Deterministic Simulation of Junctionless Nanowire Field Effect Transistors

---

Von der Fakultät für Elektrotechnik und Informationstechnik
der Rheinisch-Westfälischen Technischen Hochschule Aachen

zur Erlangung des akademischen Grades eines

Doktors der Ingenieurwissenschaften

genehmigte Dissertation

vorgelegt von

Maziar Noei

aus Iran

Berichter

Univ.-Prof. Dr.-Ing. Christoph Jungemann
Prof. Dr.-Ing. Bernd Meinerzhagen

Tag der mündlichen Prüfung: 08.06.2020

I hereby declare that I have created this work completely on my own and used no other sources or tools than the ones listed, and that I have marked any citations accordingly.

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie Zitate kenntlich gemacht habe.

Aachen, June 2020
Maziar Noei

Abstract

As the feature lengths of the field–effect transistors (FETs) are scaled down to the deca–nanometer range, the commonly used macroscopic approaches such as drift–diffusion and hydrodynamic models lose their validity and a detailed description of the microscopic behavior of charge carriers becomes essential for device simulation. In this work, a fully self–consistent and deterministic solver for the system of Poisson, Schrödinger, and Boltzmann equations tailored to the specific case of gate–all–around junctionless nanowire FETs is developed. The simulation framework employs various numerical techniques such as the H–transformation and an even/odd decomposition of the distribution function on a staggered grid for stabilization of the Boltzmann equation (BE), and the equations are solved with the Newton–Raphson approach which demonstrates quadratic convergence within just a few iterations. Different inter– and intra–valley scattering mechanisms, suitable boundary conditions, and quantization effects are included, and the solver is shown to be robust and stable even in the deep subthreshold region.

In addition to the stationary simulations, small signal analysis is carried out under the sinusoidal steady state condition and important figures of merit such as the cut–off frequency, maximum oscillation frequency, and Rollet stability factor are obtained and discussed. Moreover, the Langevin–source approach is used for self–consistent calculation of noise, resulting in the first deterministic BE solver for noise analysis of nanowire FETs. Quantities such as the power spectral densities of terminal currents, the drain and gate excess noise factors, cross–correlation and the noise suppression factors are presented and compared for different gate lengths.

In the second part of this work, an alternative approach based on the characteristic curves and matrix exponentials is developed for the discretization of the BE, which is also applicable to the ballistic transport and does not suffer from the numerical deficiencies of H–transformation in 1D phase space. The results of the quasi–ballistic simulations are presented and compared to those of the moments equations obtained from the projection of the BE onto Hermite polynomials. It is shown that the predominantly ballistic phenomena cannot be treated with systems of moments equations and simplified boundary conditions. The failure of moments model in describing the ballistic modes of transport has important implications for the existence of Dyakonov–Shur terahertz instabilities in high mobility 1D devices.

Zusammenfassung

Mit zunehmender Verkleinerung der Strukturgröße von Feldeffekttransistoren (FETs) in den Deka-Nanometerbereich verlieren die üblicherweise verwendeten makroskopischen Ansätze wie Drift–Diffusion und hydrodynamische Modelle ihre Gültigkeit und eine detaillierte Beschreibung des mikroskopischen Verhaltens der Ladungsträger wird für die Bauelementesimulation unerlässlich. In dieser Arbeit wird ein vollständig selbstkonsistenter und deterministischer Löser für das System aus Poisson–, Schrödinger– und Boltzmann–Gleichung entwickelt, zugeschnitten auf den speziellen Fall des Gate–all–around Nanodraht FET. Im Simulationsframework werden verschiedene numerische Verfahren verwendet, wie die H–Transformation und eine Zerlegung der Verteilungsfunktion in gerade und ungerade Anteile auf einem gestaffelten Gitter, und die Gleichungen werden mit dem Newton–Raphson–Verfahren gelöst, welches eine quadratische Konvergenz bereits innerhalb einiger weniger Iterationen aufweist. Verschiedene Inter– und Intra–Valley Streumechanismen, geeignete Randbedingungen und Quantisierungseffekte werden mit einbezogen, und es wird gezeigt, dass der Löser sogar im Deep-Subthreshold Bereich robust und stabil ist.

Zusätzlich zu den stationären Simulationen wird eine Kleinsignalanalyse in einem sinusförmigen stationären Zustand durchgeführt und wichtige Größen wie die Grenzfrequenz, die maximale Oszillationsfrequenz und der Rollet–Stabilitätsfaktor erhalten und erörtert. Darüber hinaus wird der Langevin–Quellen–Ansatz für die selbstkonsistente Berechnung des Rauschverhaltens angewendet, wodurch der erste deterministische BE–Löser für die Rauschanalyse von Nanodraht–FETs resultierte. Größen wie die spektrale Leistungsdichte der Kontaktströme, die Drain und Gate Überschussrauschfaktoren, Kreuzkorrelations– und die Rauschunterdrückungsfaktoren werden vorgestellt und für verschiedene Gatelängen verglichen.

Im zweiten Teil dieser Arbeit wird ein alternativer Ansatz für die Diskretisierung der BE auf der Grundlage von charakteristischen Kurven und des Matrixexponentials entwickelt, welcher auch auf den ballistischen Transport anwendbar ist und nicht von den numerischen Defiziten der H–Transformation im 1D Phasenraum betroffen ist. Die Ergebnisse dieser quasi–ballistischen Simulationen werden vorgestellt und mit solchen aus Momentengleichungen verglichen, die durch die Projektion der BE auf Hermite–Polynome erhalten werden. Es wird gezeigt, dass die vorwiegend ballistischen Phänomene nicht mit Systemen von Momentengleichungen unter vereinfachten Randbedingungen behandelt werden können. Das Versagen des Momentenmodells bei der Beschreibung des ballistischen Transports hat wichtige Implikationen für die Existenz von Dyakonov–Shur Terahertz–Instabilitäten in 1D high-mobility Bauteilen.

Contents

Chapter 1

Introduction

1.1 Evolution of Device Simulation

The aggressive miniaturization of the CMOS transistors in the past 50 years has enabled higher performance per unit area, lower power requirements, and lower manufacturing costs [1]. Today, nanoscale field–effect transistors (FETs) with a channel length of only 10 nm are in production [1, 2], and there is a general consensus in the semiconductor industry that the traditional planar CMOS transistors will gradually become obsolete and novel multigate architectures such as FinFETs and nanowire FETs will replace them for the foreseeable future [3]. Consequently, device modeling has evolved into sophisticated tools able to optimize transistor layouts, accurately predict device characteristics, and troubleshoot the available designs. In the area of device simulation, the operation regime of ultrashort FET designs has moved towards the quasi–ballistic transport, and the transport phenomena are better expressed in terms of the microscopic distribution of carriers rather than their mobility, diffusion constant, etc. As the feature lengths are scaled down to the deca–nanometer range, the commonly used macroscopic approaches such as drift–diffusion and hydrodynamic models are losing their validity even in the linear regime [4–7] and a detailed description of the microscopic behavior of charge carriers (e.g. with incorporation of complicated scattering mechanisms, realistic boundary conditions, complex device geometries, etc.) becomes essential for device simulation.

The Boltzmann transport equation (BE), as the master equation for the carriers' distribution function, is considered to be the best classical description of electrons and provides a proper balance between the computational effort and the solution accuracy [8–10]. Although the BE incorporates several quantum mechanical concepts such as the band structure and includes probabilistic scattering rates based on the Fermi's Golden Rule, it basically describes the evolution of the trajectory of a particle in the phase space using Newtonian mechanics of

motion. An introduction to the BE can be found in [11].

The integro–differential nature of the BE makes it extremely challenging to solve the equation for general problems. In the early years, approximate analytical methods based on the Legendre polynomial expansion were used to simplify the BE and obtain analytical solutions in simple cases [12]. Iterative integration techniques were presented for low–field transport [13], while other approaches such as the matrix method [14] and cellular automata method [15] have been useful for special problems. However, stochastic approaches based on the Monte Carlo (MC) method [9,10,16–26] have been the conventional choice for over 30 years as they can include various effects at a very fundamental physical level and simulate complicated device geometries with full band structures [22,27]. The results of the MC simulations often demonstrate excellent agreement with the experimental measurements [28] and are frequently used as a benchmark for the simpler macroscopic models.

Despite their advantages, there are several major drawbacks associated with the MC methods. Firstly, they suffer from statistical noise which is inherent to the method and cannot be completely avoided. Moreover, due to the transient nature of the MC method, simulation of phenomena on very different time scales is practically impossible. One example is the investigation of SOI devices with floating bodies, where the charging time of the body is several orders of magnitude larger than the time steps required for numerical stability of the system of equations (see e.g. [29]). Hence, no successful MC simulation of the hysteresis effects in SOI transistors has been reported yet. Similarly, the small signal and noise analyses are limited to frequencies larger than the higher GHz range [30], and the CPU time necessary for simulation of rare events or subthreshold currents is prohibitive. As the number of samples is increased by the factor N, only a $\sqrt{\mathrm{N}}$–fold reduction in the statistical uncertainty is obtained. Therefore, while the MC solvers can be straightforward and accessible ways to solve complicated equations for small-to-moderate uncertainty, they become computationally intensive as a higher accuracy is required. In addition, although the MC solvers are able to include the Pauli exclusion principle [31], this does not come naturally and requires very large particle ensembles [32].

Due to the above mentioned disadvantages of the MC solvers and higher computational power of the workstation computers, in the last decade interest in deterministic BE solvers has increased significantly and various techniques for improving their numerical stability and memory requirements have been presented. In [33–35] the discrete system of equations is constructed using the expansion of phase space in spherical harmonics and box integration is employed to achieve exact current continuity at the algebraic level. Stabilization schemes such as H–transformation [33], maximum entropy dissipation scheme [36], and upwind discretization [35] make the simulation of complicated devices possible. In [37], the full band structure including the valence bands was incorporated in the

calculations, and the areas of application has been expanded to magnetotransport simulation [38, 39], electrothermal simulations and degradation [40, 41], coupled hot carrier and phonon systems [42], and avalanche breakdown of pn–junctions and devices [43, 44]. It is expected that deterministic BE solvers will play an increasingly important role in the semiconductor device simulation.

1.2 Dissertation Outline

In this work, an efficient numerical framework for simulation of gate–all–around (GAA) nanowire transistors is presented, which is based on the deterministic solution of the BE. Numerical investigation of GAA nanowires is important because as the gate lengths of the conventional FETs are further scaled down, the cross–sectional size of their channels is also shrinking to the sub–$10 \times 10\,\mathrm{nm}^2$ range for acceptable electrostatic control and it is reasonable to assume that GAA FETs with nanowire channels will replace FinFETs from approximately the 10 nm node [45].[1] In the recent years, the electron transport in nanowire transistors has been the subject of many investigations. Atomistic computations in [46] and numerical solution of the open–boundary Schrödinger equation in [47–49] provide a description of current transport under ballistic conditions and steady state operation. However, a realistic description of the scattering rates is of utmost importance in assessment of the device characteristics even for ultrashort channels. Moreover, when the carriers are confined by the small cross–section and their streaming is restricted to the axial direction, quantization of the energy states manifests and electrons occupy subbands. Although the pseudo–potential corrections can be used to modify the distribution of carriers according to the confinement, calculation of the discrete energy eigenstates in the 2D perpendicular planes is necessary. Therefore, an approach based on the self–consistent solution of the 3D Poisson equation coupled to a 2D Schrödinger equation in the transverse planes and the multi–subband BE along the transport direction is suitable since it offers a compromise between full quantum transport approaches [50–54] and a classical BE framework [22]. The developed solver in this PhD project incorporates various inter– and intra–valley scattering mechanisms, suitable boundary conditions, and quantization effects. It is shown to be robust and stable even in deep subthreshold simulations, and can be easily generalized

[1] In June 2017, IBM Research (in collaboration with GlobalFoundries and Samsung) announced that they had developed a breakthrough process to build transistors for chips at the 5 nm node. In order to achieve this feat, new architectures and improvements in the fabrication process have been developed, with the primary technique being to use silicon nanosheets with the GAA configuration effectively covering all four sides of the ultrathin active channel. The novel technology makes it possible to construct processors with 600 million transistors per mm^2, and is expected to be ready for mass–production in 10–15 years after the research prototype stage.

to provide insight in key phenomena occuring in various one–dimensional devices with confinement in two spatial dimensions. The combined system of equations is solved using the Newton–Raphson method, which apart from being quadratically converging, paves the way towards a deterministic small signal analysis covering the full frequency range. In Chap. 2, the constituent equations are recalled and their discretization in steady state simulations is explained. Sections 2.2, 2.3, and 2.4 are dedicated to the stationary Poisson, Schrödinger, and Boltzmann equation while the discretization of the small signal equations are discussed in Sect. 2.6.

In addition to the steady state and small signal analysis of the N^+NN^+ transistors, the presented numerical framework also allows for the deterministic simulation of RF noise. Despite the fact that the electronic noise sets a fundamental limit on the performance and sensitivity of RF circuits, there are very limited numerical investigations on the subject and to date no numerical analysis on the noise behavior of junctionless N^+NN^+ nanowires has been reported. The high–frequency noise performance of novel devices is usually evaluated using MC methods, equivalent circuits, or simulations based on the drift–diffusion and hydrodynamic models [55–60], whereas a detailed microscopic and physics–based description can provide better understanding of the noise sources and higher accuracy. In Sect. 2.7 the Langevin–source approach is used to obtain an equation system which can be directly solved for the fluctuations of random variables. The idea, suggested by Langevin in his theory of Brownian motion [61] and developed further by Kogan and Shul'man in [62], offers an elegant way of formulating noise and unlike the MC methods, allows the investigation of the spatial origin of the terminal current fluctuations. The Langevin-Boltzmann equation has been solved in [63–68] for other structures, and our noise analysis uses some of the ideas developed for a double–gate device in [69, 70] .

Investigation of quasi–ballistic phenomena is another interesting area. Moving towards the ballistic limit, the distribution function gets strongly asymmetric and discontinuous with respect to the wave number, as discussed in [4, 71–73]. Although the energy–based stabilization schemes of Chapter 2 are useful for evaluating the performance of conventional devices and mobilities, they fail at numerical analysis of transport in the quasi–ballistic regime where interesting phenomena such as plasma oscillations occur. Therefore, a robust BE solver capable of handling the ballistic limit is desirable. In Chapter 4, the BE is discretized directly in the phase space, with the particular focus on demonstrating a stabilization method based on the method of characteristic curves and matrix exponentials. The results of the phase space discretization are compared to the results of moments equations, which are obtained in Chapter 3 by projection of the BE onto Hermite polynomials.

Chapter 5 presents the simulation results. Nanowire transistors with N^+NN^+ doping and gate lengths of 10–100 nm are compared in terms of their DC, AC, and

noise behavior in both diffusive and ballistic regimes and the results are discussed. For shorter gate lengths ($L_G < 5\,nm$), quantum transport becomes important and other approaches such as non–equilibrium Green's functions [74,75], Wigner functions [76], or the Pauli master equation [77, 78] are preferred which are beyond the scope of this work.

The results of this dissertation have been reported in several publications. In Ref. [79] we have presented our first implementation of the Newton–Raphson solver for the stationary solution of the system of constituent equations. In Refs. [80,81] we showed that our in–house simulator is capable of small signal and noise analysis of the junctionless nanowire FETs, and in Ref. [82] we have used the developed solver for assessing the effect of downscaling on the RF and noise performance of these devices.

Chapter 2

The Simulation Framework

2.1 Overview

The methods and techniques used in this work are specifically tailored to the problem of electronic transport in gate-all-around (GAA) N^+NN^+ FETs with nanowire channels, as shown schematically in Fig. 2.1. The carriers are confined in the $x-y$ plane, and transport happens along the z–direction from source to drain. We assume that the cross-section is quite small, i.e. L_x and L_y are in the nanometer range.

Since the selected device geometry is not translational invarient along any direction, a 3D Poisson equation (PE) needs to be solved on the entire domain. This way, the details of the 3D charge density and the applied bias translate to the electric potential profile. In order to model the confinement effects, the dimensional splitting technique employed in [32, 83, 84] is used, where the 2D Schrödinger equation (SE) is evaluated in the $x-y$ planes, for each position in z–direction. This leads to a set of eigenstates (also called subbands), each with a discrete energy and a probability amplitude function. Once the energy states of the confined carriers are known, their gradient in the z–direction acts as the driving force that is exerted on the particles from source to drain. The transport is modeled separately for each subband using the Boltzmann transport equation (BE) evaluated in z–direction. Different energy subbands are coupled by the inter-subband scattering rates, which describe the hopping of carriers between these states and are proportional to the overlap integral over the initial and final subbands.

The PE, SE, and BE are tightly coupled and need to be solved together until full self-consistency is obtained. Their interdependency is schematically described in Fig. 2.2, which can be numerically implemented in a Gummel-type iteration scheme. Starting with the electric potential, the SE is used to evaluate the subband energies and wavefunctions which are used as input for the BE in

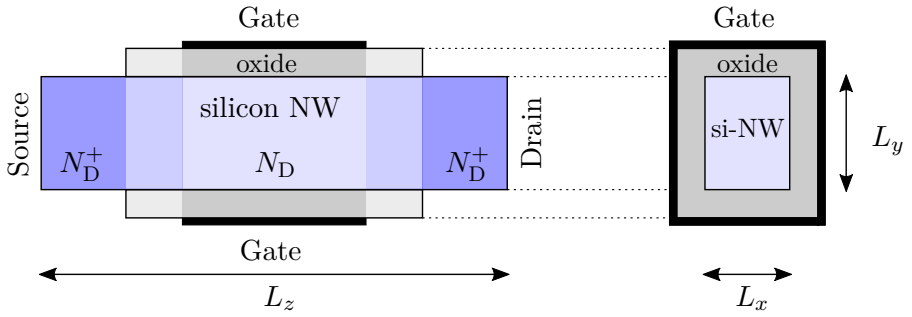

Figure 2.1: Cross-sections of the Gate-All-Around SiNW transistor.

order to compute the carrier densities. These densities are then fed into the PE and the electric potential gets updated, hence completing the dependence of the equations on each other and concluding one iteration step.

This chapter presents a brief discussion of the physical models and constituent equations which describe the transport phenomena in our simulation framework. The discretization of these equations is also explained. First, the solution of the stationary equations is considered, which are obtained if the boundary conditions are time invariant. Sections 2.2, 2.3, and 2.4 are dedicated to the stationary PE, SE, and BE, respectively. The Newton-Raphson method, which allows for simultaneous solution of the linearized system of equations in each iteration step, is presented in Sect. 2.5. Finally, Sections 2.6 and 2.7 present the numerical framework for small signal and noise analyses, respectively.

2.2 The Poisson Equation

The Poisson equation (PE), which is one of the basic equations in electrostatics, is directly derived from the eddy-current-free Maxwell's equation and the material relation $\boldsymbol{D} = \epsilon \boldsymbol{E}$, where $\boldsymbol{D}$ stands for the electric displacement field and $\boldsymbol{E}$ denotes the electric field. The quasi-stationary approximation holds up to frequencies where the wavelengths are much larger than the spatial dimensions of the device. For silicon-based transistors (with a phase velocity of about $10^8 \mathrm{m\,s^{-1}}$), this corresponds to frequencies higher than 100 THz for nanoscale devices. In this section, we will discuss how the PE is set up and discretized for our device.

