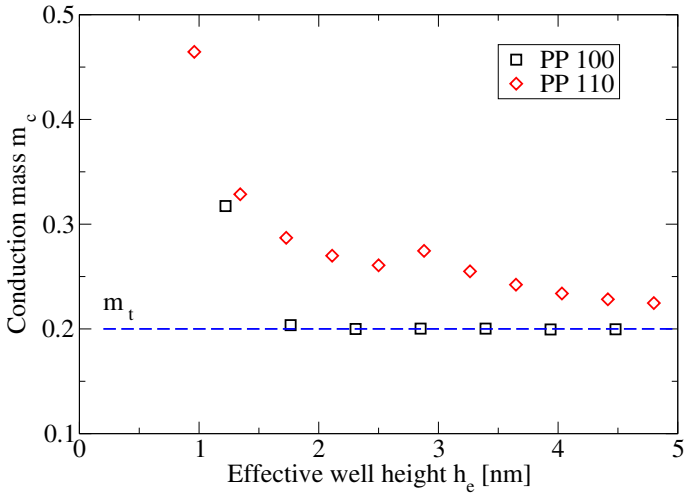

Figure 3.17: Conduction masses of the quantum wells described in Sec. 3.6.4. The masses are extracted by means of Eq. (3.64). The dashed line denotes the bulk value $m_t = 0.2$ from the Wa(3.5) case in Tab. 3.1.

transport framework employed in this work is not straightforward. It is therefore desirable to preserve the simplicity of the effective mass approximation while being able to capture the main band structure effects. For this purpose, the effective mass equation (3.36) is written in the more general form [23]

$$\left[\epsilon_b(-i\vec{\nabla}) + U^{ext}(\vec{r})\right] F_b(\vec{r}) = EF_b(\vec{r}), \tag{3.71}$$

where ϵ_b is the band containing the most relevant contributions to the full solution $\Psi(\vec{r})$ (see Sec. 3.5). Using a more sophisticated approximation for $\epsilon_b(\vec{k})$ than a simple parabola is expected to reduce the energy overestimation when Bloch states far from the minimum become relevant. For $U^{ext} \equiv 0$, the Bloch states $\psi_b(\vec{k},\vec{r})$ are solutions of Eq. (3.71) to the eigenvalue $\epsilon_b(\vec{k})$. Equation. (3.71) can be efficiently solved in momentum space by fast Fourier transforms for arbitrary potentials [38, 39]. A widely used nonparabolicity (NP) model for the conduction bands of covalent semiconductors [40, 41] is given by the dispersion

$$\epsilon^{NP}(\vec{k}) = \frac{1}{2\alpha}\left[\sqrt{1 + 4\alpha\frac{\hbar^2}{2m_e}\vec{k}^{\mathrm{T}}\,\mathbf{M}\,\vec{k}} - 1\right] + E_{edge}, \tag{3.72}$$

where α is referred to as the NP coefficient having the dimension of an inverse energy. Using the dispersion (3.72) in Eq. (3.71) yields a Schrödinger equation of the form

$$\left[\frac{1}{2\alpha}\left(\sqrt{1 - 4\alpha\frac{\hbar^2}{2m_e}\vec{\nabla}^{\mathrm{T}}\,\mathbf{M}\,\vec{\nabla}} - 1\right) + E_{edge} + U^{ext}(\vec{r})\right] F(\vec{r}) = EF(\vec{r}). \tag{3.73}$$

[23] This form appears during the derivation of the effective mass equation [10]. Using the parabolic expansion (3.33) one recovers Eq. (3.36).

Equation (3.73) can be further simplified for the case of silicon nanowires grown along the ⟨100⟩ direction [42, ?], where $\mathbf{M}$ is diagonal and $U^{ext}(\vec{r}) \to U^{ext}(y,z)$ is independent on x. Using the simple model potential

$$U^{ext}(y,z) = \begin{cases} 0, & (y,z) \in [0,D]^2 \\ \infty, & \text{otherwise} \end{cases} \tag{3.74}$$

in Eq. (3.73) yields the spectrum

$$E(n,m,k_x,\alpha) \equiv \frac{1}{2\alpha}\left[\sqrt{1+4\alpha\left(\epsilon_{\parallel}+\epsilon_{\perp}\right)}-1\right]+E_{edge} \tag{3.75}$$

with $n,m = 1,2,\ldots$, $\epsilon_{\parallel}(k_x) \equiv \hbar^2 k_x^2/(2m_e m_x)$, and $\epsilon_{\perp}(n,m) \equiv \hbar^2\pi^2(n^2/m_y + m^2/m_z)/(2m_e D^2)$. The key assumption is that in small nanowires, the transverse confinement prevails against the longitudinal energy, i.e. $\epsilon_{\parallel} \ll \epsilon_{\perp}$. This allows to expand the spectrum (3.75) in a Taylor series

$$\begin{aligned} E(n,m,k_x,\alpha) &= \frac{1}{2\alpha}\left[\sqrt{1+4\alpha\epsilon_{\perp}}-1\right]+E_{edge} \\ &+\frac{\epsilon_{\parallel}}{\sqrt{1+4\alpha\epsilon_{\perp}}}+O\left[\left(\frac{\epsilon_{\parallel}}{\epsilon_{\perp}}\right)^2\right]. \end{aligned} \tag{3.76}$$

The term $\epsilon_{\parallel}/\sqrt{1+4\alpha\epsilon_{\perp}}$ on the right hand side of Eq. (3.76) corresponds to a longitudinal kinetic energy with a renormalized mass

$$m_x \to m_x\sqrt{1+4\alpha\epsilon_{\perp}}. \tag{3.77}$$

This finding is consistent with the observations made in Sec. 3.6.3 and 3.6.4, where the conduction masses of the atomistic dispersions increase when the effective diameter is reduced. Neglecting terms of order $O[(\epsilon_{\parallel}/\epsilon_{\perp})^2]$ and fixing $\epsilon_{\perp}$ to $\epsilon_{\perp}^g \equiv \epsilon_{\perp}(1,1)$ in the longitudinal kinetic energy term in Eq. (3.76), yields

$$E(n,m,k_x,\alpha) \approx \frac{1}{2\alpha}\left[\sqrt{1+4\alpha\epsilon_{\perp}}-1\right]+E_{edge}+\frac{\epsilon_{\parallel}}{\sqrt{1+4\alpha\epsilon_{\perp}^g}}. \tag{3.78}$$

A modified Schrödinger equation which reproduces the spectrum (3.78) reads

$$\begin{aligned} &\Bigg\{-\frac{\hbar^2}{2m_e m_c}\frac{\partial^2}{\partial x^2} \\ &+\frac{1}{2\alpha}\left[\sqrt{1-4\alpha\frac{\hbar^2}{2}\left(\frac{\partial}{\partial y}\frac{1}{m_y}\frac{\partial}{\partial y}+\frac{\partial}{\partial z}\frac{1}{m_z}\frac{\partial}{\partial z}\right)}-1\right] \\ &+E_{edge}+U^{ext}(y,z)\Bigg\}F(\vec{r}) = EF(\vec{r}), \end{aligned} \tag{3.79}$$

where

$$m_c \equiv m_x\sqrt{1+4\alpha\epsilon_{\perp}^g}. \tag{3.80}$$

The form (3.79) notably simplifies the implementation of NP [24] within the transport framework described in chapter 4. The derivation of Eq. (3.79) can be repeated for Dirichlet conditions on arbitrary boundaries with appropriate modifications of $\epsilon_{\parallel}$, $\epsilon_{\perp}$, and consequently $\epsilon_{\perp}^g$ and m_c.

[24] The Hamiltonian in Eq. (3.79) is solely used within the single-material approximation in this work.

In summary, the NP models derived in this section are able to capture the two bandstructure effects observed in Secs. 3.6.3 and 3.6.4, namely the overestimation of the conduction band edge and the underestimation of the conduction mass. Both effects can be contemporarily accounted for by means of the NP coefficient, i.e. for an increasing α the conduction band edge decreases and the conduction mass increases while for $\alpha \to 0$ the results from the EMA are recovered. The extraction of NP coefficients α by means of atomistic nanowire band structures is thouroughly described in Sec. 5.2. Conversely, the use of bulk density of states for this extraction is shown in Sec. 3.8.

A simplified Schrödinger equation including NP can be derived for the case of silicon quantum wells with $\langle 100 \rangle$ oriented surfaces

$$\begin{aligned}
&\left\{ -\frac{\hbar^2}{2m_e\sqrt{1+4\alpha\epsilon_\perp^g}} \left(\frac{1}{m_x}\frac{\partial^2}{\partial x^2} + \frac{1}{m_y}\frac{\partial^2}{\partial y^2} \right) \right. \\
&+\frac{1}{2\alpha}\left[\sqrt{1-4\alpha\frac{\hbar^2}{2m_e}\left(\frac{\partial}{\partial z}\frac{1}{m_z}\frac{\partial}{\partial z} \right)} - 1 \right] \\
&\left. +E_{edge} + U^{ext}(z) \right\} F(\vec{r}) = EF(\vec{r}). \qquad (3.81)
\end{aligned}$$

Similar to the wire case, the effective masses in the free directions are increased due to NP. The extraction of α for the well case is described in Sec. 5.2.

A Taylor expansion of the nonparabolic dispersion (3.72) has already been employed by Jungemann et. al. [43]. They simplify the nonparabolic dispersion to obtain a fully parabolic Schrödinger equation for the case of silicon quantum wells. More recently, the unaltered dispersion (3.72) has been used for quantum wires [44] and dots [45]. Beside the present NP models, improved effective mass theories [46] which employ a finite number of terms to approximate the silicon conduction band have been successfully applied to well, wires, and dots.

3.8 Calculation of Charge Densities

3.8.1 Ideal Semiconductors

The electron density in an ideal crystal reads [10]

$$\begin{aligned}
\rho^{ideal} &= \frac{2}{|\Omega_{\vec{N}}|} \sum_{n,\vec{k}\in\widetilde{\mathbf{\Gamma}}^{\mathrm{red}}_{\vec{N}}} \frac{1}{1+e^{\frac{\epsilon_n(\vec{k})-E_F}{k_B T}}} \\
&= \int_{\mathbb{R}} \mathcal{G}^{ideal}(E) f(E) dE \qquad (3.82)
\end{aligned}$$

with $\widetilde{\mathbf{\Gamma}}^{\mathrm{red}}_{\vec{N}}$ from Eq. (3.19). The density of states

$$\mathcal{G}^{ideal}(E) = \frac{2}{|\Omega_{\vec{N}}|} \sum_{n,\vec{k}\in\widetilde{\mathbf{\Gamma}}^{\mathrm{red}}_{\vec{N}}} \delta(E-\epsilon_n(\vec{k})) \qquad (3.83)$$

contains informations about the possible quantum states of the system while the information about the statistical occupation of these states is exclusively contained in the Fermi distribution

function $f(E)$. In the macroscopic limit, i.e. for $\vec{N} \to \infty$, the density of states becomes [14]

$$\mathcal{G}^{ideal}(E) = \frac{1}{4\pi^3} \sum_n \int_{\widetilde{\Omega}_c} \delta(E - \epsilon_n(\vec{k})) d\vec{k} \tag{3.84}$$

$$= \frac{1}{4\pi^3} \sum_n \int_{S_n(E)} \frac{1}{|\nabla_{\vec{k}} \epsilon_n(\vec{k})|} dS, \tag{3.85}$$

where $S_n(E)$ is a surface defined by the condition $\epsilon_n(\vec{k}) = E$, dS is the corresponding surface element, and $\widetilde{\Omega}_c$ is the unit cell (3.14) in the reciprocal lattice.
Within the EMA, the valleys of the conduction (c) and valence (v) bands of a semiconductor are approximated by parabolic dispersions

$$\epsilon_{c/v}^{\mathrm{EMA}}(\vec{k}) = E_{c/v} \pm \frac{\hbar^2}{2m_e} \vec{k}^T \mathbf{M}_{c/v}^{(i)} \vec{k}, \tag{3.86}$$

where the $\mathbf{M}_{c/v}^{(i)}$ are the effective mass tensors of the i^{th} valley and are generally dense. In the case of silicon, the valence bands are warped in the vicinity of the maximum and an average is necessary to obtain a parabolic approximation, i.e. the $\mathbf{M}_v^{(i)}$, as mentioned in appendix A.0.2. The densities of states of the dispersions (3.86) can be calculated analytically

$$\begin{aligned} \mathcal{G}_{c/v}^{\mathrm{EMA},i}(E) &= \frac{2}{|\Omega_{\vec{N}}|} \sum_{\vec{k} \in \widetilde{\mathbf{\Gamma}}_{\vec{N}}} \delta(E - \epsilon_{c/v}^{\mathrm{EMA},i}(\vec{k})) \\ &\to \frac{1}{4\pi^3} \int_{\mathbb{R}^3} \delta\left(E - E_{c/v} \mp \frac{\hbar^2}{2m_e} \vec{k}^T \mathbf{M}_{c/v}^{(i)} \vec{k}\right) d\vec{k} \\ &= \frac{1}{2\pi^2 \sqrt{\det(|\mathbf{M}_{c/v}^{(i)})|}} \left(\frac{2m_e}{\hbar^2}\right)^{\frac{3}{2}} \sqrt{|E - E_{c/v}|}. \end{aligned} \tag{3.87}$$

In the presence of $\#_c$ conduction band and $\#_v$ valence band valleys, the total density of states reads

$$\begin{aligned} \mathcal{G}_{c/v}^{\mathrm{EMA}}(E) &= \sum_{i=1}^{\#_{c/v}} \mathcal{G}_{c/v}^{\mathrm{EMA},i}(E) \\ &\equiv \frac{1}{2\pi^2} \left(\frac{2m_e m_{c/v}}{\hbar^2}\right)^{\frac{3}{2}} \sqrt{|E - E_{c/v}|} \end{aligned} \tag{3.88}$$

The *density of states effective mass* $m_{c/v}$ collects the contributions of all valleys. Similarly, the nonparabolic dispersion (3.72) can be used to compute the density of states of the conduction band

$$\begin{aligned} &\mathcal{G}_c^{\mathrm{NP},i}(E) \\ &= \frac{1}{4\pi^3} \int_{\mathbb{R}^3} \delta\left(E - E_c - \frac{1}{2\alpha}\left[\sqrt{1 + 4\alpha \frac{\hbar^2}{2m_e} \vec{k}^T \mathbf{M}_c^{(i)} \vec{k}} - 1\right]\right) d\vec{k} \\ &= \frac{1}{2\pi^2 \sqrt{\det(\mathbf{M}_c^{(i)})}} \left(\frac{2m_e}{\hbar^2}\right)^{\frac{3}{2}} [1 + 2\alpha(E - E_c)] \\ &\quad \times \sqrt{(E - E_c)[1 + \alpha(E - E_c)]} \end{aligned} \tag{3.89}$$

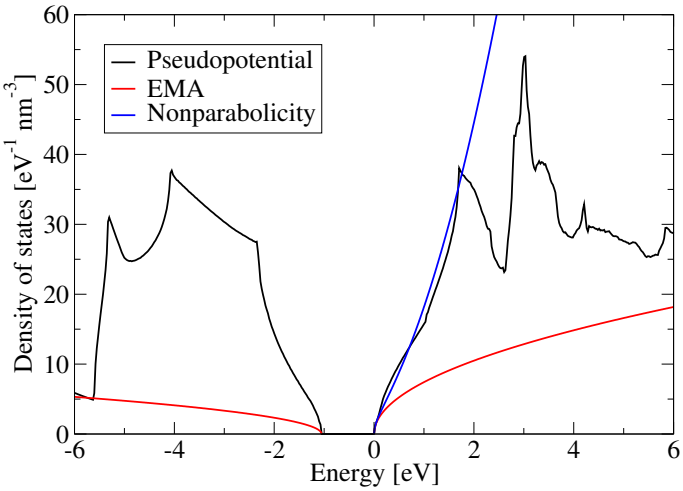

Figure 3.18: Density of states (B.4) of silicon obtained by means of the pseudopotential of Chelikowski et al. described in appendix A.0.2. The corresponding parabolic (3.88) and nonparabolic (3.90) densities of states are shown as well.

and consequently

$$\begin{aligned} \mathcal{G}_c^{\mathrm{NP}}(E) &= \sum_{i=1}^{\#_c} \mathcal{G}_c^{\mathrm{NP},i}(E) \\ &= \frac{1}{2\pi^2}\left(\frac{2m_e m_c}{\hbar^2}\right)^{\frac{3}{2}} [1+2\alpha(E-E_c)] \\ &\quad \times\sqrt{(E-E_c)[1+\alpha(E-E_c)]}. \end{aligned} \tag{3.90}$$

A plot of $\mathcal{G}, \mathcal{G}_{c/v}^{\mathrm{EMA}}, \mathcal{G}_c^{\mathrm{NP}}$ from Eqs. (B.4), (3.88), and (3.90) respectively can be found in Fig. 3.18 for the case of silicon. The pseudopotential of Chelikowski et al. described in appendix A.0.2 is used to compute the band structure $\epsilon_n(\vec{k})$ and consequently $\mathcal{G}(E)$ for a reciprocal space discretization of $M = 70$ (see appendix B). These results are in quantitative agreement with the data given in Ref. [47]. To be consistent with the pseudopotential calculations, $m_c = 1.06$ and $m_v = 0.5$ are used for $\mathcal{G}_c^{\mathrm{EMA}}$ and $\mathcal{G}_v^{\mathrm{EMA}}$ respectively. The NP coefficient α is obtained by fitting the expression (3.88) to the approximate density of states (3.90). Employing a least square minimization procedure on the range $[E_c, E_c + 0.6\mathrm{eV}]$ yields $\alpha = 0.51\mathrm{eV}^{-1}$.

3.8.2 Doped Semiconductors

A fraction of the semiconductor atoms are now replaced by dopands providing single-ionizible donor and acceptor levels E_A and E_D. With N_A^- and N_D^+ being the concentrations of the acceptors and donors respectively, the neutrality condition which fixes the Fermi energy E_F reads [10]

$$0 = n_c - p_v - N_D^+ + N_A^- \tag{3.91}$$

with

$$n_c = \int_{E_c}^{\infty} \mathcal{G}_c^{\text{real}}(E) f(E) dE \tag{3.92}$$

$$p_v = \int_{-\infty}^{E_v} \mathcal{G}_v^{\text{real}}(E)[1 - f(E)] dE. \tag{3.93}$$

Complete ionization of the dopands is assumed in Eq. (3.91). The density of states $\mathcal{G}_{c/v}^{\text{real}}$ of the doped (real) semiconductor appearing in Eqs. (3.92) and (3.93) can be approximated [10] by $\mathcal{G}_{c/v}^{\text{EMA}}$ from Eq. (3.88). In this case, the electron density in the conduction band n_c and the hole density in the valence band p_v become

$$n_c = 2 \left(\frac{m_e m_c k_B T}{2\pi\hbar} \right)^{3/2} \mathcal{F}_{1/2} \left(\frac{E_F - E_c}{k_B T} \right) \tag{3.94}$$

and

$$p_v = 2 \left(\frac{m_e m_v k_B T}{2\pi\hbar} \right)^{3/2} \mathcal{F}_{1/2} \left(\frac{E_v - E_F}{k_B T} \right). \tag{3.95}$$

The complete *Fermi-Dirac integral* $\mathcal{F}_j$ is defined as

$$\mathcal{F}_j(x) = \frac{1}{\Gamma(j+1)} \int_0^{\infty} \frac{t^j dt}{\exp(t - x) + 1} \tag{3.96}$$

with Γ being Euler's gamma function [48]. For values $j < -1$ the integral (3.96) does not converge but analytic continuations can be defined. For further details on Fermi-Dirac integrals and routines to compute these integrals to almost any precision see Ref. [49] and references therein. Analytic expressions for $\mathcal{F}_j$ exist for integers $j \leq 0$. The $j = 0$ case reads

$$\mathcal{F}_0(x) = \ln[1 + \exp(x)] \tag{3.97}$$

and the application of the property

$$\mathcal{F}_j'(x) = \mathcal{F}_{j-1}(x) \tag{3.98}$$

allows one to write $\mathcal{F}_j$ for negative integers as a rational function of $\exp(x)$. Rational approximations for efficient implementations for the cases $j \in \{-3/2, -1/2, 1/2\}$ can be found in Refs. [50, 51] while pre-assembled routines are provided by the GNU scientific library [52].

The densities n_c and p_v are derived under the assumption of spatial isotropy. Under real circumstances, inhomogeneities due to the presence of different materials and externally applied fields have to be considered. In the simplest case, local equilibrium is assumed and spatial changes of the material properties are accounted for by introducing a position dependence in the quantities $m_{c/v}$ and $E_{c/v}$ [25] as in the effective mass description of nanostructures explained in Sec. 3.5. However, when quantization effects become relevant, the local equilibrium approximation is expected to fail. In this case the envelope functions of the entire nanostructure have to

[25] In the presence of an external perturbation U^{ext}, the band edges are replaced by $U^{eff} = E_{c/v} + U^{ext}$.

be calculated. For nanowires, the densities read

$$n_c^{1DEG}(\vec{r}) = \sqrt{\frac{2m_e m_x k_B T}{\pi\hbar^2}} \sum_n |\psi_n(y,z)|^2 \mathcal{F}_{-1/2}\left(\frac{E_F - \epsilon_n}{k_B T}\right) \tag{3.99}$$

$$p_v^{1DEG}(\vec{r}) = \sqrt{\frac{2m_e m_x k_B T}{\pi\hbar^2}} \sum_n |\psi_n(y,z)|^2 \mathcal{F}_{-1/2}\left(\frac{-E_F - \epsilon_n}{k_B T}\right). \tag{3.100}$$

The expressions (3.99) and (3.100) are valid for arbitrary mass tensors. In this case, the mass m_x has to be chosen as described in appendix D. For the case of quantum wells, the densities read

$$n_c^{2DEG}(\vec{r}) = \frac{\sqrt{m_x m_y} m_e k_B T}{\pi\hbar^2} \sum_n |\psi_n(z)|^2 \mathcal{F}_0\left(\frac{E_F - \epsilon_n}{k_B T}\right) \tag{3.101}$$

$$p_v^{2DEG}(\vec{r}) = \frac{\sqrt{m_x m_y} m_e k_B T}{\pi\hbar^2} \sum_n |\psi_n(z)|^2 \mathcal{F}_0\left(\frac{-E_F - \epsilon_n}{k_B T}\right). \tag{3.102}$$

The densities n_c^{1DEG}, p_v^{1DEG}, n_c^{2DEG}, and p_v^{2DEG} are always related to a single valley. In a multi-valley material, the total density is obtained by collecting the contributions from all valleys [26] The quantities $\{\psi_n, \epsilon_n\}$ in Eqs. (3.99), (3.100), (3.101), and (3.102) are the eigensolutions of the Schrödinger equation for the confined direction. Several examples related to the conduction band minimum of silicon have been presented so far. Given a valence band maximum described by the tensor $\mathbf{M}_v = \mathrm{diag}(-1/m_x, -1/m_y, -1/m_z)$, the envelope equation for the confined direction in a quantum well [27] reads

$$\left[-\frac{\hbar^2}{2m_e}\frac{\partial}{\partial z}\frac{1}{m_z}\frac{\partial}{\partial z} - U^{eff}(z)\right]\psi(z) = \epsilon\psi(z). \tag{3.103}$$

The resulting eigensystem $\{\psi_n, \epsilon_n\}$ is used within Eq. (3.102).
Generally, the near-gap energy bands are degenerate around the extremal points such as for instance the valence bands of silicon. As mentioned in Sec. 3.5, the $k \cdot p$ method can be used to compute the states and thus the densities in this case. However, the single-band picture, and thus the effective mass equation, can be retained by means of a very rough approximation, i.e. by lumping the valence bands into a single effective mass tensor $\mathbf{M}_v = \mathrm{diag}(-1/m_v, -1/m_v, -1/m_v)$ with m_v being the density of states effective mass.
The neutrality conditions for the wire and well cases are discussed in chapter 4.

3.9 Summary

An review of the basic theory and an overview of different methods for the calculation of band structures have been given at the beginning of this chapter with a particular focus on the empirical pseudopotential (EPM) and the ETB. While the ETB is attractive from a computational

[26] In this way, the coupling between diffent valleys is fully suppressed as already mentioned in Sec. 3.5.
[27] The wire case follows analogously.

point of view, the EPM is numerically more involved but is supposed to describe the charge density in a more physical way. For the bulk case, both methods yield a satisfactory agreement. Unfortunately, such an agreement can not be observed for the case of nanostructures, whereas the discrepancy is more pronounced for nanowires. A possible source for this discrepancy has been attributed to different treatment of the surface. Comparisons with more advanced ab-initio methods could shed more light on this behaviour.
Electronic states in nanostructures have been calculated by the EMA and compared to results from the ETB and the EPM. The shortcomings of the EMA compared to the atomistic approaches are referred to as band structure effects. Two such effects, which turn out to be the most relevant in qunatum transport simulations (see chapter 4), are the overestimation of the conduciton band edge and the underestimation of the conduction mass. Simple NP models being parametrized by a single coefficient incorporate these two effects and can be used to improve the EMA. The chapter is concluded with some considerations on particle densities in nanostructures at thermodynamic equilibrium.

Chapter 4

Simulation of Ballistic Transport

4.1 Introduction

With the EMA described in the previous chapter, the basis for an efficient quantum-mechanical treatment of electronic systems in semiconductor devices has been laid. Sticking to the level of quantum mechanics, the transport of electrons is then most simply treated in the absence of phase breaking scattering processes, being referred to as *quantum-ballistic* transport (QBT). The validity of the ballistic approximation strongly depends on the dimensions of the devices which should not exceed the coherence length of the charge carriers. Ignoring this prerequisite, QBT can still be used to investigate the theoretical performance limit of a device. For given material and contact specifications, the Schrödinger and Poisson equation of the device are repeatedly solved until the electrostatic potential and the charge density become consistent. Upon successful exit, the current is calculated.
The Landauer-Büttiker formalism, which is widely used to describe QBT, is briefly outlined at the beginning of this chapter. Subsequently, some historical notes on the simulator for nanodevices SIMNAD are given. Fundamental numerical issues are addressed such as the discretization of the Schrödinger and Poisson equation as well as iterative schemes to achieve self-consistency. In this context, the box integration method and the finite element method (FEM) are described. In order to improve the EMA, a spectral method to include the nonparabolicity models elaborated in the previous chapter, is presented. Finally, the impact of band structure effects on current characteristics is investigated.

4.2 Devices

Two types of devices are considered in this work, namely nanowire and planar field effect transistors (FETs). Since a nanowire FET hosts a quasi one-dimensional electron gas (1DEG), the terms 1DEG and nanowire will be used interchangeably. Analogously, the term 2DEG is related to a planar FET. Both types of devices are schematically shown in Fig. 4.1. The gate contact is described by a Dirichlet boundary condition for the Poisson equation while zero von Neumann conditions are imposed on the remaining surface. The prismatically shaped semiconductor region is fully enclosed by the dielectric and is divided in three segments along the transport direction x, i.e. the source, channel, and drain region. Note that the overall prismatic shape of the devices is not mandatory for the present simulation framework, i.e. the source and drain regions can be flared out for instance. In particular, the oxide thickness t_{ox} does not have to be uniform along

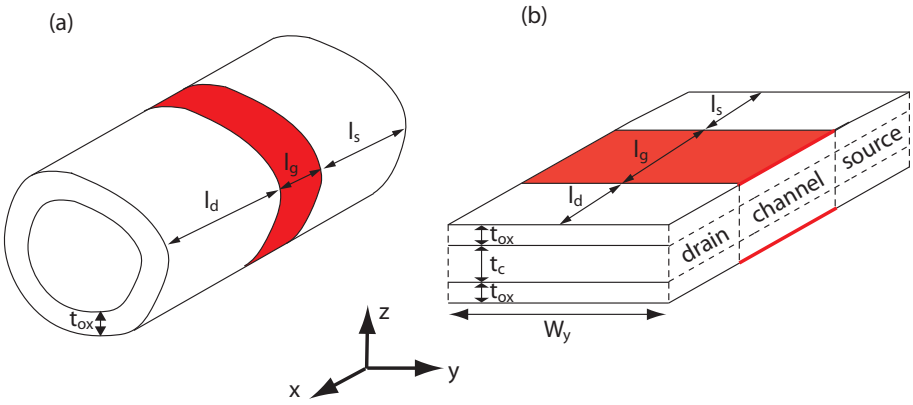

Figure 4.1: Schematic representation of the devices considered in this work, i.e. nanowire FETs (a) and planar FETs (b). The red region denotes the gate contact of length l_g which is placed precisely above the channel region and coincides with the Dirichlet boundary $\partial\Omega_D$. The remaining surface is referred to as the von Neumann boundary $\partial\Omega_N$, i.e. $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N$ is the total boundary of the simulation domain Ω. The lengths of the source and drain regions are denoted by l_s and l_d respectively. Finally, the width and the channel thickness of the planar FET are denoted by W_y and t_c, respectively and t_{ox} is the oxide thickness.

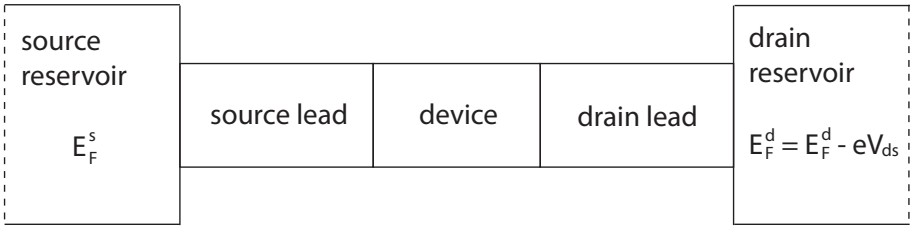

Figure 4.2: A device coupled to semi-infinite leads which are in turn coupled to reservoirs being at thermodynamic equilibrium. This is the picture used within the Landauer-Büttiker formalism.

the surface.

In the following, the semiconductor region consists of silicon grown along the $\langle 100 \rangle$ direction and silicon dioxide is used as dielectric. The confinement direction in the planar FETs coincides with the $\langle 100 \rangle$ direction.

4.3 Landauer-Büttiker Formalism

In this formalism, the current flowing through a conductor is related to the probability that a charge carrier can be transmitted through this conductor. A detailed review is given in Refs.[53, 54]. For a two-terminal device, the schematic situation is shown in Fig. 4.2. The device is coupled to semi-infinite terminals (leads) which are assumed to be perfect wave guides. These leads are

coupled to particle reservoirs being in thermodynamic equilibrium [1]. The Fermi energies are E_F^s and $E_F^s - eV_{ds}$ for the source (s) and drain (d) reservoir respectively, where V_{ds} is the applied forward bias.
In the following, the potential $U^{eff}(\vec{r})$ describes the device and the affiliated leads. The effective mass Schrödinger equation is assumed to be of the form

$$\left[-\frac{\hbar^2}{2m_e}\vec{\nabla}\mathbf{M}\vec{\nabla}^T + U^{eff}(\vec{r})\right] F(\vec{r}) = EF(\vec{r}) \tag{4.1}$$

with $\mathbf{M} = \mathrm{diag}\{1/m_x(x), 1/m_y(\vec{r}), 1/m_z(\vec{r})\}$. The mass m_x is independent on (y,z) within the device [2] and fully position-independent in the leads. In the leads of a nanowire, the potential U^{eff} depends only on (y,z). The eigenstates of the Hamiltonian (4.1) for this case have been thoroughly described in Sec. 3.5. The general form reads

$$F(\vec{r}) = \sum_i \left(a_i e^{ik_i x} + b_i e^{-ik_i x}\right) \psi_i(y,z) \tag{4.2}$$

with $k_i = \sqrt{2m_e m_x(E-\epsilon_i)}/\hbar$ and $\{\epsilon_i, \psi_i\}$ are the eigensolutions of the transverse Schrödinger equation [3]. The solution (4.2) is continued into the device until reaching the opposite terminal, where the form of the solution equals the one given in Eq. (4.2). A suitable basis set for the lead-device-lead system is given by the so called *scattering states* [55]. Such a state $F_j^{(t)}(E,\vec{r})$ is injected into a single mode j from a given terminal $t \in \{s,d\}$ at a fixed total energy E [4]. If $a_i = \delta_{i,j}$ in the source for instance, b_i has to be zero in the drain lead for all i. The density of electrons in the conduction band is given by

$$\begin{aligned} n_c^{1DEG}(\vec{r}) &= \frac{\sqrt{m_e}}{\sqrt{2}\pi\hbar} \sum_\nu \sum_t \sqrt{m_x^{(t,\nu)}} \sum_j \int_{\epsilon_j^{(t,\nu)}}^{\infty} |F_j^{(t,\nu)}(E,\vec{r})|^2 \\ &\quad \times \frac{1}{1+e^{(E-E_F^{(t)})/k_BT}} \frac{dE}{\sqrt{E-\epsilon_j^{(t,\nu)}}} \end{aligned} \tag{4.3}$$

where ν is a label for the different valleys, if present, of the conduction band giving rise to different mass tensors in the Hamiltonian (4.1) and $\epsilon_j^{(t,\nu)}$ is the j-th subband energy of the valley ν in terminal t. The Landauer-Büttiker formalism gives rise to a somewhat unfamiliar picture of non-equilibrium as the Fermi level is not strictly defined throughout the device. Instead, each charge carrier is linked to a single reservoir and thus a single Fermi level. The charge density (4.3) is the sum of the source and drain populations. The current through the nanowire FET is given by

$$\begin{aligned} I_d^{1DEG} &= \frac{e}{\pi\hbar} \sum_\nu \int_{\min_j(\epsilon_j^{(s,\nu)})} \mathcal{T}^{(1DEG,\nu)}(E) \\ &\times \left[\frac{1}{1+e^{(E-E_F^{(d)})/k_BT}} - \frac{1}{1+e^{(E-E_F^{(s)})/k_BT}}\right] dE, \end{aligned} \tag{4.4}$$

[1] The lead is not semi-infinite in the strict sense but merely much longer than the device. Furthermore, the leads are assumed to be perfect prolongations of the source and drain regions of the device. More involved injection conditions, i.e. from 2DEGs into nanowire FETs, have been studied within the Green's functions formalism [17].
[2] This assumption notably simplifies the implementation of the scattering matrix approach (see Sec. 4.5).
[3] The energies ϵ_i are also referred to as *subband* energies.
[4] The parameters j, t, and E dictate the *scattering boundary conditions* which are used within the scattering matrix formalism from Sec. 4.5 to compute the state $F_j^{(t)}(E,\vec{r})$ inside the device.

where $\mathcal{T}^{(1DEG,\nu)}(E)$ is the total transmission probability contributed from valley ν. This probability can be explained by means of the current

$$\mathcal{J}_x[F](x) \equiv \int_{\mathbb{R}^2} \underbrace{\frac{i\hbar}{2m_e m_x}\left[\left(\frac{\partial F}{\partial x}\right)^* F - \left(\frac{\partial F}{\partial x}\right) F^*\right]}_{\{\vec{J}[F](\vec{r})\}_x} dydz \tag{4.5}$$

where $\{\vec{J}[F](\vec{r})\}_x$ is the x-component of the current

$$\vec{J}[F](\vec{r}) = \frac{i\hbar}{2m_e m_x}\left[\left(\vec{\nabla}F\right)^* F - \left(\vec{\nabla}F\right) F^*\right]. \tag{4.6}$$

Applying the operator (4.5) to the scattering state $F_j^{(s,\nu)}(E,\vec{r})$ for instance, yields $\hbar k_j^{(s,\nu)}(E)/(m_e m_x^{(s,\nu)})$ in the source and

$$\frac{\hbar}{m_e m_x^{(d,\nu)}} \sum_i k_i^{(\nu)}(E)|a_i^{(\nu)}(E)|^2 \tag{4.7}$$

in the drain, where $\{k_i^{(\nu)}(E), a_i^{(\nu)}(E)\}_i$ are the parameters belonging to $F_j^{(s,\nu)}(E,\vec{r})$ in the drain [5]. The transmission probability $\mathcal{T}_j^{(1DEG,\nu)}(E)$ provided by this state is given by the quotient of the drain and source current and the total probability is then obtained by summing up the contributions from all states, i.e. $\mathcal{T}^{(1DEG,\nu)}(E) = \sum_j \mathcal{T}_j^{(1DEG,\nu)}(E)$ [6]. For the 2DEG case, the density and current expressions can be computed analogously [54]

$$\begin{aligned} n_c^{2DEG}(\vec{r}) &= \frac{m_e\sqrt{k_B T}}{\hbar^2\pi^{3/2}} \sum_\nu \sum_t \sqrt{m_x^{(t,\nu)} m_y^{(t,\nu)}} \int_{\min_j(\epsilon_j^{(t,\nu)})} \\ &\quad \sum_j \mathcal{F}_{-1/2}\left(\frac{E_F^{(t)} - \widetilde{E}}{k_B T}\right) |F_j^{(t,\nu)}(\widetilde{E},x,z)|^2 \frac{d\widetilde{E}}{\sqrt{\widetilde{E} - \epsilon_j^{(t,\nu)}}} \end{aligned} \tag{4.8}$$

$$\begin{aligned} I_d^{2DEG} &= \frac{W_y\sqrt{2k_B T}}{\hbar^2\pi^{3/2}} \sum_\nu \sqrt{m_y^{(s,\nu)}} \int_{\min_j(\epsilon_j^{(s,\nu)})} \mathcal{T}^{(2DEG,\nu)}(E) \\ &\quad \times \left[\mathcal{F}_{-1/2}\left(\frac{E_F^{(s)} - \widetilde{E}}{k_B T}\right) - \mathcal{F}_{-1/2}\left(\frac{E_F^{(d)} - \widetilde{E}}{k_B T}\right)\right] d\widetilde{E}, \end{aligned} \tag{4.9}$$

where W_y is the width of the 2DEG as shown in Fig. 4.1.b. Note that $\widetilde{E}$ in Eqs. (4.8) and (4.9) does not denote the total energy any longer [54].

Finally, some remarks on the role of zero von Neumann conditions for the Poisson equation shall be given. A device is described by means of the conduction and valence band edge profiles, the effective masses, the dielectric constants, and the doping profiles [7]. In the semi-infinite leads of the device, charge-neutrality is mandatory. This requirement fixes the electrostatic potential in the leads once the Fermi levels are specified. While charge neutrality is naturally included in the local equilibrium approximation mentioned in Sec. 3.8.2, more effort is needed when the nanostructure is considered as a whole. This is the point, where the zero von Neumann conditions come into play. From Gauss' theorem it follows that the total charge in a box, here the leads, has

[5] Only transverse modes are considered with subband energies being strictly smaller than E in source and drain.

[6] The transmission probability is not defined for $E < \epsilon_j^{(s,\nu)}$ and is set to zero in this regime.

[7] The doping is uniformly distributed in a specified region. See Sec. 4.9 for more details.

to vanish if these conditions are applied. Irrespective of the electrostatic potential, the source-to-gate voltage (V_{gs}) is defined as the difference between the Dirichlet condition on the gate and the Fermi level in the source while the source-to-drain voltage (V_{ds}) is the difference between the source and drain Fermi levels. These specifications are of crucial importance for comparisons between different simulators. It has to be noted that only electrons are considered, i.e. the source and drain regions and thus the corresponding leads are n-doped. The calculation of the hole density in the valence band p_v is not addressed as it is negligible.

4.4 SIMNAD

The implementation of the SIMNAD software started in 1997 [56] with the intention to simulate single electron transistors (SET). The promising future attributed to SETs relied on the idea that the switching procedure in such devices is based on adding or removing a single electron from a quantum dot (island) and thus being in principle the physical limit of miniaturization. The conventional FET is controlled by means of a gate contact which opens or closes the channel capacitively and allows the current flow (ON-state).
In a subsequent work [54], starting around 2000, the simulator has been widely extended. Amongst other features, the ability to simulate 2DEG devices has been provided. In this context, tunneling effects and self-consistency are fully taken into account. The equivalent extension followed for the case of nanowires, where the issue of self-consistency is much more involved. So far, the EMA has been employed in conjunction with a tensorial grid discretization for the Schrödinger and Poisson equation. The scattering matrix formalism served as basis for the calculation of transmission probabilities and charge densities.
In particular for the wire case, the tensorial grid discretization puts stringent limits for the choice of the cross-sectional shape. This motivated the implementation of the FEM. Furthermore, the most relevant band structure effects are now accounted for by a suitable nonparabolicity models in case of nanowire and planar FETs. Since 2005 several other features have been implemented, beside this work, by Martin Frey. The most important ones being the Green's functions formalism for charge transport and the inclusion of inelastic scattering processes. A conceptual comparison between the scattering matrix formalism and the Greens functions formalism can be found in Ref. [54].

4.5 Scattering Matrix Formalism

The simulation domain, i.e. the device, is sliced along the transport direction as illustrated in Fig. 4.3. For this purpose, a set of points $\{x_n\}_{n=0,\ldots,N}$ is distributed on the interval $[0, L]$ with $x_0 = 0$, $x_N = L$, and L being the length of the device. Each slice is delimited by an interval

$$\mathcal{I}_n^x = \begin{cases} [0, x_1/2], & n = 0 \\ ((x_{N-1} + L)/2, L), & n = N \\ ((x_{n-1} + x_n)/2, (x_n + x_{n+1})/2), & \text{otherwise} \end{cases} . \tag{4.10}$$

The external potential $U^{eff}(\vec{r})$ is approximated by $U_n^{eff}(y, z) \equiv U^{eff}(x_n, y, z)$ on the interval $\mathcal{I}_n^x$. Consequently, the general form of the envelope function on $\mathcal{I}_n^x$ is formally identical to the expression (4.2) in the leads, i.e.

$$F^{(n)}(\vec{r}) \equiv \sum_i \left(a_i^{(n)} e^{ik_i^{(n)} x} + b_i^{(n)} e^{-ik_i^{(n)} x} \right) \psi_i^{(n)}(y, z), \tag{4.11}$$

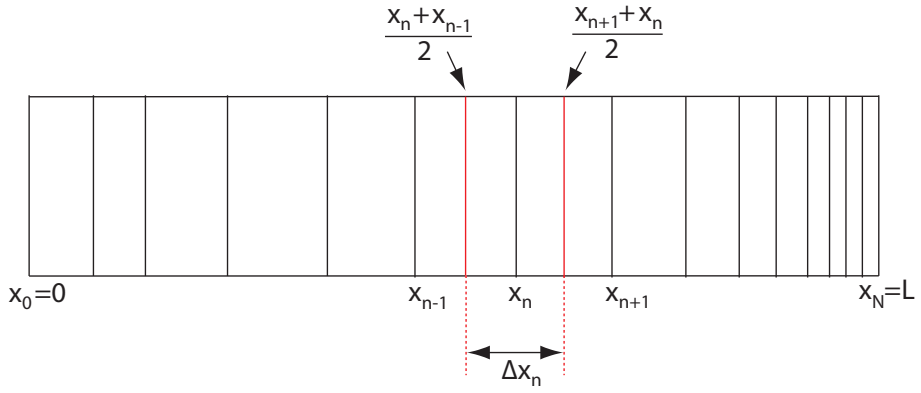

Figure 4.3: Slicing of the simulation domain along the transport direction.

where $\psi_i^{(n)}(y,z) = \langle y,z|i\rangle^{(n)}$ is the eigensolution of the transverse Hamiltonian [8]

$$\mathbf{H}_\perp^{(n)} = -\frac{\hbar^2}{2m_e}\left(\frac{\partial}{\partial y}\frac{1}{m_y}\frac{\partial}{\partial y} + \frac{\partial}{\partial z}\frac{1}{m_z}\frac{\partial}{\partial z}\right) + U_n^{eff}(y,z) \tag{4.12}$$

to the eigenvalue $\epsilon_i^{(n)}$ and

$$k_i^{(n)} = \begin{cases} \sqrt{2m_e m_x(E-\epsilon_i^{(n)})}/\hbar, & E \geq \epsilon_i^{(n)} \\ i\sqrt{2m_e m_x(\epsilon_i^{(n)}-E)}/\hbar, & \text{otherwise} \end{cases}. \tag{4.13}$$

The position-dependence of the effective masses is defined in Sec. 4.3. Note that separability of the Schrödinger problem (4.1) on a slice $\mathcal{I}_n^x$ is partially due to the (y,z)-independence of m_x beside the enforced x-dependence of U^{eff}. For instance, the mass m_x can be taken from the material which hosts the majority of the charge density on $\mathcal{I}_n^x$, i.e. usually the channel material. The envelope function on the whole simulation domain reads

$$F(\vec{r}) = \sum_{n=0}^{N} \chi_n(x) F_n(\vec{r}), \tag{4.14}$$

where $\chi_n(x)$ is unity on $\mathcal{I}_n^x$ and zero otherwise. The continuity of the current (4.5) and the charge densities (4.3) and (4.8) at the interface between two slices is ensured by the conditions

$$F_n((x_n+x_{n+1})/2, y, z) \stackrel{!}{=} F_{n+1}((x_n+x_{n+1})/2, y, z), \quad (y,z) \in \mathbb{R}^2 \tag{4.15}$$

and

$$\frac{\partial}{\partial x}F_n(\vec{r})\Big|_{x=(x_n+x_{n+1})/2} \stackrel{!}{=} \frac{\partial}{\partial x}F_{n+1}(\vec{r})\Big|_{x=(x_n+x_{n+1})/2}, \quad (y,z) \in \mathbb{R}^2 \tag{4.16}$$

[8] The transverse wave functions $\psi_i^{(n)}(y,z)$ are subject to Dirichlet boundary conditions. For instance, the surface of an arbitrary prismatic hull of the device oriented along the transport direction can serve to define the boundary for the Dirichlet conditions.

Equations (4.15) and (4.16) provide connection rules between the coefficients on two neighboring slices

$$\begin{aligned}
a_j^{(n+1)} &= \frac{1}{2} e^{ik_j^{(n+1)}\Delta x_n/2} \\
&\times \sum_i \left[a_i^{(n)} e^{ik_i^{(n)}\Delta x_n/2} \left(1 + \frac{m_x^{(n+1)} k_i^{(n)}}{m_x^{(n+1)} k_j^{(n+1)}} \right) \right. \\
&+ \left. b_i^{(n)} e^{-ik_i^{(n)}\Delta x_n/2} \left(1 - \frac{m_x^{(n+1)} k_i^{(n)}}{m_x^{(n+1)} k_j^{(n+1)}} \right) \right] {}^{(n+1)}\langle j|i\rangle^{(n)} \\
b_j^{(n+1)} &= \frac{1}{2} e^{-ik_j^{(n+1)}\Delta x_n/2} \\
&\times \sum_i \left[a_i^{(n)} e^{ik_i^{(n)}\Delta x_n/2} \left(1 - \frac{m_x^{(n+1)} k_i^{(n)}}{m_x^{(n+1)} k_j^{(n+1)}} \right) \right. \\
&+ \left. b_i^{(n)} e^{-ik_i^{(n)}\Delta x_n/2} \left(1 + \frac{m_x^{(n+1)} k_i^{(n)}}{m_x^{(n+1)} k_j^{(n+1)}} \right) \right] {}^{(n+1)}\langle j|i\rangle^{(n)}
\end{aligned}$$

which can be cast as matrix vector product

$$\begin{pmatrix} \mathbf{a}^{(n+1)} \\ \mathbf{b}^{(n+1)} \end{pmatrix} = \mathbb{T}^{(n)} \begin{pmatrix} \mathbf{a}^{(n)} \\ \mathbf{b}^{(n)} \end{pmatrix} = \left(\begin{array}{c|c} \mathbb{T}_{00}^{(n)} & \mathbb{T}_{01}^{(n)} \\ \hline \mathbb{T}_{10}^{(n)} & \mathbb{T}_{11}^{(n)} \end{array} \right) \begin{pmatrix} \mathbf{a}^{(n)} \\ \mathbf{b}^{(n)} \end{pmatrix}$$

with $\mathbf{a}^{(n)} = (a_0^{(n)}, a_1^{(n)}, \ldots, a_{N_s}^{(n)})^T$ and $\mathbf{b}^{(n)} = (b_0^{(n)}, b_1^{(n)}, \ldots, b_{N_s}^{(n)})^T$. The number of subband energies is denoted by N_s. Given the initial boundary conditions

$$F(0, y, z) = f(y, z) \tag{4.17}$$

$$\frac{\partial}{\partial x} F(\vec{r})\Big|_{x=0} = g(y, z), \tag{4.18}$$

one can compute the coefficients $\{\mathbf{a}^{(0)}, \mathbf{b}^{(0)}\}$ and consequently $\{\mathbf{a}^{(i)}, \mathbf{b}^{(i)}\}_{i=1,\ldots,N}$ by means of the partial transfer matrices $\{\mathbb{T}^{(i)}\}_{i=0,\ldots,N-1}$. However, as mentioned in Sec. 4.3, the portion of the wave function in the waveguide is determined by means of scattering boundary conditions. In contrast to the Cauchy boundary conditions (4.17) and (4.18), the components of the wave function propagating into the device, $\mathbf{a}^{(0)}$ and $\mathbf{b}^{(N)}$, are used to compute the outward propagating components $\mathbf{a}^{(N)}$ and $\mathbf{b}^{(0)}$ [9]. This is achieved by means of the scattering matrix

$$\begin{pmatrix} \mathbf{a}^{(N)} \\ \mathbf{b}^{(0)} \end{pmatrix} = \mathbb{S}^{(0,N)} \begin{pmatrix} \mathbf{a}^{(0)} \\ \mathbf{b}^{(N)} \end{pmatrix}$$

The forward construction scheme for $\mathbb{S}(0, N)$ employs the total transfer matrix $\mathbb{T} = \prod_{n=0}^{N-1} \mathbb{T}^{(n)}$ and the relation (4.19) to compute the scattering matrix

$$\mathbb{S}^{(0,N)} = \left(\begin{array}{c|c} \mathbb{T}_{00} - \mathbb{T}_{01}\mathbb{T}_{11}\mathbb{T}_{10} & \mathbb{T}_{01}\mathbb{T}_{01}^{-1} \\ \hline -\mathbb{T}_{11}^{-1} & \mathbb{T}_{11}^{-1} \end{array} \right)$$

However, the forward construction scheme is numerically unstable [57, 54]. This instability resides in the exponentially increasing and decreasing terms in the partial transfer matrices.

[9] Note that the scattering boundary conditions contain the same amount of information as given by the Cauchy boundary conditions. Uniqueness of the solution is therefore guaranteed for a fixed total energy E.

Depending on the device length, the entries of $\mathbb{T}$ may differ by several orders of magnitude and cause numerical problems when the scattering matrix is finally assembled. A remedy is provided by a recursive construction of the scattering matrix. In this scheme, the matrices $\mathbb{S}^{(0,n)}$ which fulfill

$$\begin{pmatrix} \mathbf{a}^{(n)} \\ \mathbf{b}^{(0)} \end{pmatrix} = \mathbb{S}^{(0,n)} \begin{pmatrix} \mathbf{a}^{(0)} \\ \mathbf{b}^{(n)} \end{pmatrix},$$

are recursively calculated starting from $n = 1$. The $\mathbb{S}^{(0,1)}$ is obtained by replacing $\mathbb{T}$ with $\mathbb{T}^{(0)}$ in Eq. (4.19). The step from n to $n+1$ is accomplished by the pair of matrices $\mathbb{S}^{(0,n)}$ and $\mathbb{T}^{(n)}$, i.e.

$$\begin{aligned}
\mathbb{S}_{00}^{(0,n+1)} &= \left(\mathbb{B}_{00}^{(n)} - \mathbb{S}_{01}^{(0,n)}\mathbb{B}_{10}^{(n)}\right)^{-1} \mathbb{S}_{00}^{(0,n)} \\
\mathbb{S}_{01}^{(0,n+1)} &= \left(\mathbb{B}_{00}^{(n)} - \mathbb{S}_{01}^{(0,n)}\mathbb{B}_{10}^{(n)}\right)^{-1} \left(\mathbb{S}_{01}^{(0,n)}\mathbb{B}_{11}^{(n)} - \mathbb{B}_{01}^{(n)}\right) \\
\mathbb{S}_{10}^{(0,n+1)} &= \mathbb{S}_{10}^{(0,n)} + \mathbb{S}_{11}^{(0,n)}\mathbb{B}_{10}^{(n)}\mathbb{S}_{00}^{(n+1)} \\
\mathbb{S}_{11}^{(0,n+1)} &= \mathbb{S}_{11}^{(0,n)} \left(\mathbb{B}_{11}^{(n)} + \mathbb{B}_{10}^{(n)}\mathbb{S}_{01}^{(n+1)}\right)
\end{aligned} \tag{4.19}$$

where $\mathbb{B}^{(n)} = (\mathbb{T}^{(n)})^{-1}$. Alternatively, the $\mathbb{S}^{(0,1)}$ can be calculated using $\mathbb{S}^{(0,0)} = \mathbb{1}$ and $\mathbb{T}^{(0)}$ in Eq. (4.19). The advantage of this scheme, compared to the forward construction, is that the partial transfer matrices are included slice by slice in the scattering matrix. In this way, exponentially growing and decreasing terms are mixed at each recursion step.
The calculation of the charge density inside the device requires the computation of the entire set of coefficients $\{\mathbf{a}^{(n)}, \mathbf{b}^{(n)}\}_{n=0,\ldots,N}$. This can be accomplished by transforming the scattering boundary conditions $\{\mathbf{a}^{(0)}, \mathbf{b}^{(N)}\}$ in Cauchy conditions $\{\mathbf{a}^{(0)}, \mathbf{b}^{(0)}\}$ by means of $\mathbb{S}^{(0,N)}$. The partial transfer matrices can then be employed to compute the coefficients inside the device. However, this approach suffers from instabilities similar to the forward construction of the scattering matrix. Again, this problem can be circumvented by using a slice-wise incorporation of the partial transfer matrices [57, 58]. For this purpose, the scattering matrix $\mathbb{S}^{(n,N)}$ with the property

$$\begin{pmatrix} \mathbf{a}^{(N)} \\ \mathbf{b}^{(n)} \end{pmatrix} = \mathbb{S}^{(n,N)} \begin{pmatrix} \mathbf{a}^{(n)} \\ \mathbf{b}^{(N)} \end{pmatrix},$$

is introduced. The construction of the set $\{\mathbb{S}^{(n,N)}\}_{n=0,\ldots,N}$ starts with $\mathbb{S}^{(N,N)} = \mathbb{1}$. Then, the scattering matrix $\mathbb{S}^{(n-1,N)}$ is computed by means of $\{\mathbb{T}^{(n-1)}, \mathbb{S}^{(n,N)}\}$, i.e.

$$\begin{aligned}
\mathbb{S}_{00}^{(n-1,N)} &= \mathbb{S}_{00}^{(n,N)}(\mathbb{T}_{00}^{(n-1)} + \mathbb{T}_{01}^{(n-1)}\mathbb{S}_{10}^{(n-1,N)}) \\
\mathbb{S}_{01}^{(n-1,N)} &= \mathbb{S}_{01}^{(n,N)} + \mathbb{S}_{00}^{(n,N)}\mathbb{T}_{01}^{(n-1)}\mathbb{S}_{11}^{(n-1,N)} \\
\mathbb{S}_{10}^{(n-1,N)} &= \left(\mathbb{T}_{11}^{(n-1)} - \mathbb{S}_{10}^{(n,N)}\mathbb{T}_{01}^{(n-1)}\right)^{-1} \left(\mathbb{S}_{10}^{(n,N)}\mathbb{T}_{00}^{(n-1)} - \mathbb{T}_{10}^{(n-1)}\right) \\
\mathbb{S}_{11}^{(n-1,N)} &= \left(\mathbb{T}_{11}^{(n-1)} - \mathbb{S}_{10}^{(n,N)}\mathbb{T}_{01}^{(n-1)}\right)^{-1} \mathbb{S}_{11}^{(n,N)}.
\end{aligned} \tag{4.20}$$

Combining Eqs. (4.19) and (4.20) allows to express the coefficients on each slice as a function of the scattering boundary conditions

$$\begin{aligned}
\mathbf{a}^{(n)} &= \left(\mathbb{1} - \mathbb{S}_{01}^{(0,n)}\mathbb{S}_{10}^{(n,N)}\right)^{-1} \\
&\quad \times \left(\mathbb{S}_{00}^{(0,n)}\mathbf{a}^{(0)} + \mathbb{S}_{01}^{(0,n)}\mathbb{S}_{11}^{(n,N)}\mathbf{b}^{(N)}\right), \qquad (4.21) \\
\mathbf{b}^{(n)} &= \left(\mathbb{1} - \mathbb{S}_{10}^{(n,N)}\mathbb{S}_{01}^{(0,n)}\right)^{-1} \\
&\quad \times \left(\mathbb{S}_{10}^{(n,N)}\mathbb{S}_{00}^{(0,n)}\mathbf{a}^{(0)} + \mathbb{S}_{11}^{(n,N)}\mathbf{b}^{(N)}\right). \qquad (4.22)
\end{aligned}$$

The algorithm for the computation of $\{\mathbf{a}^{(n)}, \mathbf{b}^{(n)}\}_{n=0,\ldots,N}$ can be summarized as follows. First, calculate the partial transfer matrices $\{\mathbb{T}^{(n)}\}_{n=0,\ldots,N-1}$ defined in Eq. (4.17) and store them in memory. Then, compute the scattering matrices $\{\mathbb{S}^{(n,N)}, \mathbb{S}^{(0,n)}\}_{n=0,\ldots,N-1}$ by means of Eqs. (4.19) and (4.20) and store them as well. The coefficients $\mathbf{a}^{(n)}$ and $\mathbf{b}^{(n)}$ can be finally computed by means of Eqs. (4.21) and (4.22), respectively. The number of matrix-matrix multiplications in this algorithm is proportional to the number of slices N. In a more memory-saving approach, a loop over the slices can be performed and the required transfer and scattering matrices are calculated anew for each slice. In this case the number of matrix-matrix multiplications scales like N^2.

For the numerical calculation of the electrostatic potential and the density, the envelope function (4.14) has to be sampled at positions whose x-components are part of the set $\{x_i\}_{i=0,\ldots,N}$ from Fig. 4.3 [10]. The evaluation of $F_n(x_n, y, z)$ is rather simple since the exponential terms are omitted. More information from the entire slice can be included by averaging the wave function over the slice

$$\begin{aligned}\widetilde{F}_n(y,z) &= \frac{1}{\Delta x_n}\int_{(x_n+x_{n-1})/2}^{(x_n+x_{n+1})/2} F_n(\vec{r})dx \\ &= \sum_i \phi_i^{(n)}(y,z)(a_i^{(n)}+b_i^{(n)})\frac{2\sin\left(k_i^{(n)}\Delta x_n/2\right)}{k_i^{(n)}\Delta x_n}. \end{aligned} \tag{4.23}$$

Note that the terms containing purely complex wave vectors $k_i^{(n)} = i\widetilde{k}_i^{(n)}$ with $\widetilde{k}_i^{(n)} \in \mathbb{R}$ in Eq. (4.23) diverge exponentially as a function of $\Delta x_n \widetilde{k}_i^{(n)}$, i.e.

$$(a_i^{(n)}+b_i^{(n)})\frac{2\sinh\left(\widetilde{k}_i^{(n)}\Delta x_n/2\right)}{\widetilde{k}_i^{(n)}\Delta x_n}. \tag{4.24}$$

Especially in the presence of high barriers, the grid spacing has to be chosen such that the corresponding $\Delta x_n \widetilde{k}_i^{(n)}$ does not excessively exceed unity. This is a crucial prerequisite for the achievement of self-consistency.

In this section, some details on the SMA have been explicitly described for the case of nanowire FETs. The planar case can be treated analogously [54]. The usage of the SMA for arbitrary effective mass tensors is described in appendix D. Neglecting the off-diagonal entries in the overlap matrices ${}^{(n+1)}\langle j|i\rangle^{(n)}$ is referred to as the *uncoupled mode space* approach. This approximation notably reduces the computational burden but in some cases this goes at the expense of important physical insights [59, 60]. In this work, the mode coupling is fully taken into account.

Finally, some notes on the charge density computation shall be given. Instead of evaluating the time-demanding *injected* densities (4.3) and (4.8), a local equilibrium like approach can be employed by using the expressions (3.99) and (3.101) on each slice $\mathcal{I}_n^x$. In this approximation, tunneling and reflections are completely discarded. However, this *adiabatic* density can be calculated very efficiently and can serve as an initial guess to a self-consistency algorithm (see Sec. 4.7.6). Details on the implementation of the adiabatic density can be found in Ref. [54]. In this work, current calculation are always based on the injected variants (4.3) and (4.8).

[10] This restriction for the sampling points is required for the numerical approaches used in this work.

4.6 Discretization

4.6.1 Fundamental Equations

In d dimensions, the general form of the stationary Schrödinger equation within the EMA is given by

$$\left[-\nabla \mathbf{M}(\underline{x})\nabla^T + U^{eff}(\underline{x})\right] F(\underline{x}) = EF(\underline{x}), \tag{4.25}$$

and the Poisson equation reads

$$-\frac{1}{8\pi}\underline{\nabla}\boldsymbol{\epsilon}(\underline{x})\nabla^T F(\underline{x}) = \rho(\underline{x}) \tag{4.26}$$

with

$$\rho(\underline{x}) = -\left[n_c(\underline{x}) - p_v(\underline{x}) + N_A^-(\underline{x}) - N_D^+(\underline{x})\right]. \tag{4.27}$$

Both equations are expressed in atomic units, i.e. lengths in Bohr radii, energies in Rydberg, and potentials in Ry/e. The effective mass tensor $\mathbf{M}(\underline{x})$ is generally dense while the dielectricity $\boldsymbol{\epsilon}(\underline{x})$ is assumed to be of the form $\epsilon(\underline{x})\mathbb{1}$. The self-consistent solution of Eqs. (4.25) and (4.26), i.e. with $n_c(\underline{x})$ computed by means of the solutions of the Schrödinger equation (4.25), partially accounts for many-particle interactions in the sense of the Hartree approximation described in Sec. 2.3.

4.6.2 Box Integration Method

Finite difference methods are widely used for the solution of partial differential equations. The differential operators are expressed by means of differences between point evaluations of the solution, thus allowing to cast the differential equation as an algebraic matrix problem. These point evaluations are then obtained by solving this, generally sparse, matrix problem[11]. A straightforward discretization of the simulation domain, which is particularly suited for the finite difference method, is provided by a tensorial grid. Starting with the set $s_x = \{x_n\}_{n=0,\ldots,N}$ used for the slicing of the device (cf. Fig. 4.3), the tensorial grid $\mathcal{T}G = s_x \otimes s_y \otimes s_z$ is constructed by means of two further sets $s_y = \{y_j\}_{j=0,\ldots,J}$ and $s_z = \{z_k\}_{k=0,\ldots,K}$ in three dimensions [12]. Intervals $\{\mathcal{I}_j^y\}_{j=0,\ldots,J}$ and $\{\mathcal{I}_k^z\}_{k=0,\ldots,K}$ can be constructed for s_y and s_z respectively, according to Eq. (4.10).

Within the box integration method in three dimensions ($\underline{x} = \vec{r}$), the operator $-\vec{\nabla}\mathbf{B}(\vec{r})\vec{\nabla}^T$, with $\mathbf{B} = \mathrm{diag}(b^x, b^y, b^z)$ being either $\mathbf{M}$ or $\boldsymbol{\epsilon}$, is integrated over the box $\Omega_{n,j,k} = \mathcal{I}_n^x \times \mathcal{I}_j^y \times \mathcal{I}_k^z$

$$-\int_{\Omega_{n,j,k}} \vec{\nabla}\mathbf{B}(\vec{r})\vec{\nabla}^T F(\vec{r})d\vec{r} = -\oint_{\partial\Omega_{n,j,k}} \vec{n}(\vec{r})(\mathbf{B}(\vec{r})\vec{\nabla}^T F(\vec{r}))dS \tag{4.28}$$

The same operation has to be performed on the remaining terms of the partial differential equation. The surface integral in Eq. (4.28) can be divided in six integrals corresponding to the six faces of $\Omega_{n,j,k}$. Each of these integrals is approximated by evaluating the integrand at the middle of the face and multiplying this value with the area of the face. The integrands itself are

[11] A simple linear or eigenvalue problem has to be solved in the case of the Poisson or Schrödinger equation, respectively.