2.2.1 Main Equation

The PE reads,

$$\mathrm{F_{PE}} := \nabla_{\boldsymbol{r}} \cdot (\epsilon(\boldsymbol{r}) \nabla_{\boldsymbol{r}} \varphi(\boldsymbol{r})) + \rho(\boldsymbol{r}) = 0 \,, \tag{2.1}$$

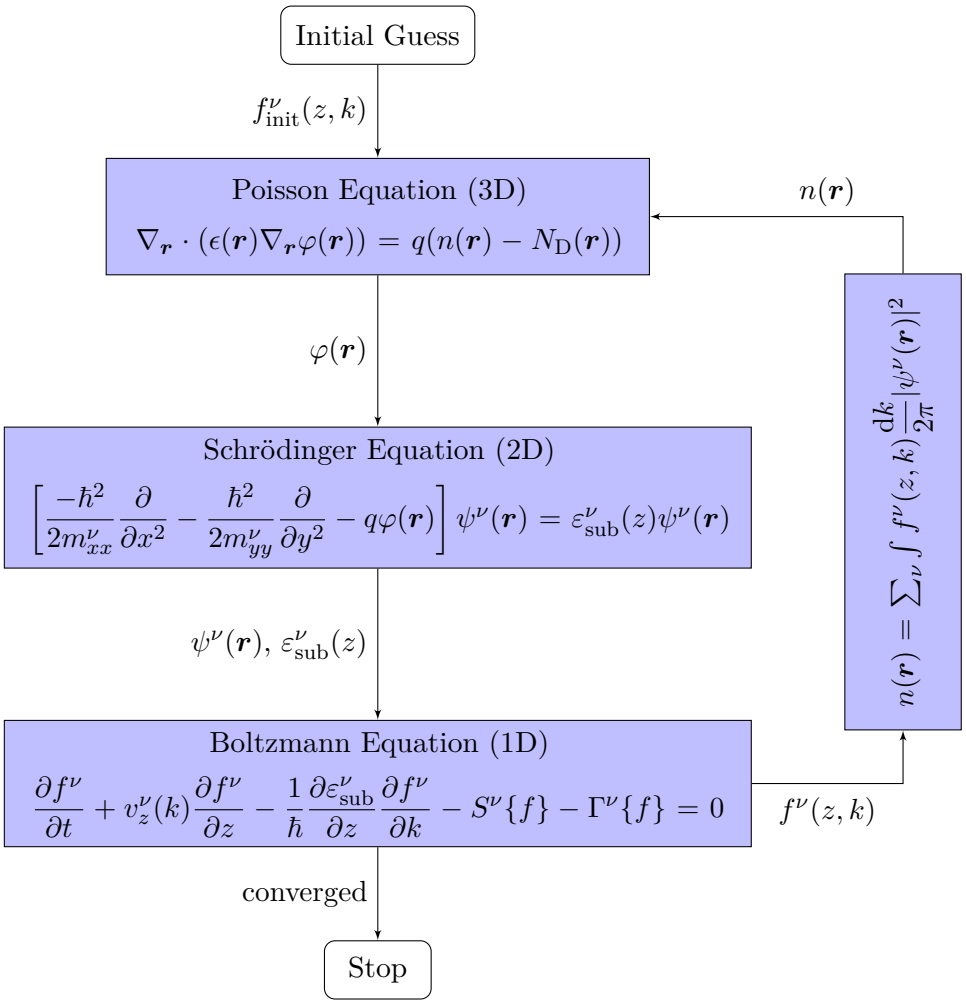

Figure 2.2: Interdependency of the PE, SE, and BE.

where $\nabla_{\boldsymbol{r}}$ is the nabla operator, φ is the electrostatic potential, ρ is the charge density, and ϵ is the dielectric constant which is assumed to be a scalar quantity[1]. The vector $\boldsymbol{r} = (x, y, z)$ is three-dimensional in space since the electric field varies in all dimensions for the nanowire structures under study. The net charge density ρ is commonly broken apart into the ionized donor concentration N_{D} and the

[1] In principle, permittivity needs to be expressed as a non-diagonal tensor of rank two. However, the conventional materials used in the semiconductor technology do not exhibit a significant anisotropy of the permittivity and inhomogeneity effects [85], e.g. the properties that can be represented by tensors of second rank are isotropic for the cubic crystal structure of silicon.

ionized acceptor concentration N_A, as well as the free charges which are electrons in the conduction band n_3D and holes in the valence band p_3D. In our $\mathrm{N^+NN^+}$ device, the acceptor density and hole concentration can be safely set to zero, which gives:

$$\rho(\boldsymbol{r}) = q\left(N_\mathrm{D}(\boldsymbol{r}) - n_\mathrm{3D}(\boldsymbol{r}) - N_\mathrm{A}(\boldsymbol{r}) + p_\mathrm{3D}(\boldsymbol{r})\right) \approx q\left(N_\mathrm{D}(\boldsymbol{r}) - n_\mathrm{3D}(\boldsymbol{r})\right) , \quad (2.2)$$

where q is the positive elementary charge. A substantial trick for stabilizing the PE is to use the nonlinear transformation of the carrier density (see e.g. [85])

$$n_\mathrm{3D}(\boldsymbol{r}) = n_i \exp\left(\frac{\varphi(\boldsymbol{r}) - \phi_n(\boldsymbol{r})}{V_\mathrm{T}}\right) , \quad (2.3)$$

where n_i is the intrinsic carrier density, ϕ_n is the quasi–Fermi potential of electrons, and $V_\mathrm{T} = k_\mathrm{B}T/q$ denotes the thermal voltage with the Boltzmann constant k_B and the absolute temperature T. In the Gummel iterations, the nonlinear PE is solved for the potential φ with fixed quasi–Fermi levels. Note that (2.3) is used solely as a transformation to quantities with the same dimensions, and the fact that its form assumes a Maxwell–Boltzmann distribution of electrons does not invalidate this mathematical procedure for the degenerate case. When the Fermi–Dirac statistics is considered, the effects of the Pauli exclusion principle are implicitly included in the quasi–Fermi potential ϕ_n.

The PE, being a second order elliptical PDE, is a boundary value problem and needs to be supplemented with suitable boundary conditions. For the device shown in Fig. 2.1, the bias V_G is directly applied to the gate electrode to modulate the number of free electrons. Hence, Dirichlet boundary conditions are assumed on the gate contact

$$\varphi(\boldsymbol{r})\Big|_{\boldsymbol{r}\in\partial D_\mathrm{G}} = V_\mathrm{G} + \varphi_\mathrm{MS} , \quad (2.4)$$

with φ_MS representing the metal-semiconductor work function difference. Everywhere else on the boundary of the device, Neumann boundary conditions are imposed. It should be noted, that the source and drain biases are incorporated through the BE, as will be discussed later on in Sect. 2.4.

2.2.2 Numerical Implementation

Equation (2.1) is discretized using the finite volume method[2] [89]. The grid may consist of 2D unstructured triangular meshes on the transverse $x-y$ planes,

[2] Although the finite element method can easily discretize complex geometries, it does not satisfy the relevant conservation principles within individual elements since the equations of equilibrium and boundary conditions are replaced by their weighted means [86, 87]. In practice, this drawback is usually addressed by mesh refinement in order to ensure that a sufficient number of elements is present so that the inequilibrium of individual elements does not influence the overall solution. The finite volume method, on the other hand, provides a strong physical representation of the conservation laws [88] and has become the major technique for discretization of the Poisson and Boltzmann transport equations.

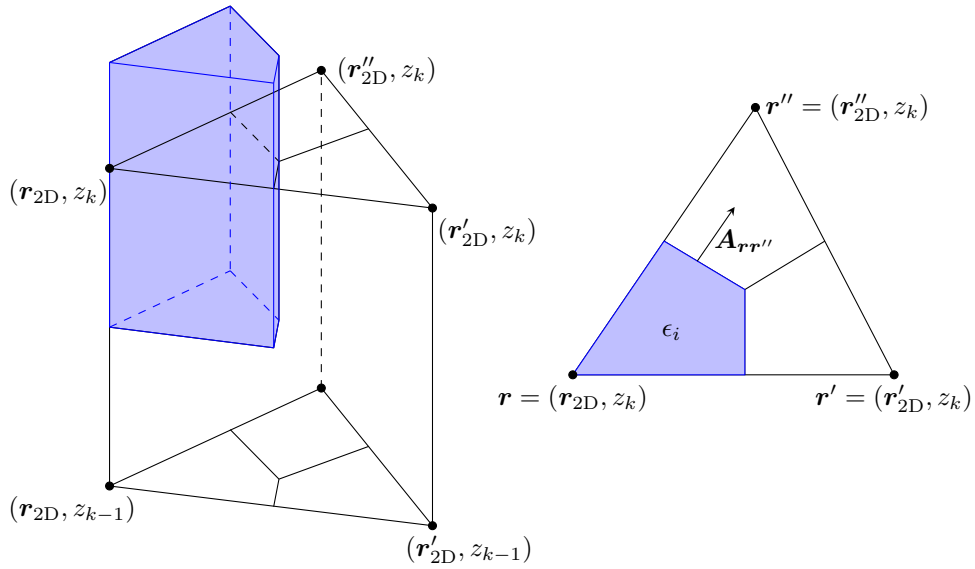

Figure 2.3: Schematic of the control volume (left) and surface vector (right) in the finite volume method.

repeated n_z times in the transport direction. These triangular meshes are assumed to be non-obtuse (i.e. they form a Delauney grid) and Voronoi polygons split the simulation domain into prismatic control volumes by the perpendicular bi-sectors of neighboring primary element edges [90–92]. In Fig. 2.3, part of the control volume of node $(\boldsymbol{r}_{2D}, z_k)$ with the respective edges of the Delauney grid are shown in blue. Each control volume l has $n_l - 2$ faces with vertical orientation and two with horizontal orientation, and a finite volume discretization can be obtained by integration over the control volumes. The electric flux through each of the n_l surfaces is given by the permittivity of the corresponding grid primitive times the electric field times the surface vector. Defining $A_{\boldsymbol{r}\boldsymbol{r}''}$ as the area of the box side lying between nodes $\boldsymbol{r}$ and $\boldsymbol{r}''$, the surface vector is:

$$\boldsymbol{A}_{\boldsymbol{r}\boldsymbol{r}''} = \frac{\boldsymbol{r} - \boldsymbol{r}''}{\|\boldsymbol{r} - \boldsymbol{r}''\|_2} A_{\boldsymbol{r}\boldsymbol{r}''} \tag{2.5}$$

and only the component of the electric field along the edge is required:

$$\boldsymbol{D}_{\boldsymbol{r}\boldsymbol{r}''} \cdot \boldsymbol{A}_{\boldsymbol{r}\boldsymbol{r}''} = -\epsilon_{\boldsymbol{r}\boldsymbol{r}''} \nabla\varphi \frac{\boldsymbol{r} - \boldsymbol{r}''}{\|\boldsymbol{r} - \boldsymbol{r}''\|_2} A_{\boldsymbol{r}\boldsymbol{r}''} \approx -\epsilon_{\boldsymbol{r}\boldsymbol{r}''} A_{\boldsymbol{r}\boldsymbol{r}''} \frac{\varphi(\boldsymbol{r}) - \varphi(\boldsymbol{r}'')}{\|\boldsymbol{r} - \boldsymbol{r}''\|_2}\,, \tag{2.6}$$

where the potential φ is defined directly on the grid nodes and assumed to be constant within the corresponding finite volume. If the two grid nodes $\boldsymbol{r}$, $\boldsymbol{r}''$ are interchanged, we get the flux from node $\boldsymbol{r}''$ to $\boldsymbol{r}$,

$$\boldsymbol{D}_{\boldsymbol{r}''\boldsymbol{r}} \cdot \boldsymbol{A}_{\boldsymbol{r}''\boldsymbol{r}} = -\epsilon_{\boldsymbol{r}\boldsymbol{r}''} A_{\boldsymbol{r}\boldsymbol{r}''} \frac{\varphi(\boldsymbol{r}'') - \varphi(\boldsymbol{r})}{\|\boldsymbol{r} - \boldsymbol{r}''\|_2} = -\boldsymbol{D}_{\boldsymbol{r}\boldsymbol{r}''} \cdot \boldsymbol{A}_{\boldsymbol{r}\boldsymbol{r}''}\,, \tag{2.7}$$

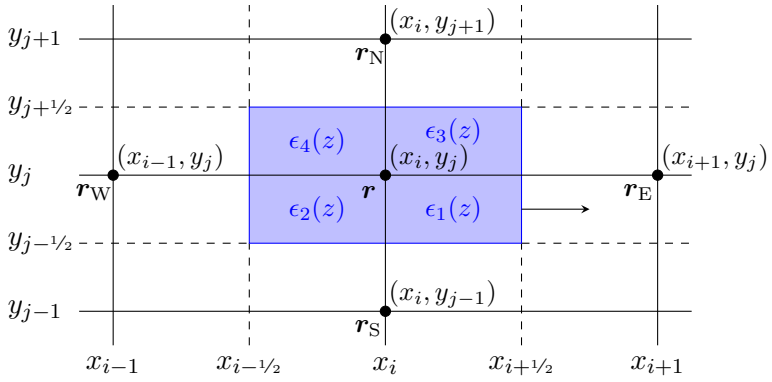

Figure 2.4: Cross-section of the finite volume for grid node $(\boldsymbol{r}, z_k)$ on the $z = z_k$ plane (blue area). The space charge density ρ and the electric potential φ are defined on direct grid points, indicated by black circles. The permittivities ϵ_i are assumed to be constant within the grid primitives.

which equals the flux from $\boldsymbol{r}$ to $\boldsymbol{r}''$ with an inverted sign, because the surface areas are the same and their vectors antiparallel.

Since the unstructured grids introduce unnecessary computational cost for no additional physical insight, the special case of a Cartesian tensor product grid is chosen for this work. This results in cuboid control volumes, as shown in Fig. 2.4. A composite index is calculated based on the indices i, j, and k:

$$l = (k-1)n_x n_y + (j-1)n_x + i \;, \tag{2.8}$$

where n_x and n_y are the number of nodes in the x and y directions, respectively. Using the notation in Fig. 2.4 and defining $\partial\mathrm{D}(\boldsymbol{r}, z_k)$ and $\mathrm{D}(\boldsymbol{r}, z_k)$ as the surface and volume of the box for the generic grid node $(\boldsymbol{r}, z_k)$, the sum over all surface parts yields

$$\begin{aligned}
&\oint_{\partial\mathrm{D}(\boldsymbol{r},z_k)} \boldsymbol{D}\cdot d\boldsymbol{A} \approx \\
&\frac{y_j - y_{j-1}}{4(x_{i+1}-x_i)}\Big[\epsilon_1(z_{k+})(z_{k+1}-z_k) + \epsilon_1(z_{k-})(z_k - z_{k-1})\Big]\left(\varphi(\boldsymbol{r}, z_k) - \varphi(\boldsymbol{r}_\mathrm{E}, z_k)\right) \\
&+ \frac{y_{j+1} - y_j}{4(x_{i+1}-x_i)}\Big[\epsilon_3(z_{k+})(z_{k+1}-z_k) + \epsilon_3(z_{k-})(z_k - z_{k-1})\Big]\left(\varphi(\boldsymbol{r}, z_k) - \varphi(\boldsymbol{r}_\mathrm{E}, z_k)\right) \\
&+ \frac{y_j - y_{j-1}}{4(x_i-x_{i-1})}\Big[\epsilon_2(z_{k+})(z_{k+1}-z_k) + \epsilon_2(z_{k-})(z_k - z_{k-1})\Big]\left(\varphi(\boldsymbol{r}, z_k) - \varphi(\boldsymbol{r}_\mathrm{W}, z_k)\right) \\
&+ \frac{y_{j+1} - y_j}{4(x_i-x_{i-1})}\Big[\epsilon_4(z_{k+})(z_{k+1}-z_k) + \epsilon_4(z_{k-})(z_k - z_{k-1})\Big]\left(\varphi(\boldsymbol{r}, z_k) - \varphi(\boldsymbol{r}_\mathrm{W}, z_k)\right) \\
&+ \frac{x_i - x_{i-1}}{4(y_{j+1}-y_j)}\Big[\epsilon_4(z_{k+})(z_{k+1}-z_k) + \epsilon_4(z_{k-})(z_k - z_{k-1})\Big]\left(\varphi(\boldsymbol{r}, z_k) - \varphi(\boldsymbol{r}_\mathrm{N}, z_k)\right)
\end{aligned}$$

$$
\begin{aligned}
&+\frac{x_{i+1}-x_i}{4(y_{j+1}-y_j)}\Big[\epsilon_3(z_{k+})(z_{k+1}-z_k)+\epsilon_3(z_{k-})(z_k-z_{k-1})\Big]\left(\varphi(\boldsymbol{r},z_k)-\varphi(\boldsymbol{r}_{\mathrm{N}},z_k)\right)\\
&+\frac{x_i-x_{i-1}}{4(y_j-y_{j-1})}\Big[\epsilon_2(z_{k+})(z_{k+1}-z_k)+\epsilon_2(z_{k-})(z_k-z_{k-1})\Big]\left(\varphi(\boldsymbol{r},z_k)-\varphi(\boldsymbol{r}_{\mathrm{S}},z_k)\right)\\
&+\frac{x_{i+1}-x_i}{4(y_j-y_{j-1})}\Big[\epsilon_1(z_{k+})(z_{k+1}-z_k)+\epsilon_1(z_{k-})(z_k-z_{k-1})\Big]\left(\varphi(\boldsymbol{r},z_k)-\varphi(\boldsymbol{r}_{\mathrm{S}},z_k)\right)\\
&+\frac{(x_{i+1}-x_{i-1})(y_{j+1}-y_{j-1})}{4(z_k-z_{k-1})}\epsilon_{\mathrm{avg}}(z_k)\left(\varphi(\boldsymbol{r},z_k)-\varphi(\boldsymbol{r},z_{k-1})\right)\\
&+\frac{(x_{i+1}-x_{i-1})(y_{j+1}-y_{j-1})}{4(z_{k+1}-z_k)}\epsilon_{\mathrm{avg}}(z_k)\left(\varphi(\boldsymbol{r},z_k)-\varphi(\boldsymbol{r},z_{k+1})\right)\\
&=\frac{(x_{i+1}-x_{i-1})}{2}\frac{(y_{j+1}-y_{j-1})}{2}\frac{(z_{k+1}-z_{k-1})}{2}\rho(\boldsymbol{r},z_k)\approx\int_{\mathrm{D}(\boldsymbol{r},z_k)}\rho\,dV\ ,
\end{aligned}
\tag{2.9}
$$

where the charge density acts as the source term in the control volume, and the right–hand side is simply given by the box volume times the average charge density in the box. In this case of a 3D tensor-product grid the Laplace operator involves the potential on seven grid nodes, i.e. the central grid node $(\boldsymbol{r},z_k)$ and the six nearest neighbors. Interface conditions are automatically included in (2.9), because the permittivity can change from primitive to primitive. On the boundary of the solution domain, where Neumann boundary conditions apply, the integrals over the corresponding surface sections vanish. This corresponds to the assumption, that the permittivity is zero outside of the solution domain.

Although the differential terms are linear, the space charge density ρ depends nonlinearly on the potential and the PE cannot be solved in one step. In order to solve the discretized set of equations, the Newton-Raphson method can be employed. The derivative of the charge density is calculated as

$$
\frac{\partial\rho(\boldsymbol{r},z_k)}{\partial\varphi(\boldsymbol{r}',z_{k'})}=\frac{-qn_i}{V_{\mathrm{T}}}\exp\left(\frac{\varphi(\boldsymbol{r},z_k)-\phi_n(\boldsymbol{r},z_k)}{V_{\mathrm{T}}}\right)\delta_{\boldsymbol{r}\boldsymbol{r}'}\delta_{kk'}\ . \tag{2.10}
$$

It should be noted that the Jacobian is an M–matrix [93] and the Newton-Raphson method converges for all initial values. The typical convergence behavior is shown in Fig. 2.5. Faraway from the solution, the potential changes at most by V_{T}. As the potential gets closer to the solution, the rate of convergence becomes quadratic and the solution is obtained within a few more iteration steps.

2.3 The Schrödinger Equation

As the cross-sectional area of the nanowires shrink to below $10\times10\,\mathrm{nm}^2$, confinement of the carriers in the perpendicular directions begins to play an important role. The quantum effects are manifested in the development of discrete energy subbands, which in turn leads to variations of the bandgap with the cross-sectional

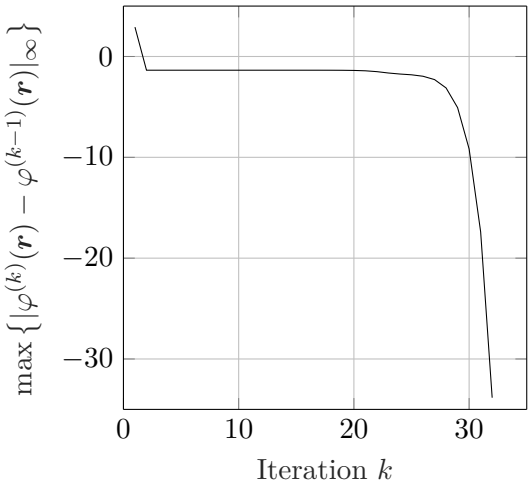

Figure 2.5: Maximum change of the potential of the Newton-Raphson method for the nonlinear Poisson equation.

area and reduction of the conduction channels available for charge transport [94]. Moreover, the carriers get pushed away from the semiconductor/oxide interface, thereby increasing the effective oxide thickness and reducing the capacitance of the gate dielectric. In order to capture the quantization of charge carriers, it is necessary to include the SE into a self–consistent computation with the PE. In this section, numerical implementation of the 2D time independent SE is briefly discussed.

2.3.1 Main Equation

In principle, the SE describes the coherent motion of an electron subject to internal potentials, i.e. the potentials induced by the crystal to whom the electron belongs ($\mathcal{V}_\mathrm{C}$), as well as the effect of the external forces ($\mathcal{V}_\mathrm{E}$),

$$\frac{-\hbar^2}{2m_\mathrm{e}}\nabla^2\psi(\boldsymbol{r}) + [\mathcal{V}_\mathrm{C}(\boldsymbol{r}) + \mathcal{V}_\mathrm{E}(\boldsymbol{r})]\psi(\boldsymbol{r}) = E\psi(\boldsymbol{r}) \tag{2.11}$$

where $\hbar$ is the reduced Planck constant, ψ is the wave function and E the energy of the particle. No time-dependence is assumed in (2.11), i.e. the relaxation time for the eigenvalues and eigenfunctions of (2.11) is short enough and the Hamiltonian shifts instantaneously by time variations of the applied bias and perturbations in the steady state. The effect of the crystal is expressed through the band structure,

$$\varepsilon(-\mathrm{i}\nabla)\psi(\boldsymbol{r}) + \mathcal{V}_\mathrm{E}(\boldsymbol{r})\psi(\boldsymbol{r}) = E\psi(\boldsymbol{r}) \tag{2.12}$$

where the notation $\varepsilon(-\mathrm{i}\nabla)$ is adopted to show that the band structure function $\varepsilon(k_x, k_y, k_z)$ reads $\varepsilon(-\mathrm{i}\frac{\partial}{\partial x}, -\mathrm{i}\frac{\partial}{\partial y}, -\mathrm{i}\frac{\partial}{\partial z})$ [95]. Hence, the effects of the internal potential $\mathcal{V}_\mathrm{C}(\boldsymbol{r})$ are taken into account by the differential operator $\varepsilon(-\mathrm{i}\nabla)$. The eigenstates of this so-called envelope function equation do not oscillate at the atomic scale of the crystal, but rather describe the behavior of electrons subject to slowly varying external potentials.

The direct application of sophisticated full band structure calculations is time consuming. Assuming parabolic mass approximation in $x-$ and $y-$directions, the electron is represented around the energy minimum of the crystal band structure by,

$$\frac{-\hbar^2}{2}\nabla\cdot\left(\hat{\mathrm{M}}^{-1}\nabla\psi(\boldsymbol{r})\right)+\mathcal{V}_\mathrm{E}(\boldsymbol{r})\psi(\boldsymbol{r})=E\psi(\boldsymbol{r})\,, \tag{2.13}$$

where the mass tensor $\hat{\mathrm{M}}$ parameterizes the inertia resulting from the compound interaction of the underlying crystal lattice[3]. Corrections to account for the non-parabolicity of the bands can extend the validity of the model to higher energies [9, 99].

Under the assumption of a negligible electric field in the longitudinal direction, the separation of variables leads to a 2D SE in the $x-y$ plane. The procedure of breaking the full 3D problem into its confined and non–confined parts is well known and widely used in the numerical investigation of low–dimensional structures. For detailed discussions on the rationale of this framework, the reader is referred to, e.g., Refs. [11, 83, 100, 101].

2.3.2 Numerical Implementation

The stationary SE for the envelope function is solved on $x-y$ planes perpendicular to the $z-$axis, in each mesh point of the $z-$axis. The eigenvalues are the quantized subband energies, represented by $\varepsilon^\nu_\mathrm{sub}(z)$ throughout this work and ψ^ν is the wave function corresponding to the eigenvalue $\varepsilon^\nu_\mathrm{sub}$. The composite superscript $\nu=(v,s)$ comprises the valleys of the band structure v and the subbands s emerging from the carrier confinement. For the Cartesian tensor product grid of Fig. 2.4, the 2D Hamiltonian gives

$$\left[\frac{-\hbar^2}{2m^\nu_{xx}}\frac{\partial}{\partial x^2}-\frac{\hbar^2}{2m^\nu_{yy}}\frac{\partial}{\partial y^2}+\mathcal{V}_\mathrm{E}(\boldsymbol{r})\right]\psi^\nu(\boldsymbol{r})=\varepsilon^\nu_\mathrm{sub}(z)\psi^\nu(\boldsymbol{r})\,, \tag{2.14}$$

[3] Treating the 2D Schrödinger equation within the envelope function and effective mass framework has proven to be an adequate approximation down a wire diameter of 3 nm [48]. For example, [48] and [96] have shown that valley splitting is rather insignificant for [001] nanowires (only 10 meV even for 4 nm^2 cross-sections). However, it should be noted that the effective mass approximation has certain limitations compared to the more rigorous band structure calculations, and effects such as the band coupling, Brillouin zone folding and energy dependence of the non-parabolicity cannot be considered within the EMA framework [47, 48, 96–98].

where m^{ν}_{xx} and m^{ν}_{yy} are the electron effective masses in x and y direction, respectively. At this point, it is important to note that although the SE is solved on 2D planes, the eigenfunctions will be functions of $\boldsymbol{r}$ since the electrostatic potential also depends on the z–coordinate. The natural way to solve (2.14) is to replace the differential operator with its finite differences equivalent,

$$\frac{-\hbar^2}{2m^{\nu}_{xx}}\left[\frac{\psi^{\nu}_{i-1,j}-2\psi^{\nu}_{i,j}+\psi^{\nu}_{i+1,j}}{(\Delta x)^2}+\mathcal{O}(\Delta x^2)\right] - \frac{\hbar^2}{2m^{\nu}_{yy}}\left[\frac{\psi^{\nu}_{i,j-1}-2\psi^{\nu}_{i,j}+\psi^{\nu}_{i,j+1}}{(\Delta y)^2}+\mathcal{O}(\Delta y^2)\right]+\mathcal{V}_{\mathrm{E}}(\boldsymbol{r})\psi^{\nu}_{i,j}=\varepsilon^{\nu}_{\mathrm{sub}}(z)\psi^{\nu}_{i,j} \quad (2.15)$$

where (i,j) are the indices of a generic node in the 2D space grid. The potential energy is simply given by $\mathcal{V}_{\mathrm{E}}(\boldsymbol{r})=-q\varphi(\boldsymbol{r})$ and shifts the main diagonal elements of the Hamiltonian matrix. The grid nodes are defined as:

$$\begin{aligned} x_i &= x_0+i\Delta x \qquad \text{for } i\in\{0,1,\cdots,n_x-1\} \\ y_j &= y_0+j\Delta y \qquad \text{for } j\in\{0,1,\cdots,n_y-1\}\,, \end{aligned} \quad (2.16)$$

which is the same grid used for discretization of the PE, and the wave function is represented by a single column vector. The resulting Hamiltonian is a sparse square matrix with dimension $n_x\times n_y$. One can try to choose the optimal grid constants $(\Delta x,\Delta y)$ per dimension, but if all energy levels up to some energy are needed it is profitable to use mass-weighted coordinates ($x'=\sqrt{m^{\nu}_{xx}}x$, $y'=\sqrt{m^{\nu}_{yy}}y$), since the same grid constant can be used for every dimension. Another idea is to transform the Hamiltonian matrix into a real symmetric matrix using a similarity transform. Having the symmetric matrix $\hat{H}'=\hat{D}^{-1}\hat{H}\hat{D}$, it follows that:

$$\hat{H}'(\hat{D}^{-1}\boldsymbol{\psi})=(\hat{D}^{-1}\hat{H}\hat{D})\hat{D}^{-1}\boldsymbol{\psi}=\hat{D}^{-1}(E\boldsymbol{\psi})=E(\hat{D}^{-1}\boldsymbol{\psi})\,. \quad (2.17)$$

Hence, the required eigenvalues E are also the eigenvalues of the symmetric matrix $\hat{H}'$. Once the eigenfunctions $\boldsymbol{\psi}'=\hat{D}^{-1}\boldsymbol{\psi}$ of the matrix $\hat{H}'$ have been found, the eigenfunctions $\boldsymbol{\psi}$ of the physical problem are found simply from the inversion $\boldsymbol{\psi}=\hat{D}\boldsymbol{\psi}'$. As for the boundary conditions, the potential barriers formed by the insulating materials are assumed to be infinitely high. Thus, if we solve the SE on a domain $[0,L_x]\times[0,L_y]$, the boundary condition reads,

$$\begin{aligned} \psi(x,0)&=\psi(x,L_y)=0\,, \\ \psi(0,y)&=\psi(L_x,y)=0\,. \end{aligned} \quad (2.18)$$

The discretized SE is fed into the FEAST eigensolver package [102], and its solutions are subjected to the normalization condition

$$\int_0^{L_y}\mathrm{d}y\int_0^{L_x}\mathrm{d}x\,|\psi^{\nu}(\boldsymbol{r})|^2=1\,. \quad (2.19)$$

The elapsed time to solve the SE on all n_z slices remains well below 1 % of what the BE needs. The numerical solutions, some of them plotted in Fig. 2.6, are standing waves which correspond to a set of increasing energy levels. A closer inspection of these functions verifies the necessity of including the SE in our simulation framework. As shown in Fig. 2.6, the electrons associated with the lowest energy (i.e. $s = 1$) are concentrated mostly in the center (bulk) of the nanowire, rather than at the surface. In subthreshold operation, most of the electrons are in the lowest subband and thus located in the center of the nanowire. By increasing the gate voltage, more subbands become populated and contribute to the carrier transport. Hence, peaks of electron concentration closer to the edges and corners of the channel appear for higher V_{GS}. It is, however, important to notice that a vase majority of electrons are still concentrated near the center of the nanowire, which leads to a different conduction mechanism compared to the conventional FETs in which the carriers form a thin inversion layer at the surface of the semiconductor. This phenomenon, called "volume inversion", is a salient feature of the low-dimensional architectures such as thin SOI films and nanowires [103]. Calculation of the quantum capacitance, a parameter that encompasses variations in channel depth and shape is important in modeling of nanowire transistors [104].

The SE, being an eigenvalue problem, cannot be directly included into a Newton–Raphson solver for the system of constituent equations. However, we can use the first–order time independent perturbation theory in order to express how the subband energies and wave functions change by small perturbations of the potential, and calculate the corresponding derivatives. For the non-degenerate case, this gives [95]

$$\delta\varepsilon^{\nu}_{\mathrm{sub}}(z) = -q \iint \mathrm{d}x\,\mathrm{d}y\,|\psi^{\nu}(\boldsymbol{r})|^2 \delta\varphi(\boldsymbol{r}) \tag{2.20}$$

$$\delta\psi^{\nu}(\boldsymbol{r}) = -q \sum_{\substack{s'\neq s \\ v'=v}} \frac{\iint \mathrm{d}x'\,\mathrm{d}y'\,\psi^{\nu'}(\boldsymbol{r}')\delta\varphi(\boldsymbol{r}')\psi^{\nu}(\boldsymbol{r}')\delta(z-z')}{\varepsilon^{\nu}_{\mathrm{sub}}(z) - \varepsilon^{\nu'}_{\mathrm{sub}}(z)} \psi^{\nu'}(\boldsymbol{r})\,, \tag{2.21}$$

where the summation is performed over valley indices v and subbands s, included in the aggregate index $\nu = (v, s)$. At this juncture, it is important to note that although only a few subbands up to a certain energy suffice for the evaluation of the BE, the numerical computation of the SE cannot be truncated at some subband index since Eqs. (2.20) and (2.21) require the summation over all subbands.

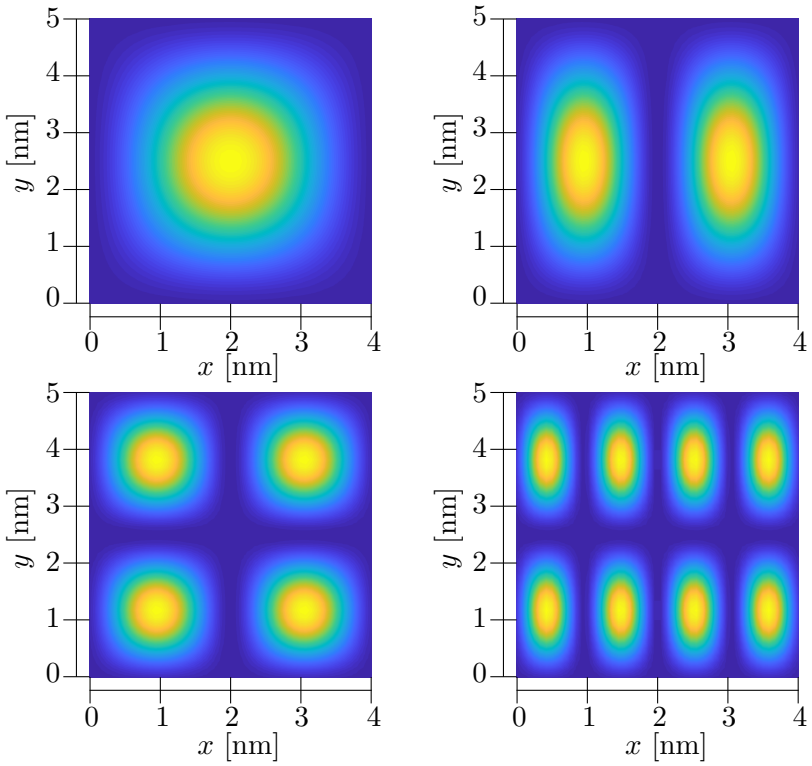

Figure 2.6: Wave functions squared $|\psi(x,y)|^2$ for $\nu = (1,1)$ (upper left), $\nu = (1,3)$ (upper right), $\nu = (1,5)$ (lower left), and $\nu = (1,10)$ (lower right).

2.4 The Boltzmann Transport Equation

The multi–subband BE plays the pivotal role in our simulation framework and is the basic statistical description from which the drift–diffusion or hydrodynamic approximations are derived. Although the validity of the BE already implies several simplifying assumptions[4], it is in general an integro–differential equation with several independent variables and proves to be more complicated compared to the PE and SE. Hence, specialized numerical techniques are required to transform it into a manageable form.

This section covers the discretization of the 1D BE along the transport axis (see Fig. 2.1). Several deterministic approaches to the BE such as discrete velocity models [105], particle methods [106], spectral approximations [107] and power-series discretization [108] have been proposed, and the reader can refer to [109, 110] for more details and references. The numerical schemes in this work

[4]For example, it is assumed that the external forces are almost constant over a length comparable to the physical dimensions of the carrier's wave packet, and the band theory and the effective mass approximation apply to the semiconductor under consideration.

are developed in [33, 84, 111, 112] and have been tested on general problems and higher dimensionalities [33, 34, 42–44, 84, 111, 113–115].

2.4.1 Main Equation

The 1D Boltzmann equation for the distribution function $f_k^\nu = f^\nu(z,k,t)$ reads,

$$\mathrm{F_{BE}} := \frac{\partial f_k^\nu}{\partial t} + v_z^\nu(k)\frac{\partial f_k^\nu}{\partial z} + \frac{1}{\hbar}F^\nu(z,t)\frac{\partial f_k^\nu}{\partial k} - S^\nu\{f_k\} - \Gamma^\nu\{f_k\} = 0 \qquad (2.22)$$

where $S^\nu\{f\}$ is the scattering integral for single electron interactions, and $\Gamma^\nu\{f\}$ represents the contacts' boundary conditions. $F^\nu(z,t)$ is the force which acts on the carriers in transport direction, and $v_z^\nu(k)$ is the electrons' group velocity in transport direction. The subscript k is a reminder that the representations are in the phase space, and the superscript $\nu = (v,s)$ indicates that transport can happen along different channels which, in this work, comprise six ellipsoidal parabolic valleys of the band structure of silicon v and a set of subbands s emerging from the electron confinement. We have

$$F^\nu(z,t) = -\frac{\partial}{\partial z}\varepsilon_{\mathrm{sub}}^\nu(z,t)\,, \qquad v_z^\nu(k) = \frac{1}{\hbar}\frac{\partial}{\partial k}\tilde{\varepsilon}^\nu(k)\,. \qquad (2.23)$$

Unlike the free streaming term, the scattering integral $S\{f\}$ couples different valleys and subbands, and its explicit description depends on the nature of the microscopic interactions. Since the spin–flipping scattering events are not included in this work, the spin index is dropped as a convention, and a twofold spin degeneracy is assumed in the calculation of macroscopic quantities of interest. Having calculated the distribution function $f^\nu(z,k,t)$, the quantities that characterize the macroscopic state of the transport can be easily obtained. The 1D electron and current densities are given by:

$$n(z) = 2\sum_\nu \int_{-\infty}^{\infty} f_k^\nu(z,k)\frac{\mathrm{d}k}{2\pi}\,, \qquad (2.24)$$

$$j(z) = 2\sum_\nu \int_{-\infty}^{\infty} v_z^\nu(k) f_k^\nu(z,k)\frac{\mathrm{d}k}{2\pi}\,. \qquad (2.25)$$

where the pre factor 2 stands for the spin degeneracy.

Direct discretization of (2.22) with the finite volume method results in spurious oscillations and numerical instabilities. Hence, the stabilization scheme in [36] which is based on entropy principles and dual meshes is adopted here. First, we note that the steady state free streaming term gives a natural even/odd decomposition of the function space. That is, for an even dispersion relation

$\tilde{\varepsilon}(k) = \tilde{\varepsilon}(-k)$, it maps even functions of k into odd functions and vice versa[5]. Assuming that the band structure is parabolic along the z–axis, splitting the distribution function $f^\nu(z,k)$ into its even and odd parts in wave vector k

$$f_k^\nu(z,k) = f_k^{\nu\mathrm{e}}(z,k) + f_k^{\nu\mathrm{o}}(z,k) \tag{2.26}$$

and projection of (2.22) onto equi–energy surfaces, i.e. $\int \frac{\mathrm{d}k}{2\pi} \mathrm{F_{BE}}(k)\delta(\tilde{\varepsilon}(k)-\varepsilon)$, yields,

$$\begin{aligned}
D_t :=& \int \frac{\mathrm{d}k}{2\pi} \left(\frac{\partial}{\partial t} f_k^\nu(z,k)\right) \delta\left(\tilde{\varepsilon}(k)-\varepsilon\right) \\
=& \frac{\partial}{\partial t} \int \frac{\mathrm{d}k}{2\pi} f_k^\nu(z,k)\, \frac{1}{\hbar}\sqrt{\frac{m_{zz}^\nu}{2\varepsilon}} \left[\delta\Big(k - \frac{\sqrt{2m_{zz}^\nu\varepsilon}}{\hbar}\Big) + \delta\Big(k + \frac{\sqrt{2m_{zz}^\nu\varepsilon}}{\hbar}\Big)\right] \\
=& \frac{\partial}{\partial t} \frac{1}{2\pi\hbar}\sqrt{\frac{m_{zz}^\nu}{2\varepsilon}} \left[f_k^\nu\Big(z, \frac{\sqrt{2m_{zz}^\nu\varepsilon}}{\hbar}\Big) + f_k^\nu\Big(z, -\frac{\sqrt{2m_{zz}^\nu\varepsilon}}{\hbar}\Big)\right] \\
=& \frac{\partial}{\partial t}\Big(Z^\nu(\varepsilon) f_\varepsilon^{\nu\mathrm{e}}(z,\varepsilon)\Big)
\end{aligned} \tag{2.27}$$

$$\begin{aligned}
D_z :=& \int \frac{\mathrm{d}k}{2\pi} v_z^\nu(k) \left(\frac{\partial}{\partial z} f_k^\nu(z,k)\right) \delta\left(\tilde{\varepsilon}(k)-\varepsilon\right) \\
=& \frac{\partial}{\partial z} \int \frac{\mathrm{d}k}{2\pi} v_z^\nu(k) f_k^\nu(z,k)\, \frac{1}{\hbar}\sqrt{\frac{m_{zz}^\nu}{2\varepsilon}} \left[\delta\Big(k - \frac{\sqrt{2m_{zz}^\nu\varepsilon}}{\hbar}\Big) + \delta\Big(k + \frac{\sqrt{2m_{zz}^\nu\varepsilon}}{\hbar}\Big)\right] \\
=& \frac{\partial}{\partial z}\Big(Z^\nu(\varepsilon) v_z^\nu(\varepsilon) f_\varepsilon^{\nu\mathrm{o}}(z,\varepsilon)\Big)
\end{aligned} \tag{2.28}$$

$$\begin{aligned}
D_k :=& \int \frac{\mathrm{d}k}{2\pi} \frac{1}{\hbar} F^\nu(z,t) \left(\frac{\partial}{\partial k} f_k^\nu(z,k)\right) \delta\left(\tilde{\varepsilon}(k)-\varepsilon\right) \\
=& -\int \frac{\mathrm{d}k}{2\pi} \frac{1}{\hbar} F^\nu(z,t) f_k^\nu(z,k) \left(\frac{\partial}{\partial k}\delta\Big(\frac{\hbar^2k^2}{2m_{zz}^\nu} - \varepsilon\Big)\right) \\
=& -\int \frac{\mathrm{d}k}{2\pi} F^\nu(z,t) f_k^\nu(z,k) \Big[v_z^\nu(k) \frac{\partial}{\partial\tilde{\varepsilon}}\delta(\tilde{\varepsilon}(k)-\varepsilon)\Big]
\end{aligned} \tag{2.29}$$

$$= \frac{\partial}{\partial\varepsilon}\Big(Z^\nu(\varepsilon) v_z^\nu(\varepsilon) F^\nu(z,t) f_\varepsilon^{\nu\mathrm{o}}(z,\varepsilon)\Big)\,. \tag{2.30}$$

[5]Note that the gradient of an even function is an odd function and vice versa. Now, let $f^{\mathrm{even}} \in C^{\mathrm{even}}$. This results in

$$\mathrm{L}\{f^{\mathrm{even}}\} = \underbrace{\boldsymbol{v}(\boldsymbol{k})}_{\mathrm{odd}} . \underbrace{\nabla_{\boldsymbol{r}} f^{\mathrm{even}}}_{\mathrm{even}} + \frac{1}{\hbar}\boldsymbol{F}(\boldsymbol{r}) . \underbrace{\nabla_{\boldsymbol{k}} f^{\mathrm{even}}}_{\mathrm{odd}} \in C^{\mathrm{odd}}$$

$$\mathrm{L}\{f^{\mathrm{odd}}\} = \underbrace{\boldsymbol{v}(\boldsymbol{k})}_{\mathrm{odd}} . \underbrace{\nabla_{\boldsymbol{r}} f^{\mathrm{odd}}}_{\mathrm{odd}} + \frac{1}{\hbar}\boldsymbol{F}(\boldsymbol{r}) . \underbrace{\nabla_{\boldsymbol{k}} f^{\mathrm{odd}}}_{\mathrm{even}} \in C^{\mathrm{even}}$$

where $\mathrm{L}\{\cdot\}$ denotes the free streaming operator of the BE.