[12] The two-dimensional case is constructed analogously.

approximated by a central difference, i.e. the integrand pointing along the positive x-direction for instance is approximated by

$$\begin{aligned} &- b^x(x,y_j,z_k)\frac{\partial}{\partial x}F(x,y_j,z_k)\Big|_{x=(x_i+x_{i+1})/2} = \\ &-b^x_{n+1/2,j,k}\left(\frac{F_{n+1,j,k}-F_{n,j,k}}{x_{i+1}-x_i}+\mathcal{O}((x_{i+1}-x_i)^3)\right) \end{aligned} \tag{4.29}$$

where $F_{n,j,k}=F(x_n,y_j,z_k)$ and $b^x_{n+1/2,j,k}=b^x((x_{i+1}-x_i)/2,y_j,z_k)$. Note that the matrix $\mathbf{B}$ has to be evaluated at artificial points lying between the sites of the tensorial grid. The approximation of the surface integral (4.28) is formulated for each lattice site and cast as a matrix vector multiplication $\mathbf{A}\underline{\mathbf{F}}$ with $\underline{\mathbf{F}}=(\ldots,F_{n,j,k},\ldots)^T$. [13] For the inclusion of Dirichlet and von Neumann Boundary conditions see Ref. [56]. With $\mathbf{\Omega}=\mathrm{diag}(\ldots,|\Omega_{n,j,k}|,\ldots)$, $\mathbf{U}=\mathrm{diag}(\ldots,U^{eff}_{n,j,k},\ldots)$, and $\underline{\boldsymbol{\rho}}=(\ldots,\rho_{n,j,k},\ldots)^T$, the application of the box integration method to the Schrödinger (4.25) and Poisson equation (4.26) yields the algebraic problems

$$(\mathbf{A}+\mathbf{U})\underline{\mathbf{F}}=E\underline{\mathbf{F}} \tag{4.30}$$

and

$$\mathbf{A}\underline{\mathbf{F}}=\underline{\boldsymbol{\rho}} \tag{4.31}$$

respectively. In the 1DEG case for instance, the Poisson equation is solved on the three-dimensional tensorial grid yielding the electrostatic potential at the grid points. At each position $x_i \in s_x$, a two-dimensional Schrödinger equation expressed on the grid points perpendicular to the transport direction is solved employing the electrostatic potential.

Solution of the linear Problem

In one, two, and three dimensions, the matrix $\mathbf{A}$ contains at most three, five, and seven entries per row respectively, and is therefore rather sparse. The Poisson problem related to the 1DEG case for instance has to be solved in three dimensions. In this case, typical sizes of the tensorial grid, and therefore of the matrix problem, are of the order of ~ 500000 degrees of freedom. The solution of sparse linear systems up to these sizes can be accomplished by several well-established algorithms and data structures. A common way to store large sparse matrices is provided by the *column row sparse* (CRS) format for instance. The solvers for sparse linear systems are generally subdivided in two classes, i.e. *direct* [14] and *iterative* [15] methods. A survey on direct and iterative approaches for the solution of sparse linear systems can be found in Refs. [62] and [63], respectively, with references to corresponding software packages. A powerful framework which combines efficient data structures and an unified access to state-of-the-art linear solvers, even for distributed computing, is provided by the PETSC[64] software.

Solution of the Eigenvalue Problem

The eigenvalue problems which have to be solved within the 1DEG case are sparse as well, but of more moderate size, i.e. ~ 10000 degrees of freedom. However, these problems are still not optimally suited for direct eigenvalue solvers such as LAPACK. More appropriate approaches

[13] The hermiticity of the matrix $\mathbf{A}$ is ensured by the particular choice of integration (4.28). Non-hermitian matrices may arise from a direct finite difference discretization of the Laplacian on a non-uniform grid.

[14] The PARDISO (`www.pardiso-project.org`), MUMPS (`mumps.enseeiht.fr/`), and UMFPACK (`www.cise.ufl.edu/research/sparse/umfpack/`) libraries are well-known direct solvers.

[15] The iterative solver ILS[61] is well suited for computational electronics.

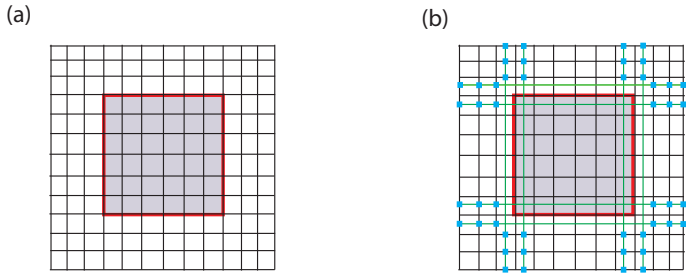

Figure 4.4: (a) Tensorial grid discretization of a rectangular domain. The boundary of the dark region, denoted by the red line, is refined (green lines) in (b). The redundant lattice sites are highlighted by the blue dots.

are Krylov subspace iteration methods such as the implicitly restarted Arnoldi method. This algorithm is implemented in the ARPACK [65] software. A further iterative approach is given by the Jacobi Davidson QR method [66]. Both, the Arnoldi and Jacobi Davidson method search for some eigenvalues around a specified parameter λ. As only the energetically lowest eigenvalues are interesting, the λ can be set to the minimum of the effective potential in the Schrödinger equation. A detailed survey on methods for the solution of eigenvalue problems is given in Ref. [62]. At the beginning, the device simulator SIMNAD was equipped with the JDQR and Slip90 software for eigenvalue and linear problems respectively. The ARPACK solver has been added operating in shift-invert mode by means of the UMFPACK.

Remarks

Finally, some major shortcomings of the box integration method used in conjunction with tensorial grids shall be mentioned. The tensorial grid itself implies two main disadvantages. First, refining the grid in regions where strong variations of the potential or density for instance are expected, such as at a material interface, generates a notable amount of undesired grid points as illustrated in Fig. 4.4. Secondly, the treatment of von Neumann conditions on arbitrarily shaped surfaces is not straightforward. Major complication can also be expected when dense effective mass or dielectric tensors are employed. The application of the box integration on Delaunay triangulations and related theory can be found in Ref. [67].

4.6.3 Finite Element Method

When partial differential equations have to be solved on complicated domains such as in aerodynamics or when the desired accuracy of the solution notably varies within the domain such as in car crash simulations, the finite element method is a popular choice.
In this work, the Poisson and Schrödinger equation for the 1DEG case with arbitrarily shaped boundaries[16] are solved by means of the FEM on a tetrahedrization and triangulation, respectively. Again, the set $s_x = \{x_n\}_{n=0,\ldots,N}$ from Fig. 4.3 is used to construct the grid. As the device is assumed to be prismatic, i.e. the transverse shape does not change along the transport direction, a two-dimensional grid[17] is generated, according to the cross-sectional shape and refinement

[16] The box integration method is exclusively used for rectangular cross sections.
[17] The term *grid* denotes a set of points while a *mesh* contains informations on how the grid points are connected. Within the finite difference method, the meshing is trivial and thus knowing the grid is sufficient.

regions, and replicated [18] along the positions $\{x_i\}_{i=0,\ldots,N}$ as shown in Fig. 4.5. The two- and three-dimensional grid contain N_p^{2D} and N_p^{3D} points respectively. The QHULL [19] software is employed to generate a Delaunay tetrahedrization from the resulting grid in three dimensions. This grid is used for the Poisson problem. Accordingly, a Delaunay triangulation is created from the 2D grid for the Schrödinger problem [20]. The FEM discretization of the Schrödinger and Poisson equation is accomplished by means of piecewise linear functions (hat functions) $b_i(\underline{x})$ located at the lattice sites (see appendix C). First, the Schrödinger problem is considered. In this case, the solution F is subject to zero Dirichlet conditions on the boundary $\partial\Omega_t$ of the transverse domain Ω_t which is the same for all slices. Hence, only hat functions which do not peak at $\partial\Omega_t$ are considered. This set of hat functions is referred to as $\mathcal{S}_t$. Multiplying Eq. (4.25) on the left by a test function $v \in \mathcal{S}_t$ and integrating (by parts) over Ω_t yields

$$\begin{aligned}\int_{\Omega_t} (\underline{\nabla} v)(\mathbf{M}\underline{\nabla} F) d\underline{x} - \oint_{\partial\Omega_t} v(\underline{x})(\underline{\nabla} v \mathbf{M}\underline{n}(\underline{x}))dS \\ + \int_{\Omega_t} v(\underline{x})U(\underline{x})F(\underline{x}))d\underline{x} = E\int_{\Omega_t} v(\underline{x})F(\underline{x}))d\underline{x}.\end{aligned} \tag{4.32}$$

The condition (4.32) has to be fulfilled for each $v \in \mathcal{S}_t$. Using the ansatz

$$F(\underline{x}) = \sum_{b_i \in \mathcal{S}_t} f_i b_i(\underline{x}) \tag{4.33}$$

in Eq. (4.25) yields an algebraic problem of the form (4.30) with

$$\mathbf{A}_{i,j} = \int_{\Omega_t} (\underline{\nabla} b_i)(\mathbf{M}\underline{\nabla} b_j) d\underline{x}, \tag{4.34}$$

$$\mathbf{U}_{i,j} = \int_{\Omega_t} b_i(\underline{x})U(\underline{x})b_j(\underline{x})d\underline{x}, \tag{4.35}$$

and $\underline{\mathbf{F}} = (\ldots, f_i, \ldots)^T$. Details on the assembly of $\mathbf{A}$ and $\mathbf{U}$ within the hat function basis are given in appendix C.

The solution of the Poisson problem is more involved as the data at the Dirichlet boundary $\partial\Omega_D$ is not necessarily zero. In this case, the problem is recasted as follows. Instead of searching for F with $F|_{\partial\Omega_D} = g$, the inhomogeneous Dirichlet conditions are eliminated by introducing a new function $w = F - F_0$, where F_0 is a given function with $F_0|_{\partial\Omega_D} = g(\underline{x})$ and $\underline{\mathbf{n}}\,\underline{\nabla} F_0|_{\partial\Omega_N} = 0$. The modified Poisson problem reads

$$\begin{aligned} -\frac{1}{8\pi}\underline{\nabla}\boldsymbol{\epsilon}(\underline{x})\underline{\nabla}^T w(\underline{x}) &= \rho(\underline{x}) + \frac{1}{8\pi}\underline{\nabla}\epsilon(\underline{x})\underline{\nabla}^T F_0(\underline{x}), \qquad \underline{x} \in \Omega \\ w &= 0, \qquad \underline{x} \in \partial\Omega_D \\ \underline{\mathbf{n}}\,\underline{\nabla} w &= 0, \qquad \underline{x} \in \partial\Omega_N. \end{aligned}$$

and is solved within the set $\mathcal{S}$ of hat functions which do not peak at $\partial\Omega_D$. The solution

$$w(\underline{x}) = \sum_{b_i \in \mathcal{S}} w_i b_i(\underline{x}) \tag{4.36}$$

[18] Conversely, different 2D grids for different x_n could be employed to further reduce the number of grid points. However, this modification notably complicates the calculation of the overlap integrals.

[19] `www.qhull.org`

[20] Similar to the box integration method, the grids for the Schrödinger problems are subsets of the 3D grid. This is not necessarily the case for the corresponding 2D and 3D meshes within the finite element method.

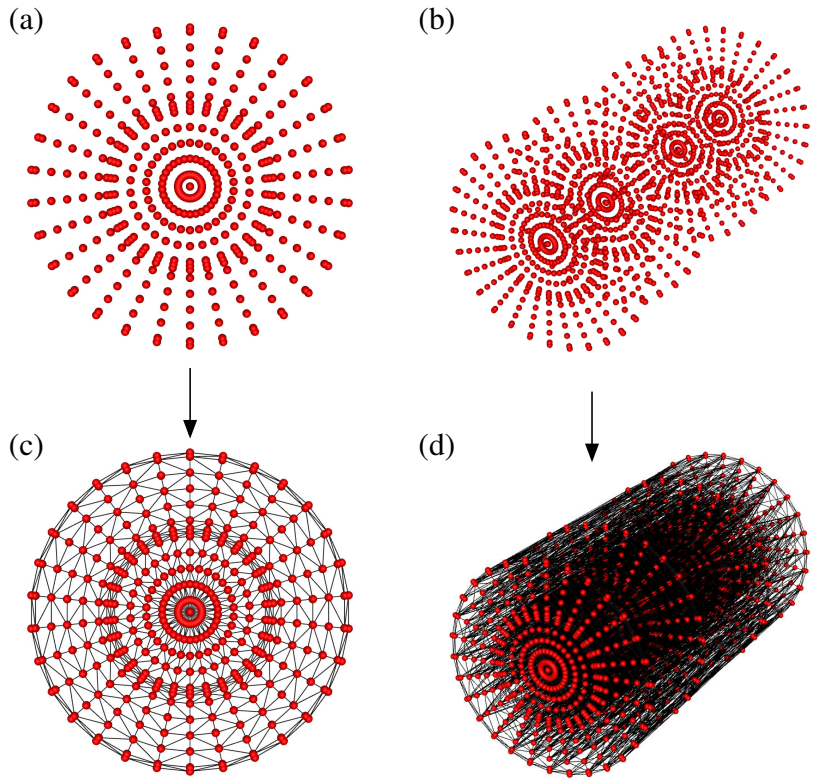

Figure 4.5: Meshing strategy for the finite element method described by means of a circular nanowire. A two-dimensional grid (a) for Ω_t is replicated along the set s_x consisting of four points in this case. The resulting three-dimensional grid (b) is then meshed as shown in (d). Accordingly, a triangulation (c) is generated from the two-dimensional grid (a). Note that the mesh (c) does not necessarily coincide with the mesh (d).

has to fulfill the condition [21]

$$\int_\Omega (\underline{\nabla} v)(\epsilon \underline{\nabla} w) d\underline{x} = \int_\Omega v(\underline{x})\rho(\underline{x}) d\underline{x} - \int_\Omega (\underline{\nabla} v)(\epsilon \underline{\nabla} F_0) d\underline{x} \tag{4.37}$$

for each $v \in \mathcal{S}$. This yields an algebraic problem of the form $\mathbf{A}\underline{\mathbf{w}} = \underline{\boldsymbol{\rho}}$ with

$$\mathbf{A}_{i,j} = \int_\Omega (\underline{\nabla} b_i)(\boldsymbol{\epsilon} \underline{\nabla} b_j) d\underline{x}, \tag{4.38}$$

$$\underline{\boldsymbol{\rho}}_i = \int_\Omega b_i(\underline{x})\rho(\underline{x}) - \int_\Omega (\underline{\nabla} b_i)(\boldsymbol{\epsilon} \underline{\nabla} F_0) d\underline{x} \tag{4.39}$$

$\underline{\mathbf{w}}_i = (\ldots, w_i, \ldots)^T$. A suitable boundary function F_0 is given by $F_0 = \sum_{b_i \in \mathcal{S}_0} g_i b_i$, where $\mathcal{S}_0$ is the set of hat functions which peak at $\partial\Omega_D$ and the g_i are the evaluations of $g(\underline{x})$ at the lattice sites i within $\partial\Omega_D$. This gives a piecewise linear approximation of the boundary condition $g(\underline{x})$. The Poisson equation is solved on the tetrahedrization yielding the piecewise linear electrostatic potential. In this form, the potential is included in the transverse Schrödinger equations on each slice. The resulting wave functions are used to compute the total charge density which is in turn absorbed in piecewise linear form in the Poisson equation for the next iteration step.
As already mentioned previously, the main advantage of the FEM is the ability to easily include von Neumann boundary conditions on arbitrarily shaped boundaries. This benefit, compared to the box method, vanishes in the 2DEG case which is the main reason why the FEM is used for the 1DEG exclusively. The agreement between the FEM and the box method, in terms of current computations, is found to be excellent for comparable rectangular nanowires.
Finally, some notes on further promising discretization methods shall be given. A hybridization between the FEM and finite difference method, referred to as the finite difference element method (FDEM) [68, 69], combines the advantages from both approaches. This method has been successfully applied to several important industrial problems (see Ref. [70] for instance). Sparse grids [71] are a further promising utility which have not received much attention so far. Amongst other applications, when higher dimensional problems are considered in conjunction with the FEM, these grids can yield a substantial reduction of memory consumption and computational burden. The presence of atomic-like, i.e. singular, potentials or densities in a device simulation often hinders the achievement of self-consistency. An improvement can be obtained by splitting the potential in a long-range and a short-range part [72], whereas the latter is treated analytically.

4.7 Self-Consistency

4.7.1 Main Difficulties

A straightforward iteration between the Poisson and Schrödinger equation, i.e. a direct exchange of the charge density and the electrostatic potential, rarely yields a satisfactory convergence behavior [22]. Mostly, the desired quantities exhibit an oscillatory behavior or diverge after some iterations. This shortcoming can be partially ameliorated by tracking the history of the observable and perform an appropriate mixing before the observable is handed over to the next equation. A popular mixing scheme is described in Sec. 4.7.2. Especially when the quantities are

[21] This variational formulation is derived as in the Schrödinger case using integration by parts. The boundary terms vanish due to the properties $v|_{\partial\Omega_D} = 0$, $\underline{\mathbf{n}} \underline{\nabla} F_0|_{\partial\Omega_N} = 0$, and $\underline{\mathbf{n}} \epsilon \underline{\nabla} w|_{\partial\Omega_N} = 0$.

[22] Similar observations can be expected within other frameworks for device simulations such as the drift-diffusion model or more generally when computing the self consistent charge density within the density functional theory.

still far from the final solution, the mixing prevents overshoots and therefore inhibits divergence. A further enhancement is provided by the so called predictor-corrector method which employs an approximate non-linear Poisson equation to obtain a better prediction for the electrostatic potential. The solution of this non-linear equation (predictor step) is referred to as the inner loop while in the outer loop (corrector step), the potential is used within the Schrödinger equation to compute the wave functions and consequently the charge density for the next predictor step. Details on the simulation flow are given in Sec. 4.7.5.

4.7.2 Kerker Mixing

With $\phi_{out}^{(i)}$ being the output from the Poisson solver, the input to the Schrödinger solver can be computed by means of Kerker [23] mixing

$$\phi_{in}^{(i)} = \gamma\phi_{out}^{(i)} + (1-\gamma)\phi_{in}^{(i-1)}, \tag{4.40}$$

where $\phi_{in}^{(i-1)}$ is the Schrödinger input from the previous iteration and $\gamma \in [0,1]$ is referred to as the damping factor. This factor can be modified during the iteration procedure, i.e. $\gamma \rightarrow \gamma^{(i)}$, depending on the behavior of the potential error

$$\mathcal{P}_{err}[\phi_{in}^{(i)}, \phi_{in}^{(i-1)}] = ||\phi_{in}^{(i)} - \phi_{in}^{(i-1)}||_p \tag{4.41}$$

and the residual

$$\mathcal{R}[\phi_{in}^{(i)}, \rho^{(i)}] = ||\underline{\nabla}\epsilon\underline{\nabla}^T\phi_{in}^{(i)} + \rho^{(i)}||_r \tag{4.42}$$

A small residual is an indicator for self-consistency whereas the potential criterion (4.41) acts as an additional control parameter. It has to be noted that the difference $\phi_{in}^{(i)} - \phi_{in}^{(i-1)}$ is directly proportional to γ. Therefore, only if the damping factor is close to unity at the end of the iteration procedure, the potential satisfactorily fulfills the specified error criterion. A common choice for $||.||_p$ is given by the maximum norm $||f||_p = \max_{\underline{x}\in\Omega}$ whereas a possible norm for the residual $||.||_r$ is described in appendix C for the present FEM. Typical convergence criterions, which can be fullfilled within reasonable time limits, are given by $\mathcal{P}_{err} = 10^{-3}$V and $\mathcal{R} = 10^{16}\text{cm}^{-3}$ for doping concentrations of $\sim 10^{20}\text{cm}^{-3}$. However, the poor convergence behavior in large structures often requires a loosening of these criterions.

A successful strategy for the Kerker mixing of the outer loop consists of keeping the γ constant during some iterations as long as both $\mathcal{P}_{err}$ and $\mathcal{R}$ decrease. If one of the latter conditions is not fulfilled any longer, the γ is instantly decreased. Otherwise, the factor is increased and kept constant for the next few iterations.

The ability to modify the damping factor is particularly interesting when the quantities are close to the self-consistent solution. In this regime, a large damping factor notably reduces the number of required iterations. However, in cases where the increment of γ does not yield a satisfactory improvement of the convergence near the self-consistent solution, an acceleration could be obtained by means of Pulay mixing [73]. In this approach, a linear combination of all previous iterations is used in order to construct the input for the Schrödinger routine.

4.7.3 Predictor-Corrector Method

Knowing the exact dependence of the total charge density ρ on the electrostatic potential ϕ allows to formulate a non-linear Poisson equation which can be solved by means of a Newton-Raphson

[23] The term *linear* mixing or *underrelaxation* is often encountered in the literature for this scheme. In this work, the mixing is only applied to the electrostatic potential.

scheme, for instance, as described in Sec. 4.7.4. This approach fully circumvents a Schrödinger-Poisson type iteration procedure to obtain the self-consistent charges and potentials. However, an explicit expression for $\rho[\phi]$ for general quantized structures is not known. Nevertheless, an approximate dependence $\widetilde{\rho}[\phi]$ could be employed in the non-linear Poisson equation

$$-\vec{\nabla}\epsilon\vec{\nabla}^T\phi = \frac{1}{\epsilon_0}\widetilde{\rho}[\phi] \tag{4.43}$$

to compute an improved guess for the potential within a Schrödinger-Poisson scheme. For instance, a semi-classical form [74] for the charge density could be used. A more elaborate expression for the case of nanowires has been presented by Trellakis et al. [75]. The approximate charge density in this case reads

$$\widetilde{\rho}[\phi] = -e\{\widetilde{n}_c[\phi] - p_v + N_A^- - N_D^+\} \tag{4.44}$$

with

$$\begin{aligned}\widetilde{n}_c[\phi] &= \sqrt{\frac{2m_e m_x k_B T}{\pi\hbar^2}}\sum_n |\psi_n^{(i)}(y,z)|^2 \\ &\quad\times\mathcal{F}_{-1/2}\left[\frac{E_F - \epsilon_n^{(i)} + e(\phi - \phi_{in}^{(i)})}{k_B T}\right],\end{aligned} \tag{4.45}$$

where $\{\psi_n^{(i)}, \epsilon_n^{(i)}\}$ were computed in the previous Schrödinger step, i.e. by means of the potential $\phi_{in}^{(i)}$. This potential remains fixed during the inner loop. Once the non-linear problem is solved, the resulting potential $\phi_{out}^{(i+1)}$) is handed over to the mixing routine for the outer loop. The dopand concentrations are assumed to be independent on ϕ. Note that the approximate electron density (4.45) is similar to the adiabatic density (3.99) with the difference that the eigenvalues are shifted by the potential offset, i.e. $\epsilon_n^{(i)} \rightarrow \epsilon_n^{(i)} - e(\phi - \phi^{(i)})$. Following this idea [24], the density (4.3) can be appropriately modified to obtain an approximate expression for the injected case. The expressions for the 2DEG case follows analogously. Note that the computation of the injected density within the predictor loop is much more involved than the evaluation of the conventional expression (4.3). This is due to the potential term in the Fermi-Dirac integral.

4.7.4 Solution of the Non-Linear Poisson Equation

Given a non-linear equation of the form $\mathcal{F}(\mathbf{x}) = 0$, the Newton-Raphson method starts by linearizing the function $\mathcal{F}$ in the vicinity of the solution $\mathbf{x}$ by means of a Taylor expansion

$$\mathcal{F}(\mathbf{x}+\delta\mathbf{x}) = \mathcal{F}(\mathbf{x}) + \underbrace{\frac{\delta\mathcal{F}}{\delta\mathbf{x}}}_{\mathcal{J}(\mathbf{x})}\delta\mathbf{x} + \mathcal{O}(\delta\mathbf{x}^2). \tag{4.46}$$

Neglecting terms of order $\delta\mathbf{x}^2$ and requiring $\mathcal{F}(\mathbf{x}+\delta\mathbf{x}) = 0$, yields the condition

$$\mathcal{J}(\mathbf{x})\delta\mathbf{x} = -\mathcal{F}(\mathbf{x}) \tag{4.47}$$

for the shift $\delta\mathbf{x}$. Equation (4.47) is solved for an initial guess $\mathbf{x}_0$ and the resulting correction $\delta\mathbf{x}_0$ is used to obtain a new position

$$\mathbf{x}_1 = \mathbf{x}_0 + \delta\mathbf{x}_0. \tag{4.48}$$

[24] In the original work only the adiabatic density (3.99) has been considered.

By means of an iterative application of this procedure, the $\mathbf{x}_i$ is assumed to converge towards the solution of $\mathcal{F}(\mathbf{x}) = 0$.
Conversely, instead of seeking for the correction $\delta\mathbf{x}_0$ in Eq. (4.47), the relation (4.48) could be employed to recast Eq. (4.47) into a linear equation for $\mathbf{x}_1$, i.e.

$$\mathcal{J}(\mathbf{x}_0)\mathbf{x}_1 = \mathcal{J}(\mathbf{x}_0)\mathbf{x}_0 - \mathcal{F}(\mathbf{x}_0). \tag{4.49}$$

The non-linear Poisson problem (4.43) can be recast in the form $\mathcal{F}[\phi](\vec{r}) = 0$

$$\mathcal{F}[\phi](\vec{r}) = -\vec{\nabla}\epsilon(\vec{r})\vec{\nabla}^T\phi(\vec{r}) - \widetilde{\rho}[\phi](\vec{r}) \tag{4.50}$$

and the Jacobian

$$\mathcal{J}[\phi](\vec{r}) = -\vec{\nabla}\epsilon(\vec{r})\vec{\nabla}^T - \underbrace{\frac{\delta\widetilde{\rho}[\phi]}{\delta\phi}(\vec{r})}_{\widetilde{\rho}_D[\phi](\vec{r})} \quad . \tag{4.51}$$

With ϕ_0 being the initial electrostatic potential, the linear problem (4.47) reads

$$\left\{-\vec{\nabla}\epsilon\vec{\nabla}^T - \widetilde{\rho}_D[\phi_0]\right\}\delta\rho_0 = -\vec{\nabla}\epsilon\vec{\nabla}^T\phi_0 + \widetilde{\rho}[\phi_0] \tag{4.52}$$

or alternatively

$$\left\{-\vec{\nabla}\epsilon\vec{\nabla}^T - \widetilde{\rho}_D[\phi_0]\right\}\phi_1 = \widetilde{\rho}[\phi_0] - \widetilde{\rho}_D[\phi_0]\phi_0 \tag{4.53}$$

using Eq. (4.49). The form (4.53) is more convenient from an implementational point of view since the set up of an additional linear problem due to the boundary conditions for $\delta\rho$ is circumvented. As in the outer loop, handing over the potential directly to the next Newton step results in a poor convergence behavior. Very elaborate damping schemes such as the Broyden or Bank-Rose method (see Ref. [56] and references therein) could be employed to improve the convergence. A more manageable approach is provided by the Kerker mixing described in Sec. 4.7.2. The residual (4.42) in this case is $||\mathcal{F}[\phi](\vec{r})||_r$. Contemporarily, the convergence of the potential ϕ is monitored. The functionality of the predictor-corrector scheme within the entire self-consistency simulation flow is schematically plotted in Fig. 4.6.

4.7.5 Simulation Flow

The interplay between the Schrödinger-Poisson iteration (outer loop) and the solution of the non-linear Poisson equation (inner loop) outlined in Secs. 4.7.2 and 4.7.3 respectively, is illustrated in Fig. 4.6. The i-th iteration starts with the potential $\phi_{in}^{(i)}$ and comprises the solution of the Schrödinger problems, the calculation of the charge density $\rho^{(i)}$, as well as the computation of a new potential $\phi_{out}^{(i+1)}$ by means of the inner loop. The $\phi_{in}^{(i+1)}$ for the $(i+1)$-th iteration is obtained by Kerker mixing. The convergence criterions for the outer loop are given in Eqs. (4.41) and (4.42) while the termination of the inner loop is described in Sec. 4.7.4.

4.7.6 Initial Guess

The efficiency and, under circumstances, the success of simple or sophisticated damping strategies strongly depend on the choice of the starting solution, i.e. the initial guess. An elementary, and rather generic, approach to compute this initial guess is provided by the local equilibrium approximation described in Sec. 3.8.2 with $U^{ext}(\vec{r}) = -e\phi(\vec{r})$. However, especially in small

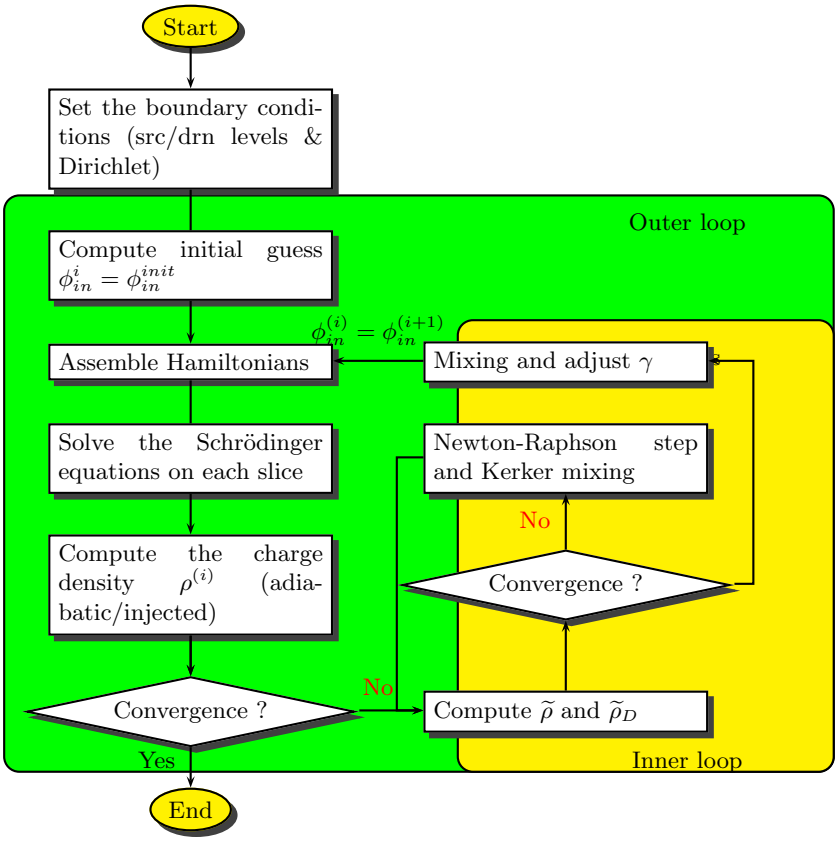

Figure 4.6: Simulation flow for self-consistency.

nanostructures, the resulting potential $\phi(\vec{r})$ is rather far away from the self-consistent solution due to quantization effects. In this regime, an improved initial guess can be obtained by evaluating the densities (3.99) or (3.101), for nanowire or planar FETs respectively, on each slice [25]. The required subband energies and wave functions $\{\epsilon^{(n)}, \psi^{(n)}\}$ are computed by means of the potential $U^{eff} = E_{c/v}(\vec{r}) - e\phi(\vec{r})$ with $\phi(\vec{r})|_{x\in\mathcal{I}_n^x} = \phi^{(n)}$ being piecewise constant on a slice. For the 1DEG case, the values $\{\phi^{(n)}\}_n$ are determined by the condition

$$0 = \int_\sigma \left(n_c^{1DEG} + p_v^{1DEG} + N_A - N_D\right) dydz \tag{4.54}$$

on each slice n. The 2DEG case is treated analogously.

For the calculation of transfer characteristics for instance, the self-consistent solution for a given gate voltage can serve as input for the successive voltage point. When the solutions to several voltage points are available, a suitable extrapolation scheme could be employed to further improve the initial guess for the next point.

4.7.7 Computation of the Injected Charge Density

The integrands in the injected charge densities (4.3) and (4.8) diverge for energies equal to the subband energies in the terminal. Furthermore, a transfer matrix is not defined, i.e. divergent, if the injection energy is equal to a subband energies on the corresponding slice. Therefore, in order to avoid further numerical instabilities, the energy grid has to account for these singularities. A straightforward approach would be to project the subband energies of all slices onto the energy axis. The allowed energy range is then obtained by removing an interval $[s - \Delta E_1/2, s + \Delta E_1/2]$ around each singularity s. The integrand is sampled on the resulting energy range and the integral can be evaluated by means of a conventional Riemann sum. A typical sampling interval is given by $\Delta E_2 \simeq 2 \cdot 10^{-5}$ eV while $\Delta E_1 \simeq 2 \cdot 10^{-6}$ eV. It remains to specify the integration boundaries for the injected densities. The lowest subband energy in the terminal is the natural choice for the lower boundary as states with lower transport energies are exponentially suppressed. The exponential decay of the Fermi distribution function suggests the higher of the source or drain Fermi energy plus a cutoff ΔE_c, being typically a few $k_B T$, as upper limit. If the Fermi level is lower than the lowest subband energy, the latter quantity shall be taken as upper bound (plus ΔE_c). Typically, the total amount of energy points with ΔE_1 and ΔE_2 as specified above is of the order of ~ 1000.

The energy grid for the computation of the currents (4.4) and (4.9) has to account for the same singularities as in the density calculation. Furthermore, in some cases the transmission probability may grow faster than the Fermi function decreases (thermally assisted tunneling). While the lowest subband energy in the source remains a proper lower boundary for the integration domain, the upper limit might be raised to the maximum of the lowest subband energy barrier (plus ΔE_c) in order to capture thermally assisted tunneling effects. However, when the barrier starts to disappear, the upper bound should be kept above $E_F^s + \Delta E_c$.

Conversely, instead of building the energy grid explicitly for densities and currents, an adaptive Simpson quadrature rule could be employed which automatically accounts for singularities. This approach is technically more involved but more advantageous for transport frameworks where the injection singularities are not known a priori.

[25] This is equivalent to the calculation of the adiabatic density described at the end of Sec.4.5.

4.7.8 Implementation

The major computational burden within the self-consistency algorithm resides in the calculation of the injected charge densities and the solution of the transverse eigenvalue problems. The effort related to these tasks can be fortunately divided into independent subtasks, i.e. by splitting the energy integral in the density calculation or subdivide the slices on which the eigenvalue problems are solved. Parallelization can be achieved by means of the open multi-processing (OMP) programming interface and the message passing interface (MPI) [26]. The two programming models are basically different in that OMP requires a connection of the compute nodes to a common memory (shared memory parallelism) while the MPI is not bound to this restriction (distributed memory parallelism). For an excellent overview see Ref. [76]. The OMP programming interface is rather simple to employ but computing machines featuring shared memory capabilities typically contain only a small amount of compute nodes. On the other hand, scaling a code by means of the MPI is more involved.
Both programming models have been employed for the calculation of the charge density and the eigenvalue problems and in both cases the scaling of the individual routines is found to be excellent. However, a major drawback of the MPI is the overhead caused by the communication between the nodes. The solution of the Poisson problem and the collection of the transverse modes for instance require substantial communication. The simulation time spent for the total self-consistency routine can therefore be notably affected by a slow node interconnect within the MPI model.

4.8 Inclusion of Nonparabolicity

The nonparabolictiy (NP) models derived in Sec.3.7 can be implemented within the SMA from Sec. 4.5. [27] The main consequences for the SMA are twofold, namely, the conduction mass m_x and the transverse Hamiltonian (4.12) are affected. In the wire case, the transverse Hamiltonian is replaced by [28]

$$\mathbf{H}^{(n)}_{NP,\perp} = \frac{1}{2\alpha}\left[\sqrt{1-4\alpha\frac{\hbar^2}{2}\left(\frac{1}{m_y}\frac{\partial^2}{\partial y^2}+\frac{1}{m_z}\frac{\partial^2}{\partial z^2}\right)}-1\right]+U_n^{eff}(y,z) \tag{4.55}$$

A real space discretization of the Hamiltonian (4.55) is not as straight-forward as in the EMA case. Instead, a spectral method can be employed, i.e. the solution $F^{(n)}$ is expanded in terms of basis functions which are used to build the Hamilton matrix [29]. For rectangular nanowires a suitable choice consists of sine waves with nodes at the boundary of the transverse domain, i.e.

$$F^{(n)}(y,z) = \sum_{r=1}^{C_r}\sum_{s=1}^{C_s} f_{r,s} \underbrace{\frac{2}{\sqrt{L_y L_z}}\sin\left(\frac{\pi r y}{L_y}\right)\sin\left(\frac{\pi s z}{L_z}\right)}_{\sigma_{r,s}(y,z)} , \tag{4.56}$$

where $\Omega_t = [0, L_y] \times [0, L_z]$ is the transverse domain in this case and C_r and C_s are the cutoffs for the basis set.

[26] See `www.openmp.org` and `www.mcs.anl.gov/research/projects/mpich2` for instance.
[27] The present NP models are also compatible with the Green's functions mode space approach for quantum transport [60, 59, 77].
[28] The single-material approximation will be used in conjunction with NP.
[29] Further works which employ spectral methods for device simulations can be found in Refs. [78, 39, 38].

Within the box method from Sec. 4.6.2 the effective potential U_n^{eff} is given by samples $U_{n,j,k}^{eff}$ on the sites of the tensorial grid. A piecewise constant analytic continuation of the potential on Ω_t reads

$$U_n^{eff}(y,z) = \sum_{j=0}^{J} \sum_{k=0}^{K} U_{n,j,k}^{eff} \chi_{jk}(y,z), \tag{4.57}$$

where χ_{jk} is unity on $\mathcal{I}_j^y \times \mathcal{I}_k^z$ and zero elsewhere. The Hamilton matrix elements

$$\langle r', s' | \mathbf{H}_{NP,\perp}^{(n)} | r, s \rangle = \int_0^{L_y} \int_0^{L_z} \sigma_{r',s'} \mathbf{H}_{NP,\perp}^{(n)} \sigma_{r,s} dy dz \tag{4.58}$$

can be computed very efficiently since the kinetic part is diagonal and the terms $\langle r', s' | \chi_{jk} | r, s \rangle$ can be calculated and stored a-priori. Typically, $C_r = C_s = 30$ is sufficient and the resulting dense eigenvalue problems can be solved by means of LAPACK. The planar FETs are treated analogously.

In the FEM the situation is more involved since the effective potential is piecewise linear on an arbitrary triangulation. An apparently rough, but effective, simplification consists of sampling the piecewise linear potential on a tensorial grid and adopt the procedure from the tensorial case, i.e. Eq. (4.57). For this purpose, a bounding box surrounding the arbitrary transverse domain Ω_t is constructed and the originated vacuum region is mimicked by raising the conduction band edge to a sufficient high value. For $J = K = 100$ an excellent agreement with the conventional FEM is found for several transverse shapes.

Arbitrary transverse kinetic operators, including position-dependent effective masses, can be tackled by means of spectral methods such as the one described in this section. The sine waves in the expansion (4.56) can be replaced for instance by the eigenfunctions of the transverse kinetic operator subjected to zero Dirichlet conditions on Ω_t.

4.9 Simulation results

4.9.1 Overview

The NP models from Sec.3.7 are now used to investigate the impact of band structure effects on transfer characteristics of silicon nanowire and planar FETs (cf. Fig. 4.1). The ballistic quantum transport formalism outlined in this chapter is employed to compute the currents. In particular, the spectral method from Sec. 4.8 is used for the inclusion of NP [30]. The penetration of the wave function in the oxide is neglected, i.e. zero Dirichlet boundary conditions are employed at the channel oxide interface [31]. Furthermore, NP is solely used for the energetically lowest valleys, since particularly in the presence of strong confinements, only these valleys notably contribute to quantum transport. On the other hand, for weak confinements all valleys become relevant but band structure effects are less important. Charge densities are computed self-consistently as described in Sec. 4.7.

As mentioned in Sec.3.7, two band structure effects, i.e. the overestimation of the conduction band edge and the underestimation of the conduction mass, are captured by the present NP models. These two effects can be accounted for individually in order to distinguish between the respective contributions to the currents [32]. In the following, increasing the conduction mass alone is referred to as the MC model while increasing the NP coefficient alone is referred to

[30] Note that the expressions for the currents (4.4) and (4.9) are left formally unaltered by the present NP models.
[31] In this way, only the effective masses of the channel material, i.e. silicon, become relevant.
[32] The investigations are closely related to the strategy employed by Ref. [79].

as the ALPHA model. Neglecting both effects yields the EMA case and the inclusion of both effects is still referred to as the NP case. The numerical values of the NP coefficients and modified conduction masses for the simulation of the FETs are given in Sec. 5.2 and are of minor importance in this section since the attention is paid to qualitative considerations.
The impact of band structure effects on transfer characteristics $I_{\mathrm{d}}(V_{\mathrm{gs}})$ of FETs operating under a fixed forward bias V_{ds} is investigated by means of specific quantities. One of these quantities is the threshold voltage V_{th} defined by the condition

$$I_{\mathrm{d}}(V_{\mathrm{th}}) \stackrel{!}{=} I_{\mathrm{th}}. \tag{4.59}$$

The threshold voltage V_{th} divides the transfer characteristic in two parts, the subthreshold regime ($V_{\mathrm{gs}} \leq V_{\mathrm{th}}$) and the on-current regime. The subthreshold regime is characterized by the subthreshold slope (SS)

$$\mathrm{SS} = 1000 \times \ln(10) \times \min_{V_{\mathrm{gs}} \in [V_{\mathrm{th}} - \Delta V_{\mathrm{SS}}, V_{\mathrm{th}}]} \left[\frac{d}{dV_{\mathrm{gs}}} \ln(\widetilde{I}_{\mathrm{d}}) \right]^{-1}, \tag{4.60}$$

where $\widetilde{I}_{\mathrm{d}}$ is the drain current in arbitrary units. The theoretically minimal value for the SS is given by $1000 \times \ln(10) \times k_B T/e = 59.637\mathrm{mV/dec}$[33]. Finally, the on-current I_{on} is defined by

$$I_{\mathrm{on}} = I_{\mathrm{d}}(V_{\mathrm{th}} + \Delta V_{\mathrm{on}}). \tag{4.61}$$

The drain current in ideal FETs subject to a fixed bias V_{ds} is solely controlled by the gate voltage, i.e. the width of the barrier underneath the gate in the subthreshold regime is comparable to the length of the gate. Applying strong bias conditions in short channel FETs leads to a deformation (thinning) of the barrier being referred to as drain induced barrier thinning (DIBT). This phenomenon is accompanied by a reduction of the barrier height, i.e. drain induced barrier lowering (DIBL) [54]. Consequently, the tunneling component to the drain current I_{d} increases, leading to a shift of the threshold voltage V_{th} and a degradation of the subthreshold slope depending on V_{ds}. Non-ideal effects such as the deformation of the barrier appear both in planar and nanowire FETs whereas the electrostatic control of the gate contact in the latter case is more efficacious [80, 3]. In the following, non-ideal effects are abetted by using a rather high forward bias of $V_{\mathrm{ds}} = 0.6V$ and the behavior of these effects under the influence of NP is investigated.

4.9.2 Results for Nanowire FETs

The nanowire FETs considered in this section have a prismatic shape with profiles shown in Fig. 4.7. The silicon core and the circumjacent silicon oxide layer are characterized by the quantities t_c and t_{ox} respectively. The source and drain regions are both n-doped with a concentration of $N_D = 10^{20}\mathrm{cm}^{-3}$ and the lengths are $l_s = l_d = 10$ nm. The gate contact surrounds the FET as shown in Fig. 4.1 (gate-all-around structure) and the length takes the values $l_g[\mathrm{nm}] \in \{5, 10, 15, 20\}$. The various channel thicknesses of the triangular FETs are $t_c[\mathrm{nm}] \in \{4, 5, 6, 7\}$ while $t_c[\mathrm{nm}] \in \{2, 3, 4, 5\}$ are the ones of the square and circular FETs. For the calculation of V_{th}, I_{on}, and SS the values $I_{\mathrm{th}} = 10^{-7}$ A, $\Delta V_{\mathrm{on}} = 0.2$ V, and $\Delta V_{\mathrm{SS}} = 0.2$ V, respectively, are employed [34].
Transfer characteristics of the square FETs are given in Fig. 4.8. For $t_c = 2$ nm and $l_g = 5$ nm, the SS seems to be mainly affected by the increment of the conduction mass while for the remaining cases the SS is essentially left unaltered by the various models. A more detailed analysis of the SS for different geometries is given in Fig. 4.9. For the smallest t_c, the SS diminishes for

[33] A temperature of $T = 300K$ is assumed for all current calculations in this work.
[34] These values are also used for the planar FETs in Sec. 4.9.3.

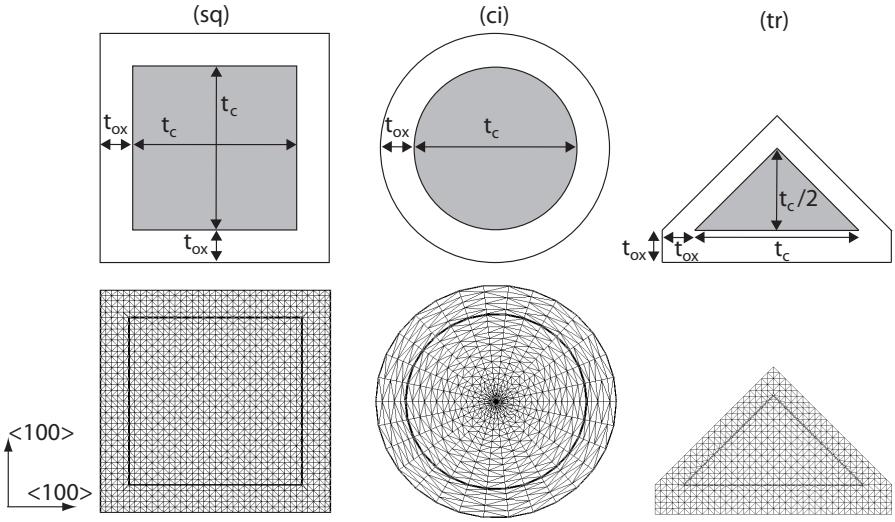

Figure 4.7: Profile specifications of the square (sq), circular (ci), and triangular (tr) nanowire FETs from Sec. 4.9.2. Typical triangulations are shown as well. The oxide thickness is always $t_{ox} = 1\text{nm}$.

an increasing l_g and approaches the theoretical minimal value of 59.637 mV/dec. Particularly evident is the influence of the conduction mass on the SS in the square and circular cases with the smallest l_g. In this case, the tunneling contribution to the total current is notable and is further enhanced by the strong forward bias as already mentioned in Sec. 4.9.1. Increasing the conduction mass leads to a suppression of tunneling and thus an improvement of the SS. The discrepancy between the various models is negligible beyond $l_g = 10$ nm since tunneling is generally suppressed. For an increasing channel thickness t_c, the electrostatic control of the gate contact diminishes leading again to an increased tunneling and thus a degradation of the SS. Contemporarily, the band structure effects decrease which explains the smaller discrepancy between the various models compared to the $l_g = 5$ nm cases for the smallest t_c.

The threshold voltage V_{th} is plotted in Fig. 4.10 as a function of both l_g and t_c. The major impact on V_{th} can be attributed to the increment of α. This finding is best explained by means of Fig. 4.11.a showing the lowest subband energy profile of a square nanowire FET at $V_{\text{gs}} = 0$ V. The transverse eigenvalue problems generally yield a smaller eigenvalue when $\alpha > 0\text{eV}^{-1}$ and thus reduce the barrier below the gate. The decrease of the barrier causes a shift of the transmission coefficient and thus an increment of the spectral current as shown in Fig. 4.11.c. This leads to a shift of the transfer characteristics and thus V_{th} towards smaller values. Apparently, this shift of the barrier due to α does not notably influence the SS (cf. Fig. 4.9). The situation in the source and drain regions is essentially left unaltered by α (inset) while the increment of the conduction mass lifts the subband energy (see discussion about the on-current) having, however, a minor impact on V_{th}. For the smallest t_c in Fig. 4.10, the V_{th} approaches a constant value for an increasing l_g. A weak dependence on the conduction mass can be observed for $l_g = 5$ nm, where the V_{th} is generally smaller compared to the remaining l_g. Again, this can be attributed to an increased contribution of the tunneling component to the current for short channels, i.e. a larger subthreshold current yields a smaller V_{th} while an increased conduction mass suppresses

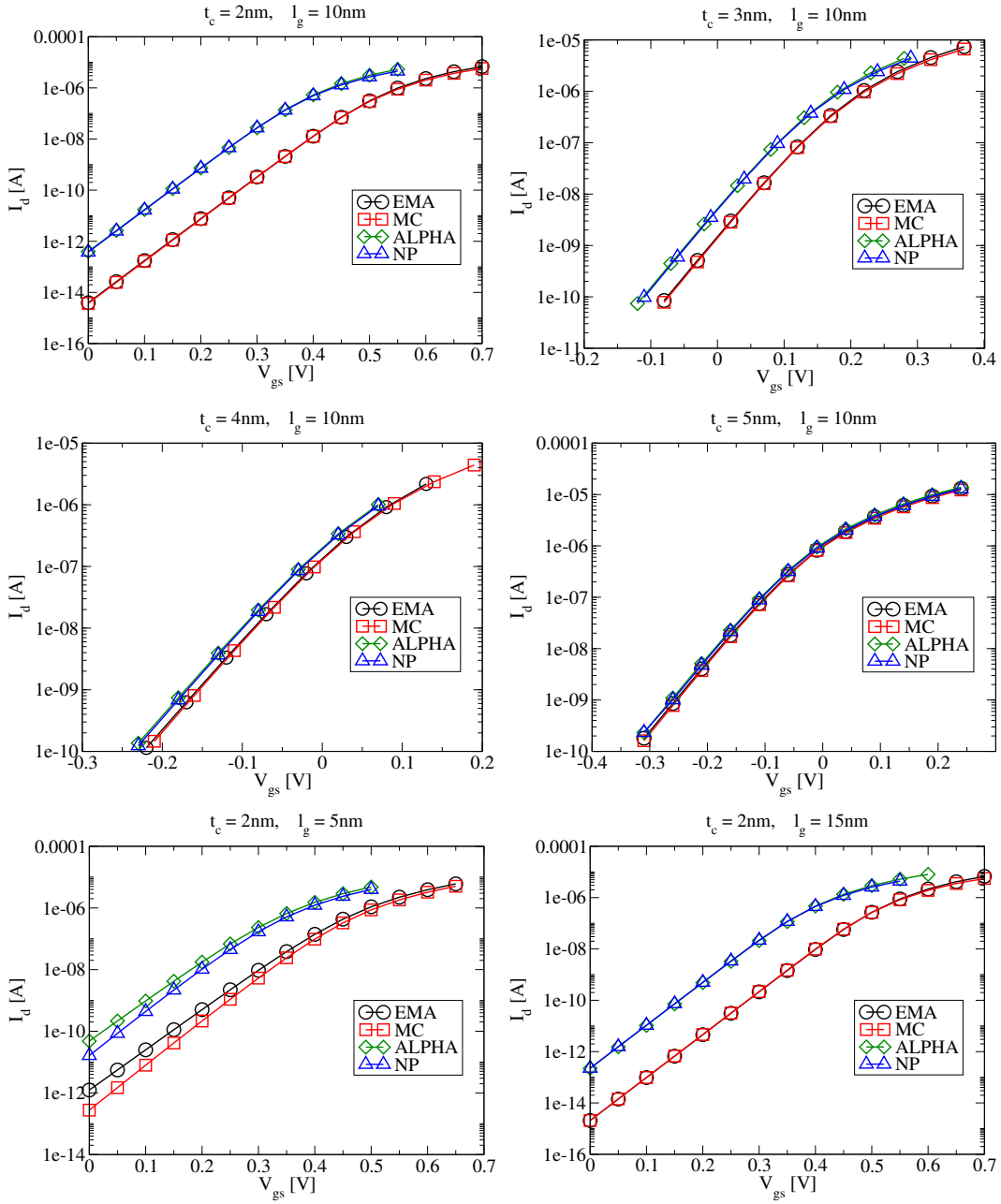

Figure 4.8: Transfer characteristics (subthreshold regime) of square nanowire transistors by various models. Gate lengths l_g and channel thicknesses t_c are specified in the graphs.

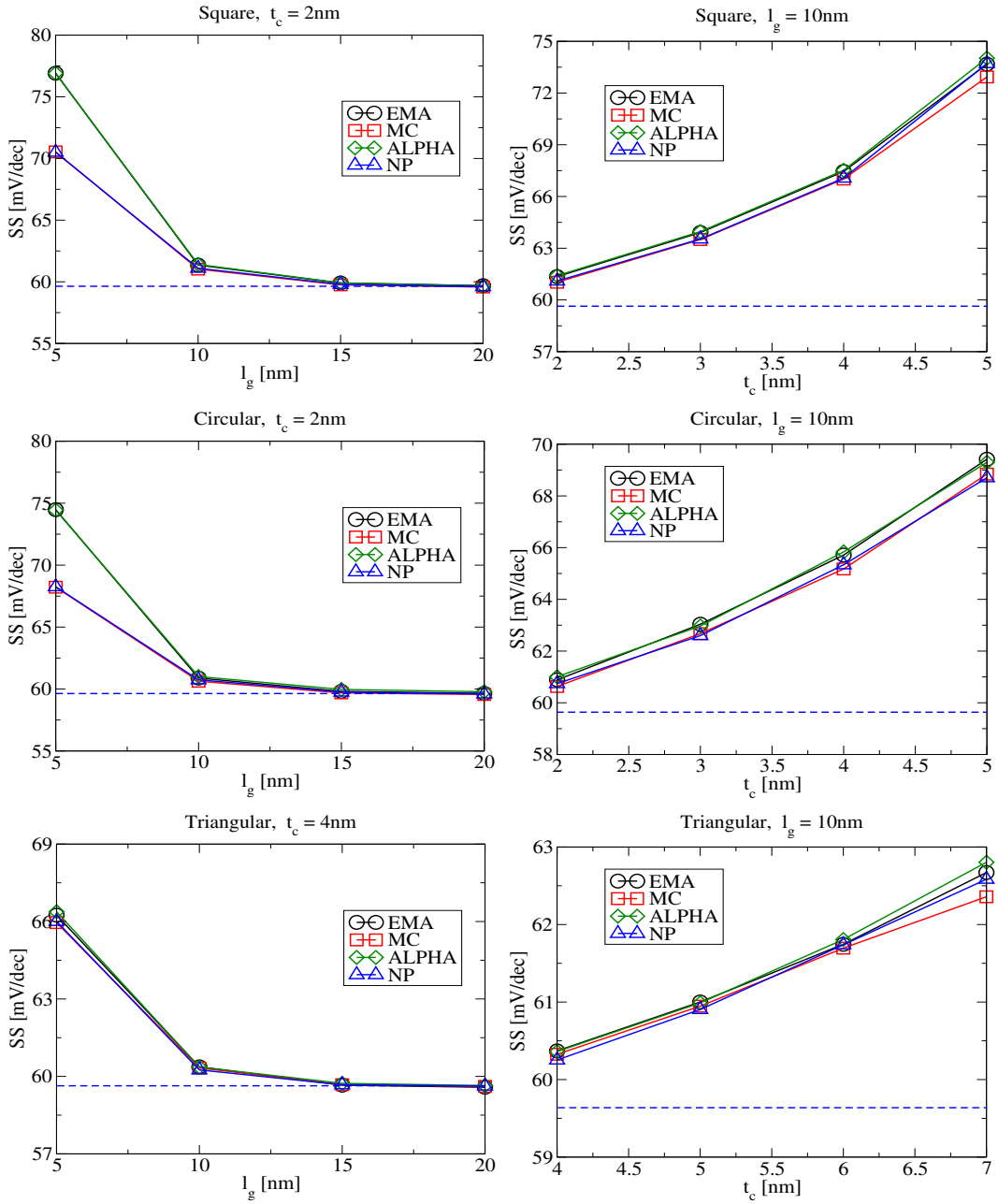

Figure 4.9: Subthreshold slope (SS) of nanowire FETs. The gate lengths l_g, channel thicknesses t_c, and the cross-sectional shapes are specified in the graphs.

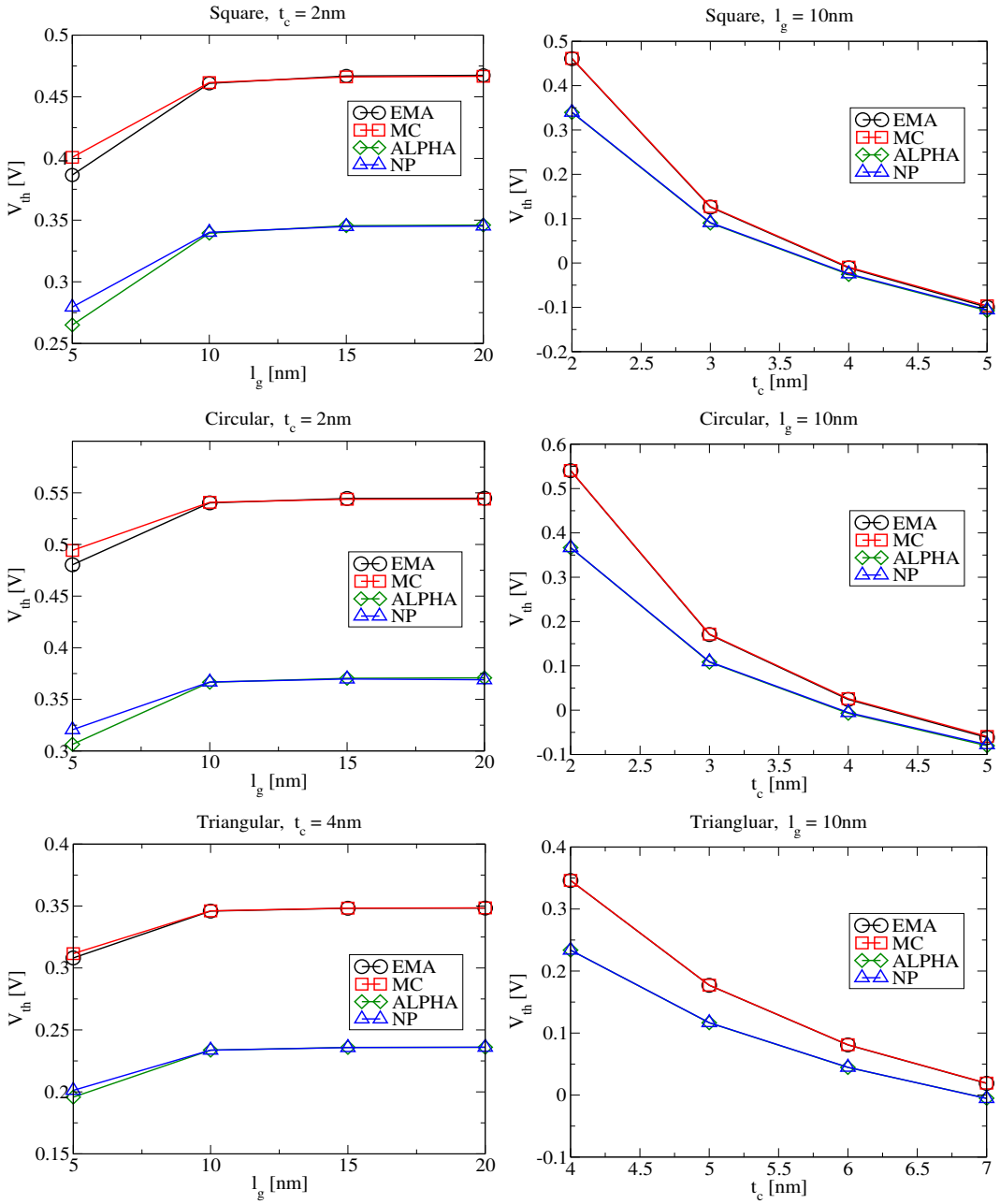

Figure 4.10: Threshold voltage V_{th} of nanowire FETs. The gate lengths l_g, channel thicknesses t_c, and the cross-sectional shapes are specified in the graphs.

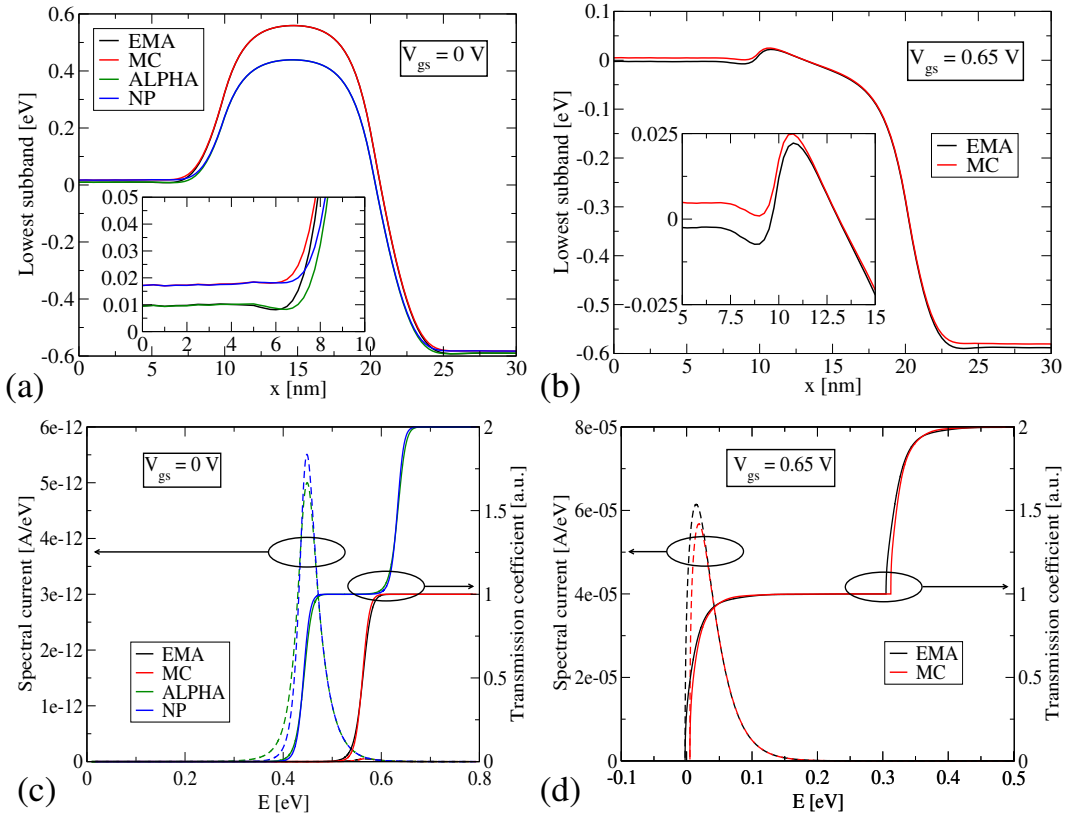

Figure 4.11: Lowest subband energy profile (a & b), spectral current (dashed lines), and transmission coefficient (c & d) of a square nanowire FET with $t_c = 2$ nm and $l_g = 10$ nm. The data belongs to one of the four Δ_4 valleys which, in this case, contribute approximately in the same manner to the current. The gate voltages are specified in the graphs and the source Fermi level is $E_F^s = 0$ eV.

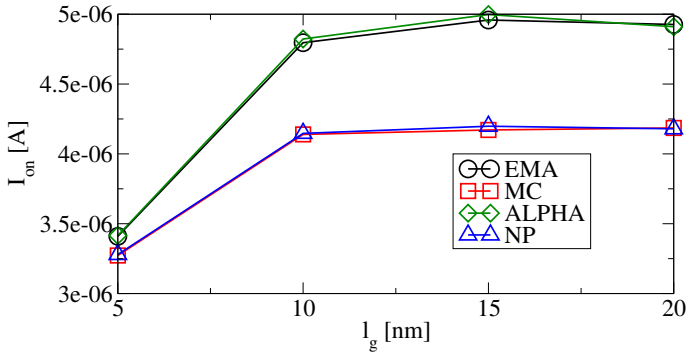

Figure 4.12: ON-currents of square nanowire transistors by various models. The channel thickness is $t_c = 2$nm.

the tunneling and thus diminishes the decrease of V_{th}. Beyond $l_g = 10$ nm the V_{th} is not influenced by the tunneling any longer. For increasing channel thickness t_c, the subthreshold current generally increases leading to a reduction of V_{th}. Furthermore, the discrepancy between various models vanishes since band structure effects become less important.
The on-current of square nanowire FETs is plotted in Fig. 4.12 as a function of the gate length l_g. The I_{on} increases for an increasing l_g and becomes approximately constant beyond $l_g = 10$ nm. The small V_{th} at $l_g = 5$ nm due to tunneling shifts the evaluation point for the I_{on} and is mainly responsible for the smaller on-currents compared to the remaining gate lengths. However, in each case, the conduction mass seems to have the leading impact on I_{on}. This finding is somehow counterintuitive since the gate barrier, and thus tunneling, essentially vanishes in the on-current regime and the conduction mass can not influence the current via the tunneling contribution any longer. Figure 4.11.b sheds more light on this situation. The gate voltage is close to the voltage point used for the evaluation of I_{on}. In the source and channel region, the subband energy is raised when the conduction mass increases leading to a smaller spectral current as shown in Fig. 4.11.d. This influence of the conduction mass on the subband energy in the source region is due to charge neutrality. Since the density of states in a wire is directly proportional to the conduction mass, a larger mass mandates a larger subband energy in order to obtain the same number of conduction band electrons.
Finally, the V_{th} and the SS of the various nanowire FETs are plotted in Fig. 4.13 as a function of the effective diameter d_e (see Sec. 5.2). While the various V_{th} are scattered approximately in the same manner, the SS of the triangular FETs shows a distinct behavior compared to the square and circular case.
Further reading on ballistic transport is given in Refs. [81, 60, 82, 59, 83, 84, 85, 86, 87] and particular attention to band structure effects is paid in Refs. [88, 89, 90, 31, 79, 91, 92, 93, 94].