Similarly, the subscript ε reminds us that the representations are in the energy space. Equation (2.29) is obtained using the following relation:

$$\frac{\partial}{\partial k}\delta(\tilde{\varepsilon}(k)-\varepsilon^{\nu}) = \hbar v_z^{\nu}(\varepsilon)\frac{\partial}{\partial \tilde{\varepsilon}}\delta(\tilde{\varepsilon}(k)-\varepsilon^{\nu})\,, \tag{2.31}$$

and we have derived (2.30) using:

$$\int \mathrm{d}x\, h(x)\Big(\frac{\partial}{\partial x}\delta(x-x_0)\Big) = -\int \mathrm{d}x\, \frac{\partial h(x)}{\partial x}\delta(x-x_0) = -\left(\frac{\partial h}{\partial x}\right)(x_0)\,. \tag{2.32}$$

In the above equations, $Z^{\nu}(\varepsilon) = \dfrac{1}{\pi\hbar}\sqrt{\dfrac{m_{zz}^{\nu}}{2\varepsilon}}$ is the 1D density of states.

Since the free streaming part of the BE contains two distinct partial derivatives with respect to position and energy, we might have numerical problems in handling the equation. In order to overcome these difficulties and cancel the term containing the derivative with respect to energy, we can use the H–transformation [34, 84] and introduce a variable transformation from (z,ε) to $(z,\varepsilon_{\mathrm{sub}}^{\nu}(z)+\varepsilon)$,

$$D_t = \frac{\partial}{\partial t}\Big(Z^{\nu}(\varepsilon)f_{\varepsilon}^{\nu}(z,\varepsilon)\Big) = \frac{\partial}{\partial t}\Big(Z^{\nu}(z,H)f_{\mathrm{H}}^{\nu}(z,H)\Big) \tag{2.33}$$

$$D_z = \frac{\partial}{\partial z}\Big(Z^{\nu}(\varepsilon)v_z^{\nu}(\varepsilon)f_{\varepsilon}^{\nu}(z,\varepsilon)\Big) = \Big(\frac{\partial}{\partial z} + \frac{\partial \varepsilon_{\mathrm{sub}}^{\nu}}{\partial z}\frac{\partial}{\partial H}\Big)\Big(Z^{\nu}(z,H)v_z^{\nu}(z,H)f_{\mathrm{H}}^{\nu}(z,H)\Big) \tag{2.34}$$

$$D_k = \frac{\partial}{\partial \varepsilon}\Big(Z^{\nu}(\varepsilon)v_z^{\nu}(\varepsilon)F^{\nu}(z,t)f_{\varepsilon}^{\nu}(z,\varepsilon)\Big) = F^{\nu}(z,t)\frac{\partial}{\partial H}\Big(Z^{\nu}(z,H)v_z^{\nu}(z,H)f_{\mathrm{H}}^{\nu}(z,H)\Big) \tag{2.35}$$

and the term containing the energy derivative gets cancelled conveniently. The free streaming term is then transformed to,

$$\begin{aligned}&\mathrm{T}\{f_{\mathrm{H}}^{\nu}(z,H)\} + \mathrm{L}\{f_{\mathrm{H}}^{\nu}(z,H)\} := \\ &\frac{\partial}{\partial t}\Big(Z^{\nu}(z,H)f_{\mathrm{H}}^{\nu}(z,H)\Big) + \frac{\partial}{\partial z}\Big(Z^{\nu}(z,H)v_z^{\nu}(z,H)f_{\mathrm{H}}^{\nu}(z,H)\Big)\end{aligned} \tag{2.36}$$

with

$$Z^{\nu}(z,H) = \frac{1}{\pi\hbar}\sqrt{\frac{m_{zz}^{\nu}}{2\big(H-\varepsilon_{\mathrm{sub}}^{\nu}(z)\big)}}\,, \quad v_z^{\nu}(z,H) = \sqrt{\frac{2\big(H-\varepsilon_{\mathrm{sub}}^{\nu}(z)\big)}{m_{zz}^{\nu}}}\,. \tag{2.37}$$

This simplification, however, comes at the price of a curvilinear simulation domain, which complicates the solution of the coupled PE–BE system. As for the scattering term, it can be approximated using a relaxation time τ,

$$S_{\mathrm{RTA}}^{\nu}\{f_k\} = -\frac{f_k^{\nu}(z,k) - f_{k\,\mathrm{eq}}^{\nu}(z,k)}{\tau}\,. \tag{2.38}$$

where the single–particle equilibrium distribution function $f^{\nu}_{k\,\mathrm{eq}}$ needs to be constructed in such a way to conserve the charge:

$$f^{\nu}_{\mathrm{eq}}(z,k) = \frac{\hbar}{\sqrt{2\pi m^{\nu}_{zz} k_{\mathrm{B}} T}} \exp\left(\frac{-\hbar^2 k^2}{2 m^{\nu}_{zz} k_{\mathrm{B}} T}\right) \int_{-\infty}^{\infty} f^{\nu}_{k}(z,k)\,\mathrm{d}k \ . \tag{2.39}$$

In this work, we assume a relaxation time constant with energy. The extension to an energy-dependent relaxation time is, however, streightforward. Projection of (2.38) onto constant energy surfaces yields

$$\int \frac{\mathrm{d}k}{2\pi}\Big[S^{\nu}_{\mathrm{RTA}}\{f_k\}\Big]\,\delta\left(\tilde{\varepsilon}(k)-\varepsilon\right) = -\frac{Z^{\nu}(\varepsilon) f^{\nu}_{\varepsilon}(z,\varepsilon) - Z^{\nu}(\varepsilon) f^{\nu}_{\varepsilon\,\mathrm{eq}}(z,\varepsilon)}{\tau}\,, \tag{2.40}$$

and the H–transformation results in

$$S^{\nu}_{\mathrm{RTA}}\{f_{\mathrm{H}}\} = -\frac{Z^{\nu}(z,H) f^{\nu}_{\mathrm{H}}(z,H) - Z^{\nu}(z,H) f^{\nu}_{\mathrm{H\,eq}}(z,H)}{\tau}\,. \tag{2.41}$$

The above relaxation time approximation (RTA) has a rather simple form and can be used in generic studies where the details of scattering events are not important. For device simulation, however, a more accurate description of the scattering mechanisms is required. Phonon scattering can be expressed by the following integral [9],

$$\begin{aligned} S^{\nu}_{\mathrm{PH}}\{f_k\} =& L_{\mathrm{sys}} \sum_{\nu'} \int \frac{\mathrm{d}k'}{2\pi}\Big[f^{\nu'}_{k}(z,k',t)\big(1-f^{\nu}_{k}(z,k,t)\big)W^{\nu\nu'}(z;k|k') \\ & - f^{\nu}_{k}(z,k,t)\big(1-f^{\nu'}_{k}(z,k',t)\big)W^{\nu'\nu}(z;k'|k)\Big] \end{aligned} \tag{2.42}$$

where L_{sys} is the 1D system volume and the proportionality factor $W^{\nu\nu'}(z;k|k')$ is the transition rate from state (k',ν') into (k,ν), calculated by Fermi's Golden Rule [8]. The first term in the integrand describes the number of electrons scattered from volume element $(\mathrm{d}k',\nu')$ into the state (k,ν) per unit time, and the second term equals the number of electrons scattered from state (k,ν) into the volume element $(\mathrm{d}k',\nu')$ per unit time. The factors $(1-f^{\nu}(z,k))$ and $(1-f^{\nu'}(z,k'))$ take into account the Pauli principle, which does not allow transitions into occupied states. The intravalley acoustic phonon scattering in this work is approximated as to be elastic, with its scattering rate given by [9]:

$$W^{\nu\nu'}_{\mathrm{intra}}(z;k|k') = \frac{2\pi k_{\mathrm{B}} T |\Xi^{\nu}|^2}{\hbar\rho L_{\mathrm{sys}} v^2_{\lambda}} \mathrm{F}_{\nu\nu'}(z)\delta\left(\varepsilon^{\nu}_{\mathrm{tot}}(z,k) - \varepsilon^{\nu'}_{\mathrm{tot}}(z,k')\right)\delta_{v,v'} \tag{2.43}$$

where the Kronecker delta $\delta_{v,v'}$ ensures that the electrons do not scatter to other valleys. The overlap integral $\mathrm{F}_{\nu\nu'}(z)$, defined as [116]

$$\mathrm{F}_{\nu\nu'}(z) = \iint \mathrm{d}x\,\mathrm{d}y\,|\psi^{\nu}(\boldsymbol{r})|^2 |\psi^{\nu'}(\boldsymbol{r})|^2 \tag{2.44}$$

quantity	value	description
k_B	$8.617330 \times 10^{-5}\,\mathrm{eV\,K^{-1}}$	Boltzmann constant
T	$300\,\mathrm{K}$	temperature
Ξ	$20\,\mathrm{eV}$	effective deformation potential
ρ	$2.33 \times 10^{3}\,\mathrm{kg\,m^{3}}$	mass density of the silicon crystal
v_λ	$9.05\,\mathrm{m\,s^{-1}}$	sound velocity of acoustic phonons

Table 2.1: Phonon scattering parameters used in this work.

Phonon process	η	$\hbar\omega_\eta$ [meV]	D_η [10^{10}eVm^{-1}]
g–type, transversal acoustic	g_1	12.1	0.47
g–type, longitudinal acoustic	g_2	18.5	0.74
g–type, longitudinal optical	g_3	62.0	10.23
f–type, transversal acoustic	f_1	19.0	0.28
f–type, longitudinal acoustic	f_2	47.4	1.86
f–type, longitudinal optical	f_3	58.6	1.86

Table 2.2: Phonon energies and deformation potentials for different f-type and g-type intervalley scattering processes in silicon [117–120].

is determined by the SE, and represents an effective distance over which electrons in different states ν and ν' interact. Other parameters in (2.43) are given in Table 2.1. As for the intervalley phonon scattering, thorough studies and approximations are given in [100, 117, 121]. The transition rate reads,

$$W^{\nu\nu'}_{\mathrm{inter},\eta}(z;k|k') = \frac{\pi D_\eta^2}{\rho L_\mathrm{sys}\omega_\eta}\mathrm{F}_{\nu\nu'}(z)\Big[n(\hbar\omega_\eta)\delta\big(\varepsilon^{\nu}_\mathrm{tot}(z,k) - \varepsilon^{\nu'}_\mathrm{tot}(z,k') + \hbar\omega_\eta\big) + \big(n(\hbar\omega_\eta)+1\big)\delta\big(\varepsilon^{\nu}_\mathrm{tot}(z,k) - \varepsilon^{\nu'}_\mathrm{tot}(z,k') - \hbar\omega_\eta\big)\Big](1-\delta_{v,v'})\,, \tag{2.45}$$

where η indicates the type of transition (f– or g–type) and the type of phonon (acoustic or optical), and $n(\hbar\omega_\eta)$ is the phonon number of a phonon with energy $\hbar\omega_\eta$, which needs to be evaluated without the equipartition approximation since the phonon energies are non–negligible compared to the thermal energy $k_\mathrm{B}T$. The intervalley transitions obviously have to occur between different valleys, which is shown by the last term in (2.45).

For further convenience, the factors in front of the δ–distributions in Eqs. (2.43) and (2.45) are abbreviated to give a general expression,

$$W^{\nu\nu'}_{\eta\sigma}(z;k|k')\Big|_{\substack{\eta=f_i,g_i,\mathrm{intra}\\ \sigma=\pm1,0}} = \frac{1}{L_\mathrm{sys}}c^{\nu\nu'}_{\eta\sigma}(z)\delta\left(\varepsilon^{\nu}_\mathrm{tot}(z,k) - \varepsilon^{\nu'}_\mathrm{tot}(z,k') + \sigma\hbar\omega_\eta\right)\,. \tag{2.46}$$

For silicon η runs over six different types of transitions: f_1, f_2, f_3 as well as g_1, g_2 and g_3. Furthermore, σ has two possible values, i.e. $\sigma = \pm 1$, which alter the sign of the phonon energy $\hbar\omega_\eta$. Hence, the value of σ ultimately decides whether $W_{\eta\sigma}$ is an emission or an absorption rate. The phonon scattering integral (2.42) is then given by

$$\begin{aligned} &S^\nu_{\mathrm{PH}}\{f_k\} = \\ &\sum_{\nu'}\sum_{\eta,\sigma}\int \frac{\mathrm{d}k'}{2\pi}\Big[f_k^{\nu'}(z,k')\big(1-f_k^{\nu}(z,k)\big)c_{\eta\sigma}^{\nu\nu'}(z)\delta\left(\varepsilon^\nu_{\mathrm{tot}}(z,k)-\varepsilon^{\nu'}_{\mathrm{tot}}(z,k')+\sigma\hbar\omega_\eta\right) \\ &-f_k^{\nu}(z,k)\big(1-f_k^{\nu'}(z,k')\big)c_{\eta\sigma}^{\nu\nu'}(z)\delta\left(\varepsilon^\nu_{\mathrm{tot}}(z,k)-\varepsilon^{\nu'}_{\mathrm{tot}}(z,k')-\sigma\hbar\omega_\eta\right)\Big], \end{aligned} \tag{2.47}$$

where the time arguments are dropped from now on for the sake of legibility. Projection of these scattering terms onto the constant-energy space yields,

$$\begin{aligned} &\int \frac{\mathrm{d}k}{2\pi}\Big[S^\nu_{\mathrm{PH}}\{f_k\}\Big]\delta\left(\tilde{\varepsilon}(k)-\varepsilon\right) \\ &=\sum_{\nu',\eta,\sigma}\iint \frac{\mathrm{d}k}{2\pi}\frac{\mathrm{d}k'}{2\pi}\Big[f_k^{\nu'}(z,k')\big(1-f_k^{\nu}(z,k)\big)c_{\eta\sigma}^{\nu\nu'}(z)\delta\left(\varepsilon^\nu_{\mathrm{tot}}(z,k)-\varepsilon^{\nu'}_{\mathrm{tot}}(z,k')+\sigma\hbar\omega_\eta\right) \\ &\quad -f_k^{\nu}(z,k)\big(1-f_k^{\nu'}(z,k')\big)c_{\eta\sigma}^{\nu'\nu}(z)\delta\left(\varepsilon^\nu_{\mathrm{tot}}(z,k)-\varepsilon^{\nu'}_{\mathrm{tot}}(z,k')-\sigma\hbar\omega_\eta\right)\Big]\delta\left(\tilde{\varepsilon}(k)-\varepsilon\right) \\ &=\sum_{\nu',\eta,\sigma}\int \frac{\mathrm{d}k'}{2\pi}Z^\nu(\varepsilon)\Big[f_k^{\nu'}(z,k')\big(1-f_\varepsilon^{\nu}(z,\varepsilon)\big)c_{\eta\sigma}^{\nu\nu'}(z)\delta\left(\varepsilon^\nu_{\mathrm{tot}}(z,\varepsilon)-\varepsilon^{\nu'}_{\mathrm{tot}}(z,k')+\sigma\hbar\omega_\eta\right) \\ &\quad -f_\varepsilon^{\nu}(z,\varepsilon)\big(1-f_k^{\nu'}(z,k')\big)c_{\eta\sigma}^{\nu'\nu}(z)\delta\left(\varepsilon^\nu_{\mathrm{tot}}(z,\varepsilon)-\varepsilon^{\nu'}_{\mathrm{tot}}(z,k')-\sigma\hbar\omega_\eta\right)\Big] \end{aligned} \tag{2.48}$$

and the integral over $\mathrm{d}k'$ can be simplified by considering the δ–function, and projection of (2.48) onto constant energy surfaces followed by an H–transformation gives:

$$\begin{aligned} S^\nu_{\mathrm{PH}}\{f_{\mathrm{H}}\} = \sum_{\nu',\eta,\sigma} &Z^\nu(z,H)Z^\nu(z,H+\sigma\hbar\omega_\eta)\big(1-f^\nu_{\mathrm{H}}(z,H)\big)c^{\nu\nu'}_{\eta\sigma}(z)f^{\nu'}_{\mathrm{H}}(z,H+\sigma\hbar\omega_\eta) \\ &-Z^\nu(z,H)Z^\nu(z,H-\sigma\hbar\omega_\eta)f^\nu_{\mathrm{H}}(z,H)c^{\nu'\nu}_{\eta\sigma}(z)\big(1-f^{\nu'}_{\mathrm{H}}(z,H-\sigma\hbar\omega_\eta)\big). \end{aligned} \tag{2.49}$$

As for the surface roughness (SR) scattering, many publications have focused on modeling this term in nanowires with rectangular cross–sections. In [122] an ideal metallic nanowire with periodic BCs in its transport direction is considered, and SR scattering is modeled in the absence of transverse potential profile. Ref. [123] models the SR scattering in rectangular germanuium nanowires using atomistic simulations, and in [99] the SR scattering for ultrathin–body SOI MOSFETs was studied. Ref. [124] extends [99] to take into account the

modifications of Hamiltonian caused by anisotropic and nondiagonal effective mass tensor for rectangular nanowires with arbitrary orientations. In all of the above mentioned works, the roughness on different faces of the nanowire cross–section are assumed to be uncorrelated and the total scattering rate is calculated as the sum of the scattering rates by each surface. In this work a simplified model for SR scattering is implemented, in which the infinitely high potential barrier at the oxide regions reduces the matrix elements of [99] to the Prange–Nee form [125] and we have

$$\left|M_{\mathrm{SR}}^{\nu\nu'}(z;k|k')\right| = \frac{1}{L_{\mathrm{sys}}}\langle|\Delta q|^2\rangle\left|\Gamma_{\mathrm{PN}}^{\nu\nu'}(z)\right|, \tag{2.50}$$

where the power spectrum of surface roughness is given by the exponential model[6]

$$\langle|\Delta q|^2\rangle = \pi\Delta^2 L^2\left(1+\frac{|q|^2L^2}{2}\right)^{-\gamma}. \tag{2.51}$$

Here, Δ and L are the root mean square height and correlation length of the roughness, respectively. Moreover, $q=k'-k$ denotes the change in wave number. The statistical nature of non-idealities depends on the nanowire fabrication methods, and the parameters Δ and L of (2.51) are usually adjusted to fit the low field mobility of measurements. Furthermore, the Prange–Nee term for an edge perpendicular to the x–direction is given by

$$\Gamma_{\mathrm{PN}}^{\nu\nu'}(z) = \frac{\hbar^2}{2m_{xx}^{\nu}}\frac{\partial\psi^{\nu}}{\partial x}\frac{\partial\psi^{\nu'}}{\partial x}\Big|_{x=x_0}, \tag{2.52}$$

where x_0 denotes the position of the edge. The total transition rate for the SR scattering on one surface can be expressed with Fermi's Golden Rule as

$$\begin{aligned} W_{\mathrm{SR}}^{\nu\nu'}(z;k|k') &= \delta_{\nu\nu'}\frac{2\pi}{\hbar}\left|M_{\mathrm{SR}}^{\nu\nu'}(z;k|k')\right|^2\delta\big(\varepsilon_{\mathrm{tot}}^{\nu}(z,k)-\varepsilon_{\mathrm{tot}}^{\nu'}(z,k')\big) \\ &= \delta_{\nu\nu'}\frac{\pi\hbar^3}{2(m_{xx}^{\nu})^2L_{\mathrm{sys}}}\langle|\Delta q|^2\rangle\left|\frac{\partial\psi^{\nu}}{\partial x}\frac{\partial\psi^{\nu'}}{\partial x}\right|^2\delta\big(\varepsilon_{\mathrm{tot}}^{\nu}(z,k)-\varepsilon_{\mathrm{tot}}^{\nu'}(z,k')\big) \end{aligned} \tag{2.53}$$

where the scattering is considered to be elastic and intra–valley. The SR scattering at different interfaces is assumed to be uncorrelated with no edge effects [127], and the respective scattering rates are additive. Although the SR scattering is very serious in the conventional planar devices, narrow nanowire FETs have limited number of conducting modes due to strong quantum confinement and the surface roughness is not important in these structures [128–131]. Ref. [132]

[6] According to [126], the exponential power spectrum better captures the underlying physics compared to the Gaussian one.

confirms that for the doping concentrations and cross–sectional areas chosen in this work, the SR scattering is 10^3 times weaker compared to electron–phonon interactions. Therefore, the SR scattering is neglected in Chapter 5 for the sake of computational cost although it is implemented in our solver.

Finally, the boundary conditions for the source and drain contacts need to be carefully specified. While the thermal equilibrium assumption is reasonable at the inflow contacts, it leads to boundary layers at the outflow contacts at higher bias, forcing a heated electron distribution function to thermal equilibrium. This deficiency is addressed by a generation/recombination process

$$\Gamma^{\nu}_{\mathrm{GR}}\{f_k\} = v_{\mathrm{GR}}\left(f^{\nu}_k(z,k) - f^{\nu}_{k\,\mathrm{eq}}(z,k)\right)\left(\delta(z-z_{\mathrm{S}}) + \delta(z-z_{\mathrm{D}})\right) \tag{2.54}$$

$$\Rightarrow \Gamma^{\nu}_{\mathrm{GR}}\{f_{\mathrm{H}}\} = v_{\mathrm{GR}} Z^{\nu}(z,H)\left(f^{\nu}_{\mathrm{H}}(z,H) - f^{\nu}_{\mathrm{H\,eq}}(z,H)\right)\left(\delta(z-z_{\mathrm{S}}) + \delta(z-z_{\mathrm{D}})\right), \tag{2.55}$$

where v_{GR} is the recombination velocity, which provides control over the difference between thermal equilibrium and the computed solution. In the limit $v_{\mathrm{GR}} \to \infty$, the source and drain distribution functions are fixed to the equilibrium value. An improvement to (2.55) is made by assuming thermal equilibrium only for the electrons entering the device. These "thermal bath" BCs are given by

$$\begin{aligned}\Gamma^{\nu}_{\mathrm{TH}}\{f_k\} =& \left[f^{\nu}_k(z,k)\Theta(-k) + f_{k\,\mathrm{eq}}(z,k)\Theta(k)\right] v^{\nu}_z(k)\delta(z-z_{\mathrm{S}}) \\ &- \left[f^{\nu}_k(z,k)\Theta(k) + f_{k\,\mathrm{eq}}(z,k)\Theta(-k)\right] v^{\nu}_z(k)\delta(z-z_{\mathrm{D}})\,.\end{aligned} \tag{2.56}$$

where $\Theta(x)$ is the Heaviside step function. Projection of $\Gamma^{\nu}_{\mathrm{TH}}\{f_k\}$ for the source contact onto the equi-energy surfaces gives:

$$\begin{aligned}&\int_{-\infty}^{\infty}\frac{\mathrm{d}k}{2\pi}\Big[f^{\nu}_k(z,k)\Theta(-k) + f^{\nu}_{k\,\mathrm{eq}}(z,k)\Theta(k)\Big] v^{\nu}_z(k)\delta\left(\tilde{\varepsilon}(k)-\varepsilon\right) \\ &= \frac{1}{2\pi\hbar}\Big[-f^{\nu}_k\big(z,-\frac{\sqrt{2m^{\nu}_{zz}\varepsilon}}{\hbar}\big) + f^{\nu}_{k\,\mathrm{eq}}\big(z,\frac{\sqrt{2m^{\nu}_{zz}\varepsilon}}{\hbar}\big)\Big] \\ &= \frac{1}{2\pi\hbar}\Big[-f^{\nu\mathrm{e}}_k\big(z,\frac{\sqrt{2m^{\nu}_{zz}\varepsilon}}{\hbar}\big) + f^{\nu\mathrm{o}}_k\big(z,\frac{\sqrt{2m^{\nu}_{zz}\varepsilon}}{\hbar}\big) + f^{\nu}_{k\,\mathrm{eq}}\big(z,\frac{\sqrt{2m^{\nu}_{zz}\varepsilon}}{\hbar}\big)\Big] \\ &= \frac{1}{2\pi\hbar}\Big[-f^{\nu\mathrm{e}}_{\varepsilon}(z,\varepsilon) + f^{\nu\mathrm{o}}_{\varepsilon}(z,\varepsilon) + f^{\nu}_{\varepsilon\,\mathrm{eq}}(z,\varepsilon)\Big],\end{aligned} \tag{2.57}$$

and H–transformation results in

$$\Gamma^{\nu}_{\mathrm{TH}}\{f\}\Big|_{\mathrm{source}} = \frac{1}{2\pi\hbar}\Big[-f^{\nu\mathrm{e}}_{\mathrm{H}}(z,H) + f^{\nu\mathrm{o}}_{\mathrm{H}}(z,H) + f^{\nu}_{\mathrm{H\,eq}}(z,H)\Big]\delta(z-z_{\mathrm{S}})\,. \tag{2.58}$$

Similarly, thermal BCs for the drain contact yields:

$$\Gamma^{\nu}_{\mathrm{TH}}\{f\}\Big|_{\mathrm{drain}} = \frac{1}{2\pi\hbar}\Big[-f^{\nu\mathrm{e}}_{\mathrm{H}}(z,H) - f^{\nu\mathrm{o}}_{\mathrm{H}}(z,H) + f^{\nu}_{\mathrm{H\,eq}}(z,H)\Big]\delta(z-z_{\mathrm{D}})\,. \tag{2.59}$$

2.4.2 Numerical Implementation

The discretization scheme is based on the entropy principle and box integration method, as developed in [36]. First, *direct* and *adjoint* spatial grids are defined in the transport direction, with densities represented on the direct grid nodes and fluxes assigned to the adjoint grid nodes. The spatial and energy direct grid nodes form a Cartesian tensor product grid (z_i, H_j) with n_z grid nodes in z direction and n_{H} in total energy. The adjoint spatial grid is given by[7]

$$z_{i\pm½} = \frac{z_i + z_{i\pm1}}{2} \qquad \text{for } i \in \{1, \cdots, n_z\}\,, \tag{2.60}$$

on which the electrostatic potential and the subband energies are linearly interpolated, i.e.

$$\varepsilon^{\nu}_{\mathrm{sub}}(z_{i\pm½}) = \frac{\varepsilon^{\nu}_{\mathrm{sub}}(z_{i\pm1}) + \varepsilon^{\nu}_{\mathrm{sub}}(z_i)}{2}\,. \tag{2.61}$$

For the other quantities such as electrons' group velocity and density of states, a constant interpolation scheme is used where the quantities are defined on the grid points and considered constant in the corresponding boxes.

Fig. 2.7 shows a schematic illustration of the grid, where some subband energy $\varepsilon^{\nu}_{\mathrm{sub}}(z)$ is drawn. The H–grid is static, meaning it is constant throughout the whole simulation and does not change between different iterations. Choosing suitable $H_{\min}$ and $H_{\max}$ values, this grid is defined globally for a given ΔH. For the box integration, the boxes around the direct and adjoint grid nodes in transport direction are given by,

$$\begin{aligned} \Delta z_i &= z_{i+½} - z_{i-½}\,, \qquad i \in \{1, \cdots, n_z\} \\ \Delta z_{i+½} &= z_{i+1} - z_i\,, \qquad i \in \{1, \cdots, n_z\}\,. \end{aligned} \tag{2.62}$$

As for the boxes of the H–grid, the support of H–space integration is limited by the subband energy and this must be reflected in the definition of boxes. Hence, the H–boxes are given by

$$\Delta H^{\nu}_j(z) = \begin{cases} 0 & \text{if } H_{j+½} \le \varepsilon^{\nu}_{\mathrm{sub}}(z) \\ H_{j+½} - \varepsilon^{\nu}_{\mathrm{sub}}(z) & \text{if } H_{j-½} < \varepsilon^{\nu}_{\mathrm{sub}}(z) < H_{j+½} \\ H_{j+½} - H_{j-½} & \text{if } \varepsilon^{\nu}_{\mathrm{sub}}(z) \le H_{j-½} \end{cases} \tag{2.63}$$

and the intermediary H–grid points are added similar to the adjoint spatial points. The H–grid is truncated at high energies, where the contribution of the distribution function to the observables is well below the numerical precision of the computations.

[7] Obviously, the notation of (2.60) gives duplicate definitions since for each grid node we have, for example, $z_{i+½} = z_{(i+1)-½}$. However, this redundancy does not introduce numerical issues.

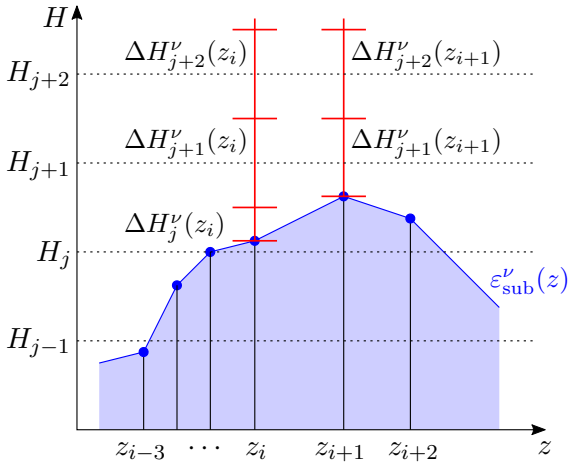

Figure 2.7: Schematic representation of the tensor product grid used in the discretization of the BE. The grid is non-equidistant in transport direction, but equidistant in energy. Only the energy boxes above the subband energy are included in the simulation domain, and the lowest box is truncated by the subband energy.

In this section, we only present the discretization of the BE in steady state conditions, i.e. constant applied bias and time independent generation/recombination mechanisms. In this case, the time derivative will vanish. The structure of the free streaming term is chosen based on [36], which gives a numerically sound representation and reproduces the desirable entropy properties[8]. For the direct grid nodes, we have:

$$\begin{aligned}\tilde{\mathrm{L}}\{f^\nu_{\mathrm{H}}(z_i,H_j)\} &:= \int_{\Delta H^\nu_j(z)} \mathrm{d}H \int_{z_{i-1/2}}^{z_{i+1/2}} \mathrm{d}z\, \mathrm{L}\{f^{\nu\mathrm{e}}_{\mathrm{H}}(z,H)\} \\ &= \int_{\Delta H^\nu_j(z_i)} \mathrm{d}H \int_{z_{i-1/2}}^{z_{i+1/2}} \mathrm{d}z\, \frac{\partial}{\partial z}\Big(Z^\nu(z,H)v^\nu_z(z,H)f^{\nu\mathrm{o}}_{\mathrm{H}}(z,H)\Big) \\ &= \Big[\int_{H^\nu_{j\min}(z_{i+1/2})}^{H^\nu_{j\max}(z_{i+1/2})} \mathrm{d}H\, Z^\nu(z_{i+1/2},H)v^\nu_z(z_{i+1/2},H)\Big] f^\nu_{\mathrm{H}}(z_{i+1/2},H_j) \\ &\quad - \Big[\int_{H^\nu_{j\min}(z_{i-1/2})}^{H^\nu_{j\max}(z_{i-1/2})} \mathrm{d}H\, Z^\nu(z_{i-1/2},H)v^\nu_z(z_{i-1/2},H)\Big] f^\nu_{\mathrm{H}}(z_{i-1/2},H_j)\end{aligned} \tag{2.64}$$

where the integrations over energy have to be performed analytically and special care needs to be taken with the integration boundaries. Since the free streaming term only describes the movement of carriers along constant total energy lines and is not necessarily aligned with the spatial changes of subband energy, certain

[8] A detailed treatment of this subject is beyond the scope of our work and the reader can refer to [36, 37, 84, 133–135] for more details.

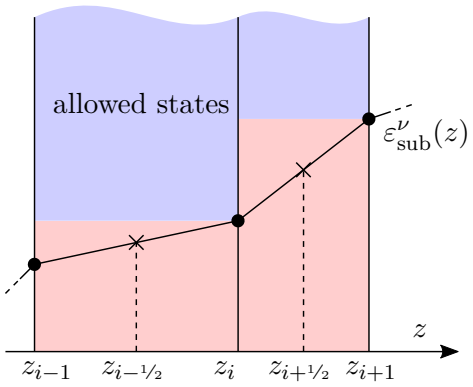

Figure 2.8: Allowed (blue) and forbidden (red) energies for transitions in the discretized free streaming term.

energies might be included in the simulation domain at some position z_i but be forbidden states at $z_{i\pm1}$ since $H < \varepsilon^{\nu}_{\mathrm{sub}}(z_{i\pm1})$. These forbidden transitions have to be carefully considered in our integrations, i.e. we need to exclude all transitions to energies lying below the subband energy profile (see Fig. 2.8). Hence, the lower and upper integral limits in (2.64) are calculated as,

$$H^{\nu}_{j\mathrm{min}}(z_{i\pm1/2}) = \begin{cases} H_{j-1/2} & \text{if } \max\{\varepsilon^{\nu}_{\mathrm{sub}}(z_i), \varepsilon^{\nu}_{\mathrm{sub}}(z_{i\pm1})\} \leq H_{j-1/2} \\ \max\{\varepsilon^{\nu}_{\mathrm{sub}}(z_i), \varepsilon^{\nu}_{\mathrm{sub}}(z_{i\pm1})\} & \text{else} \end{cases}$$

$$H^{\nu}_{j\mathrm{max}}(z_{i\pm1/2}) = \begin{cases} H_{j+1/2} & \text{if } \max\{\varepsilon^{\nu}_{\mathrm{sub}}(z_i), \varepsilon^{\nu}_{\mathrm{sub}}(z_{i\pm1})\} \leq H_{j+1/2} \\ \max\{\varepsilon^{\nu}_{\mathrm{sub}}(z_i), \varepsilon^{\nu}_{\mathrm{sub}}(z_{i\pm1})\} & \text{else} \end{cases} \tag{2.65}$$

Regarding the discretization of free streaming term on the adjoint grid nodes, it must be noted that the same formulation as in (2.64) does not preserve the entropy for non–interacting particles, and therefore cannot be used. A discretization in accordance with the maximum entropy dissipation scheme is given by [36],

$$\begin{aligned} \tilde{\mathrm{L}}\{f^{\nu}_{\mathrm{H}}(z_{i+1/2}, H_j)\} &:= \int_{\Delta H^{\nu}_j(z_{i+1/2})} \mathrm{d}H \int_{z_i}^{z_{i+1}} \mathrm{d}z\, \mathrm{L}\{f^{\nu\mathrm{o}}_{\mathrm{H}}(z, H)\} \qquad (2.66) \\ &= \int_{\Delta H^{\nu}_j(z_{i+1/2})} \mathrm{d}H \int_{z_i}^{z_{i+1}} \mathrm{d}z\, \frac{\partial}{\partial z}\Big(Z^{\nu}(z,H) v^{\nu}_z(z,H) f^{\nu\mathrm{e}}_{\mathrm{H}}(z,H)\Big) \\ &= \Big[\int_{H^{\nu}_{j\mathrm{min}}(z_{i+1/2})}^{H^{\nu}_{j\mathrm{max}}(z_{i+1/2})} \mathrm{d}H\, Z^{\nu}(z_{i+1/2}, H) v^{\nu}_z(z_{i+1/2}, H)\Big] f^{\nu}_{\mathrm{H}}(z_{i+1}, H_j) \\ &\quad - \Big[\int_{H^{\nu}_{j\mathrm{min}}(z_{i+1/2})}^{H^{\nu}_{j\mathrm{max}}(z_{i+1/2})} \mathrm{d}H\, Z^{\nu}(z_{i+1/2}, H) v^{\nu}_z(z_{i+1/2}, H)\Big] f^{\nu}_{\mathrm{H}}(z_i, H_j)\,. \end{aligned}$$

The scattering term is local in the spatial direction and its discretization is straightforward. Under the RTA, $S\{f\}$ can be split into the relaxation of $f^{\nu\mathrm{e}}$ on the direct grid nodes towards the equilibrium distribution function, and relaxation of $f^{\nu\mathrm{o}}$ on the adjoint grid nodes towards zero. That is,

$$\begin{aligned}\tilde{S}_{\mathrm{RTA}}\{f^{\nu}_{\mathrm{H}}(z_i,H_j)\} &= \int_{\Delta H^{\nu}_{j}(z)} \mathrm{d}H \int_{z_{i-1/2}}^{z_{i+1/2}} \mathrm{d}z\, S_{\mathrm{RTA}}\{f^{\nu\mathrm{e}}_{\mathrm{H}}(z,H)\} \qquad (2.67)\\ &= \Delta z_i \Big[\int_{H^{\nu}_{j\min}(z_i)}^{H^{\nu}_{j\max}(z_i)} \mathrm{d}H\, Z^{\nu}(z_i,H)\Big]\left(f^{\nu}_{\mathrm{H}}(z_i,H_j) - f^{\nu}_{\mathrm{H\,eq}}(z_i,H_j)\right),\\ \tilde{S}_{\mathrm{RTA}}\{f^{\nu}_{\mathrm{H}}(z_{i+1/2},H_j)\} &= \int_{\Delta H^{\nu}_{j}(z)} \mathrm{d}H \int_{z_i}^{z_{i+1}} \mathrm{d}z\, S_{\mathrm{RTA}}\{f^{\nu\mathrm{o}}_{\mathrm{H}}(z,H)\} \qquad (2.68)\\ &= \Delta z_{i+1/2} \Big[\int_{H^{\nu}_{j\min}(z_{i+1/2})}^{H^{\nu}_{j\max}(z_{i+1/2})} \mathrm{d}H\, Z^{\nu}(z_{i+1/2},H)\Big] f^{\nu}_{\mathrm{H}}(z_{i+1/2},H_j)\,.\end{aligned}$$

The treatment of the scattering integral (2.42) is more involved. On the direct grid nodes we have

$$\begin{aligned}\tilde{S}_{\mathrm{PH}}\{f^{\nu}_{\mathrm{H}}(z_i,H_j)\} = &\sum_{\nu'}\sum_{\eta,\sigma} \Delta z_i \Big[\int_{H^{\nu}_{j\min}(z_i)}^{H^{\nu}_{j\max}(z_i)} \mathrm{d}H\, Z^{\nu}(z_i,H_j) Z^{\nu'}(z_i,H_j+\sigma\hbar\omega_\eta)\Big]\\ &\times (1-f^{\nu}_{\mathrm{H}}(z_i,H_j))\, c^{\nu\nu'}_{\eta\sigma}(z_i) f^{\nu'}_{\mathrm{H}}(z_i,H_j+\sigma\hbar\omega_\eta)\\ &- \sum_{\nu'}\sum_{\eta,\sigma} \Delta z_i \Big[\int_{H^{\nu}_{j\min}(z_i)}^{H^{\nu}_{j\max}(z_i)} \mathrm{d}H\, Z^{\nu}(z_i,H_j) Z^{\nu'}(z_i,H_j-\sigma\hbar\omega_\eta)\Big]\\ &\times f^{\nu}_{\mathrm{H}}(z_i,H_j) c^{\nu'\nu}_{\eta\sigma}(z_i)\left(1-f^{\nu'}_{\mathrm{H}}(z_i,H_j-\sigma\hbar\omega_\eta)\right)\end{aligned} \qquad (2.69)$$

and on the adjoint grid nodes the in–scattering term vanishes. It must be noted that for the detailed balance condition to still hold for the discretized scattering integral, the phonon energy appearing in $n(\hbar\omega_\eta)$ and the phonon energy appearing in $H_j \pm \hbar\omega_\eta$ must exactly cancel. This can only occur for energies $H_j \pm \hbar\omega_\eta$ which lie exactly on the H–grid. Hence, the phonon energies $\hbar\omega_\eta$ need to be multiples of the grid spacing for an equidistant H–grid.

The boundary generation/recombination term of Eq. (2.55) only exists on the first and last direct grid nodes due to the δ–distributions. Therefore we find

$$\begin{aligned}&\tilde{\Gamma}_{\mathrm{GR}}\{f^{\nu}_{\mathrm{H}}(z_i,H_j)\} \qquad (2.70)\\ &= v_{\mathrm{GR}} \left[\int_{H^{\nu}_{j\min}(z_i)}^{H^{\nu}_{j\max}(z_i)} \mathrm{d}H\, Z^{\nu}(z_i,H)\right]\Big(f^{\nu}_{\mathrm{H}}(z_i,H_j) - f^{\nu}_{\mathrm{H\,eq}}(z_i,H_j)\Big)(\delta_{i,1}+\delta_{i,n_z})\,.\end{aligned}$$

which transforms the singular function of Eq. (2.55) to a volume generation and recombination rate. For the adjoint grid nodes, we have:

$$\tilde{\Gamma}_{\mathrm{GR}}\{f^{\nu}_{\mathrm{H}}(z_{i+½}, H_j)\} = 0\,. \tag{2.71}$$

Finally, discretization of the thermal bath BCs on the source contact yields

$$\tilde{\Gamma}^{\nu}_{\mathrm{TH}}\{f\}\Big|_{\mathrm{source}} = \frac{1}{2\pi\hbar}\Big[-f^{\nu}_{\mathrm{H}}(z_i, H_j) + f^{\nu}_{\mathrm{H}}(z_{i+½}, H_j) + f^{\nu}_{\mathrm{H\,eq}}(z_i, H_j)\Big]\delta_{i,1}\,, \tag{2.72}$$

and on the drain contact we have

$$\tilde{\Gamma}^{\nu}_{\mathrm{TH}}\{f\}\Big|_{\mathrm{drain}} = \frac{1}{2\pi\hbar}\Big[-f^{\nu}_{\mathrm{H}}(z_i, H_j) - f^{\nu}_{\mathrm{H}}(z_{i-½}, H_j) + f^{\nu}_{\mathrm{H\,eq}}(z_i, H_j)\Big]\delta_{i,n_z}\,. \tag{2.73}$$

2.5 System of Equations

The PE/SE/BE system of equations is highly nonlinear and is solved with the Newton-Raphson approach which is beneficial in several ways. Firstly, the quadratic convergence of a full Newton-Raphson solver makes it possible to reach errors as low as the machine precision (e.g. $10^{-16}\,\mathrm{V}$ in electric potential) with just a few iterations. In Fig. 2.9, the convergence behavior of the Gummel iteration[9] and full Newton-Raphson approches is shown for $V_{\mathrm{GS}} = 0.7\,\mathrm{V}$ and $V_{\mathrm{DS}} = 0.5\,\mathrm{V}$. While the Gummel type iteration converges linearly and takes 35 iterations to get to the accuracy of $10^{-9}\,\mathrm{V}$, the full Newton-Raphson solver demonstrates rapid convergence with just a few iterations. Moreover, the Newton-Raphson approach provides the necessary setup for small signal and noise analyses, as explained in Sections 2.6 and 2.7. In this section, a brief explanation of the Newton-Raphson method for our system of equations is presented.