4.9.3 Results for Planar FETs

The discussion in this section is kept rather brief since several explanations from the wire case apply also to planar FETs. Transfer characteristics were obtained from hypothetical double-gate FETs similar to the one employed by Lundström et al. [95]. Doping and geometry specifications are given in Fig. 4.14. The gate length l_g and channel thickness t_c are variable while the source and drain regions are fixed.
Transfer characteristics are plotted in Fig. 4.15. The currents are given in A/μm, i.e. the current

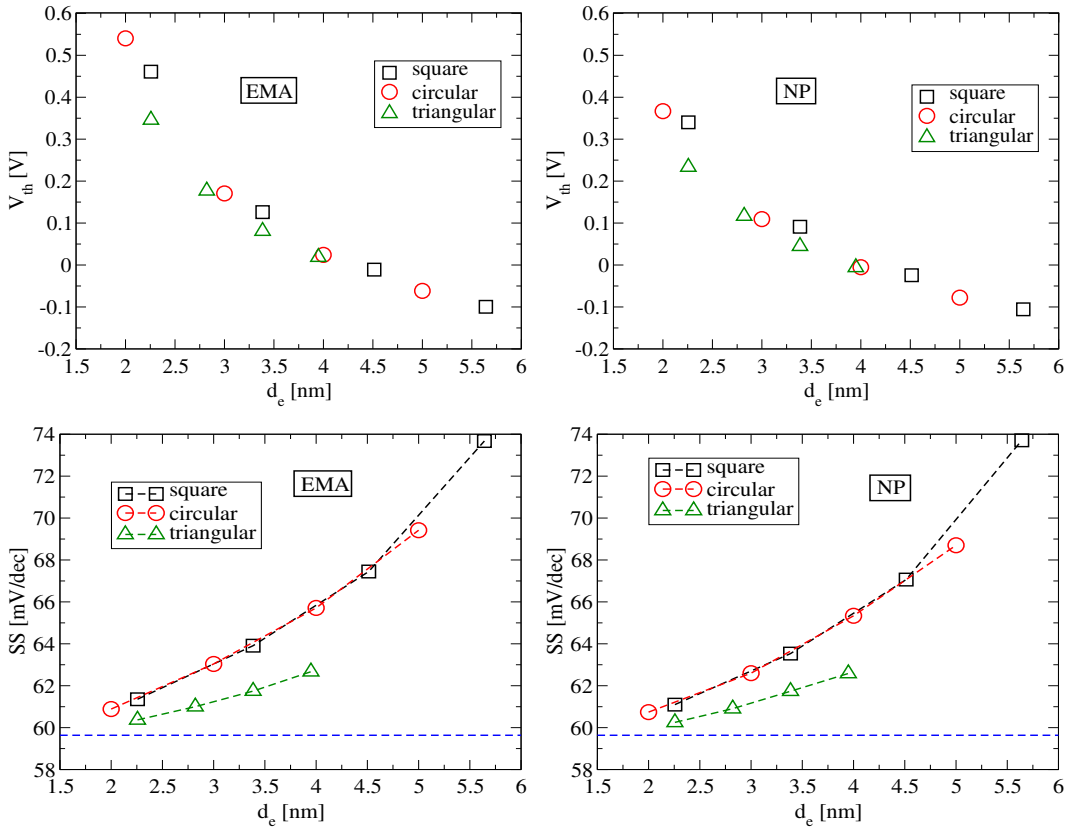

Figure 4.13: Threshold voltage $V_{\rm th}$ and subthreshold slope (SS) of nanowire transistors by various cross-sectional shapes. The gate length is $l_g = 10$nm and the models are specified in the graphs.

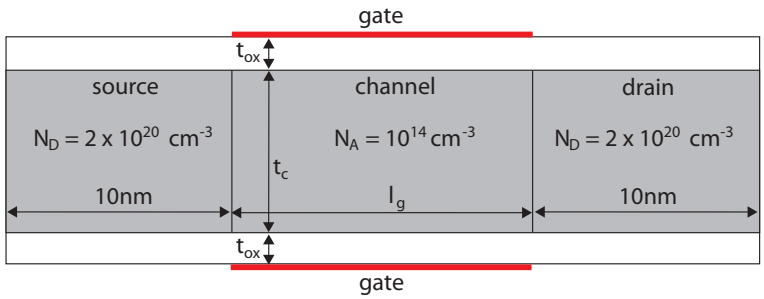

Figure 4.14: Double-gate FET studied in Sec. 4.9.3. The oxide thickness is always $t_c = 0.6$ nm.

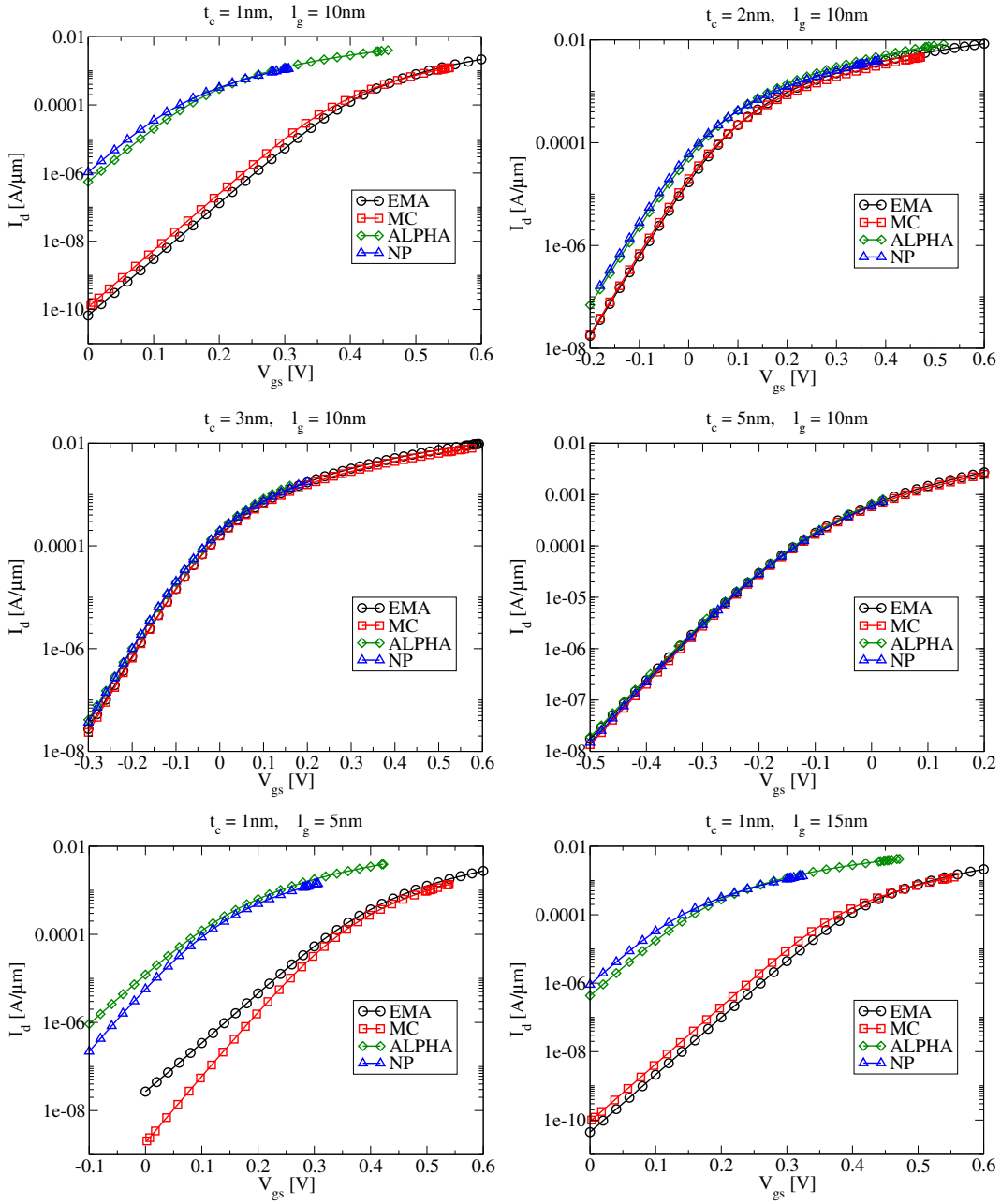

Figure 4.15: Transfer characteristics (subthreshold regime) of planar FETs by various models. Gate lengths l_g and channel thicknesses t_c are specified in the graphs.

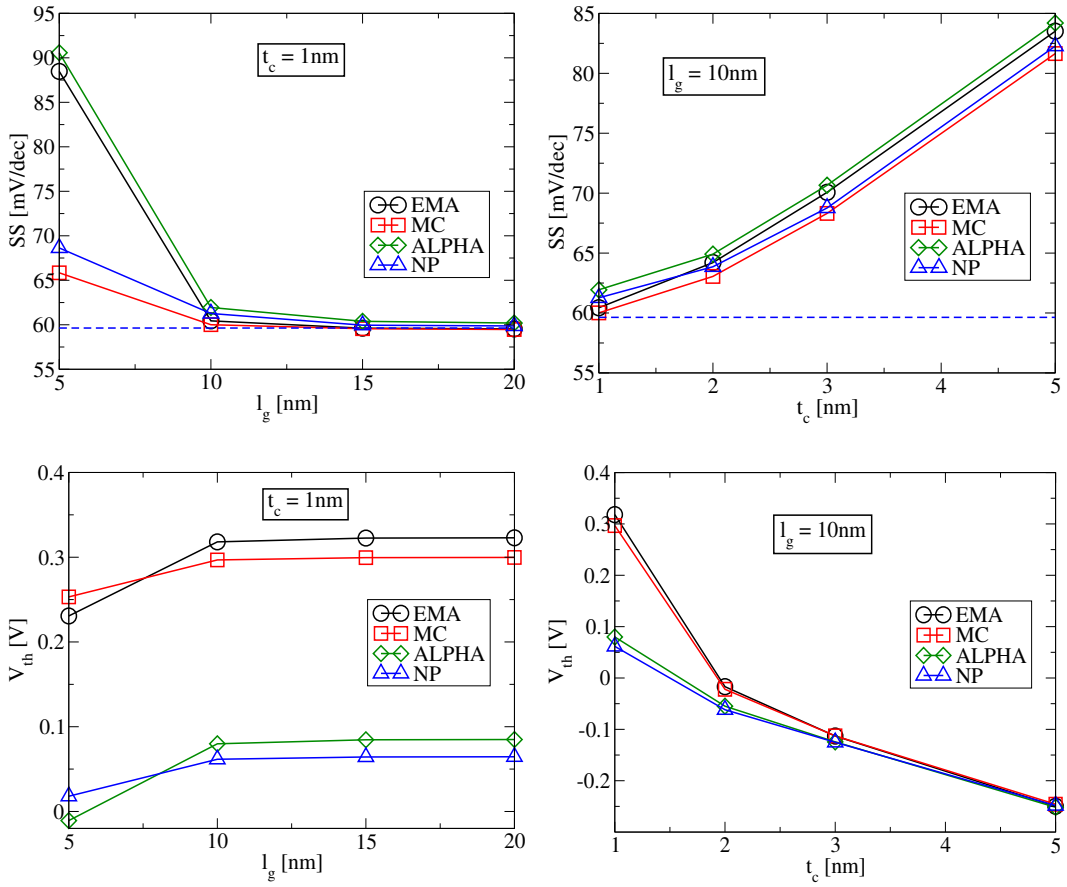

Figure 4.16: Threshold voltage (V_{th}) and subthreshold slope (SS) of planar FETs as a function of the gate length l_g and channel thickness t_c by various models.

(4.9) has to be divided by W_y. A point which is worth noting is the pronounced [35] influence of the conduction mass on the V_{th}, compared to the wire case, even for large l_g as can be seen from Fig. 4.16. A closer look at the current formula (4.9) reveals that even in the absence of tunneling, the transverse mass m_y, which is increased in the same manner as the conduction mass, is able to increment the current. The inferior electrostatic control of the gate contact compared to the wire case becomes evident, for instance, for $\{t_c = 5\text{nm}, l_g = 10\text{nm}$, where SS = 83.512 mV/dec (EMA case) compared to the SS = 73.675 mV/dec of the square nanowire FET with $\{t_c = 5\text{nm}, l_g = 10\text{nm}$ (EMA case) from Sec. 4.9.2. Apart from that, the qualitative behavior of SS and V_{th} as a function of t_c and l_g is very similar to the one observed in the wire case.

Further reading on ballistic transport and band structure effects for the case of planar FETs can be found in Ref. [96, 97].

[35] The impact of α on V_{th} is still dominant.

4.10 Summary

The scattering matrix formalism for the implementation of quantum ballistic transport within the Landauer-Büttiker formalism has been outlined at the beginning of this chapter with a particular focus on devices hosting two- and one-dimensional electron gases, i.e. planar and nanowire FETs. For the discretization of nanowire FETs, the FEM is found to be superior compared to the box integration method in that the treatment of arbitrary surfaces is more straightforward. Beside the consideration of different discretization schemes for the Poisson and Schrödinger equation, a detailed survey on techniques to achieve self-consistency has been given. The inclusion of the present NP models within the SMA by means of spectral methods allows to investigate the impact of band structure effects on the device characteristics. Nanowire FETs with various cross sectional shapes have been employed in order to include the effect of the geometry. While square and circular FETs show similarities, the triangular FET shows a distinct behavior concerning the SS. Short-channel effects such as the degradation of the SS is found to be alleviated by the increment of the conduction mass in both planar and nanowire FETs. Beside the expected inferior electrostatic control in planar FETs, the qualitative impact of band structure effects depending on geometry variations is found to be similar in both types of devices. Finally, it has to be reminded that the conclusions elaborated in this chapter are based on quantum ballistic transport simulations. The presence of scattering, in particular surface roughness, might might notably influence the impacts of band structure effects.

Chapter 5

Comparison with Tight-Binding

5.1 Introduction

Simulation of quantum transport can be accomplished on various levels. The Landauer Büttiker formalism in conjunction with the EMA and SMA has been outlined in the previous chapter for the ballistic limit. An emerging class of simulators make use of the non-equilibrium Green's functions formalism (NEGF) in conjunction with atomistic methods, i.e. the tight-binding [98, 99] or fully ab-initio techniques [100, 101] for instance. These methods fully account for band structure effects. A good overview of related methods is given in Ref. [102]. Further approaches to atomistic quantum transport can be found in Refs. [103, 104, 105, 106, 107, 108]. The common denominator in atomistic quantum transport simulators is the notable computational burden compared to the EMA case. Thus, simplified methods being able to capture the main impacts of band structure effects on quantum transport are desirable in order to maintain simulation times within a reasonable limit. Extensions of the EMA within semi-classical transport frameworks have been proposed [88, 31] as well as quantum mechanical approaches [93, 94] including scattered comparisons to fully atomistic approaches. In this chapter the present NP models are appropriately calibrated and transfer characteristics are compared to results obtained by a TB-NEGF simulator [98] being able to tackle devices of considerable size. Extensive comparisons for circular and square nanowire FETs are carried out for different diameters and gate lengths.

5.2 Extraction of α

The square and circular silicon wires used to build the channels of the FETs in this section have been thoroughly studied in Secs. 3.4 and 3.6. The termination at the surfaces which mimics hydrogen passivation is consistent with the neglect of the penetration of the wave function in the oxide. This penetration will be neglected for the comparisons in this section. Given a square or circular nanowire FET, the transfer characteristics computed by the present transport framework (EMA case) are supposed to differ from the corresponding TB-NEGF results due to band structure effects. The NP models from Sec. 3.7 aimed at improving the EMA results but no parameterizations have been specified so far in this work [1]. For each wire cross section, the NP coefficient is extracted by means of tight-binding band structures based on the parametrization

[1] The qualitative considerations from Sec. 4.9 about the impact of band structure effects on transfer characteristics by means of NP are based on the parameters extracted in this section.

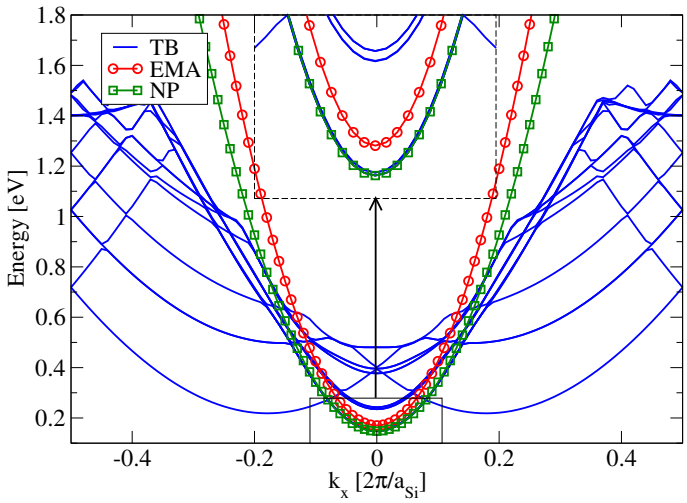

Figure 5.1: Tight-binding (TB) band structure of a square silicon nanowire (see text) and nonparabolic (NP) dispersion (3.78). The corresponding dispersion obtained by the EMA is plotted as well. Data is given for $d_e = 4.139$ nm, $\alpha = 0.841\text{eV}^{-1}$, $m_c = 0.252$. The inset highlights the agreement between the nonparabolic and tight-binding dispersion.

in Refs.[109, 110], i.e. the TB cases in Figs. 3.14 and 3.15 and Tab. 3.2 [2]. A widely used approach to determine the NP coefficient in silicon is based on the condition [3] [31, 79]

$$\frac{1}{2\alpha}\left[\sqrt{1+4\alpha\epsilon_\perp^g}-1\right] \stackrel{!}{=} \epsilon_c^a \tag{5.1}$$

which states that the conduction band edge ϵ_c^a of the nanowire computed by the tight-binding approach is equal to the band edge obtained by means of the EMA, i.e. Eq. (3.78) in the square case[4]. Once the α is known, the modified conduction mass is given by $m_l\sqrt{1+4\alpha\epsilon_\perp^g}$ [5]. For a square nanowire, the tight-binding and corresponding nonparabolic dispersion (3.78), after fitting the α, are compared in Fig. 5.1. The extraction procedure (5.1) is employed for both circular and square nanowires. As mentioned in Sec. 3.6.3, an analytic expression for the EMA band edge $\epsilon_\perp^g$ in the circular case is not straight-forward. The spectral method from Sec. 4.8 is used to compute the band edges of the circular wires (diameters dictated by the TB cases) from Sec. 3.6.3. The fit model (3.63) is then used to obtain a continuous form for the band edge, yielding $K_c = 2.807$ eVnm^2, $a_c = 0.135$ nm, and $b_c = -0.070$ nm^2. The NP coefficients and related conduction masses are reported in Tabs. 5.1 and 5.2 for the square and circular case, respectively. Using the fits from Tab. 3.2, the NP coefficient, and therefore the

[2] This parametrization is employed by the TB-NEGF simulator used for the present comparisons. The bulk effective masses for this parametrization are $m_l = 0.891$ and $m_t = 0.201$.

[3] In the following, the notation from Sec. 3.7 is used. An other approach to determine the NP coefficient in silicon consists in fitting the density of states obtained by means of the dispersion (3.72) to data obtained by atomistic approaches or experiments as outlined in Sec. 3.8.

[4] Note that the same α and m_c are obtained when the full spectrum (3.75) is fitted to the atomistic band edge.

[5] As mentioned in Sec. 4.9.1 , the NP model is only employed for the energetically lowest conduction band valleys, i.e. the Δ_4 valleys in this case. In this connection, the band splitting effects mentioned in Sec. 3.6.3 are fully ignored.

Table 5.1: NP coefficients α, conduction masses m_c, and effective diameters d_e for the present comparison to tight-binding (square case). The m_c for the case of a single (sg) α is given as well.

d_e [nm]	2.451	2.914	3.370	3.831	4.139	4.595	5.057
$\alpha[\mathrm{eV}^{-1}]$	0.625	0.688	0.748	0.806	0.841	0.895	0.945
m_c	0.299	0.280	0.267	0.257	0.252	0.246	0.241
m_c (sg)	0.319	0.290	0.270	0.256	0.250	0.241	0.234

Table 5.2: NP coefficients α, conduction masses m_c, and effective diameters d_e for the present comparison to tight-binding (circular case). The m_c for the case of a single (sg) α is given as well.

d_e [nm]	2.044	2.536	3.010	3.500	4.000	4.529	4.926
$\alpha[\mathrm{eV}^{-1}]$	0.636	0.755	0.888	1.044	1.222	1.427	1.595
m_c	0.326	0.303	0.289	0.279	0.272	0.267	0.264
m_c (sg)	0.334	0.295	0.271	0.255	0.244	0.235	0.230

conduction masses, can be evaluated at arbitrary positions [6]. The conduction masses from the NP model are compared to atomistic data in Fig. 3.15. A satisfactory quantitative agreement with the conduction masses extracted from the tight-binding band structure can be observed. The continuity of α has been used to extract the parameters for the calculations in Sec. 4.9.2 from the previous chapter. The parameters for the various nanowires are given in Tab 5.3. With the specifications from Fig. 4.7, the effective diameters for the square and triangular case are $2t_c/\sqrt{\pi}$ and $t_c/\sqrt{\pi}$, respectively [7].
The NP coefficient α exhibits a notable dependence on the effective diameter d_e. This dependence suggests that there is no overall agreement between the actual silicon conduction band and

[6] It has to be noted that for large d_e, the α diverges and becomes very sensitive for variations of the fitting parameters. However, the analytic expression for the NP coefficient and the conduction mass are restricted to the range of diameters used to fit the atomistic band edges.

[7] Note that for the triangular FETs, the α from the circular case has been employed. Conversely, atomistic nanowires with a triangular shape can be used to provide the correct α.

Table 5.3: NP coefficients α, conduction masses m_c, and effective diameters d_e for the various nanowire FETs from Sec. 4.9.2.

t_c [nm]	2	3	4	5	6	7
$\alpha[\mathrm{eV}^{-1}]$ (sq)	0.598	0.750	0.886	1.003	-	-
m_c (sq)	0.309	0.267	0.247	0.235	-	-
d_e [nm] (sq)	2.257	3.385	4.514	5.642	-	-
$\alpha[\mathrm{eV}^{-1}]$ (ci)	0.627	0.885	1.221	1.627	-	-
m_c (ci)	0.329	0.289	0.272	0.263	-	-
d_e [nm] (ci)	2	3	4	5	-	-
$\alpha[\mathrm{eV}^{-1}]$ (tr)	-	-	0.685	0.833	1.006	1.203
m_c (tr)	-	-	0.314	0.293	0.281	0.273
d_e [nm] (tr)	-	-	2.257	2.821	3.385	3.949

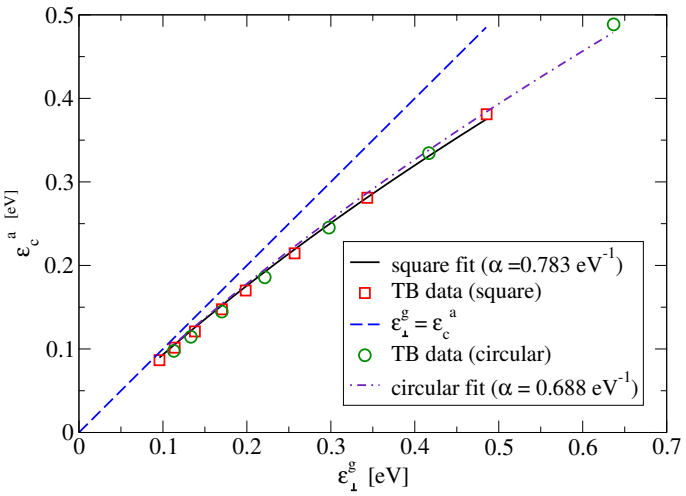

Figure 5.2: Tight-binding band edges (cf. Fig. 3.14) versus the corresponding value computed by means of the EMA (square and circular wires). The solid and dash-dotted lines denote fits of the relation (5.1) to the tight-binding data for a single α.

Table 5.4: NP coefficients α and conduction masses m_c for the various planar FETs from Sec. 4.9.3.

t_c [nm]	1	2	3	5
$\alpha[\mathrm{eV}^{-1}]$	7.731	9.567	12.294	18.696
m_c	0.706	0.423	0.344	0.285

the nonparabolic dispersion (3.72) for a single α. Even in the case of a perfect agreement, when further energy bands become relevant, the α has to be modified to account for this effect. However, it is possible to extract a single α for a set of atomistic and corresponding EMA band edges [31, 79] by fitting the relation (5.1) to the entire dataset. Figure 5.2 shows the fits to the TB sq100 band and TB ci edges from Fig. 3.14 [8]. The resulting NP coefficients differ from the value obtained by the bulk method described in Sec. 3.8. This finding is not surprising since the present fit employs the band edge and not the density of states. Neglecting the d_e-dependence of α is investigated in Sec. 5.3. The modified conduction masses when using the single α from Fig. 5.2 are reported in Tabs. 5.1 and 5.2. In the following, using a single α is referred to as the NPsingle model.

Finally, NP coefficients are extracted for the planar FETs simulated in Sec. 4.9.3. Again, this is accomplished by means of tight-binding band edges. The parametrization by Niquet et al. [30] is employed, i.e. the 100 case from Fig 3.16, with the bulk masses $m_l = 0.918$ and $m_t = 0.191$. As in the wire case, the condition (5.1) is employed to extract the α. In the well case, the transverse mass m_y is increased in the same way as the conduction mass m_x. The NP coefficients and conduction masses are reported in Tab. 5.4 for the various channel thicknesses t_c.

[8] The EMA band edges from Fig. 3.14 are employed for the fit.

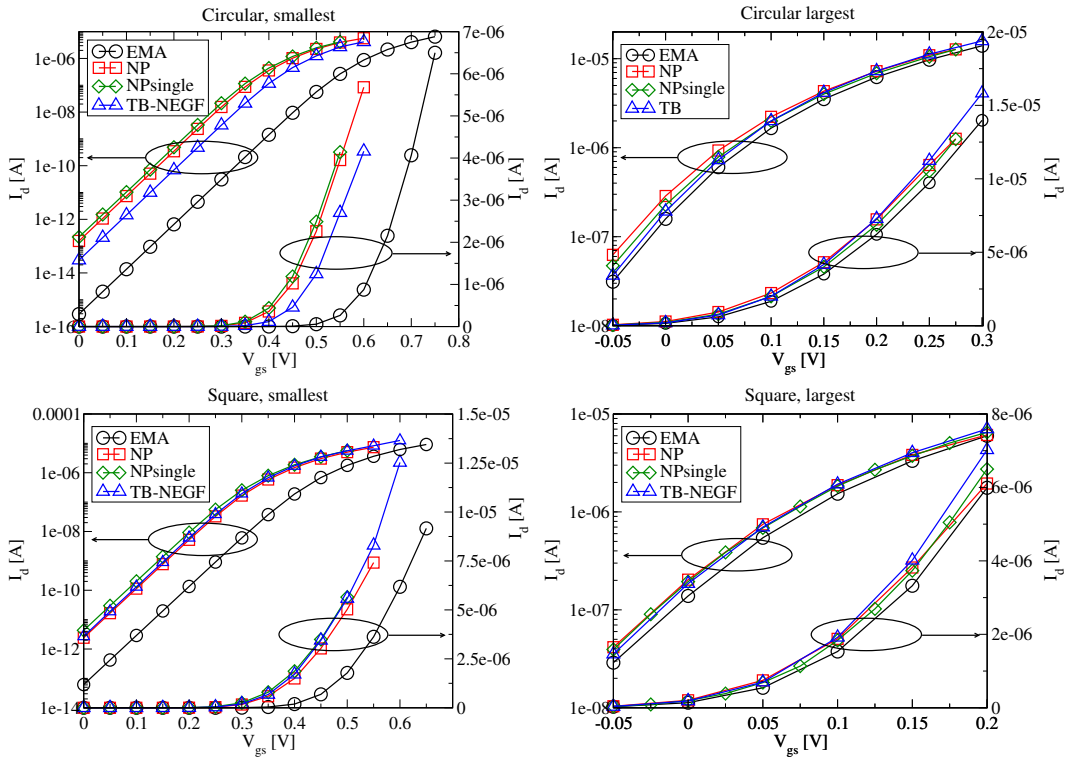

Figure 5.3: Transfer characteristics of the smallest and largest square and circular nanowire FETs by various models. The gate length is $l_g = 15$ nm.

5.3 Simulation Results

The circular and square nanowire FETs used for the comparison have the same structure as the FETs from Sec. 4.9, i.e. gate-all-around, $V_{\rm ds} = 0.6$ V, $l_s = l_d = 10$ nm, and the source and drain regions are n-doped with 10^{20} cm^{-3}. The diameters are dictated by the atomic structure of the cross section and are reported in Tabs. 5.1 and 5.2. For the smallest diameter, the gate lengths take the values $l_g[\text{nm}] = \{5, 10, 15, 20\}$ while $l_g = 15$ nm is used for the remaining diameters.