We start with the formal description for an N–dimensional multivariate equation system:

$$\mathbf{F}_{\mathrm{sys}}(\boldsymbol{x}) = \mathbf{0}\,, \tag{2.74}$$

where $\mathbf{F}_{\mathrm{sys}}$ defines a vector function of $\boldsymbol{x}$. The Newton-Raphson iteration scheme seeks a solution $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \delta\boldsymbol{x}^{(k)}$ at the $(k+1)$th iteration step, where the correction term $\delta\boldsymbol{x}^{(k)}$ is given by the linear approximation

$$\hat{J}^{(k)}\delta\boldsymbol{x}^{(k)} = -\mathbf{F}_{\mathrm{sys}}(\boldsymbol{x}^{(k)}) \tag{2.75}$$

and the elements of the Jacobian matrix are calculated as

$$\big[\hat{J}^{(k)}\big]_{ij} = \frac{\partial \mathrm{F}_i(\boldsymbol{x}^{(k)})}{\partial x_j^{(k)}}\,. \tag{2.76}$$

[9] In the Gummel iteration approach, the equations are solved successively (as shown in Fig. 2.2) until the maximum change in the electrostatic potential becomes lower than a given limit.

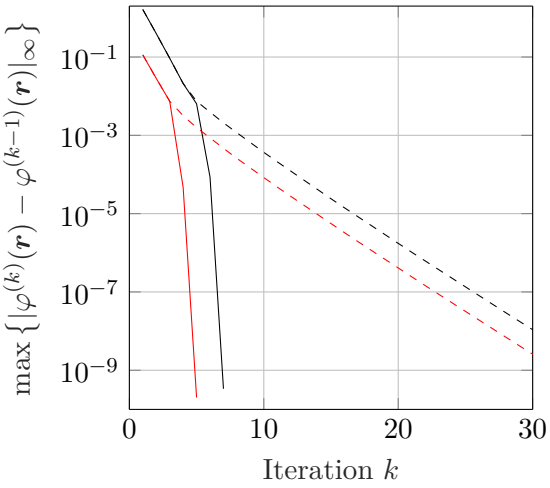

Figure 2.9: Convergence given by the maximum in absolute change of the potential at each iteration step for $V_{\mathrm{DS}} = 0.5\,\mathrm{V}$ (black) and $V_{\mathrm{DS}} = 0.05\,\mathrm{V}$ (red) of the Newton-Raphson approach (solid lines) and the Gummel type iteration (dashed lines) for $V_{\mathrm{GS}} = 0.7\,\mathrm{V}$. The threshold where the Newton-Raphson approach is turned on is $10^{-2}\,\mathrm{V}$.

In this work, $\mathbf{F}_{\mathrm{sys}}$ is a vector function of rank two which itself consists of the vector functions $\mathbf{F}_{\mathrm{PE}}$ and $\mathbf{F}_{\mathrm{BE}}$, corresponding to the discrete approximations for the PE and BE, respectively. The vector of unknowns $\boldsymbol{x}$ is also comprised by two vectors which are formed by the values of electrostatic potential and distribution function at discrete points of the simulation geometry. That is,

$$\begin{pmatrix} \mathbf{F}_{\mathrm{PE}}(\boldsymbol{\varphi}, \boldsymbol{f}) \\ \mathbf{F}_{\mathrm{BE}}(\boldsymbol{\varphi}, \boldsymbol{f}) \end{pmatrix} = \mathbf{0} \tag{2.77}$$

The solution, which can be interpreted as the intersection of $\mathbf{F}_{\mathrm{PE}} = \mathbf{0}$ and $\mathbf{F}_{\mathrm{BE}} = \mathbf{0}$ planes, is then obtained by the following linearization of $\mathbf{F}_{\mathrm{PE}}$ and $\mathbf{F}_{\mathrm{BE}}$ in terms of the primary variables (φ, f):

$$\sum_{b \in \tilde{\mathrm{D}}_{\mathrm{PE}}} \frac{\partial \mathrm{F}_{\mathrm{PE}}(a)}{\partial \varphi(b)} \delta\varphi(b) + \sum_{\beta \in \tilde{\mathrm{D}}_{\mathrm{BE}}} \frac{\partial \mathrm{F}_{\mathrm{PE}}(a)}{\partial f(\beta)} \delta f(\beta) = -\mathrm{F}_{\mathrm{PE}}(a)\,, \tag{2.78}$$

$$\sum_{b \in \tilde{\mathrm{D}}_{\mathrm{PE}}} \frac{\partial \mathrm{F}_{\mathrm{BE}}(\alpha)}{\partial \varphi(b)} \delta\varphi(b) + \sum_{\beta \in \tilde{\mathrm{D}}_{\mathrm{BE}}} \frac{\partial \mathrm{F}_{\mathrm{BE}}(\alpha)}{\partial f(\beta)} \delta f(\beta) = -\mathrm{F}_{\mathrm{BE}}(\alpha)\,, \tag{2.79}$$

for every $a \in \tilde{\mathrm{D}}_{\mathrm{PE}}$ and $\alpha \in \tilde{\mathrm{D}}_{\mathrm{BE}}$. In the above expressions and throughout this work, D_{PE} and D_{BE} stand for aggregate indices covering the domain of definition for the PE and BE, respectively. For example, by $a \in \mathrm{D}_{\mathrm{PE}}$ we mean:

$$a \in \left\{(x, y, z)\,|\,0 \le x \le L_x,\, 0 \le y \le L_y,\, 0 \le z \le L_z\right\} \tag{2.80}$$

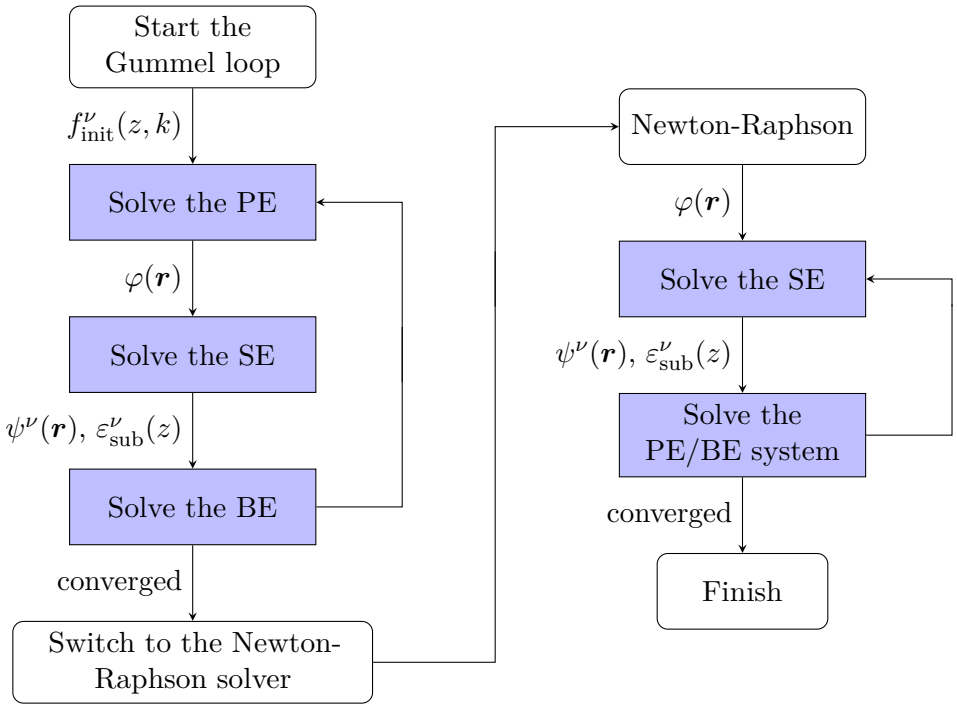

Figure 2.10: Newton-Raphson approach for solving the PE, SE, and BE. After the Gummel iterations have converged to, say, $10^{-2}\,\text{V}$, we switch to the Newton-Raphson solver and update the subband energies ε_{sub} and wave functions ψ. Thereafter we solve the combined system of PE and BE and return to the SE, until our solution has sufficiently converged.

and similarly, in the discretized space $a \in \tilde{\text{D}}_{\text{PE}}$ represents

$$a \in \left\{(x_i, y_j, z_k)\,|\,i \in \{1,\cdots,n_x\},\, j \in \{1,\cdots,n_y\},\, k \in \{1,\cdots,n_z\}\right\}. \tag{2.81}$$

Analogously, D_{BE} runs over the spatial, energy and time domains, as well as different valleys and subbands in the BE.

As mentioned in Section 2.3, the SE is an eigenvalue problem and cannot be directly cast into the Newton-Raphson matrix equation. Moreover, the BE does not explicitly depend on the potential but only on the wave functions and subband energies resulting from the SE. Using the first order time-independent perturbation theory, changes in subband energies and wave functions can be expressed in terms of small perturbations in the potential and the first terms in (2.78) and (2.79) can be reformulated as

$$\frac{\partial \text{F}_{\text{PE}}(a)}{\partial \varphi(b)} = \frac{\partial \text{F}^{(\Delta)}_{\text{PE}}(a)}{\partial \varphi(b)} + \left(\sum_k \frac{\partial \text{F}^{(\rho)}_{\text{PE}}(a)}{\partial \varepsilon_{\text{sub}}(k)} \frac{\partial \varepsilon_{\text{sub}}(k)}{\partial \varphi(b)} + \sum_l \frac{\partial \text{F}^{(\rho)}_{\text{PE}}(a)}{\partial \psi(l)} \frac{\partial \psi(l)}{\partial \varphi(b)}\right) \tag{2.82}$$

$$\frac{\partial \mathrm{F}_{\mathrm{BE}}(\alpha)}{\partial \varphi(b)} = \sum_k \frac{\partial \mathrm{F}_{\mathrm{BE}}(\alpha)}{\partial \varepsilon_{\mathrm{sub}}(k)} \frac{\partial \varepsilon_{\mathrm{sub}}(k)}{\partial \varphi(b)} + \sum_l \frac{\partial \mathrm{F}_{\mathrm{BE}}(\alpha)}{\partial \psi(l)} \frac{\partial \psi(l)}{\partial \varphi(b)} \,. \tag{2.83}$$

where the terms $\mathrm{F}_{\mathrm{PE}}^{(\Delta)}$ and $\mathrm{F}_{\mathrm{PE}}^{(\rho)}$ represent the Laplacian and the charge density part of the PE, respectively. The indices k and l run over the corresponding domain of definition for $\varepsilon_{\mathrm{sub}}$ and ψ. Note that the linearization of the PE and BE needs to be performed with utmost care, having in mind that the lowest H–box strongly depends on the electrostatic potential. Such complications would not arise in the case of kinetic-energy-based schemes.

The iteration process as it stands can be too harsh in the sense that consecutive values of $\boldsymbol{x}^{(k)}$ can oscillate wildly about the exact solution. In order to avoid this overshoot, the solution updates can be damped like

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \alpha(\hat{\boldsymbol{J}}^{(k)})^{-1}\mathbf{F}(\boldsymbol{x}^{(k)}) \,, \tag{2.84}$$

where $0 < \alpha \leq 1$ is some pre-assigned value which can be taken small in the fragile initial stages of the iteration and is increased as the iteration progresses. Another approach (shown in Fig. 2.10) is to start the simulations with Gummel-type iterations to provide a better initial guess for the Newton-Raphson iterations and switch to the Newton-Raphson solver when the error in the potential is less than, say, 10 meV.

The equation system is solved using ILUPACK [136], which is a fast, memory efficient, and reliable sparse linear system solver based on an incomplete LU–decomposition [137].

2.6 Small Signal Analysis

Ascertainment of the small signal behavior of transistors is a vital component of semiconductor device simulation. Unlike the MC method, exact small signal analysis can be performed in the presented deterministic framework. Among the various techniques (e.g. transient excitations followed by Fourier decomposition [138], incremental charge partitioning [138], equivalent circuits [139, 140], etc.) sinusoidal steady state analysis is generally considered as the superior approach since it is directly applied in the frequency domain and provides accurate, rigorously correct results for reasonable computational cost. In this section, the numerical framework for small signal analysis of the given device under sinusoidal steady state condition (SSSC) is presented. Starting with the PE and BE,

$$\mathrm{F}_{\mathrm{PE}} := \nabla_{\boldsymbol{r}} \cdot \big(\epsilon(\boldsymbol{r})\nabla_{\boldsymbol{r}}\varphi(\boldsymbol{r},t)\big) - q\big[n(\boldsymbol{r},t) - N_{\mathrm{D}}(\boldsymbol{r})\big] = 0 \tag{2.85}$$

$$\mathrm{F}_{\mathrm{BE}} := \frac{\partial f_k^\nu}{\partial t} + v_z^\nu(k)\frac{\partial f_k^\nu}{\partial z} + \frac{1}{\hbar}F^\nu(z,t)\frac{\partial f_k^\nu}{\partial k} - S^\nu\{f_k\} - \Gamma^\nu\{f_k\} = 0 \tag{2.86}$$

all variables of interest can be split into a large signal part, which is time independent, and a small signal part which depends on time:

$$\varphi(\boldsymbol{r},t) = \varphi_{\mathrm{DC}}(\boldsymbol{r}) + \delta\varphi(\boldsymbol{r},t)\,, \tag{2.87}$$

$$f^\nu(z,k,t) = f^\nu_{\mathrm{DC}}(z,k) + \delta f^\nu(z,k,t)\,. \tag{2.88}$$

The DC subscript shows that the large signal parts are solutions of the steady state model, for which the small signal quantities are zero. Starting with the PE and ignoring the BCs for now, substitution of (2.87) into (2.85) gives:

$$\begin{aligned}\mathrm{F}_{\mathrm{PE}} :=& \nabla_{\boldsymbol{r}} \cdot \Big[\epsilon(\boldsymbol{r})\nabla_{\boldsymbol{r}}\big(\varphi_{\mathrm{DC}}(\boldsymbol{r}) + \delta\varphi(\boldsymbol{r},t)\big)\Big] - q\Big[\big(n_{\mathrm{DC}}(\boldsymbol{r}) + \delta n(\boldsymbol{r},t)\big) - N_{\mathrm{D}}(\boldsymbol{r})\Big] \\ =& \nabla_{\boldsymbol{r}} \cdot \big[\epsilon(\boldsymbol{r})\nabla_{\boldsymbol{r}}\varphi_{\mathrm{DC}}(\boldsymbol{r})\big] - q\big[n_{\mathrm{DC}}(\boldsymbol{r}) - N_{\mathrm{D}}(\boldsymbol{r})\big] \\ &+ \nabla_{\boldsymbol{r}} \cdot \big[\epsilon(\boldsymbol{r})\nabla_{\boldsymbol{r}}\delta\varphi(\boldsymbol{r},t)\big] - q\delta n(\boldsymbol{r},t) = 0\,,\end{aligned} \tag{2.89}$$

and we get

$$\delta\mathrm{F}_{\mathrm{PE}} := \nabla_{\boldsymbol{r}} \cdot \big[\epsilon(\boldsymbol{r})\nabla_{\boldsymbol{r}}\delta\varphi(\boldsymbol{r},t)\big] - q\delta n(\boldsymbol{r},t) = 0\,. \tag{2.90}$$

Since the small signal variables of interest are $\delta\varphi$ and δf, the small signal electron density $\delta n(\boldsymbol{r},t)$ needs to be expressed in terms of these two variables. After the H–transformation and discretization via the box integration method, the linear coefficients are obviously the same values which were calculated when constructing the Jacobian of the Newton-Raphson approach.

Regarding the small signal BCs, Dirichlet boundary conditions are imposed on the contacts, with the AC potential normalized to the value one. We thus obtain the following discretized presentation in SSSC:

$$\underline{\mathrm{F}}_{\mathrm{PE}}(a) := \sum_{b\in\mathrm{D}_{\mathrm{PE}}} \frac{\partial \mathrm{F}_{\mathrm{PE}}(a)}{\partial\varphi(b)}\underline{\varphi}(b) + \sum_{\beta\in\mathrm{D}_{\mathrm{BE}}} \frac{\partial \mathrm{F}_{\mathrm{PE}}(a)}{\partial f(\beta)}\underline{f}(\beta) = \Gamma_{\mathrm{PE}}(a)\,, \tag{2.91}$$

where $\Gamma_{\mathrm{PE}}(a)$ specifies the AC bias on the contact while being zero elsewhere. From now on, $\underline{x}$ denotes the phasor of the small sinusoidal perturbation

$$x(t) = x_{\mathrm{DC}} + \mathrm{Re}\{\underline{x}\mathrm{e}^{\mathrm{i}\omega t}\} \tag{2.92}$$

and ω is the angular frequency. Substitution of (2.88) into the time dependent BE (2.86), followed by projection onto constant energy surfaces and H−transformation yields:

$$\begin{aligned}\mathrm{F}_{\mathrm{BE}} :=& \frac{\partial}{\partial t}\left(Z^\nu(z,H)\delta f^\nu(z,H,t)\right) + \frac{\partial}{\partial z}\left(v_z^\nu(z,H)Z^\nu(z,H)f_{\mathrm{DC}}^\nu(z,H)\right)\\ &+\frac{\partial}{\partial z}\left(v_z^\nu(z,H)Z^\nu(z,H)\delta f^\nu(z,H,t)\right) - S^\nu\{f_{\mathrm{DC}}+\delta f\} - \Gamma^\nu\{f_{\mathrm{DC}}+\delta f\} = 0\,.\end{aligned} \tag{2.93}$$

After cancelling out the terms which satisfy the stationary Boltzmann equation, what remains will be:

$$\begin{aligned}\delta\mathrm{F}_{\mathrm{BE}} :=& \frac{\partial}{\partial t}\left(Z^\nu(z,H)\delta f^\nu(z,H,t)\right) + \frac{\partial}{\partial z}\left(v_z^\nu(z,H)Z^\nu(z,H)\delta f^\nu(z,H,t)\right)\\ &- \left(\frac{\partial S^\nu}{\partial f}\delta f^\nu(z,H,t) + \frac{\partial S^\nu}{\partial \varphi}\delta\varphi(\boldsymbol{r},t)\right)\\ &- \left(\frac{\partial \Gamma^\nu}{\partial f}\delta f^\nu(z,H,t) + \frac{\partial \Gamma^\nu}{\partial \varphi}\delta\varphi(\boldsymbol{r},t)\right) = 0\,.\end{aligned} \tag{2.94}$$

It is important to notice that apart from the time derivative, the other terms in (2.94) are already calculated since we have constructed the Jacobian for our full Newton-Raphson system earlier. Hence, for $\alpha \in \tilde{\mathrm{D}}_{\mathrm{BE}}$ the discretized equations in SSSC are written down as:

$$\underline{\tilde{\mathrm{F}}}_{\mathrm{BE}}(\alpha) := \sum_{b\in\tilde{\mathrm{D}}_{\mathrm{PE}}} \frac{\partial\tilde{\mathrm{F}}_{\mathrm{BE}}(\alpha)}{\partial\varphi(b)}\bigg|_{\mathrm{DC}} \underline{\varphi}(b) + \sum_{\beta\in\tilde{\mathrm{D}}_{\mathrm{BE}}} \frac{\partial\tilde{\mathrm{F}}_{\mathrm{BE}}(\alpha)}{\partial f(\beta)}\bigg|_{\mathrm{DC}} \underline{f}(\beta) + \tilde{\mathrm{T}}_\alpha\{\underline{\varphi},\underline{f}\} = \tilde{\Gamma}'_{\mathrm{BE}}(\alpha) \tag{2.95}$$

The last term on the left-hand side is the discretized form of time derivative.

$$\begin{aligned}\mathrm{T}\{\underline{\varphi},\underline{f}\} =& Z^\nu(z,H)\frac{\partial}{\partial t}f^\nu(z,H^\nu(z,\varepsilon,t),t)\\ =& Z^\nu(z,H)\frac{\partial}{\partial t}\left[f^\nu(z,H^\nu(z,\varepsilon,t),t) + \mathrm{Re}\left\{\underline{f}^\nu(z,H^\nu(z,\varepsilon,t))\mathrm{e}^{\mathrm{i}\omega t}\right\}\right]\\ =& \mathrm{Re}\left\{\mathrm{i}\omega Z^\nu(z,H)\left(\frac{\partial f^\nu(z,H)}{\partial H}\frac{\partial\varepsilon_{\mathrm{sub}}^\nu(z)}{\partial\varphi}\underline{\varphi}(\boldsymbol{r}) + \underline{f}^\nu(z,H)\right)\mathrm{e}^{\mathrm{i}\omega t}\right\}\\ &+ \mathcal{O}\left(\underline{f}^\nu(z,H)\underline{\varphi}(\boldsymbol{r})\right)\end{aligned} \tag{2.96}$$

where in the last step, we have used

$$H^\nu(z,\varepsilon,t) = \varepsilon_{\mathrm{sub}}^\nu(z,t) + \varepsilon^\nu$$

$$\Rightarrow \frac{\partial H^\nu(z,\varepsilon,t)}{\partial t} = \mathrm{Re}\{\mathrm{i}\omega \underline{\varepsilon}^\nu_{\mathrm{sub}}(z)\mathrm{e}^{\mathrm{i}\omega t}\} = \mathrm{Re}\{\mathrm{i}\omega \frac{\partial \varepsilon^\nu_{\mathrm{sub}}(z)}{\partial \varphi}\underline{\varphi}(\boldsymbol{r})\mathrm{e}^{\mathrm{i}\omega t}\}\,. \tag{2.97}$$

and the first order dependency of the subband energies on the electric potential are calculated using the perturbation theory, as explained in Section 2.3. Moreover, we have used the fact that the density of states and velocity are functions of the kinetic energy and their derivatives with respect to time vanishes. The additional term Γ'_{BE} in (2.95) represents the linearization of the source/drain boundary conditions with respect to the small signal phasor of the applied bias $\underline{V}_{\mathrm{S/D}}$. Since this term does not depend on the parimary variables, it is located on the right-hand side of (2.95). In the case of generation/recombination BCs, we have for $\alpha = (z, H; \nu)$:

$$\begin{aligned}\Gamma'_{\mathrm{BE}}(\alpha) &= \frac{\partial \Gamma^\nu_{\mathrm{GR}}(z,H)}{\partial V_{\mathrm{S}}}\underline{V}_{\mathrm{S}} + \frac{\partial \Gamma^\nu_{\mathrm{GR}}(z,H)}{\partial V_{\mathrm{D}}}\underline{V}_{\mathrm{D}} \\ &= \frac{v_{\mathrm{GR}}}{V_{\mathrm{T}}} Z^\nu(z,H) f^\nu_{\mathrm{eq}}(z,H)\big(f^\nu_{\mathrm{eq}}(z,H)-1\big)\big[\underline{V}_{\mathrm{S}}\delta(z-z_{\mathrm{S}}) + \underline{V}_{\mathrm{D}}\delta(z-z_{\mathrm{D}})\big]\,.\end{aligned} \tag{2.98}$$