Transfer characteristics of the nanowire FETs with the smallest and largest effective diameter are given in Fig. 5.3. The discrepancy between the EMA and TB-NEGF models is most pronounced for the smallest diameters. For this case, a detailed analysis of the $V_{\rm th}$, SS, and $I_{\rm on}$ can be found in Fig. 5.4. The parameters for the extraction of these quantities are $I_{\rm th} = 10^{-7}$ A, $\Delta V_{\rm on} = 0.2$ V, and $\Delta V_{\rm SS} = 0.2$ V, i.e. the same as in Secs. 4.9.2 and 4.9.3. For $l_g = 5$ nm, the SS computed by the TB-NEGF model is smaller than the corresponding EMA value due to the larger conduction mass. The NP and NPsingle models provide an improvement for both circular and square FETs. In the square case, the error $(V_{\rm th} - V_{\rm th}^{\rm TB})$ between the threshold voltage computed by the TB-NEGF model $V_{\rm th}^{\rm TB}$ and the remaining models, shows the superior improvement by the NP model compared to NPsingle. Furthermore, the error for the EMA case is comparable to the band edge overestimation $\epsilon_\perp^g - \epsilon_c^a$ (dash-dotted line) for large l_g. On the other hand, the overestimation is roughly equal to the $V_{\rm th}$ shift between the EMA and NP. This finding is not surprising since $\epsilon_\perp^g - \epsilon_c^a$ is used to calibrate the NP model. In the circular case, the improvement is less obvious

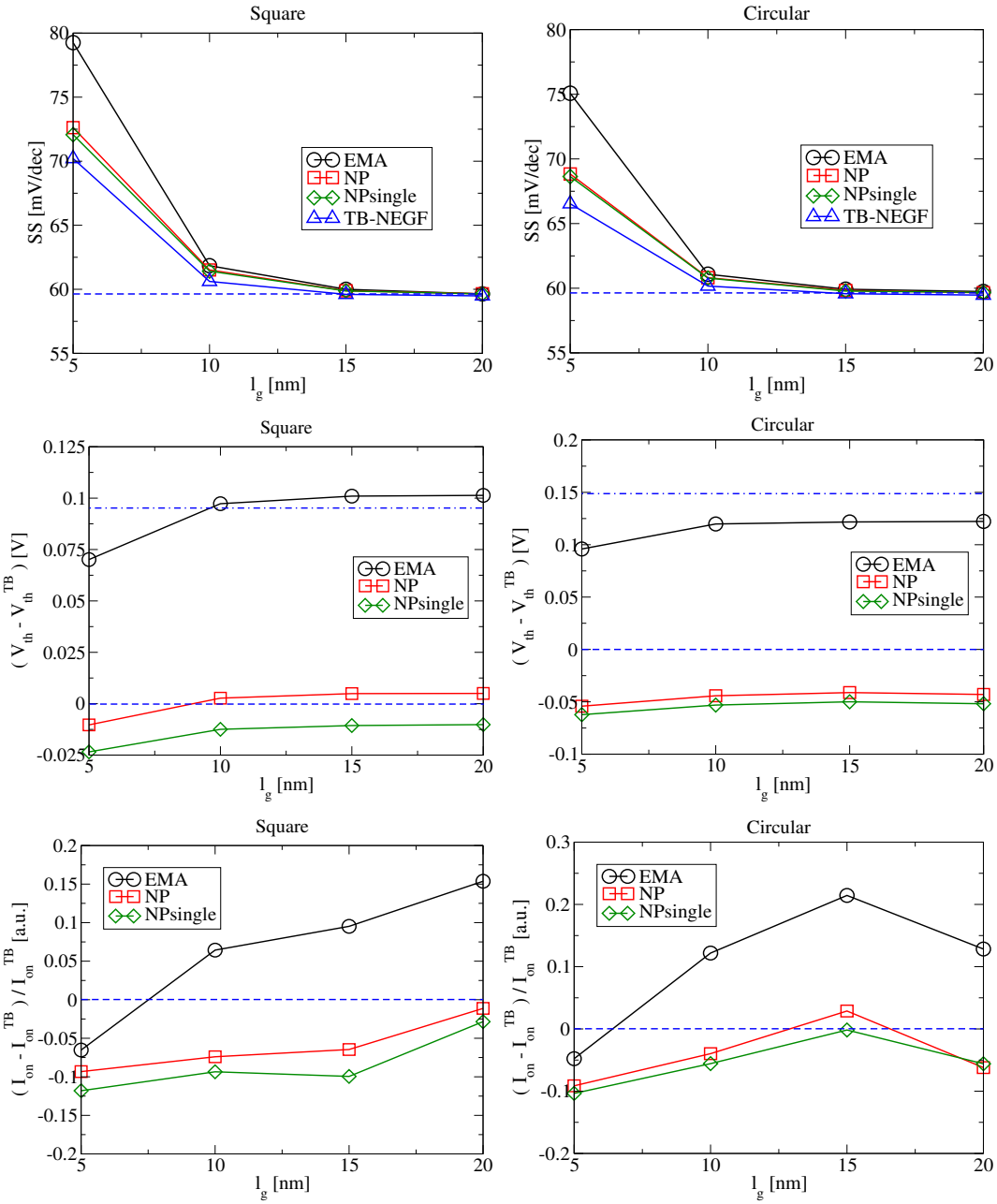

Figure 5.4: The errors $(V_{\rm th} - V_{\rm th}^{\rm TB})$ and $(I_{\rm on} - I_{\rm on}^{\rm TB})/I_{\rm on}^{\rm TB}$ as well as the subthreshold slope (SS) of the smallest square and circular nanowire FET as a function of the gate length l_g by various models. The dash-dotted lines denote the band edge overestimation $\epsilon_\perp^g - \epsilon_c^a$ by the EMA.

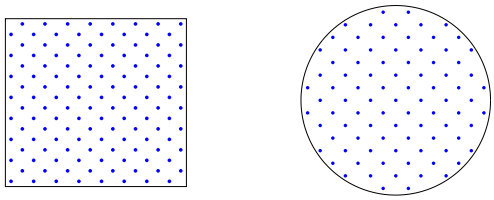

Figure 5.5: Cross sections and effective contours of the smallest square and circular nanowires.

compared to square FETs. In particular, the $(V_{\mathrm{th}} - V_{\mathrm{th}}^{\mathrm{TB}})$ for the EMA case notably differs from $\epsilon_{\perp}^{g} - \epsilon_{c}^{a}$. Several causes are possible for the reduced improvement by the NP compared to the square case. From a numerical point of view, the spectral method from Sec. 4.8 might introduce an error in the circular case. However, the method is in good agreement with the FEM and is able to satisfactorily reproduce analytic results. From a conceptual point of view, instead of using analytic fits to the TB band edges, the actual energies and conduction masses could be employed. However, in the square case, the use of analytic fits yields the expected improvement. On the other hand, the use of perfect circles to approximate the circular wire cross section might not be the best choice, particularly for small diameters, as highlighted in Fig. 5.5. The reduction of the error $(I_{\mathrm{on}} - I_{\mathrm{on}}^{\mathrm{TB}})/I_{\mathrm{on}}^{\mathrm{TB}}$ by the NP model is more evident in the circular case than in the square case for the smallest diameter. The analysis is completed by a plot of the errors $(I_{\mathrm{on}} - I_{\mathrm{on}}^{\mathrm{TB}})/I_{\mathrm{on}}^{\mathrm{TB}}$ and $(V_{\mathrm{th}} - V_{\mathrm{th}}^{\mathrm{TB}})$ as a function of the wire diameter in Fig. 5.6. The improvement of the threshold voltage by the NP model is noticeable over the whole range of diameters in the square case, as well as the agreement between $\epsilon_{\perp}^{g} - \epsilon_{c}^{a}$ and the EMA error. For the circular FETs, the improvement by the NP model remains inferior compared to the square case for an increasing diameter. However, the NPsingle model seems to perform better than NP for larger diameters. The error $(I_{\mathrm{on}} - I_{\mathrm{on}}^{\mathrm{TB}})/I_{\mathrm{on}}^{\mathrm{TB}}$ reveals that the TB-NEGF current in the on-current regime, for larger diameters, is even higher than the EMA current. This finding is counter-intuitive from a physical point of view since the conduction mass in the TB-NEGF case is still much larger than the bulk value. A better improvement for the on-current in the square case is reported in Ref. [77] which employs different effective masses and contour fits to the atomic cross sections but relies on the same NP model and comparison to the TB-NEGF approach. The results are plotted in Fig. 5.7.
Finally, the on-currents and threshold voltages computed by the TB-NEGF approach are plotted in Fig. 5.8 as a function of the effective diameter. Similar to the NP and EMA data plotted in Fig. 4.13, the V_{th} roughly follows a common line.

5.4 Summary

The calibration of the present NP models by means of atomistic band structures has been described at the beginning of this chapter. In particular, the parameters belonging to a series of square and circular nanowire FETs are reported. These parameters are used to validate the corresponding NP models by means of comparisons with a full-band TB-NEGF approach. For

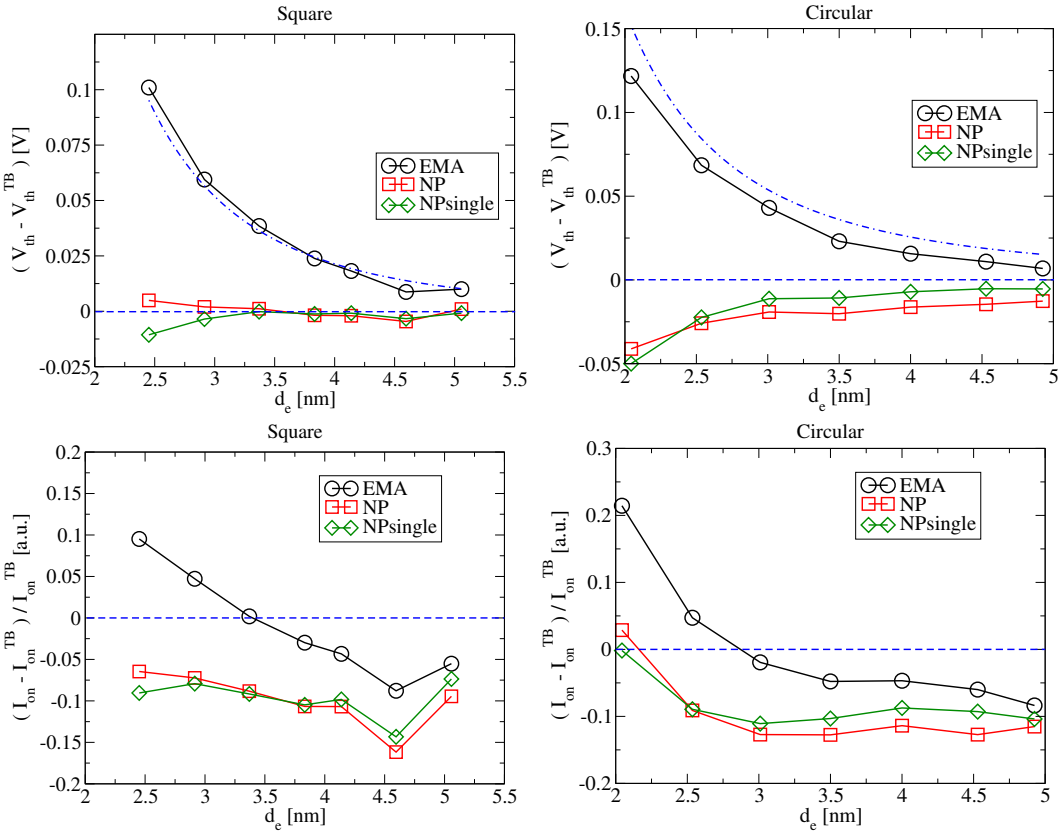

Figure 5.6: The errors $(I_{\rm on} - I_{\rm on}^{\rm TB})/I_{\rm on}^{\rm TB}$ and $(V_{\rm th} - V_{\rm th}^{\rm TB})$ of square and circular nanowire FETs as a function of the effective diameter d_e by various models. The dash-dotted lines denote the band edge overestimation $\epsilon_\perp^g - \epsilon_c^a$ by the EMA. The gate length is $l_g = 15$ nm.

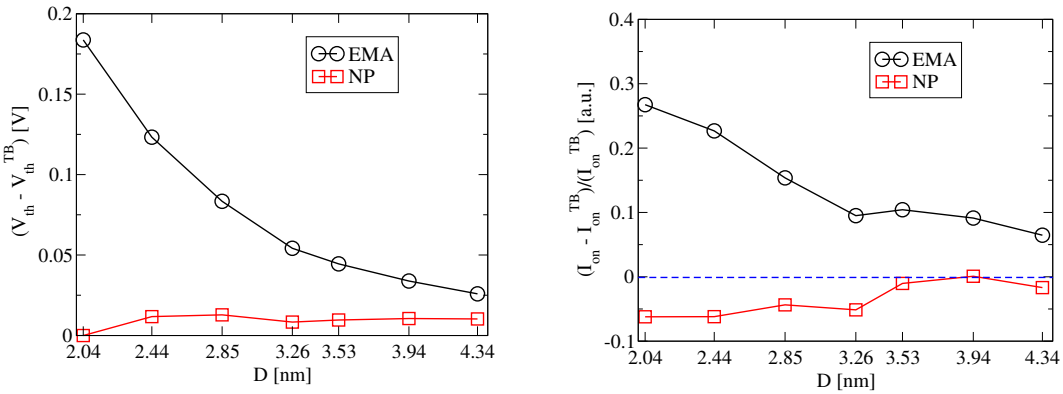

Figure 5.7: The errors $(I_{\rm on} - I_{\rm on}^{\rm TB})/I_{\rm on}^{\rm TB}$ and $(V_{\rm th} - V_{\rm th}^{\rm TB})$ of square FETs as a function of the wire width D by various models. The data is taken from Ref. [77].

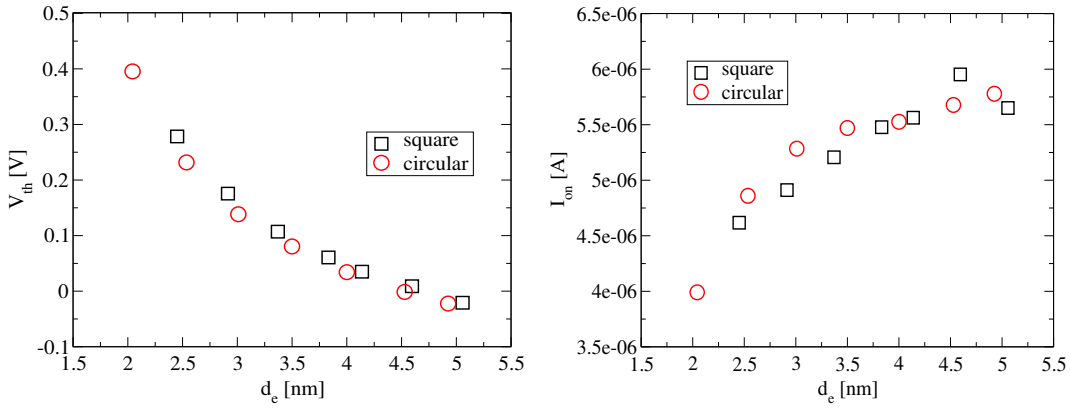

Figure 5.8: Threshold voltage $V_{\rm th}$ and on current $I_{\rm on}$ of circular and square nanowire FETs computed by the TB formalism. The gate length is $l_g = 10$nm.

the smallest diameters, the improvement provided by NP is particularly evident for the $V_{\rm th}$ of the square FETs and the SS of both types of devices. An improvement for the $V_{\rm th}$ of the circular FETs and the on-current of both types of devices is less evident. Different parameterizations provided in a previous work for the same NP models yield a more satisfactory improvement for the on-current of square FETs. Amongst other possible causes for the unexpected outcome in the circular case, it has to be investigated whether a perfect circle is an appropriate choice to approximate the cross section of, particularly small, circular nanowires. Provided that this behavior is related to the irregularity of the wire surface for very small diameters and no satisfactory improvement can be achieved by continuum type approaches such as the EMA or NP, the TB-NEGF approach could be employed since the computational burden does not notably exceed the one caused by the present transport simulator. Typically, $\sim$ 400 CPU hours are required by the TB-NEGF simulator compared to the $\sim$ 60 CPU hours needed by the present simulator for the smallest nanowires. Finally, except for the $V_{\rm th}$ of circular FETs for larger diameters, the use of a diameter-independent α does not yield a significant improvement compared to the conventional NP model.

Chapter 6

Inclusion of Scattering

6.1 Introduction

The quantum transport framework detailed in the previous two chapters does not account for scattering effects. A widely used approach, based on a phenomenological description of scattering, is given by the Büttiker probes [111, 112]. Going a step further, the inclusion of scattering effects can be accomplished by means of the NEGF [1]. In this approach, scattering is introduced in a perturbative way by means of self-energies which depend on the correlation functions, i.e. the Green's functions, and vice versa. Compared to the Büttiker probes, the NEGF is less bound to input parameters but causes a larger computational burden due to the interdependence between the quantities.
In this chapter, combined and particular impacts of NP and scattering on transfer characteristics of nanowire FETs are investigated. Scattering is treated on the level of the first self-consistent Born approximation and a novel method [113] for the treatment of the boundary conditions is employed. The result section is preceded by a short survey on NEGF for quantum transport. The interested reader can find detailed informations about the NEGF for quantum transport in Refs. [114, 115, 116, 117, 118, 119, 53, 17]

6.2 A Short Survey on NEGF

6.2.1 Steady-State Quantum Transport Equations

Before moving to the coupled-mode expansion, the continuous form of the steady state quantum transport equations will be given. The solution variables of the steady-state Dyson and Keldysh equation are the retarded (G^R) and lesser ($G^<$) Green's function [114, 115]:

$$\int \left[(E - H(\vec{r}))\delta(\vec{r} - \vec{r}_1) - \Sigma^R(\vec{r}, \vec{r}_1, E)\right] G^R(\vec{r}_1, \vec{r}', E) d\vec{r}_1 = \delta(\vec{r} - \vec{r}'),$$

$$G^<(\vec{r}, \vec{r}', E) = \int\int G^R(\vec{r}, \vec{r}_1, E)\Sigma^<(\vec{r}_1, \vec{r}_2, E)G^A(\vec{r}_1, \vec{r}', E) d\vec{r}_1 d\vec{r}_2,$$

where $H(\vec{r})$ is the Hamiltonian from Eq. (4.1) describing the device. The self-energy

$$\Sigma(\vec{r}, \vec{r}', E) = \Sigma_{int}(\vec{r}, \vec{r}', E) + \Sigma_{bc}(\vec{r}, \vec{r}', E) \tag{6.1}$$

[1] The implementation of the NEGF, inelastic scattering processes, and related approximations in SIMNAD has been provided by Martin Frey.

contains the electron-phonon interaction $\Sigma_{int}(\vec{r},\vec{r}',E)$ as well as the boundary conditions $\Sigma_{bc}(\vec{r},\vec{r}',E)$. The Hartree potential is included in $H(\vec{r})$.

6.2.2 Coupled-Mode Expansion

For devices with a well-defined transport direction, such as the ones employed in this work, the coupled-mode approach provides an attractive alternative to the real-space discretization of the Dyson and Keldysh equations from a computational point of view [81, 60] [2]. All quantities are expressed by means of the transverse modes $\{\psi_i^{(n)}(y,z)\}$ from Sec. 4.5 and restricted to the points $\{x_n\}$ shown in Fig. 4.3 along the transport direction. For instance, the retarded Green's function reads

$$G^R(\vec{r},\vec{r}',E) = \sum_{i,j} G^R_{ij}(x_n,x_m,E)\psi_i^{(n)*}(y,z)\psi_j^{(m)}(y',z'), \tag{6.2}$$

where $G^R_{ij}(x_n,x_m,E)$ is the solution of the Dyson equation formulated in mode-space

$$\begin{aligned}
&\sum_{i,n'}\Bigg[E\delta_{ji}\delta_{nn'} - \int dydz\psi_j^{(n)*}(y,z)H(\vec{r})\psi_i^{(n')}(y,z) \\
&- \int dydz\int dy'dz'\psi_j^{(n)*}(y,z)\Sigma^R(\vec{r},\vec{r}')\psi_i^{(n')}(y',z')\Bigg]G^R_{ik}(x_{n'},x_m,E) \\
&= \sum_{i,n'}\left[E\delta_{ji}\delta_{nn'} - H_{ji}(x_n,x_{n'}) - \Sigma^R_{ji}(x_n,x_{n'},E)\right]G^R_{ik}(x_{n'},x_m,E) \\
&= \delta_{jk}\delta(x_n-x_m).
\end{aligned} \tag{6.3}$$

Similarly, the expressions for the self-energies in mode-space can be derived [77]. The present NP models enter the NEGF via the transverse modes $\psi_i^{(n)}(y,z)$ which are computed by the nonparabolic transverse Hamiltonians (3.79) and (3.81).

6.2.3 Electron-Phonon Scattering

In general, scattering in quantum mechanics is spatially correlated since the particles involved are described by their own wave functions. For computational purposes, the scattering processes are assumed to occur locally in space. Furthermore, the phonon system is kept at equilibrium and the wave functions are approximated by the bulk correspondents. The electron-phonon matrix elements $|M_q|^2$ are computed by means of a perturbative approach within deformation potential theory [120]. For the steady-state case, the lesser self-energy reads [116] [3]

$$\begin{aligned}
\Sigma^<(\vec{r},\vec{r}',E) &= \frac{1}{(2\pi)^3}\int e^{i\vec{q}(\vec{r}-\vec{r}')}|M_q|^2 \\
&\times\left[N_qG^<(\vec{r},\vec{r}',E-\hbar\omega_q) + (N_q+1)G^<(\vec{r},\vec{r}',E+\hbar\omega_q)\right]d\vec{q}
\end{aligned} \tag{6.4}$$

within the self-consistent Born approximation. The retarded self-energy is generally given by

$$\begin{aligned}
\Sigma^R(\vec{r},\vec{r}',E) &= \frac{1}{2}(\Sigma^>(\vec{r},\vec{r}',E) - \Sigma^<(\vec{r},\vec{r}',E)) \\
&+ iP\frac{1}{2\pi}\int\frac{\Sigma^>(\vec{r},\vec{r}',E') - \Sigma^<(\vec{r},\vec{r}',E')}{E-E'}dE'
\end{aligned} \tag{6.5}$$

[2] The reduction of the matrix size for the transport problem is the main advantage of the coupled-mode approach. The number of matrix elements is proportional to the square of the number of modes while in the real space variant the matrix size is proportional to the number of discretization points in the transverse direction. On the other hand, accounting for gate leakage currents is not straightforward in the coupled-mode approach.

[3] The equations are given in their continuous form. For the transformation to mode-space see Sec. 6.2.2.

with $P\int dE'$ being the principal part of the integration. Conversely, the retarded self-energy can be calculated directly from the solution variables $(G^R, G^<)$ of the Keldysh and Dyson equations [118]

$$\begin{aligned}\Sigma^R(\vec{r},\vec{r}',E) &= \frac{1}{(2\pi)^3}\int e^{i\vec{q}(\vec{r}-\vec{r}')}|M_q|^2 \\ &\times\Big\{(N_q+1)G^R(\vec{r},\vec{r}',E-\hbar\omega_q)+N_qG^R(\vec{r},\vec{r}',E+\hbar\omega_q) \\ &+\frac{1}{2}\left[G^<(\vec{r},\vec{r}',E-\hbar\omega_q)-G^<(\vec{r},\vec{r}',E+\hbar\omega_q)\right] \\ &+iP\int\frac{dE'}{2\pi}\frac{G^<(\vec{r},\vec{r}',E-E')}{E'-\hbar\omega_q}-\frac{G^<(\vec{r},\vec{r}',E-E')}{E'+\hbar\omega_q}\Big\}d\vec{q}.\end{aligned} \tag{6.6}$$

In order to reduce the computational burden related to the principal part integrals in Eqs. (6.5) and (6.6) different approximations are employed. Neglecting the principle part integral in Eqs. (6.5) and (6.6) is referred to as *Approximation I* and *II*, respectively, while keeping the integrals is referred to as the *Full* model for scattering. The respective impacts are investigated in Sec. 6.3. For a discussion on the lesser self-energy for intravalley acoustic phonon and intervalley phonon scattering see Ref. [77].

6.2.4 Calculation of Density and Current

The lesser self-energies for the density and current calculation essentially consist of three contributions, i.e. coherent terms from the source and drain contacts and an incoherent term from the electron-phonon interaction. The density and current read

$$\begin{aligned}n(x_n,y,z) &= \frac{-i}{\pi}\sum_v\sum_{ij}\int G^{<,v}_{ij}(x_n,x_n,E)\psi^{(n)*}_i(y,z)\psi^{(n)*}_j(y,z)dE \\ I(x_n) &= -\frac{e}{\pi\hbar}\sum_v\sum_{ij}\int\left(2Re(H^v_{ij}(x_n,x_{n+1})G^{<,v}_{ji}(x_{n+1},x_n))\right)dE\end{aligned}$$

with v being the valley index.

6.3 Simulation Results

The nanowire FETs used for the present simulations have the same structure as the ones from Sec. 4.9.2. The main difference is that the gate contact is laid over three faces only (triple gate) compared to the gate-all-around structures employed in Sec. 4.9.2. The geometry parameters are $l_s = l_d = 9.7$ nm, $l_g = 15$ nm, $t_{ox} = 1$ nm, $t_c = 3.26$ nm, and the source and drain regions are n-doped with a concentration of 10^{20}cm^{-3}. The parameters for the extraction of I_{on} and V_{th} are $\Delta V_{\text{on}} = 0.3$ V and $I_{\text{th}} = 10^{-7}$ A. Further details such as the NP coefficients and parameters for the scattering are given in Ref. [77].
Figure 6.1 shows transfer characteristics computed by means of different approximations of the retarded self-energy (cf. Sec. 6.2.3). *Approximation I* yields an underestimation of the subthreshold current of up to 80% compared to the *Full* version. On the other hand, the discrepancy diminishes in the on-current regime. This behavior can be explained by the absence of the energy renormalization on the effective barrier height [77] due to the neglect of the principle part integral in Eq. (6.5). Conversely, *Approximation II* is found to be in excellent agreement with

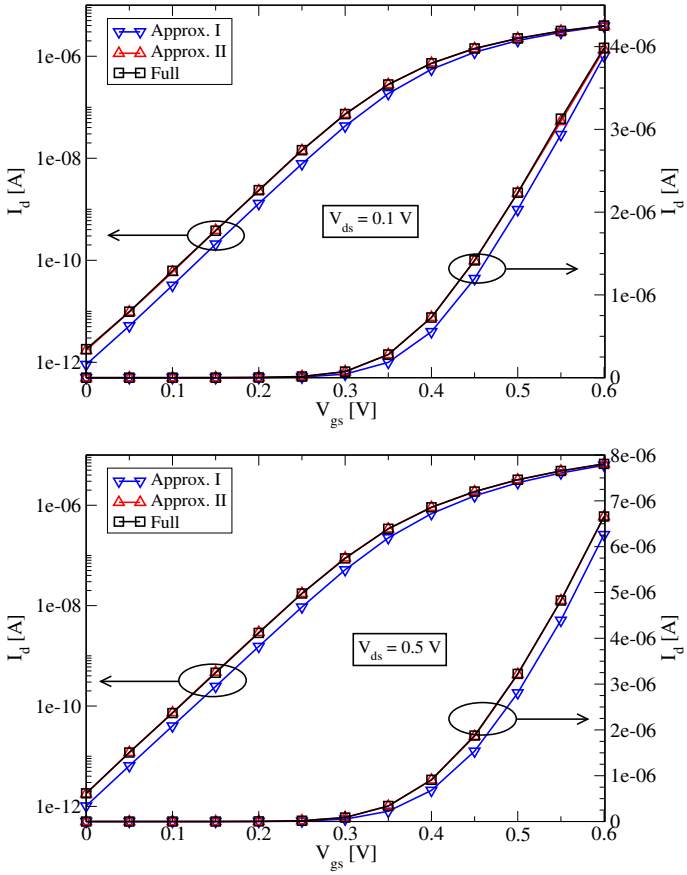

Figure 6.1: Transfer characteristics by different approximations to the retarded self-energy compared to the full solution.

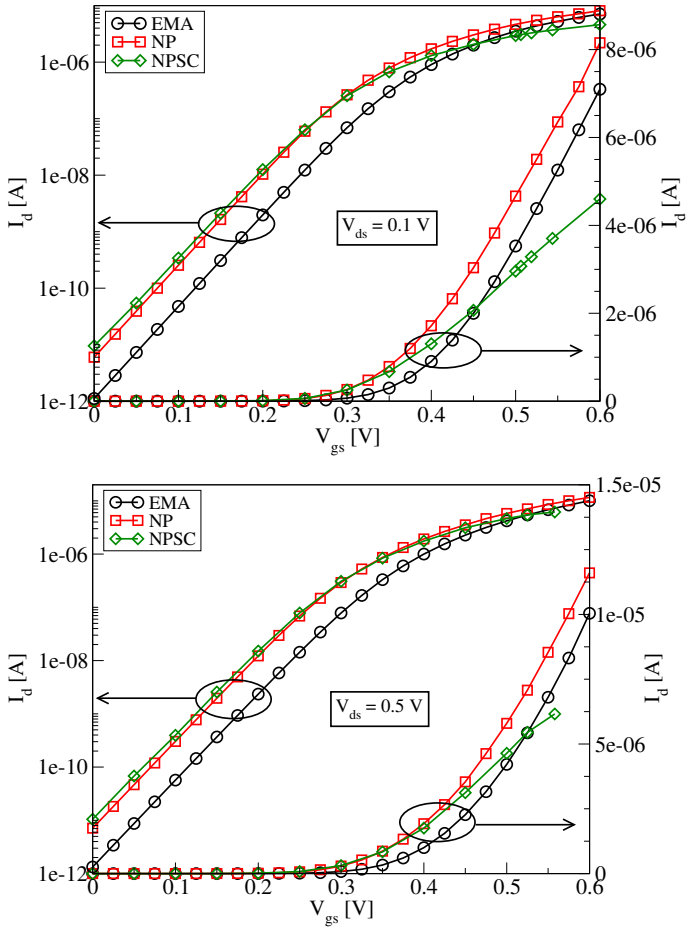

Figure 6.2: Comparison of transfer characteristics computed by means of the NP model and the combination of scattering with NP (NPSC). The ballistic EMA curve is shown as well.

the *Full* version throughout the whole voltage range. This approximation is used in the following current calculations.

Transfer characteristics resulting from simulations including both nonparabolicity and scattering (NPSC) are compared to data from simulations containing only NP, SC, or purely ballistic EMA. The currents are plotted in Figs. 6.2 and 6.3 and the extracted quantities are reported in Tab. 6.1. For both forward biases $V_{\mathrm{ds}}[V] \in \{0.1, 0.5\}$, the threshold voltage is mainly affected by the inclusion of NP while the effect of scattering is in the order of 1% and, therefore, negligible. Conversely, the on-current regime seems to be mainly dominated by scattering compared to the impact of the increased conduction mass from the NP model. The impact of the latter model on the on-current is however not negligible. Identifying the contributions from the particular models NP and SC to the results obtained by the combination NPSC is a rather difficult task. As shown in Fig. 6.2, the subthreshold current obtained by the NPSC is very similar to the one obtained by NP alone thus suggesting a minor interaction between SC and NP in this regime. The situation is different for the on-current regime, i.e. contributions from the single methods SC and NP do

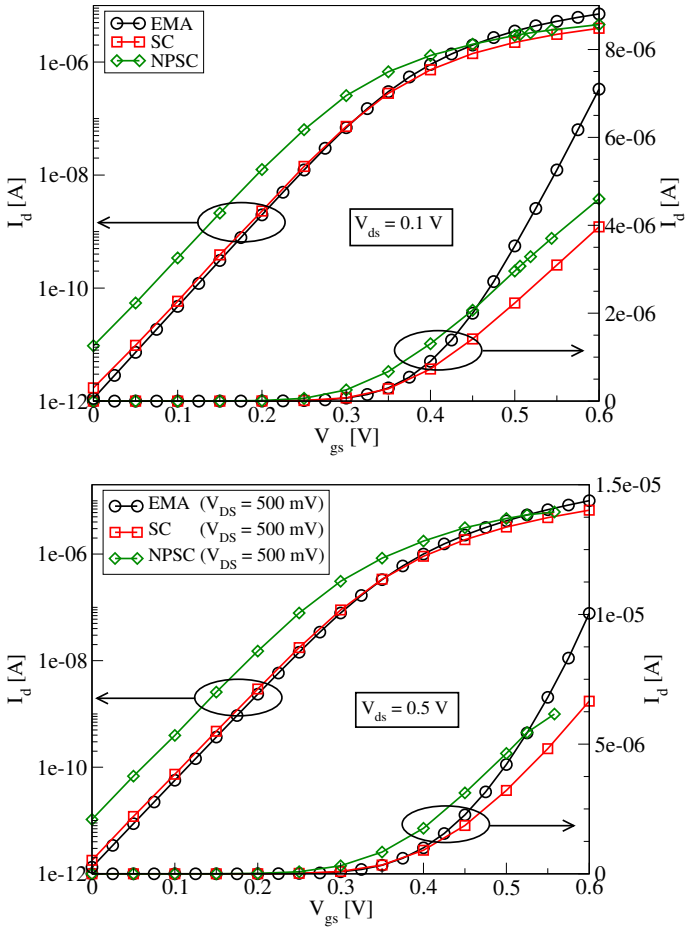

Figure 6.3: Comparison of transfer characteristics computed by means of scattering (SC) and the combination of scattering with NP (NPSC). The ballistic EMA curve is shown as well.

Table 6.1: Threshold voltages and on-currents extracted from the transfer characteristics plotted in Figs. 6.2 and Fig. 6.3. Shown are the results of the two ballistic simulations, EMA and NP, as well as of the simulations including scattering SC and the combination with nonparabolicity NPSC. Two source-to-drain biases have been considered: $V_{\mathrm{DS}} = 0.1\mathrm{V(I)}$ and $V_{\mathrm{DS}} = 0.5\mathrm{V(II)}$. The relative(REL) deviations $|I_{\mathrm{on}}^{\mathrm{EMA}} - I_{\mathrm{on}}^{\mathrm{NP,SC,NPSC}}|/I_{\mathrm{on}}^{\mathrm{EMA}}$ and $|V_{\mathrm{th}}^{\mathrm{EMA}} - V_{\mathrm{th}}^{\mathrm{NP,SC,NPSC}}|/V_{\mathrm{th}}^{\mathrm{EMA}}$ are shown as well.

mode	V_{T}(I)	I_{ON}(I)	V_{T}(II)	I_{ON}(II)
EMA	0.312 V	7.53e-06 A	0.308 V	1.06e-05 A
NP	0.266 V	6.86e-06 A	0.262 V	9.25e-06 A
NP REL	14.74 %	8.994 %	14.94 %	12.87 %
SC	0.311 V	4.15e-06 A	0.304 V	6.81e-06 A
SC REL	0.32 %	45.28 %	1.30 %	35.86 %
NPSC	0.265 V	4.04e-06 A	0.258 V	6.16e-06 A
NPSC REL	15.06 %	46.39 %	16.23 %	41.94 %

not roughly add up to the results from the combined method NPSC. A tentative explanation for this finding could be drawn from a qualitative point of view on how inelastic scattering affects the charge carriers in the effective source region. The impact of a modified conduction mass on the longitudinal velocity of an electron is more pronounced when the total energy is much higher than the corresponding conduction band edge. In a non-equilibrium system, scattering causes a redistribution of the electrons in energy and momentum. Phonon emission is the favored process [77] for electrons compared to phonon absorption, provided that the final state $E \rightarrow E - \hbar\omega_q$ is available. Therefore, the electrons tend to dissipate energy in the effective source region by relaxing toward the band edge and the reduced energy alleviates the impact of the conduction mass on the longitudinal velocity.

6.4 Summary

Inelastic scattering effects are combined with NP within the NEGF in order to investigate the respective impacts on transfer characteristics of triple-gate nanowire FETs. The shift in the threshold voltage was mainly attributed to NP while scattering dominates the on-current regime when considering the particular effects. The energy dissipation of electrons in the effective source region is assumed to alleviate the effect of NP when both effects are combined thus yielding a smaller degradation of the on-current. The findings in this chapter underline the importance of scattering. However, the inclusion of this effect goes at the expense of the simulation time which considerably increases compared to the ballistic case.

Chapter 7

Concluding Remarks

In this work, quantum transport simulators for the treatment of planar and nanowire FETs have been developed (extended). A notable effort has been dedicated to the algorithmic, numerical, and computational improvement of the pre-existing SIMNAD software. In particular, the OpenMP parallelization of time-consuming routines such as the computation of the injected charge density or the solution of the transverse eigenvalue problems for both the SMA or the coupled-model NEGF, turned out to be a decisive advancement especially for large devices. The use of state-of-the-art software libraries for both linear and eigenvalue problems not only improves the overall performance but also ensures a better supportability for future developments. For the treatment of complicated surfaces, the FEM has been implemented for both the solution of the Schrödinger and the Poisson equation. For the FEM variant, the ballistic quantum transport framework has been parallelized by means of the MPI, i.e. made suitable for large scale distributed memory compute clusters such as the Cray XT5.

The SMA is mostly used for the calculation of the transfer characteristics since several advantages compared to the coupled-mode NEGF exist for the ballistic limit. Some aspects concerning the use of the SMA for arbitrary transport directions within the EMA have been briefly addressed. However, the NEGF is an appropriate choice when scattering has to be treated beyond a phenomenological approach. Particularly the real-space variant of NEGF is attractive when complex geometries, gate leakage currents, or changes in the transport direction have to be taken into account. This flexibility goes at the expense of the computational burden which increases notably. A common denominator of the mentioned transport frameworks is the adequacy for scalability thus encouraging the consideration of high-performance languages such as CUDA (NVIDIA) for the use on GPUs.

In order to improve the EMA for the simulation of small FETs, a widely used NP model has been appropriately modified to fit within the SMA or the coupled-mode NEGF. This was accomplished by means of a Taylor expansion based on the presence of strong two- and one-dimensional confinements. Extensive simulations to investigate the impact of band structure effects have been carried out for planar and nanowire FETs of various shapes. Tight-binding band structures were used to determine the NP coefficients. Comparisons with a full-band tight-binding simulator reveal a satisfactory improvement by means of NP for the threshold voltage of square nanowire FETs while a similar improvement is missing for the circular case. However, since the present NP model shows the potential to improve the EMA (shown in previous work with different parametrizations) while preserving the same simulation efficiency, it is worth continuing the related investigations also with regard to more complex wire cross sections.

The tight-binding method is widely appreciated for the modeling of nanoelectronic devices on

the atomistic level since related simulation times are comparatively small. A part of this work was dedicated to the implementation of an EPM framework to be used in the future for quantum transport simulations. An advantage of EPM with respect to tight-binding is the more physical description of charge densities. Furthermore, the EPM is expected to be less time consuming than fully ab-initio methods. At the present stage, a band structure calculator based on the EPM is available for bulk, nanowires, and quantum wells.
A rather small part of this work was concerned with the combination of scattering [1] and NP in order to fully exploit the capabilities of the presently developed simulators. Since the related investigations are rather young compared to the effort put in the particular methods NP and scattering, the interpretations of the results are comparatively superficial and further investigative work is necessary for this topic.

[1] The implementation of the NEGF, inelastic scattering processes, and related approximations in SIMNAD has been provided by Martin Frey.

Appendix A

Band Structure Calculations

A.0.1 Empirical Tight-Binding Method

The tight-binding or LCAO method was originally proposed by Bloch in 1928. Almost a quarter of a century later, Slater and Koster [121] proposed a modified approach motivated by the fact that a rigorous evaluation of the large number of integrals involved in the LCAO method was almost impossible with the computational resources of that time. They suggested instead, that these integrals should be considered as adjustable parameters determined from results obtained by more accurate electronic structure calculations such as cellular methods or orthogonalized plane waves. Although the continuous increase of computing power has eliminated some of the reasons for developing this interpolation scheme, it has survived over several decades and evolved into a powerful tool for calculating physical properties of arbitrary systems. For a detailed survey on the Slater Koster (SK) theory and a list of applications see Ref. [122] and references therein. In particular, the SK theory provided the basis for the *empirical* tight-binding method. One of the major enhancements after the introduction of the SK theory was the addition of the excited s-like orbital, s*, by Vogl et al. [123] almost thirty years later which yielded better reproduction of the conduction band of diamond and zinc blende semiconductors along $\langle 100 \rangle$. A further improvement has been provided recently by Jancu et al. [124] for the case of X-valleys of group-IV and III-V semiconductors. Finally, parameter sets for silicon and germanium within the $sp^3d^5s^*$ empirical tight-binding model have been presented[109] allowing an accurate reproduction of the entire bulk band structure. In the following, the basic ideas related to the Slater Koster theory are summarized. Each atom is associated with a set of atomic-like orbitals $\phi_{i\sigma}$, where i denotes the position $\vec{b}_i$ of the atom in the crystal unit cell and σ is a quantum number for the atomic state. In general, orbitals related to atoms residing on different lattice sites are not orthogonal. A remedy is provided by Löwdin's method [125] to construct a set of states $\psi_{i\sigma}$ having symmetry properties similar to those of the corresponding $\phi_{i\sigma}$ but being orthogonal [1]

$$\int \psi^*_{j\tilde{\sigma}}(\vec{r} - \vec{R}_m - \vec{b}_j)\psi_{i\sigma}(\vec{r} - \vec{R}_n - \vec{b}_i)d\vec{r} = \delta_{ij}\delta_{nm}\delta_{\sigma\tilde{\sigma}}, \tag{A.1}$$

[1] An approach which avoids this orthogonalization is mentioned in Ref. [122].

where $\vec{R}_n$ and $\vec{R}_m$ belong to the reduced lattice (3.18) consisting of N unit cells Ω_c. For each atom i in the unit cell and orbital σ, a Bloch sum

$$\begin{aligned}\chi_{i\sigma}(\vec{k},\vec{r}) &= \frac{1}{\sqrt{N}}\sum_{\vec{R}_n\in\mathbf{\Gamma}^{\rm red}_{\vec{N}}} e^{i\vec{k}(\vec{R}_n+\vec{b}_i)}\psi_{i\sigma}(\vec{r}-\vec{R}_n-\vec{b}_i)\\ &= \langle\vec{r}|\vec{k}i\sigma\rangle \end{aligned} \tag{A.2}$$

can be constructed which fulfills the Bloch condition (3.10)

$$\chi_{i\sigma}(\vec{k},\vec{r}+\vec{R}) = e^{i\vec{k}\vec{R}}\chi_{i\sigma}(\vec{k},\vec{r}) \tag{A.3}$$

and orthonormality

$$\langle\vec{k}'j\widetilde{\sigma}|\vec{k}i\sigma\rangle = \delta_{\vec{k}'\vec{k}}\delta_{ij}\delta_{\widetilde{\sigma}\sigma}. \tag{A.4}$$

With N_a atoms in the unit cell and N_o orbitals per atom, the ansatz for the Bloch states reads

$$\Psi(\vec{k},\vec{r}) = \sum_{i=1}^{N_a}\sum_{\sigma=1}^{N_o} c_{i\sigma}(\vec{k})\langle\vec{r}|\vec{k}i\sigma\rangle = \langle\vec{r}|\vec{k}\rangle \tag{A.5}$$

which still fulfills the Bloch condition and orthonormality provided that

$$\sum_{i=1}^{N_a}\sum_{\sigma=1}^{N_o}|c_{i\sigma}(\vec{k})|^2 = 1. \tag{A.6}$$

Using the ansatz (A.5) for the Schrödinger problem (3.1) leads to the expression of the Hamiltonian $\mathbf{H}$ in terms of $|\vec{k}i\sigma\rangle$

$$\begin{aligned}&\langle\vec{k}\widetilde{\sigma}j|\mathbf{H}|\vec{k}\sigma i\rangle\\ &= \sum_{\vec{R}_n\in\mathbf{\Gamma}^{\rm red}_{\vec{N}}} e^{i\vec{k}(\vec{R}_n+\vec{b}_i-\vec{b}_j)}\int\psi^*_{j\widetilde{\sigma}}(\vec{r}-\vec{b}_j)\mathbf{H}\psi_{i\sigma}(\vec{r}-\vec{R}_n-\vec{b}_i)d\vec{r}\end{aligned} \tag{A.7}$$

Note that mixing between Bloch states to different $\vec{k}$ is automatically suppressed when using the ansatz (A.5), i.e. $\langle\vec{k}'|\mathbf{H}|\vec{k}\rangle = \delta_{\vec{k}\vec{k}'}\langle\vec{k}|\mathbf{H}|\vec{k}\rangle$. The potential $U(\vec{r})$ is approximated as a sum of spherically symmetric potential wells located at all the atoms of the crystal

$$U(\vec{r}) = \sum_{\vec{R}_p\in\mathbf{\Gamma}^{\rm red}_{\vec{N}}}\sum_{q=1}^{N_a} U_q(\vec{r}-\vec{b}_q-\vec{R}_p) \tag{A.8}$$

In the general case, the integral on the right hand side of Eq. (A.7) has contributions from three regions. First, the regions centered around the two atom-like wave functions $\psi_{j\widetilde{\sigma}}$ and $\psi_{i\sigma}$, and the region centered around the potentials at $\vec{b}_q+\vec{R}_p$. In the following, the two-center approximation is used, i.e. only integrals involving one (on-site) or two centers (two-center) are considered. Consequently, the remaining integrals in Eq. (A.7) have the form $\int\psi^*_{j\widetilde{\sigma}}(\vec{r}-\vec{u})\mathbf{H}\psi_{i\sigma}(\vec{r})d\vec{r}$ and depend only on the displacement $\vec{u} = \vec{b}_j-(\vec{R}_n+\vec{b}_i)$. The Hamilton matrix is now subdivided in $N_a\times N_a$ blocks of size $N_o\times N_o$ denoted by $\mathcal{H}^{ji}$. Allowing only interactions between nearest neighbors, the blocks are given by

$$\mathcal{H}^{ji} = \sum_{l=1}^{nn(j)} e^{i\vec{k}\vec{u}^l_{ji}}\mathbf{V}^{ji}(\vec{u}^l_{ji}), \tag{A.9}$$

where $\vec{u}^l_{ji} = \vec{R}_{n(l)} + \vec{b}_i - \vec{b}_j$ is a vector pointing to one of the $nn(j)$ nearest neighbors of the atom located at j and the matrix $\mathbf{V}^{ji}(\vec{u}^l_{ji})$ is related to the integrals $\int \psi^*_{j\widetilde{\sigma}}(\vec{r} - \vec{u}^l_{ji})\mathbf{H}\psi_{i\sigma}(\vec{r})d\vec{r}$. The blocks $\mathcal{H}^{ii}$ on the diagonal of the Hamilton matrix (A.7) are referred to as the on-site energy of the i^{th} atom. Within the SK theory the on-site and two-center integrals which enter the blocks (A.9) are replaced by adjustable parameters used to fit data obtained by more sophisticated methods or experiments. More details on the relation between these parameters and the corresponding integrals are given in Refs. [121, 122, 17, 10]. In the diamond or zincblende structure [14] the unit cell contains two atoms and each atom in the lattice is surrounded by four nearest neighbors. Thus, for elements crystallizing in these structures, the $sp^3d^5s^*$ tight-binding method without spin-orbit coupling leads to eigenvalue problems of size 20×20 for each wave vector $\vec{k}$.
The extension from infinite crystals to nanostructures within the tight-binding framework is straightforward [17]. If nothing else is specified, the tight-binding calculations in the remainder of this work are based on the parametrization [109] and boundary conditions [110] employed by Ref. [17].

A.0.2 Empirical Pseudopotential Method

The potential (3.1) is periodic with respect to the Bravais lattice (3.2) and can be expanded in terms of plane waves

$$U(\vec{r}) = \sum_{\vec{G}\in\widetilde{\mathbf{\Gamma}}} e^{i\vec{G}\vec{r}}\widehat{U}(\vec{G}) \tag{A.10}$$

according to Eq.(3.5). The empirical pseudopotential method involves a direct fit of the $\widehat{U}(\vec{G})$ to experimental or ab-initio band structures. Initially, Fermi surface data were used for metals and later on photoemission and reflectivity results provided the relevant informations. The method was successfully applied to silicon and germanium as well as to other important semiconductors [126]. The potential $U(\vec{r})$ is assumed to be a linear superposition of atomic potentials $V_{s(i)}$

$$U(\vec{r}) = \sum_i \sum_n V_{s(i)}(\vec{r} - \vec{R}_n - \vec{b}_i), \tag{A.11}$$

where $\vec{b}_i$ is the location of the i^{th} atom in the unit cell, $s(i)$ is the corresponding species, and $\vec{R}_n$ is the position of a site in the Bravais lattice. The coefficients of $U(\vec{r})$ become

$$\begin{aligned}
\widehat{U}(\vec{G}) &= \frac{1}{|\Omega_c|}\int_{\Omega_c} e^{-i\vec{G}\vec{r}}U(\vec{r})d\vec{r} \\
&= \sum_i \sum_n \frac{1}{|\Omega_c|}\int_{\Omega_c} e^{-i\vec{G}\vec{r}}V_{s(i)}(\vec{r} - \vec{R}_n - \vec{b}_i)d\vec{r} \\
&= \sum_i e^{-i\vec{G}\vec{b}_i}\frac{1}{|\Omega_c|}\int_{\mathbb{R}^3} e^{-i\vec{G}\vec{r}}V_{s(i)}(\vec{r} - \vec{b}_i)d\vec{r} \\
&\equiv \frac{1}{|\Omega_c|}\sum_i e^{-i\vec{G}\vec{b}_i}\widehat{V}_{s(i)}(|\vec{G}|),
\end{aligned} \tag{A.12}$$

where $V_{s(i)}(|\vec{G}|)$ depends only on the absolute value of $\vec{G}$ as the $V_{s(i)}(\vec{r})$ are assumed to be spherically symmetric. The pseudopotential enters the $\vec{k}$-dependent Schrödinger equation (3.13) which is employed to compute the Bloch states $\psi_n(\vec{k}, \vec{r})$. To account for the periodic boundary conditions (3.9), the Bloch factor $u(\vec{k}, \vec{r})$ is expanded in terms of plane waves

$$u(\vec{k}, \vec{r}) = \sum_{\vec{G}\in\widetilde{\mathbf{\Gamma}}} e^{i\vec{G}\vec{r}}\widehat{u}(\vec{G}). \tag{A.13}$$

as well. Using the ansatz (A.13) for the Schrödinger problem (3.13), multiplying on the left with $\exp(-i\vec{G}\vec{r})/(|\Omega_c|)$, and integrating over Ω_c, yields the secular equation

$$\sum_{\vec{G}'}\left[\frac{\hbar^2}{2m_e}(\vec{k}+\vec{G})^2\delta_{\vec{G},\vec{G}'}+\widehat{U}(\vec{G}'-\vec{G})\right]\widehat{u}(\vec{G}')=\epsilon(\vec{k})\widehat{u}(\vec{G}) \tag{A.14}$$

which is cast as matrix eigenvalue problem. In practice, the set of $\vec{G}$'s is delimited by a cutoff *co*

$$|\vec{G}|\leq co \tag{A.15}$$

which has to be increased until the energies $\epsilon_n(\vec{k})$ from Eq. (A.14) are well converged. This convergence strongly depends on the smoothness of the potential $U(\vec{r})$, i.e. on how fast the $\widehat{V}_{s(i)}(q)$ vanish. As in the tight-binding method, the ansatz $\psi(\vec{k},\vec{r})=\exp(i\vec{k}\vec{r})u(\vec{k},\vec{r})=\langle\vec{r}|\vec{k}\rangle$ with $u(\vec{k},\vec{r})$ from Eq. (A.13) automatically suppresses mixing between different $\vec{k}$, i.e. $\langle\vec{k}'|\mathbf{H}|\vec{k}\rangle=\delta_{\vec{k}\vec{k}'}\langle\vec{k}|\mathbf{H}|\vec{k}\rangle$.

The determination of an empirical pseudopotential starts with a guess for the $\widehat{V}_{s(i)}(|\vec{G}|)$. Next, the Schrödinger equation (A.7) is assembled and solved for a given set of wave vectors $\vec{k}$, i.e. the reduced lattice (3.18) for instance. This allows the calculation of the density of states described in Sec. 3.8.1 as well as other physical observables. These quantities are compared with experiments or ab-initio data and the $\widehat{V}_{s(i)}(q)$ are modified accordingly, if the desired agreement has not been achieved.

As an example the diamond structure is considered. The two atoms are placed at the positions $\vec{b}_1=-\vec{\tau}$ and $\vec{b}_2=\vec{\tau}$ with $\vec{\tau}=(a,a,a)/8$. For semiconductors with diamond lattice such as silicon or germanium the two atoms are identical and therefore $\widehat{V}_{s(1)}=\widehat{V}_{s(2)}$. The coefficients (A.12) become

$$\widehat{U}(\vec{G})=\cos(\vec{\tau}\vec{G})\frac{2}{|\Omega_c|}\widehat{V}_{s(1)}(|\vec{G}|) \tag{A.16}$$

which are nonzero for $\vec{G}^2\in\{0,3,8,11,\ldots\}(2\pi/a)^2$. The $\vec{G}=0$ term in Eq. (A.10) is omitted as it merely gives a constant shift to the potential $U(\vec{r})$. Chelikowski et al. [126] provided the following potential for silicon

$$\widehat{V}_{Si}^{Ch}(q)=\begin{cases}v_{Si}a_1^{Ch}/2, & (qa_{Si}/(2\pi))^2=3\\ v_{Si}a_2^{Ch}/2, & (qa_{Si}/(2\pi))^2=8\\ v_{Si}a_3^{Ch}/2, & (qa_{Si}/(2\pi))^2=11\end{cases} \tag{A.17}$$

where $a_1^{Ch}=-0.2241$Ry, $a_2^{Ch}=0.0551$Ry,$a_3^{Ch}=0.0724$Ry, and $v_{Si}=a_{Si}^3/4$ is the volume of the primitive unit cell of bulk silicon with $a_{Si}=0.543$nm. Note that the potential (A.17) is strictly bound to the choice $\vec{b}_1=-\vec{\tau}$ and $\vec{b}_2=\vec{\tau}$ for the basis vectors. Unfortunately, the definition of $\widehat{U}(\vec{G})$ on discrete points of the reciprocal lattice is insufficient for the proper description of nanostructures such as dots, wires, or wells. Wang et al.[16] presented a continuous momentum space form for silicon

$$\widehat{V}_{Si}^{Wa}(q)=v_{Si}\frac{a_1^{Wa}(q^2-a_2^{Wa})}{a_3^{Wa}e^{a_4^{Wa}q^2}-1}, \tag{A.18}$$

where $a_1^{Wa}=0.2685a_B^2$Ry, $a_2^{Wa}=2.19a_B^{-2}$, $a_3^{Wa}=2.06$, and $a_4^{Wa}=0.487a_B^2$ are determined by a fit to experiments. This atomic pseudopotential is not restricted to a specific choice for the basis $\vec{b}_1$ and $\vec{b}_2$. A plot of the form factors $\widehat{V}_{Si}^{Wa}$ and $\widehat{V}_{Si}^{Ch}$ is given in Fig. A.1. The discrete points $\widehat{V}_{Si}^{Ch}$ are very close to the continuous form $\widehat{V}_{Si}^{Wa}$ as they partially contributed to the derivation

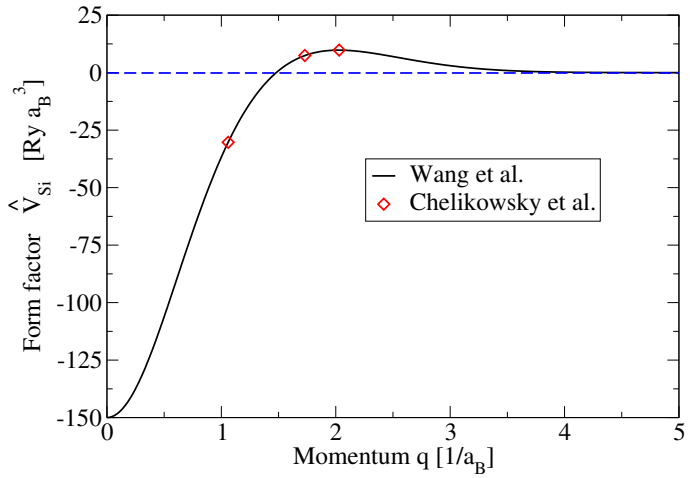

Figure A.1: Form factors $\widehat{V}_{Si}^{Wa}$ and $\widehat{V}_{Si}^{Ch}$ from Eqs. (A.18) and (A.17) respectively. The diamonds denote the values of $\widehat{V}_{Si}^{Ch}$ for $q \in \{\sqrt{3}, \sqrt{8}, \sqrt{11}\}(2\pi/a_{Si})$ and the dashed line denotes $\widehat{V}_{Si} = 0$.

of $\widehat{V}_{Si}^{Wa}$ [37]. Furthermore, as the form factor becomes very weak beyond the third $\vec{G}$ vector in Eq. (A.17) one can assume that the three $\widehat{V}_{Si}^{Ch}$ are sufficient to reproduce the bulk silicon band structure. A plot of the band structure is given in Fig. 3.1. Constant energy surfaces for the conduction and three highest valence bands, i.e. two heavy and one light hole band, are plotted in Fig. A.2. While the conduction band minima can be well approximated by ellipsoids, the situation for the valence band maxima is more complicated. In this case the warped energy surfaces [10] can be approximated by spheres obtained by means of an appropriate averaging procedure [15].

The previous considerations on the empirical pseudopotential method involved the use of purely local potentials. The extension to nonlocal potentials is straightforward [126, 127] and can lead to significant improvements. An application to strained silicon germanium alloys is given in Ref. [128]. However, nonlocality introduces some further complications and is not considered in this work. Further continuous momentum representations of empirical pseudopotentials for silicon and germanium can be found in Ref. [129].

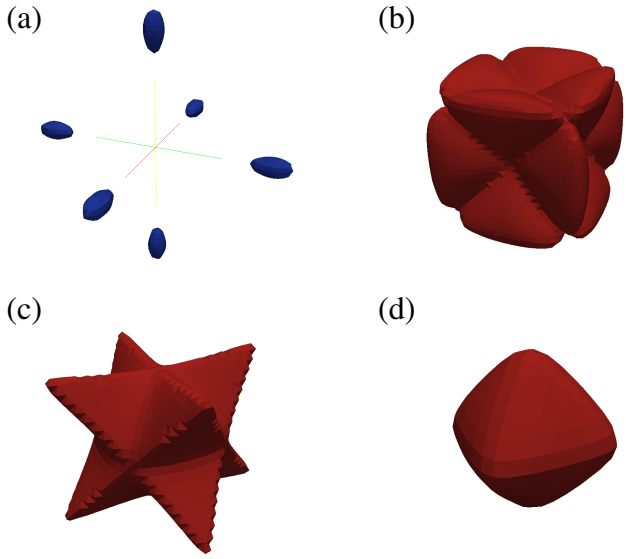

Figure A.2: Constant energy surfaces for the conduction (a), heavy hole (b & c), and the light hole band (d) of bulk silicon computed by means of the parametrization (A.17). The surfaces correspond to energies 0.1eV above or below the conduction band maximum or valence band minimum, respectively.

Appendix B

Density of States

Several methods for the numerical computation of the density of states $\mathcal{G}^{ideal}(E)$ from Eq. (3.84) are available in the literature including advanced schemes [47] which employ an adaptive sampling of the reciprocal lattice. In the following, a simple procedure is described to obtain $\mathcal{G}^{ideal}(E)$. The integration domain $\widetilde{\Omega}_c$ in (3.84) is first transformed to the unit cube $C = [0,1]^3$ in the reciprocal space

$$\mathcal{G}^{ideal}(E) = \frac{|\det(\mathbf{B})|}{4\pi^3} \sum_n \int_C \delta(E - \widetilde{\epsilon}_n(\vec{k}'))d\vec{k}' \tag{B.1}$$

$$= \frac{|\det(\mathbf{B})|}{4\pi^3} \sum_n \int_{S'_n(E)} \frac{1}{|\vec{\nabla}_{\vec{k}'}\widetilde{\epsilon}_n(\vec{k}')|} dS' \tag{B.2}$$

with $\mathbf{B}$ from Sec. 3.2, $\widetilde{\epsilon}_n(\vec{k}') = \widetilde{\epsilon}_n(\mathbf{B}^{-1}\vec{k}) = \epsilon_n(\vec{k})$, and $S'_n(E)$ is defined by the condition $\widetilde{\epsilon}_n(\vec{k}') = E$. The surface integral (B.2) is now computed numerically by subdividing the unit cube in M^3 uniformly sized cubic cells as shown in Fig. B.1.a. The energy $\widetilde{\epsilon}_n(\vec{k}')$ is sampled at the corners of these cells as illustrated in Fig. B.1.b. For a given band n the gradient $\vec{\nabla}_{\vec{k}'}\widetilde{\epsilon}_n(\vec{k}')$ is approximated on the i^{th} cell C_i by

$$\nabla_{\vec{k}'}\widetilde{\epsilon}_n(\vec{k}')|_{C_i} \simeq \frac{1}{\Delta k} \begin{pmatrix} \epsilon_1^{(i)} - \epsilon_0^{(i)} \\ \epsilon_2^{(i)} - \epsilon_0^{(i)} \\ \epsilon_3^{(i)} - \epsilon_0^{(i)} \end{pmatrix} \equiv \vec{v}_n^{(i)}, \tag{B.3}$$

where $\Delta k = 1/M$. In a next step the intersection between the i^{th} cell and the isosurface $S'_n(E)$ has to determined. For this purpose, the dispersion $\widetilde{\epsilon}_n(\vec{k}')$ is linearly interpolated on the edges of the cell by means of the samples on the corners, i.e. $\epsilon_0^{(i)} + k(\epsilon_1^{(i)} - \epsilon_0^{(i)})/\Delta k$ for instance. If for this particular edge the condition $E \in [\epsilon_0^{(i)}, \epsilon_1^{(i)}]$ holds true, the point $k = (E - \epsilon_0^{(i)})/(\epsilon_1^{(i)} - \epsilon_0^{(i)})$ is marked and applying the same procedure to the remaining eleven edges yields a polygon which approximates the desired intersection. The surface $s_n^{(i)}$ of this polygon is obtained by a suitable triangulation. Finally, the approximation to the integral (B.2) is given by the non-vanishing contributions from each cell and band

$$\mathcal{G}^{ideal}(E) \simeq \frac{|\det(\mathbf{B})|}{4\pi^3} \sum_n \sum_i \frac{s_n^{(i)}}{|\vec{v}_n^{(i)}|}. \tag{B.4}$$

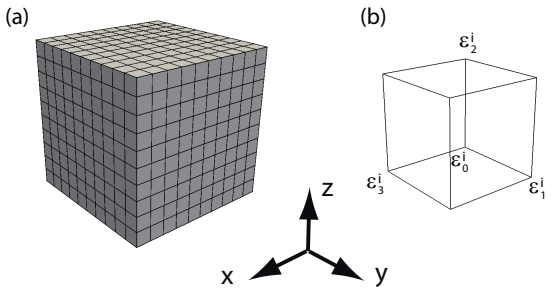

Figure B.1: (a) Uniform subdivision of the unit cube in the reciprocal space. The dispersion $\widetilde{\epsilon}_n(\vec{k}') = \epsilon_n(\vec{k})$ is sampled on the edges of the cells and labeled accordingly as shown in (b).

Appendix C

Finite Element Method

The basis for the present FEM is given by piecewise linear functions, i.e. hat functions, located at the nodes of a simplex tessellation of the simulation domain Ω as illustrated in Fig. C.1 for the two-dimensional case. A hat function $b_i(\underline{x})$ is characterized by the node i on which the function is centered. Given a two-dimensional tessellation of the form $\Omega = \cup_K T_K$ consisting of triangles T_K, each hat function is described by means of the unit triangle $D = \{(x, y)|0 \leq x, 0 \leq y, x + y \leq 1\}$. The functions $f_0(\underline{\xi}) = 1 - \xi_0 - \xi_1$, $f_1(\underline{\xi}) = \xi_0$, and $f_2(\underline{\xi}) = \xi_1$ on D are mapped to T_K by means of the transformation $\underline{x} = \Theta(\underline{\xi}) \equiv \underline{P}_0^K + \mathbf{B}_K\underline{\xi}$, where $\mathbf{B}_K = (\underline{P}_1^K - \underline{P}_0^K|\underline{P}_2^K - \underline{P}_0^K)$ and $\underline{P}_0^K$,$\underline{P}_1^K$, and $\underline{P}_2^K$ are the corners of the triangle T_K. The three-dimensional case follows analogously. The assembly of the matrix $\mathbf{A}$ from Eq. (4.34) is accomplished by summing up the contributions from each T_K. Quantities such as the charge density or the potential energy for the Schrödinger equation are expressed within the hat function basis as well. For the calculation of the right-hand side in the Schrödinger or Poisson equation, the overlap matrix

$$\mathbf{O}_{i,j} = \int_\Omega b_i(\underline{x})b_j(\underline{x})d\underline{x} \tag{C.1}$$

is employed.

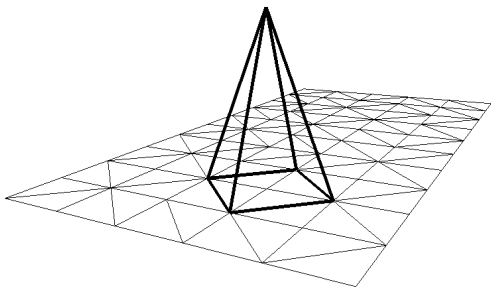

Figure C.1: A hat function b_i is equal to unity at a grid point i and decays linearly down to zero towards the edges of the neighboring simplexes.