As mentioned before, the system matrix is closely related to the Jacobian already exploited in the DC simulation. Since the discretization of the steady state equations was explained in the previous sections, it will not be repeated here except for the discretization of time derivative which will be added to the steady state Jacobian. For the direct grid node $\alpha = (z_i, H_j; \nu)$ we have

$$\begin{aligned}\tilde{\mathrm{T}}_\alpha &= \int_{H_{j-1/2}}^{H_{j+1/2}} \mathrm{d}H \int_{z_{i-1/2}}^{z_{i+1/2}} \mathrm{d}z\, \mathrm{i}\omega Z^\nu(z,H)\left(\frac{\partial f^\nu(z,H)}{\partial H}\underline{\varepsilon}^\nu_{\mathrm{sub}}(z) + \underline{f}^\nu(z,H)\right) \\ &= \mathrm{i}\omega\Delta z_i \int_{H_{j-1/2}}^{H_{j+1/2}} \mathrm{d}H\, Z^\nu(z_i,H)\left(\frac{\partial f^\nu(z_i,H)}{\partial H}\underline{\varepsilon}^\nu_{\mathrm{sub}}(z_i) + \underline{f}^\nu(z_i,H)\right)\end{aligned} \tag{2.99}$$

$$= \mathrm{i}\omega\Delta z_i\left[\left(f^\nu(z_i,H_{j+1/2}) - f^\nu(z_i,H_{j-1/2})\right)\underline{\varepsilon}^\nu_{\mathrm{sub}}(z_i) + \Delta H^\nu_j(z_i)\underline{f}^\nu(z_i,H_j)\right] \tag{2.100}$$

and similarly for the adjoint grid node $\alpha = (z_{i+1/2}, H_j; \nu)$

$$\begin{aligned}\tilde{\mathrm{T}}_\alpha = {}& \mathrm{i}\omega\Delta z_{i+1/2}\left(f^\nu(z_{i+1/2},H_{j+1/2}) - f^\nu(z_{i+1/2},H_{j-1/2})\right)\underline{\varepsilon}^\nu_{\mathrm{sub}}(z_{i+1/2}) \\ &+ \mathrm{i}\omega\Delta z_{i+1/2}\Delta H^\nu_j(z_{i+1/2})\underline{f}^\nu(z_{i+1/2},H_j)\,.\end{aligned} \tag{2.101}$$

The distribution function on the intermediate energy points $H_{j\pm1/2}$ is defined as

$$f^\nu(z,H_{j-1/2}) = \begin{cases} \frac{1}{2}\left(f^\nu(z,H_j) + f^\nu(z,H_{j-1})\right) & \text{if } \varepsilon^\nu_{\mathrm{sub}}(z) \le H_{j-1/2} \\ f^\nu(z,H_j) & \text{if } H_{j-1/2} < \varepsilon^\nu_{\mathrm{sub}}(z) < H_{j+1/2} \\ 0 & \text{if } H_{j+1/2} \le \varepsilon^\nu_{\mathrm{sub}}(z) \end{cases} \tag{2.102}$$

$$f^\nu(z,H_{j+1/2}) = \begin{cases} \frac{1}{2}\left(f^\nu(z,H_j)+f^\nu(z,H_{j+1})\right) & \text{if } \varepsilon^\nu_{\text{sub}}(z) < H_{j+1/2} \\ 0 & \text{if } H_{j+1/2} \le \varepsilon^\nu_{\text{sub}}(z) \end{cases} \tag{2.103}$$

The above discretization for the time derivative runs into problems with fulfilling the reciprocity conditions mainly because enforcing the boundary condition of (2.98) with a discretized integral as in (2.100) is inconsistent with the actual equilibrium solution of the BE, which requires that

$$\int_{H_{j-1/2}}^{H_{j+1/2}} \mathrm{d}H\, Z^\nu(z,H)\frac{\partial f^\nu_{\text{eq}}(z,H)}{\partial H} = -\left(\int_{H_{j-1/2}}^{H_{j+1/2}} \mathrm{d}H\, Z^\nu(z,H)\right) f^\nu_{\text{eq}}(z,H_j)\,. \tag{2.104}$$

In order to restore reciprocity in the discretized system of equations, the approach described in [141] is used. First, the distribution function in (2.99) is split into the multiplication of an equilibrium part f_{eq} and a non-equilibrium part f_{ne}. The box integration is then reformulated as

$$\begin{aligned} \tilde{\mathrm{T}}_\alpha =& \,\mathrm{i}\omega\Delta z Z^\nu(z,H_j)\left(f^\nu_{\text{ne}}(z,H_{j+1/2}) - f^\nu_{\text{ne}}(z,H_{j-1/2})\right) f^\nu_{\text{eq}}(z,H_j)\underline{\varepsilon}^\nu_{\text{sub}}(z_i) \\ &+ \mathrm{i}\omega\Delta z\Delta H^\nu_j(z) Z^\nu(z,H_j)\left(\frac{\partial f_{\text{eq}}(z,H)}{\partial H}\Big|_{H=H_j}\right) f^\nu_{\text{ne}}(z,H_j)\underline{\varepsilon}^\nu_{\text{sub}}(z_i) \\ &+ \mathrm{i}\omega\Delta z\Delta H^\nu_j(z) Z^\nu(z,H_j)\underline{f}^\nu(z,H_j)\,. \end{aligned} \tag{2.105}$$

This discretization satisfies (2.104) when the derivative of $f^\nu_{\text{eq}}(z,H)$ with respect to energy is calculated analytically and $f^\nu_{\text{ne}}(z,H_{j+1/2}) - f^\nu_{\text{ne}}(z,H_{j-1/2}) = 0$, which is the case for constant $f_{\text{ne}} = 1$ at equilibrium. On the other hand, simply enforcing (2.105) in the small signal time derivative term of the BE does not solve the problem since it violates the consistency of the continuity equation obtained from the BE with the electron density in the PE. If the term of the linearized density that is proportional to the derivative of the H–box with respect to the potential is replaced by the modified H–derivative in (2.104), full reciprocity of the numerical results is achieved. The resulting small signal matrix equation can be written as:

$$\left[\begin{pmatrix} \dfrac{\partial\hat{\mathrm{F}}_{\text{PE}}}{\partial\boldsymbol{\varphi}} & \dfrac{\partial\hat{\mathrm{F}}_{\text{PE}}}{\partial\boldsymbol{f}} \\ \dfrac{\partial\hat{\mathrm{F}}_{\text{BE}}}{\partial\boldsymbol{\varphi}} & \dfrac{\partial\hat{\mathrm{F}}_{\text{BE}}}{\partial\boldsymbol{f}} \end{pmatrix} + \begin{pmatrix} \hat{0} & \hat{0} \\ \dfrac{\partial\hat{\mathrm{T}}}{\partial\boldsymbol{\varphi}} & \dfrac{\partial\hat{\mathrm{T}}}{\partial\boldsymbol{f}} \end{pmatrix}\right] \begin{pmatrix} \underline{\boldsymbol{\varphi}} \\ \underline{\boldsymbol{f}} \end{pmatrix} = \boldsymbol{b} \tag{2.106}$$

where the first matrix on the left-hand side must be evaluated for steady state, and the second matrix contains the linearization of the time derivative in the BE. $\partial\hat{\mathrm{T}}/\partial\boldsymbol{f}$ is diagonal and its computation is very fast. For each bias point, the stationary solution and its corresponding Jacobian have to be calculated only once, as they can be stored and reused for different frequencies. In doing so, the computational burden of evaluating the full Jacobian for each frequency of interest is considerably alleviated.

2.6.1 Ramo-Shockley Theorem

Once the values of $\underline{\boldsymbol{f}}$ and $\underline{\boldsymbol{\varphi}}$ are obtained by solving the small signal matrix equation (2.106) for the frequency of interest, the terminal current for each contact can be calculated using the Ramo-Shockley theorem [142, 143], which states that the current induced in an electrode is due to the instantaneous change of electrostatic flux lines which end on the electrode, *not* the amount of charge received by the electrode per second.

We start this section by reminding the reader of a few relations which will prove useful in the derivations that follow. The small signal 1D continuity equation per subband, which is obtained by integrating the BE over k, reads:

$$\frac{\partial}{\partial z}\underline{j}^{\nu}(z) + \mathrm{i}\omega\underline{n}^{\nu}(z) - \underline{S}^{\nu}_{\mathrm{I}}(z) - \underline{\Gamma}^{\nu}_{\mathrm{I}}(z) = 0 \tag{2.107}$$

The subscript I in $\underline{S}^{\nu}_{\mathrm{I}}(z)$ and $\underline{\Gamma}^{\nu}_{\mathrm{I}}(z)$ indicates integration over k–space, which might result in nonzero terms since the scattering integral need not be charge conserving per subband. Moreover, linearization of the small signal 3D current in the transport direction gives for each subband:

$$\begin{aligned} &J^{\nu}_{z}(\boldsymbol{r}) = j^{\nu}(z)|\psi(\boldsymbol{r})|^2 \\ &\Rightarrow \underline{J}^{\nu}_{z}(\boldsymbol{r}) = \underline{j}^{\nu}(z)|\psi^{\nu}(\boldsymbol{r})|^2 + 2j^{\nu}(z)\psi^{\nu}(\boldsymbol{r})\underline{\psi}^{\nu}(\boldsymbol{r}) \end{aligned} \tag{2.108}$$

which gives:

$$\begin{aligned} \frac{\partial}{\partial z}\underline{J}^{\nu}_{z}(\boldsymbol{r}) = &\left(\frac{\partial}{\partial z}\underline{j}^{\nu}(z)\right)|\psi^{\nu}(\boldsymbol{r})|^2 + \underline{j}^{\nu}(z)\left(\frac{\partial}{\partial z}|\psi^{\nu}(\boldsymbol{r})|^2\right) \\ &+ 2\left(\frac{\partial}{\partial z}j^{\nu}(z)\right)\psi^{\nu}(\boldsymbol{r})\underline{\psi}^{\nu}(\boldsymbol{r}) + 2j^{\nu}(z)\frac{\partial}{\partial z}\left(\psi^{\nu}(\boldsymbol{r})\underline{\psi}^{\nu}(\boldsymbol{r})\right) . \end{aligned} \tag{2.109}$$

In addition, the 3D conduction current needs to satisfy the continuity equation[10]

$$\begin{aligned} &\nabla\cdot\underline{\boldsymbol{J}}^{\nu}(\boldsymbol{r}) + \mathrm{i}\omega\underline{n}^{\nu}(z)|\psi^{\nu}(\boldsymbol{r})|^2 + 2\mathrm{i}\omega n^{\nu}(z)\psi^{\nu}(\boldsymbol{r})\underline{\psi}^{\nu}(\boldsymbol{r}) \\ &\quad = \left(\underline{S}^{\nu}_{\mathrm{I}}(z) + \underline{\Gamma}^{\nu}_{\mathrm{I}}(z)\right)|\psi^{\nu}(\boldsymbol{r})|^2 + 2\left(S^{\nu}_{\mathrm{I}}(z) + \Gamma^{\nu}_{\mathrm{I}}(z)\right)\psi^{\nu}(\boldsymbol{r})\underline{\psi}^{\nu}(\boldsymbol{r}) \end{aligned} \tag{2.110}$$

The small signal terminal current of the k–th contact is defined as:

$$\underline{I}_k = -\int_{\partial \mathrm{D}_k}\left(\sum_{\nu}\underline{\boldsymbol{J}}^{\nu}(\boldsymbol{r}) + \mathrm{i}\omega\epsilon(\boldsymbol{r})\nabla_{\boldsymbol{r}}\underline{\varphi}(\boldsymbol{r})\right)\cdot\mathrm{d}\boldsymbol{A} \tag{2.111}$$

[10] Several terms contribute to the transverse components of the current density. For example, the two–dimensional charge density can change with redistribution of electrons in the perpendicular plane due to changes in electrostatic potential or by scattering mechanisms. Moreover, we can define a transverse current attributed to the interface of adjacent z–boxes since the wavefunctions change in z–direction. Here, we do not care for the explicit form of $\boldsymbol{J}^{\nu}(\boldsymbol{r})$ since the 3D continuity equation suffices for the derivation of small signal terminal currents.

where $\partial \mathrm{D}_k$ is defined as the surface domain of the k–th contact. The positive sign in front of the displacement current is due to $\underline{\boldsymbol{J}}^{\nu}(\boldsymbol{r})$ being the electron current and $\mathrm{d}\boldsymbol{A}$ is assumed to point inwards to the device. Let us define the fundamental function for each contact as:

$$\tau_k(\boldsymbol{r}) := \delta_{ki}\,, \quad \boldsymbol{r} \in \partial \mathrm{D}_i \tag{2.112}$$

which also obeys the Laplace equation:

$$\nabla_{\boldsymbol{r}} \cdot \big(\epsilon(\boldsymbol{r}) \nabla_{\boldsymbol{r}} \tau_k(\boldsymbol{r})\big) = 0\,. \tag{2.113}$$

The integral in (2.111) can be safely multiplied by $\tau_k(\boldsymbol{r})$, because $\tau_k(\boldsymbol{r}) = 1$ on the k–th contact:

$$\underline{I}_k = -\int_{\partial \mathrm{D}_k} \tau_k(\boldsymbol{r}) \Big(\sum_{\nu} \underline{\boldsymbol{J}}^{\nu}(\boldsymbol{r}) + \mathrm{i}\omega\epsilon(\boldsymbol{r}) \nabla_{\boldsymbol{r}} \underline{\varphi}(\boldsymbol{r}) \Big) \cdot \mathrm{d}\boldsymbol{A} \tag{2.114}$$

$$= -\oint_{\partial \mathrm{D}} \tau_k(\boldsymbol{r}) \Big(\sum_{\nu} \underline{\boldsymbol{J}}^{\nu}(\boldsymbol{r}) + \mathrm{i}\omega\epsilon(\boldsymbol{r}) \nabla_{\boldsymbol{r}} \underline{\varphi}(\boldsymbol{r}) \Big) \cdot \mathrm{d}\boldsymbol{A} \tag{2.115}$$

$$= \int_{\mathrm{D}} \nabla_{\boldsymbol{r}} \cdot \Big[\tau_k(\boldsymbol{r}) \Big(\sum_{\nu} \underline{\boldsymbol{J}}^{\nu}(\boldsymbol{r}) + \mathrm{i}\omega\epsilon(\boldsymbol{r}) \nabla_{\boldsymbol{r}} \underline{\varphi}(\boldsymbol{r}) \Big) \Big] \, \mathrm{d}V \tag{2.116}$$

In the above steps, we have extended the integration region to the complete surface of the solution domain D, because $\tau_k(\boldsymbol{r}) = 0$ for all the other contacts. We have then employed the Gauss' law to convert the surface integral into a volume integral over the divergence of the integrand and arrive at (2.116). This gives:

$$\begin{aligned} \underline{I}_k = &\int_{\mathrm{D}} \tau_k(\boldsymbol{r}) \nabla_{\boldsymbol{r}} \cdot \Big(\sum_{\nu} \underline{\boldsymbol{J}}^{\nu}(\boldsymbol{r}) + \mathrm{i}\omega\epsilon(\boldsymbol{r}) \nabla_{\boldsymbol{r}} \underline{\varphi}(\boldsymbol{r}) \Big) \, \mathrm{d}V \\ &+ \int_{\mathrm{D}} \Big(\sum_{\nu} \underline{\boldsymbol{J}}^{\nu}(\boldsymbol{r}) + \mathrm{i}\omega\epsilon(\boldsymbol{r}) \nabla_{\boldsymbol{r}} \underline{\varphi}(\boldsymbol{r}) \Big) \nabla_{\boldsymbol{r}} \tau_k(\boldsymbol{r}) \, \mathrm{d}V = \mathcal{I}_k^{(1)} + \mathcal{I}_k^{(2)} \end{aligned} \tag{2.117}$$

For the first term, we use (2.110) and the linearized Poisson equation to get:

$$\begin{aligned} \mathcal{I}_k^{(1)} = &\int_{\mathrm{D}} \tau_k(\boldsymbol{r}) \nabla_{\boldsymbol{r}} \cdot \Big(\sum_{\nu} \underline{\boldsymbol{J}}^{\nu}(\boldsymbol{r}) + \mathrm{i}\omega\epsilon(\boldsymbol{r}) \nabla_{\boldsymbol{r}} \underline{\varphi}(\boldsymbol{r}) \Big) \, \mathrm{d}V \\ = &\int_{\mathrm{D}} \tau_k(\boldsymbol{r}) \sum_{\nu} \Big(-\mathrm{i}\omega \underline{n}^{\nu}(z) |\psi^{\nu}(\boldsymbol{r})|^2 - 2\mathrm{i}\omega n^{\nu}(z) \psi^{\nu}(\boldsymbol{r}) \underline{\psi}^{\nu}(\boldsymbol{r}) \\ &\qquad + \left(\underline{S}_{\mathrm{I}}^{\nu}(z) + \underline{\Gamma}_{\mathrm{I}}^{\nu}(z)\right) |\psi^{\nu}(\boldsymbol{r})|^2 + 2\left(S_{\mathrm{I}}^{\nu}(z) + \Gamma_{\mathrm{I}}^{\nu}(z)\right) \psi^{\nu}(\boldsymbol{r}) \underline{\psi}^{\nu}(\boldsymbol{r}) \\ &\qquad + \mathrm{i}\omega \underline{n}^{\nu}(z) |\psi^{\nu}(\boldsymbol{r})|^2 + 2\mathrm{i}\omega n^{\nu}(z) \psi^{\nu}(\boldsymbol{r}) \underline{\psi}^{\nu}(\boldsymbol{r}) \Big) \, \mathrm{d}V \\ = &\int_{\mathrm{D}} \mathrm{d}V \, \tau_k(\boldsymbol{r}) \sum_{\nu} \Big(\left(\underline{S}_{\mathrm{I}}^{\nu}(z) + \underline{\Gamma}_{\mathrm{I}}^{\nu}(z)\right) |\psi^{\nu}(\boldsymbol{r})|^2 + 2\left(S_{\mathrm{I}}^{\nu}(z) + \Gamma_{\mathrm{I}}^{\nu}(z)\right) \psi^{\nu}(\boldsymbol{r}) \underline{\psi}^{\nu}(\boldsymbol{r}) \Big). \end{aligned} \tag{2.118}$$

The second term is calculated as follows:

$$\mathcal{I}_k^{(2)} = \int_{\mathrm{D}} \Big(\sum_\nu \underline{\boldsymbol{J}}^\nu(\boldsymbol{r}) \cdot \nabla_{\boldsymbol{r}} \tau_k(\boldsymbol{r}) + \mathrm{i}\omega\epsilon(\boldsymbol{r}) \nabla_{\boldsymbol{r}} \underline{\varphi}(\boldsymbol{r}) \cdot \nabla_{\boldsymbol{r}} \tau_k(\boldsymbol{r}) \Big) \,\mathrm{d}V \tag{2.119}$$
$$= \int_{\mathrm{D}} \sum_\nu \underline{\boldsymbol{J}}^\nu(\boldsymbol{r}) \cdot \nabla_{\boldsymbol{r}} \tau_k(\boldsymbol{r}) \,\mathrm{d}V + \mathrm{i}\omega \sum_i C_{ik} \underline{V}_i^C = \mathcal{I}'^{(2)}_k + \mathrm{i}\omega \sum_i C_{ik} \underline{V}_i$$

where we have made use of the capacitance coefficients[11]

$$C_{ik} := \int_{\mathrm{D}} \epsilon(\boldsymbol{r}) \nabla_{\boldsymbol{r}} \tau_i(\boldsymbol{r}) \cdot \nabla_{\boldsymbol{r}} \tau_k(\boldsymbol{r}) \,\mathrm{d}V \,. \tag{2.120}$$

What remains is to calculate $\mathcal{I}'^{(2)}_k$ in (2.119),

$$\mathcal{I}'^{(2)}_k = \sum_\nu \int_{\mathrm{D}} \mathrm{d}V \, \nabla \cdot (\underline{\boldsymbol{J}}^\nu(\boldsymbol{r}) \tau_k(\boldsymbol{r})) - \sum_\nu \int_{\mathrm{D}} \mathrm{d}V \, \tau_k(\boldsymbol{r}) \nabla \cdot \underline{\boldsymbol{J}}^\nu(\boldsymbol{r}) \tag{2.121}$$

The first integral in (2.121) is calculated as:

$$\begin{aligned}
\int_{\mathrm{D}} \mathrm{d}V \, \nabla \cdot (\underline{\boldsymbol{J}}^\nu(\boldsymbol{r}) \tau_k(\boldsymbol{r})) &= \int_{\partial\mathrm{D}} \mathrm{d}\boldsymbol{A} \cdot \underline{\boldsymbol{J}}^\nu(\boldsymbol{r}) \tau_k(\boldsymbol{r}) = \int_{\partial\mathrm{D}} \mathrm{d}\boldsymbol{A} \cdot \underline{J}_z^\nu(\boldsymbol{r}) \tau_k(\boldsymbol{r}) \\
&= \int_{\partial\mathrm{D}} \mathrm{d}\boldsymbol{A} \cdot \boldsymbol{e}_z \tau_k(\boldsymbol{r}) \Big[\underline{j}^\nu(z) |\psi^\nu(\boldsymbol{r})|^2 + 2 j^\nu(z) \psi^\nu(\boldsymbol{r}) \underline{\psi}^\nu(\boldsymbol{r}) \Big] \\
&= - \int_{\partial\mathrm{D}} \mathrm{d}\boldsymbol{A} \, \tau_k(\boldsymbol{r}) \Big[\underline{\Gamma}_{\mathrm{I}}^\nu(z) |\psi^\nu(\boldsymbol{r})|^2 + 2 \Gamma_{\mathrm{I}}^\nu(z) \psi^\nu(\boldsymbol{r}) \underline{\psi}^\nu(\boldsymbol{r}) \Big] \\
&= - \int_{\partial\mathrm{D}} \mathrm{d}\boldsymbol{A} \, \tau_k(\boldsymbol{r}) \Big[\underline{\Gamma}_{\mathrm{I}}^\nu(z) |\psi^\nu(\boldsymbol{r})|^2 \Big] \\
&= - \int_{\mathrm{D}} \mathrm{d}V \, \tau_k(\boldsymbol{r}) \Big[\underline{\Gamma}_{\mathrm{I}}^\nu(z) |\psi^\nu(\boldsymbol{r})|^2 \Big]
\end{aligned} \tag{2.122}$$

[11] Let us discuss a few important requirements that our calculated capacitance matrix must satisfy. These can be viewed as sanity checks, to make sure that our numerical computation is producing sensible results.