The evaluation of the residual (4.42) is commonly accomplished by a variant of the L^2 norm, i.e.

$$||R||_r \equiv \sqrt{\frac{1}{|\Omega|}\int_\Omega |R(\underline{x})|^2 d\underline{x}} = \sqrt{\frac{1}{|\Omega|}\sum_{i,j} R_j R_k \int_\Omega b_j(\underline{x}) b_k(\underline{x}) d\underline{x}} = \sqrt{\frac{\underline{\underline{\mathbf{R}}}\mathbf{O}\underline{\underline{\mathbf{R}}}^T}{|\Omega|}} \tag{C.2}$$

with

$$R(\underline{x}) = \underline{\nabla}\left(\epsilon \underline{\nabla}^T \phi_{in}\right) + \rho = \sum_i R_i b_i(\underline{x}) \tag{C.3}$$

and $\underline{\mathbf{R}} = (\ldots, R_i, \ldots)^T$. However, the isolation of the vector $\underline{\mathbf{R}}$ needed for the evaluation of $\sqrt{\underline{\mathbf{R}}\mathbf{O}\underline{\mathbf{R}}^T/|\Omega|}$ requires the solution of a further linear problem. This difficulty can be circumvented by using the norm $\sqrt{(\mathbf{O}\underline{\mathbf{R}}^T)(\mathbf{O}\underline{\mathbf{R}}^T)}/|\Omega|$ with the following form in position space

$$||R||_r = \frac{1}{|\Omega|}\sqrt{\sum_i \left[\int_\Omega b_i(\underline{x}) R(\underline{x}) d\underline{x}\right]^2}. \tag{C.4}$$

Appendix D

SMA for Arbitrary Directions

For silicon nanowires grown along the $\langle 100 \rangle$ direction, the envelope function (3.41) consists of a plane wave times an x-independent function. This form is a key ingredient for the quantum transport framework described in chapter 4. Unfortunately, when the growth direction of the silicon nanowire deviates from $\langle 100 \rangle$ or other materials are employed such as germanium, the symmetric effective mass tensor becomes generally full

$$\mathbf{M} = \begin{pmatrix} w_{xx} & w_{xy} & w_{xz} \\ w_{xy} & w_{yy} & w_{yz} \\ w_{xz} & w_{yz} & w_{zz} \end{pmatrix}. \tag{D.1}$$

Using an ansatz of the form (3.41) for the effective mass equation (3.40) yields

$$\begin{aligned} -\frac{\hbar^2}{2m_e} \Bigg(& -\psi(y,z)k_x^2 w_{xx} + w_{yy}\frac{\partial^2}{\partial y^2}\psi(y,z) + w_{zz}\frac{\partial^2}{\partial z^2}\psi(y,z) \\ & +2ik_x w_{xy}\frac{\partial}{\partial y}\psi(y,z) + 2ik_x w_{xz}\frac{\partial}{\partial z}\psi(y,z) \\ & +2w_{yz}\frac{\partial}{\partial y}\frac{\partial}{\partial z}\psi(y,z)\Bigg) + (U^{eff} - E)\psi(y,z) = 0. \end{aligned} \tag{D.2}$$

The solution of the transverse wave function $\psi(y,z)$ depends on k_x and notably complicates the problem. A remedy is provided by the ansatz [130]

$$\psi(y,z) = e^{ik_x(\alpha y + \beta z)}\sigma(y,z) \tag{D.3}$$

with

$$\alpha = \frac{-w_{xy}w_{zz} + w_{yz}w_{xz}}{w_{yy}w_{zz} - w_{yz}^2} \tag{D.4}$$

$$\beta = \frac{-w_{yy}w_{xz} + w_{yz}w_{xz}}{w_{yy}w_{zz} - w_{yz}^2} \tag{D.5}$$

which separates the Schrödinger problem (D.2) in a longitudinal and a k_x-independent transverse part

$$\left[-\frac{\hbar^2}{2m_e}\nabla_\perp \begin{pmatrix} w_{yy} & w_{yz} \\ w_{yz} & w_{zz} \end{pmatrix} \nabla_\perp^T + U^{eff}\right]\sigma = \epsilon\sigma, \tag{D.6}$$

where $\underline{\nabla} = (\partial/\partial y, \partial/\partial z)$, $\epsilon = E - \hbar^2 k_x^2/(2m_e\widetilde{m})$ and

$$\widetilde{m} = \frac{w_{yy}w_{zz} - w_{yz}^2}{\det(\mathbf{M})}. \tag{D.7}$$

The determinant of $\mathbf{M}$ can be written as a product of the corresponding eigenvalues $w_l = 1/m_l$ and $w_t = 1/m_t$, i.e. $\det(\mathbf{M}) = w_l w_t w_t$[1], which are available in the literature [14].
In the case of germanium for instance, the conduction band has four minima located at the boundary of the first Brillouin zone along the [111], [1$\bar{1}$1], [11$\bar{1}$], and [1$\bar{1}\bar{1}$] directions corresponding to four different effective mass tensors

$$\begin{aligned}
\mathbf{M}_1 &= \frac{1}{3m_l m_t}\begin{pmatrix} 2m_l+m_t & m_t-m_l & m_t-m_l \\ m_t-m_l & 2m_l+m_t & m_t-m_l \\ m_t-m_l & m_t-m_l & 2m_l+m_t \end{pmatrix} \\
\mathbf{M}_2 &= \frac{1}{3m_l m_t}\begin{pmatrix} 2m_l+m_t & m_l-m_t & m_t-m_l \\ m_l-m_t & 2m_l+m_t & m_l-m_t \\ m_t-m_l & m_l-m_t & 2m_l+m_t \end{pmatrix} \\
\mathbf{M}_3 &= \frac{1}{3m_l m_t}\begin{pmatrix} 2m_l+m_t & m_t-m_l & m_l-m_t \\ m_t-m_l & 2m_l+m_t & m_l-m_t \\ m_l-m_t & m_l-m_t & 2m_l+m_t \end{pmatrix} \\
\mathbf{M}_4 &= \frac{1}{3m_l m_t}\begin{pmatrix} 2m_l+m_t & m_l-m_t & m_l-m_t \\ m_l-m_t & 2m_l+m_t & m_t-m_l \\ m_l-m_t & m_t-m_l & 2m_l+m_t \end{pmatrix}.
\end{aligned} \tag{D.8}$$

The transverse Schrödinger problems (D.6) and conduction masses $\widetilde{m}$ are identical for the pairs $\{\mathbf{M}_1, \mathbf{M}_4\}$ and $\{\mathbf{M}_2, \mathbf{M}_3\}$. In this sense, two of the four valleys are two-fold degenerate.
A similar discussion for the case of arbitrarily oriented quantum wells is given in Ref. [131].
The form (4.11) of the envelope function required within the SMA is generalized for the case of arbitrary effective mass tensors to

$$F^{(n)}(\vec{r}) \equiv \sum_i \left(a_i^{(n)} \psi_i^{+,(n)}(y,z) e^{ik_i^{(n)}x} + b_i^{(n)} \psi_i^{-,(n)}(y,z) e^{-ik_i^{(n)}x} \right) \tag{D.9}$$

by means of the presently described framework. The main difference is that the transverse wave functions ψ_i depend on the total energy. However, this modification implies only a minor increase of the computational burden and does not affect the scalability of the algorithm. The continuity conditions for the current lead to connection rules between the coefficient of adiacent slices which can be expressed in terms of transfer matrices, i.e. Eq. (4.17), and thus allow the recovery of the SMA.

[1] The presence of a two-fold degenerate eigenvalue implies that the ellipsoid described by $\mathbf{M}$ is rotationally invariant.

Bibliography

[1] *International Technology Roadmap for Semiconductors 2005*. http://public.itrs.net.

[2] B. Doris, M. Ieong, T. Kanarsky, Y. Zhang, R.A. Roy, O. Dokumaci, Z. Ren, F.F. Jamin, L. Shi, W. Natzle, et al. Extreme scaling with ultra-thin Si channel MOSFETs. In *Electron Devices Meeting, 2002. IEDM'02. Digest. International*, pages 267–270, 2002.

[3] J. Appenzeller, J. Knoch, M.T. Bjork, H. Riel, H. Schmid, and W. Riess. Toward nanowire electronics. *IEEE Transactions on Electron Devices*, 55(11):2827–2845, 2008.

[4] Y. Cui, L.J. Lauhon, M.S. Gudiksen, J. Wang, and C.M. Lieber. Diameter-controlled synthesis of single-crystal silicon nanowires. *Applied Physics Letters*, 78:2214, 2001.

[5] A.I. Persson, M.W. Larsson, S. Steinström, B.J. Ohlson, L. Samuelson, and L.R. Wallenberg. Solid-phase diffusion mechanism for GaAs nanowire growth. *Nature Materials*, 3:667–681, 2004.

[6] A.B. Greytak, L.J. Lauhon, M.S. Gudiksen, and C.M. Lieber. Growth and transport properties of complementary germanium nanowire field-effect transistors. *Applied Physics Letters*, 84:4176, 2004.

[7] T. Saito, T. Saraya, T. Inukai, H. Majima, T. Nagumo, and T. Hiramoto. Suppression of Short Channel Effect in Triangular Parallel Wire Channel MOSFETs. *IEICE Transactions on Electronics*, 85(5):1073–1078, 2002.

[8] H. Majima, Y. Saito, and T. Hiramoto. Impact of quantum mechanical effects on design of nano-scale narrowchannel n-and p-type MOSFETs. In *Electron Devices Meeting, 2001. IEDM Technical Digest. International*, pages 33–3, 2001.

[9] Y. Cui, Z. Zhong, D. Wang, W.U. Wang, and C.M. Lieber. High performance silicon nanowire field effect transistors. *Nano Letters*, 3(2):149–152, 2003.

[10] R. Enderlein and N.J.M. Horing. *Fundamentals of semiconductor physics and devices*. World Scientific Pub Co Inc, 1997.

[11] G. Baym. *Lectures on quantum mechanics*. Westview Press, 1974.

[12] D. Frenkel and B. Smit. *Understanding molecular simulation: from algorithms to applications*. Academic Pr, 2002.

[13] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos. Iterative minimization techniques for ab initio total-energy calculations: molecular dynamics and conjugate gradients. *Rev. Mod. Phys.*, 64(4):1045–1097, Oct 1992.

[14] N.W. Ashcroft and N.D. Mermin. *Solid State Physics. 1976*. Brooks Cole.

[15] P.Y. Yu and M. Cardona. *Fundamentals of semiconductors: physics and materials properties*. Springer Verlag, 2005.

[16] L.W. Wang and A. Zunger. Electronic Structure Pseudopotential Calculations of Large (apprx. 1000 Atoms) Si Quantum Dots. *The Journal of Physical Chemistry*, 98(8):2158–2165, 1994.

[17] M. Luisier. *Quantum Transport Beyond the Effective Mass*. PhD thesis, SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZURICH, 2007.

[18] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, et al. *LAPACK Users' guide*. Society for Industrial Mathematics, 1999.

[19] C. Vömel, S.Z. Tomov, O.A. Marques, A. Canning, L.W. Wang, and J.J. Dongarra. State-of-the-art eigensolvers for electronic structure calculations of large scale nano-systems. *Journal of Computational Physics*, 2008.

[20] L.S. Blackford, A. Cleary, J. Choi, E. d'Azevedo, J. Demmel, I. Dhillon, J. Dongarra, S. Hammarling, G. Henry, A. Petitet, et al. *ScaLAPACK user's guide*. Society for Industrial Mathematics, 1997.

[21] J.M. Luttinger and W. Kohn. Motion of electrons and holes in perturbed periodic fields. *Physical Review*, 97(4):869–883, 1955.

[22] D. J. BenDaniel and C. B. Duke. Space-charge effects on electron tunneling. *Phys. Rev.*, 152(2):683–692, Dec 1966.

[23] I. V. Tokatly, A. G. Tsibizov, and A. A. Gorbatsevich. Interface electronic states and boundary conditions for envelope functions. *Phys. Rev. B*, 65(16):165328, Apr 2002.

[24] D. Rideau, M. Feraille, M. Michaillat, Y.M. Niquet, C. Tavernier, and H. Jaouen. On the validity of the effective mass approximation and the Luttinger kp model in fully depleted SOI MOSFETs. *Solid State Electronics*, 53(4):452–461, 2009.

[25] J.L.P.J. van der Steen, D. Esseni, P. Palestri, L. Selmi, and R.J.E. Hueting. Validity of the parabolic effective mass approximation in silicon and germanium n-MOSFETs with different crystal orientations. *IEEE Transactions on Electron Devices*, 54(8):1843–1851, 2007.

[26] F. Chirico, A. Di Carlo, and P. Lugli. Efficient self-consistent pseudopotential calculation of nanostructured devices. *Physical Review B*, 64(4):45314, 2001.

[27] L.C.L.Y. Voon and M. Willatzen. *The KP Method: Electronic Properties of Semiconductors*. Springer Verlag, 2009.

[28] H.G. Katzgraber, A. Esposito, and M. Troyer. Ramping fermions in optical lattices across a feshbach resonance. *Phys. Rev. A*, 74(4):043602, Oct 2006.

[29] A. Esposito, M. Frey, and A. Schenk. Limitations of the Effective Mass Approximation: A Specific Example. Swiss Numerics Colloquium, 2008.

[30] Y. M. Niquet, C. Delerue, G. Allan, and M. Lannoo. Method for tight-binding parametrization: Application to silicon nanostructures. *Phys. Rev. B*, 62(8):5109–5116, Aug 2000.

[31] J. Wang, A. Rahman, A. Ghosh, G. Klimeck, and M. Lundstrom. On the validity of the parabolic effective-mass approximation for the current-voltage calculation of silicon nanowire transistors. *Arxiv preprint cond-mat/0412278*, 2004.

[32] Y.M. Niquet, A. Lherbier, NH Quang, MV Fernandez-Serra, X. Blase, and C. Delerue. Electronic structure of semiconductor nanowires. *Physical Review B*, 73(16):165319, 2006.

[33] A.G.M. Neves. Eigenmodes and eigenfrequencies of vibrating elliptic membranes: a Klein oscillation theorem and numerical calculations. *Communications on Pure and Applied Analysis (CPAA)*, 9(3):611–624, 2010.

[34] G.D. Sanders and Y.C. Chang. Theory of optical properties of quantum wires in porous silicon. *Physical Review B*, 45(16):9202–9213, 1992.

[35] Chin-Yu Yeh, S. B. Zhang, and A. Zunger. Confinement, surface, and chemisorption effects on the optical properties of si quantum wires. *Phys. Rev. B*, 50(19):14405–14415, Nov 1994.

[36] D. Esseni and P. Palestri. Linear combination of bulk bands method for investigating the low-dimensional electron gas in nanostructured devices. *Physical Review B*, 72(16):165342, 2005.

[37] S. B. Zhang, Chin-Yu Yeh, and A. Zunger. Electronic structure of semiconductor quantum films. *Phys. Rev. B*, 48(15):11204–11219, Oct 1993.

[38] M.G. Pala and G. Iannaccone. A 3D solver of the Schrödinger equation in momentum space for the detailed simulation of nanostructures. *Nanotechnology*, 13:369–372, 2002.

[39] A. Trellakis and U. Ravaioli. Three-dimensional spectral solution of the Schrödinger equation for arbitrary band structures. *Journal of Applied Physics*, 92:3711, 2002.

[40] C. Jacoboni and L. Reggiani. The Monte Carlo method for the solution of charge transport in semiconductors with applications to covalent materials. *Reviews of Modern Physics*, 55(3):645–705, 1983.

[41] J.A. López-Villanueva, I. Melchor, P. Cartujo, and J.E. Carceller. Modified Schroödinger equation including nonparabolicity for the study of a two-dimensional electron gas. *Physical Review B*, 48(3):1626–1631, 1993.

[42] A. Esposito, M. Luisier, M. Frey, and A. Schenk. A nonparabolicity model compared to tight-binding: The case of square silicon quantum wires. *Solid State Electronics*, 53(3):376–382, 2009.

[43] C. Jungemann, A. Emunds, and W.L. Engl. Simulation of linear and nonlinear electron transport in homogeneous silicon inversion layers. *Solid-state electronics*, 36(11):1529–1540, 1993.

[44] A. Godoy, Z. Yang, U. Ravaioli, and F. Gámiz. Effects of nonparabolic bands in quantum wires. *Journal of Applied Physics*, 98:013702, 2005.

[45] S. Rodríguez-Bolívar, F.M. Gómez-Campos, and J.E. Carceller. An atomistic-based correction of the effective-mass approach for investigating quantum dots. *Journal of Applied Physics*, 104:104309, 2008.

[46] K. Nehari, M. Lannoo, F. Michelini, N. Cavassilas, M. Bescond, and J.L. Autran. Improved effective mass theory for silicon nanostructures. *Applied Physics Letters*, 93:092103, 2008.

[47] C. Jungemann and B. Meinerzhagen. *Hierarchical device simulation: the Monte-Carlo perspective*. Springer Verlag, 2003.

[48] I.S. Gradshteyn and I.M. Ryzhik. Table of Integrals, Series, and Products. 2000.

[49] M. Goano. Algorithm 745: computation of the complete and incomplete Fermi-Dirac integral. *ACM Transactions on Mathematical Software (TOMS)*, 21(3):221–232, 1995.

[50] H.M. Antia. Rational function approximations for Fermi-Dirac integrals. *The Astrophysical Journal Supplement Series*, 84:101, 1993.

[51] A. Trellakis, A.J. Galick, and U. Ravaioli. Rational Chebyshev approximation for the Fermi-Dirac integral. *Solid-State Electronics*, 41:771–773, 1997.

[52] M. Galassi, J. Davies, J. Theiler, B. Gough, G. Jungman, M. Booth, and F. Rossi. *GNU scientific library*. Citeseer, 2002.

[53] S. Datta. *Electronic transport in mesoscopic systems*. Cambridge Univ Pr, 1997.

[54] F.O. Heinz. *Simulation Approaches for Nano–Scale Semiconductor Devices*. PhD thesis, SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZURICH, 2004.

[55] A. M. Kriman, N. C. Kluksdahl, and D. K. Ferry. Scattering states and distribution functions for microstructures. *Phys. Rev. B*, 36(11):5953–5959, Oct 1987.

[56] A. Scholze. *Simulation of single-electron devices*. PhD thesis, SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZURICH, 2000.

[57] D.Y.K. Ko and J.C. Inkson. Matrix method for tunneling in heterostructures: Resonant tunneling in multilayer systems. *Physical Review B*, 38(14):9945–9951, 1988.

[58] A.C. Marsh and J.C. Inkson. Scattering matrix theory of transport in heterostructures. *Semiconductor Science and Technology*, 1:285–290, 1986.

[59] J. Wang, E. Polizzi, and M. Lundstrom. A three-dimensional quantum simulation of silicon nanowire transistors with the effective-mass approximation. *Journal of Applied Physics*, 96:2192, 2004.

[60] M. Luisier, A. Schenk, and W. Fichtner. Quantum transport in two-and three-dimensional nanoscale transistors: Coupled mode effects in the nonequilibrium Greens function formalism. *Journal of Applied Physics*, 100:043713, 2006.

[61] S.K. Röllin. *Parallel iterative solvers in computational electronics*. PhD thesis, SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZURICH, 2005.

[62] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. Van Der Vorst. *Templates for the solution of algebraic eigenvalue problems*. Society for Industrial Mathematics, 2000.

[63] R. Barrett, M. Berry, T. F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. Van der Vorst. *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods, 2nd Edition*. SIAM, Philadelphia, PA, 1994.

[64] S. Balay, K. Buschelman, W. D. Gropp, D. Kaushik, M. G. Knepley, L. C. McInnes, B. F. Smith, and H. Zhang. PETSc Web page, 2009. http://www.mcs.anl.gov/petsc.

[65] R.B. Lehoucq, D.C. Sorensen, and C. Yang. *ARPACK users' guide: solution of large-scale eigenvalue problems with implicitly restarted Arnoldi methods*. Siam, 1998.

[66] G. Sleijpen and H. van der Vorst. A Jacobi-Davidson iteration method for linear eigenvalue problem. *SIAM J. Matrix Anal. Appl.*, 17:401–425, 1996.

[67] B. Schmithüsen. *Grid Adaption for the Stationary Two-Dimensional Drift-Diffusion Model in Semiconductor Device Modeling*. PhD thesis, SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZURICH, 2002.

[68] W. Schönauer and T. Adolph. How we solve PDEs. *Journal of computational and applied mathematics*, 131(1-2):473–492, 2001.

[69] W. Schönauer and T. Adolph. FDEM: how we make the FDM more flexible than the FEM. *Journal of Computational and Applied Mathematics*, 158(1):157–167, 2003.

[70] T. Adolph and W. Sch
"onauer. The application of the FDEM program package with error estimate to industrial problems. *PAMM*, 7(1), 2007.

[71] H.J. Bungartz and M. Griebel. Sparse grids. *Acta Numerica*, 13:147–269, 2004.

[72] S. Kuwajima and A. Warshel. The extended Ewald method: A general treatment of long-range electrostatic interactions in microscopic simulations. *The Journal of Chemical Physics*, 89:3751, 1988.

[73] P. Pulay. Convergence acceleration of iterative sequences. The case of SCF iteration. *Chemical Physics Letters*, 73:393–398, 1980.

[74] S.E. Laux. Numerical methods for calculating self-consistent solutions of electron states in narrow channels. In *Numerical Analysis of Semiconductor Devices and Integrated Circuits, 1987. NASCODE V. Proceedings of the Fifth International Conference on the*, pages 270–275, 1987.

[75] A. Trellakis, A.T. Galick, A. Pacelli, and U. Ravaioli. Iteration scheme for the solution of the two-dimensional Schrödinger-Poisson equations in quantum structures. *Journal of Applied Physics*, 81:7880, 1997.

[76] W.P. Petersen and P. Arbenz. *Introduction to parallel computing*. Oxford University Press, USA, 2004.

[77] A. Esposito, M. Frey, and A. Schenk. Quantum transport including nonparabolicity and phonon scattering: application to silicon nanowires. *Journal of Computational Electronics*, 8(3):336–348, 2009.

[78] M. Shin. Three-dimensional quantum simulation of multigate nanowire field effect transistors. *Mathematics and Computers in Simulation*, 79(4):1060–1070, 2008.

[79] E. Gnani, S. Reggiani, A. Gnudi, P. Parruccini, R. Colle, M. Rudan, and G. Baccarani. Band-structure effects in ultrascaled silicon nanowires. *IEEE Transactions on Electron Devices*, 54(9):2243–2254, 2007.

[80] J. Knoch and H. Lüth. Semiconductor Nanostructures and Devices. *Nanoscaled semiconductor-on-insulator structures and devices*, pages 143–158.

[81] R. Venugopal, Z. Ren, S. Datta, M.S. Lundstrom, and D. Jovanovic. Simulating quantum transport in nanoscale transistors: Real versus mode-space approaches. *Journal of Applied Physics*, 92:3730, 2002.

[82] Z. Ren, R. Venugopal, S. Goasguen, S. Datta, and M.S. Lundstrom. nanoMOS 2.5: A two-dimensional simulator for quantum transport in double-gate MOSFETs. *IEEE Transactions on Electron Devices*, 50(9):1914–1925, 2003.

[83] M. Bescond, K. Nehari, J.L. Autran, N. Cavassilas, D. Munteanu, and M. Lannoo. 3D quantum modeling and simulation of multiple-gate nanowire MOSFETs. In *IEEE International Electron Devices Meeting, 2004. IEDM Technical Digest*, pages 617–620, 2004.

[84] M. Bescond, N. Cavassilas, K. Kalna, K. Nehari, L. Raymond, J.L. Autran, M. Lannoo, and A. Asenov. Ballistic transport in Si, Ge, and GaAs nanowire MOSFETs. *IEDM Tech. Dig*, 533, 2005.

[85] A. Martinez, M. Bescond, J.R. Barker, A. Svizhenko, M.P. Anantram, C. Millar, and A. Asenov. A self-consistent full 3-D real-space NEGF simulator for studying nonperturbative effects in nano-MOSFETs. *IEEE Transactions on Electron Devices*, 54(9):2213–2222, 2007.

[86] X. Shao and Z. Yu. Nanoscale FinFET simulation: A quasi-3D quantum mechanical model using NEGF. *Solid-State Electronics*, 49(8):1435–1445, 2005.

[87] D. Mamaluy, D. Vasileska, M. Sabathil, T. Zibold, and P. Vogl. Contact block reduction method for ballistic transport and carrier densities of open nanostructures. *Physical Review B*, 71(24):245321, 2005.

[88] N. Neophytou, A. Paul, M.S. Lundstrom, and G. Klimeck. Bandstructure effects in silicon nanowire electron transport. *IEEE Transactions on Electron Devices*, 55(6):1286, 2008.

[89] K. Nehari, N. Cavassilas, J.L. Autran, M. Bescond, D. Munteanu, and M. Lannoo. Influence of band structure on electron ballistic transport in silicon nanowire MOSFET's: An atomistic study. *Solid-State Electronics*, 50(4):716–721, 2006.

[90] K. Nehari, N. Cavassilas, F. Michelini, M. Bescond, J.L. Autran, and M. Lannoo. Full-band study of current across silicon nanowire transistors. *Applied Physics Letters*, 90:132112, 2007.

[91] E. Gnani, S. Reggiani, M. Rudan, and G. Baccarani. Effects of the Band-Structure Modification in Silicon Nanowires with Small Diameters. In *Solid-State Device Research Conference, 2006. ESSDERC 2006. Proceeding of the 36th European*, pages 170–173, 2006.

[92] J. Wang, A. Rahman, G. Klimeck, and M. Lundstrom. Bandstructure and orientation effects in ballistic Si and Ge nanowire FETs. In *IEEE International Electron Devices Meeting, 2005. IEDM Technical Digest*, page 4, 2005.

[93] P. Marconcini, G. Fiori, M. Macucci, and G. Iannaccone. Hierarchical simulation of transport in silicon nanowire transistors. *Journal of Computational Electronics*, 7(3):415–418, 2008.

[94] E. Gnani, A. Gnudi, S. Reggiani, M. Luisier, and G. Baccarani. Band Effects on the Transport Characteristics of Ultrascaled SNW-FETs. *IEEE Transactions on Nanotechnology*, 7(6):700–709, 2008.

[95] J.H. Rhew, Z. Ren, and M.S. Lundstrom. A numerical study of ballistic transport in a nanoscale MOSFET. *Solid-State Electronics*, 46(11):1899–1906, 2002.

[96] Y. Liu, N. Neophytou, T. Low, G. Klimeck, and M.S. Lundstrom. A tight-binding study of the ballistic injection velocity for ultrathin-body SOI MOSFETs. *IEEE Transactions on Electron Devices*, 55(3):866, 2008.

[97] Y. Liu, N. Neophytou, G. Klimeck, and M.S. Lundstrom. Band-Structure Effects on the Performance of III-V Ultrathin-Body SOI MOSFETs. *IEEE Transactions on Electron Devices*, 55(5):1116, 2008.

[98] M. Luisier, A. Schenk, W. Fichtner, and G. Klimeck. Atomistic simulation of nanowires in the sp^3 d^5s^* tight-binding formalism: From boundary conditions to strain calculations. *Physical Review B*, 74(20):205323, 2006.

[99] F. Sacconi, M.P. Persson, M. Povolotskyi, L. Latessa, A. Pecchia, A. Gagliardi, A. Balint, T. Fraunheim, and A. Di Carlo. Electronic and transport properties of silicon nanowires. *Journal of Computational Electronics*, 6(1):329–333, 2007.

[100] J. Taylor, H. Guo, and J. Wang. Ab initio modeling of quantum transport properties of molecular electronic devices. *Physical Review B*, 63(24):245407, 2001.

[101] M. Brandbyge, J.L. Mozos, P. Ordejon, J. Taylor, and K. Stokbro. Density-functional method for nonequilibrium electron transport. *Physical Review B*, 65(16):165401, 2002.

[102] G. Cuniberti, G. Fagas, and K. Richter. Introducing Molecular Electronics: A Brief Overview. *Introducing Molecular Electronics*, pages 1–10, 2005.

[103] P. Delaney and J.C. Greer. Correlated electron transport in molecular electronics. *Physical review letters*, 93(3):36805, 2004.

[104] M. Albrecht, B. Song, and A. Schnurpfeil. A wave function based ab-initio nonequilibrium Greens function approach to charge transport. *Journal of Applied Physics*, 100:013702, 2006.

[105] A. Calzolari, N. Marzari, I. Souza, and M. Buongiorno Nardelli. Ab initio transport properties of nanostructures from maximally localized Wannier functions. *Physical Review B*, 69(3):35108, 2004.

[106] P. Mavropoulos, N. Papanikolaou, and P.H. Dederichs. A KKR Green function formalism for ballistic transport. *Arxiv preprint cond-mat/0306604*, 2003.

[107] A. Bagrets, N. Papanikolaou, and I. Mertig. Conduction eigenchannels of atomic-sized contacts: Ab initio KKR Greens function formalism. *Physical Review B*, 75(23):235448, 2007.

[108] A. Garcia-Lekue and L.W. Wang. Self-consistent non-equilibrium transport using plane waves. *Computational Materials Science*, 45(4):1016–1024, 2009.

[109] T.B. Boykin, G. Klimeck, and F. Oyafuso. Valence band effective-mass expressions in the $sp^3d^5s^*$ empirical tight-binding model applied to a Si and Ge parametrization. *Physical Review B*, 69(11):115201, 2004.

[110] S. Lee, F. Oyafuso, P. von Allmen, and G. Klimeck. Boundary conditions for the electronic structure of finite-extent embedded semiconductor nanostructures. *Physical Review B*, 69(4):45316, 2004.

[111] M. Büttiker. Role of quantum coherence in series resistors. *Physical Review B*, 33(5):3020–3026, 1986.

[112] M. Büttiker. Coherent and sequential tunneling in series barriers. *IBM Journal of Research and Development*, 32(1):75, 1988.

[113] M. Frey, A. Esposito, and A. Schenk. Boundary conditions for incoherent quantum transport. *Proceedings of 13th IWCE*, page 17, 2009.

[114] L.P. Kadanoff and G. Baym. *Quantum statistical mechanics*. Westview Press, 1994.

[115] L.V. Keldysh. Diagram technique for nonequilibrium processes. *Zh. Eksp. Teor. Fiz*, 47(4):151–165, 1964.

[116] G.D. Mahan. *Many-particle physics*. Plenum Pub Corp, 2000.

[117] H. Haug and A.P. Jauho. *Quantum kinetics in transport and optics of semiconductors*. Springer Verlag, 2008.

[118] A. Wacker. Semiconductor superlattices: A model system for nonlinear transport. *Arxiv preprint cond-mat/0107207*, 2001.

[119] R.K. Lake and R.R. Pandey. Non-Equilibrium Green Functions in Electronic Device Modeling. *arXiv:cond-mat/0607219v1*, 2005.

[120] C. Jacoboni and L. Reggiani. The monte carlo method for the solution of charge transport in semiconductors with applications to covalent materials. *Rev. Mod. Phys.*, 55(3):645–705, Jul 1983.

[121] J.C. Slater and G.F. Koster. Simplified LCAO method for the periodic potential problem. *Physical Review*, 94(6):1498–1524, 1954.

[122] D.A. Papaconstantopoulos and M.J. Mehl. The Slater-Koster tight-binding method: a computationally efficient and accurate approach. *Journal of Physics Condensed Matter*, 15(10):413–413, 2003.

[123] P. Vogl, H.P. Hjalmarson, and J.D. Dow. Semi-empirical tight-binding theory of the electronic structure of semiconductors. *J. Phys. Chem. Sol.*, 44(5):265–378, 1983.

[124] J.M. Jancu, R. Scholz, F. Beltram, and F. Bassani. Empirical *spds∗* tight-binding calculation for cubic semiconductors: General method and material parameters. *Phys. Rev. B*, 57(11):6493–6507, Mar 1998.

[125] P.O. Löwdin. On the Non-Orthogonality Problem Connected with the Use of Atomic Wave Functions in the Theory of Molecules and Crystals. *The Journal of Chemical Physics*, 18:365, 1950.

[126] J.R. Chelikowsky and M.L. Cohen. Nonlocal pseudopotential calculations for the electronic structure of eleven diamond and zinc-blende semiconductors. *Physical Review B*, 14(2):556–582, 1976.

[127] M. L. Cohen and J. R. Chelikowsky. *Electronic Structure and Optical Properties of Semiconductors*. Springer, New York, 2nd edition, 1989.

[128] M.M. Rieger and P. Vogl. Electronic-band parameters in strained $si_{1-x}ge_x$ alloys on $si_{1-y}ge_y$ substrates. *Phys. Rev. B*, 48(19):14276–14287, Nov 1993.

[129] P. Friedel, M.S. Hybertsen, and M. Schlüter. Local empirical pseudopotential approach to the optical properties of Si/Ge superlattices. *Physical Review B*, 39(11):7974–7977, 1989.

[130] M. Bescond, N. Cavassilas, and M. Lannoo. Effective-mass approach for n-type semiconductor nanowire MOSFETs arbitrarily oriented. *Nanotechnology*, 18:255201, 2007.

[131] A. Rahman, M.S. Lundstrom, and A.W. Ghosh. Generalized effective-mass approach for n-type metal-oxide-semiconductor field-effect transistors on arbitrarily oriented wafers. *Journal of Applied Physics*, 97:053702, 2005.

Curriculum Vitae

Aniello Esposito was born in Zürich, Switzerland, on August 06, 1979. He studied physics at the Swiss Federal Institute of Technology (ETH) Zürich with a focus on computational methods for condensed matter physics and received his Dipl. Phys. Degree in Spring 2005. Topic of his diploma thesis was the investigation of Fermions in optical lattices. In 2005 he joined the Integrated Systems Laboratory, ETHZ, as a research and teaching assistant in the Technology CAD group. His research interests comprise the band structure calculation and the simulation of quantum transport in nanoscale devices as well as the implementation of related algorithms on high performance compute clusters. In his spare time he enjoys riding motorcycles.