1. The capacitance coefficients are symmetric $C_{ij} = C_{ji}$ and positive for $i = j$, because the permittivity is positive. For $i \neq j$ they are negative due to the maximum principle. For $i \neq j$ the fundamental solutions $\tau_j(\boldsymbol{r})$ take the value 0 on the surface of the conductor i and increase when moving away from the conductor into the solution domain. This implies $\nabla\tau_j \cdot \mathrm{d}\boldsymbol{A} > 0$ and thus $C_{ij} < 0$.
2. The sum $\sum_{j=0}^{\mathrm{N}} C_{ij}$ is 0, because

$$\sum_{j=0}^{\mathrm{N}} C_{ij} = \sum_{j=0}^{\mathrm{N}} \int_{\mathrm{D}} \epsilon(\boldsymbol{r}) \nabla_{\boldsymbol{r}} \tau_i(\boldsymbol{r}) \cdot \nabla_{\boldsymbol{r}} \tau_j(\boldsymbol{r}) \,\mathrm{d}V = \int_{\mathrm{D}} \epsilon(\boldsymbol{r}) \nabla_{\boldsymbol{r}} \tau_i(\boldsymbol{r}) \cdot \nabla_{\boldsymbol{r}} \Big(\sum_{j=0}^{\mathrm{N}} \tau_j(\boldsymbol{r}) \Big) \,\mathrm{d}V = 0 \,.$$

where the term proportional to $\psi^\nu(\boldsymbol{r})\underline{\psi}^\nu(\boldsymbol{r})$ vanishes since the source and drain contacts span the whole cross–section. For the second part, we simply use the steady state and small signal continuity equations and obtain,

$$\begin{aligned}\mathcal{I}'^{(2)}_k = &-\sum_\nu \int_\mathrm{D} \mathrm{d}V\, \tau_k(\boldsymbol{r})\underline{\Gamma}^\nu_\mathrm{I}(z)|\psi^\nu(\boldsymbol{r})|^2 + \sum_\nu \int_\mathrm{D} \mathrm{d}V\, \tau_k(\boldsymbol{r})\left(2\mathrm{i}\omega n^\nu(z)\psi^\nu(\boldsymbol{r})\underline{\psi}^\nu(\boldsymbol{r})\right)\\ &-\sum_\nu \int_\mathrm{D} \mathrm{d}V\, \tau_k(\boldsymbol{r})\left[\left(\frac{\partial}{\partial z}\underline{j}^\nu(z)\right)|\psi^\nu(\boldsymbol{r})|^2 + 2\left(\frac{\partial}{\partial z}j^\nu(z)\right)\psi^\nu(\boldsymbol{r})\underline{\psi}^\nu(\boldsymbol{r})\right]\end{aligned} \tag{2.123}$$

Inserting the calculated terms into the terminal current yields:

$$\begin{aligned}\underline{I}_k &= \int_\mathrm{D} \mathrm{d}V\, \tau_k(\boldsymbol{r})\sum_\nu\left[\left(\underline{S}^\nu_\mathrm{I}(z) - \frac{\partial}{\partial z}\underline{j}^\nu(z)\right)|\psi^\nu(\boldsymbol{r})|^2\right.\\ &\quad\left. + 2\left(S^\nu_\mathrm{I}(z) + \Gamma^\nu_\mathrm{I}(z) - \frac{\partial}{\partial z}j^\nu(z) + 2\mathrm{i}\omega n^\nu(z)\right)\psi^\nu(\boldsymbol{r})\underline{\psi}^\nu(\boldsymbol{r})\right] + \mathrm{i}\omega\sum_i C_{ik}\underline{V}_i\\ &= \int_\mathrm{D} \mathrm{d}V\, \tau_k(\boldsymbol{r})\sum_\nu\left[\left(\underline{S}^\nu_\mathrm{I}(z) - \frac{\partial}{\partial z}\underline{j}^\nu(z)\right)|\psi^\nu(\boldsymbol{r})|^2 + 2\mathrm{i}\omega n^\nu(z)\psi^\nu(\boldsymbol{r})\underline{\psi}^\nu(\boldsymbol{r})\right]\\ &\quad + \mathrm{i}\omega\sum_i C_{ik}\underline{V}_i\,.\end{aligned} \tag{2.124}$$

The admittance parameter $\underline{Y}_{ki}$ is then calculated as,

$$\underline{Y}_{ki} = \frac{\underline{I}_k}{\underline{V}_i}\Big|_{\underline{V}_j=0 \text{ for } j\neq i.} \tag{2.125}$$

2.7 Noise Analysis

Characterization and modeling of electronic noise in semiconductor devices and circuits is considered as an indispensable part of RF electronics and the microscopic noise sources of thermal origin such as the excessive drain noise in the saturation regime [144, 145] and the induced gate noise [146, 147] are becoming the limiting factors for the performance of future FETs [148]. Since the design, manufacturing technology, and operation regime of FETs can, in principle, improve their noise behavior, undestanding the physics of stochastic fluctuations allows for development of highly sensitive devices [149]. On the other hand, investigation of the irreducible noise sources which are inherent in the kinetics of charge carriers helps in determining the performance limits of CMOS and analog amplifiers. Especially for deep submicron devices and nonlinear transport, microscopic and physics–based noise models are necessary, since they allow us to investigate various noise sources within the device and gain insight into the internal stochastic processes which degrade the device's performance.

Rigorous approaches based on the drift–diffusion and hydrodynamic models [30,150–156] are shown to give unsatisfactory results considering the experimental data [153,157,158]. This implies that for ultrashort FETs, higher–order transport models are required for accurately capturing the non–equilibrium effects and noise phenomena. In the framework of the BE, noise analysis is usually performed by the MC method [9, 24, 159], which inherently contains fluctuations and allows for noise calculations with almost no additional computational cost. However, the MC simulations are in time domain and the CPU time is at least inversely proportional to the minimum frequency investigated. Even the investigation of RF noise at a limited technically relevant frequency range can prove prohibitively CPU intensive [155, 160]. Thus, a deterministic method which is based in the frequency domain and can handle arbitrary low frequencies and slow processes is the most desirable option.

In this section, deterministic noise analysis in the framework of the BE and under small signal operation is presented. An efficient Green's function approach to the Langevin–source method [153], which is a variation of direct impedence field method [150], is used for the characterization of RF noise. The discussions are limited to the noise due to scattering events and generation/recombination of the particles on the source and drain contacts, as they are shown to be the most important noise sources for RF applications. The avalanche noise observed at the breakdown of semiconductor junctions [149] and the $1/f$–noise [149,161,162] which becomes dominant only in the low frequency domain, are not included in this work.

2.7.1 Theory

Electronic noise[12] is generally characterized by the one-sided power spectral density (PSD) which, according to the Wiener-Khintchine theorem [163, 164], is the Fourier transform of the corresponding correlation function of the two microscopic quantities x and y:

$$P_{xy}(\omega) = P^*_{yx}(\omega) = 2\int_{-\infty}^{\infty} \langle x(t+\tau)y(t)\rangle \mathrm{e}^{-\mathrm{i}\omega\tau}\,\mathrm{d}\tau\ , \tag{2.126}$$

where the cross-correlation function for the time shift τ is defined as

$$\langle x(t+\tau)y(t)\rangle = \lim_{\Delta t\to\infty}\frac{1}{\Delta t}\int x(t+\tau)y(t)\,\mathrm{d}t\ . \tag{2.127}$$

and is assumed to be absolutely integrable[13]. Equation (2.127) is meaningful only if the quantum mechanical effects are insignificant and the random variables $x(t)$ and $y(t)$ may be considered as classical quantities. For discussions about the properties of auto– and cross–correlation functions and the power spectrums, the reader can refer to [149, 165].

In the Langevin-source method, the system equations are perturbed by random excitations, which are assumed to be small enough so that the connection between these random forces and macroscopic fluctuations is essentially linear and the noise is investigated in the small signal operation regime[14]. Generally speaking, the Langevin source ξ can appear as a generation rate within the transport equation and depending on the physical meaning of governing equations, can be interpreted as an injection of charge or current at some vector state $\boldsymbol{s}$. As is usual with this approach, Green's functions are used to solve the system of equations for arbitrary excitations. The Green's function $G^{\alpha,\beta}(\boldsymbol{s},t;\boldsymbol{s}',t')$ quantifies the response in variable α to a unit source $\delta(\boldsymbol{s}-\boldsymbol{s}')\delta(t-t')$ injected as β. We evaluate the fluctuation $\delta\alpha$ induced by the vector source ξ_β as the convolution integral

$$\delta\alpha(\boldsymbol{s},t) = \int_{\Omega_{\mathrm{sys}}}\int_{-\infty}^{t} G^{\alpha,\beta}(\boldsymbol{s},t;\boldsymbol{s}',t')\xi_\beta(\boldsymbol{s}',t')\,\mathrm{d}t'\,\mathrm{d}\boldsymbol{s}'\,, \tag{2.128}$$

where Ω_{sys} is the system volume. Assuming that the noiseless steady state is at least wide-sense stationary, (2.128) can be simplified and we have a time-invariant linear system:

$$G^{\alpha,\beta}(\boldsymbol{s},t;\boldsymbol{s}',t') = G^{\alpha,\beta}(\boldsymbol{s},\boldsymbol{s}';t-t')\,, \tag{2.129}$$

[12] The terms "fluctuations" and "noise" are used interchangably throughout this work.

[13] The stochastic variables $x(t)$ and $y(t)$ need not be absolutely integrable.

[14] Large signal studies are beyond the scope of this dissertation. An example for large signal noise simulations with deterministic BE solvers is presented in [166] for bulk structures, and [165, 167, 168] offer useful means for noise analysis in cyclostationary or quasiperiodic cases.

which allows for frequency-domain analysis. In the SSSC, (2.128) simplifies to

$$\underline{\alpha}(\boldsymbol{s},\omega) = \sum_{\beta} \int_{\Omega_{\mathrm{sys}}} \underline{G}^{\alpha,\beta}(\boldsymbol{s},\boldsymbol{s}';\omega)\underline{\xi}_{\beta}(\boldsymbol{s}',\omega)\,\mathrm{d}\boldsymbol{s}' \tag{2.130}$$

and the summation runs over various noise sources. Once the Green's functions have been evaluated, the PSD of the miniscule fluctuations $\delta\alpha$, $\delta\alpha'$ can be computed as:

$$\begin{aligned} &P_{\alpha\alpha'}(\boldsymbol{s},\boldsymbol{s}';\omega) = \\ &\sum_{\beta,\beta'} \int_{\Omega_{\mathrm{sys}}} \int_{\Omega_{\mathrm{sys}}} \underline{G}^{\alpha,\beta}(\boldsymbol{s},\boldsymbol{s}_0;\omega) P_{\beta\beta'}(\boldsymbol{s}_0,\boldsymbol{s}_0';\omega) \left(\underline{G}^{\alpha',\beta'}(\boldsymbol{s}',\boldsymbol{s}_0';\omega)\right)^{*} \mathrm{d}\boldsymbol{s}_0\,\mathrm{d}\boldsymbol{s}_0' \end{aligned} \tag{2.131}$$

where $P_{\beta\beta'}(\boldsymbol{s}_0,\boldsymbol{s}_0';\omega)$ is the correlation spectrum due to the macroscopic excitations $\underline{\xi}_{\beta}(\boldsymbol{s}_0,\omega)$ and $\underline{\xi}_{\beta'}(\boldsymbol{s}_0',\omega)$ within the device, and $\underline{G}^{*}$ denotes the complex conjugate of $\underline{G}$. The main problem, therefore, amounts to solving a system of (complex) linear equations for the Green's functions of the primary variables. We start with the Langevin BE, which provides a framework for dealing with the stochastic transport problem [153, 169, 170]:

$$\begin{aligned} &\Big(\frac{\partial}{\partial t} + \frac{1}{\hbar}F^{\nu}(z)\frac{\partial}{\partial k} + v_z^{\nu}(k)\frac{\partial}{\partial z}\Big)\delta f_k^{\nu}(z,k,t) \\ &\qquad + \frac{1}{\hbar}\delta F^{\nu}(z,t)\frac{\partial}{\partial k} f_k^{\nu}(z,k) - \delta S^{\nu}\{\cdot\} - \delta\Gamma^{\nu}\{\cdot\} = \xi^{\mathrm{BE}}(z,k,t) \end{aligned} \tag{2.132}$$

where ξ^{BE} is the Langevin source applied to the BE, and δf is the (time dependent) solution of the Langevin BE. The Green's functions are defined as the solution of (2.132) with $\xi^{\mathrm{BE}} = 2\pi\delta(z-z')\delta(k-k')\delta(t-t')$, i.e.

$$\begin{aligned} &\left(\frac{\partial}{\partial t} + \frac{1}{\hbar}F^{\nu}(z)\frac{\partial}{\partial k} + v_z^{\nu}(k)\frac{\partial}{\partial z}\right) G^{f,\mathrm{BE}}(z,k,t;z',k') \\ &+ \frac{1}{\hbar}G^{F,\mathrm{BE}}(z;z',k')\frac{\partial}{\partial k} f^{\nu}(z,k) - S\{\cdot\} = 2\pi\delta(z-z')\delta(k-k')\delta(t-t')\,. \end{aligned} \tag{2.133}$$

where "$\cdot$" is a placeholder as the scattering term is expressed in terms of the distribution functions for various initial and final states. The superscript f,BE of the Green's function indicates that this is the Green's function of the distribution function with a Langevin source in the BE. Note, that the Green's function $G^{F,\mathrm{BE}}$ can be calculated from the Green's function of potential in the same way that the actual fluctuation of the force δF is calculated with the fluctuation of potential using the first order perturbation theory:

$$\left(G^{F,\mathrm{BE}}\right)^{\nu,\nu'}(z;z',k') = -\frac{\partial}{\partial z}\left(G^{\varepsilon_{\mathrm{sub}},\mathrm{BE}}\right)^{\nu,\nu'}(z;z',k')\,, \tag{2.134}$$

$$\left(G^{\varepsilon_{\mathrm{sub}},\mathrm{BE}}\right)^{\nu,\nu'}(z;z',k') = -q\iint \mathrm{d}x\,\mathrm{d}y\,|\psi^{\nu}(\boldsymbol{r})|^2\left(G^{\varphi,\mathrm{BE}}\right)^{\nu'}(\boldsymbol{r};z',k')\,. \quad (2.135)$$

In a similar way, the Green's functions of scattering and contact generation and recombination terms can be further deconstructed until we have an equation described solely by the Green's functions of our primary variables.

All operators in (2.133) operate on the first set of coordinates, i.e. (z,k,t). Therefore, projection onto equi-energy surfaces and H–transformation followed by box integration, yields the same equation as in the small signal analysis. Following the notation in section 2.5, for every $\alpha\in\tilde{\mathrm{D}}_{\mathrm{BE}}$ we have:

$$\sum_{b\in\tilde{\mathrm{D}}_{\mathrm{PE}}}\frac{\partial \mathrm{F}_{\mathrm{BE}}(\alpha)}{\partial\varphi(b)}\bar{G}^{\varphi,\mathrm{BE}}(b;z',k') + \sum_{\beta\in\tilde{\mathrm{D}}_{\mathrm{BE}}}\frac{\partial \mathrm{F}_{\mathrm{BE}}(\alpha)}{\partial f(\beta)}\bar{G}^{f,\mathrm{BE}}(\beta;z',k') = \bar{E}^{\mathrm{BE}}(\alpha;z',k') \quad (2.136)$$

where ν,ν' and ω are dropped from the set of arguments for the sake of better legibility, but we should remember the transformation to the frequency domain. The coefficients in (2.136) also include the time derivative terms as derived in Sect. 2.6. Defining the transformation

$$\mathrm{T}^{\mathrm{BE}}_{ij} = \int_{H^{\nu}_{j\min}(z_i)}^{H^{\nu}_{j\max}(z_i)}\mathrm{d}H\int_{z_{i-1/2}}^{z_{i+1/2}}\mathrm{d}z\int\mathrm{d}\varepsilon\,\delta\big(\varepsilon-(H-\varepsilon^{\nu}_{\mathrm{sub}}(z))\big)\int\frac{\mathrm{d}k}{2\pi}\delta\big(\tilde{\varepsilon}^{\nu}(k)-\varepsilon\big) \quad (2.137)$$

consisting of a projection onto the constant energy surfaces, followed by an H–transformation and a subsequent box integration over H and z, the right-hand side of (2.136) is given by:

$$\begin{aligned}\bar{E}^{\mathrm{BE}}_{ij} &= \mathrm{T}^{\mathrm{BE}}_{ij}\Big(2\pi\delta(k-k')\delta(z-z')\Big)\\ &= \int_{H^{\nu}_{j\min}(z_i)}^{H^{\nu}_{j\max}(z_i)}\mathrm{d}H\int_{z_{i-1/2}}^{z_{i+1/2}}\mathrm{d}z\int\mathrm{d}\varepsilon\;\delta\big(\varepsilon-(H-\varepsilon^{\nu}_{\mathrm{sub}}(z))\big)\delta\big(\varepsilon-\tilde{\varepsilon}(k')\big)\delta(z-z')\\ &= \int_{H^{\nu}_{j\min}(z_i)}^{H^{\nu}_{j\max}(z_i)}\mathrm{d}H\int_{z_{i-1/2}}^{z_{i+1/2}}\mathrm{d}z\,\delta\big(H-H^{\nu}(z,k')\big)\delta(z-z')\,. \end{aligned} \quad (2.138)$$

Nevertheless, it is important to note that we still have a second set of coordinates, i.e. (z',k'). In order to obtain meaningful results within our coordinate framework, we need to transform both sets of coordinates in the same way. Projection and box integration of the second set of coordinates gives:

$$\begin{aligned}E^{\mathrm{BE}}_{ij} &= \mathrm{T}^{\mathrm{BE}}_{i'j'}\left(\int_{H^{\nu}_{j\min}(z_i)}^{H^{\nu}_{j\max}(z_i)}\mathrm{d}H\int_{z_{i-1/2}}^{z_{i+1/2}}\mathrm{d}z\,\delta\big(H-H(k')\big)\delta(z-z')\right)\\ &= \int_{H^{\nu}_{j'\min}(z_i)}^{H^{\nu}_{j'\max}(z_i)}\mathrm{d}H'\int_{z_{i'-1/2}}^{z_{i'+1/2}}\mathrm{d}z'\int\mathrm{d}\varepsilon'\,\delta\big(\varepsilon'-(H'-\varepsilon_{\mathrm{sub}}(z'))\big)\int\frac{\mathrm{d}k'}{2\pi}\delta\big(\varepsilon'-\tilde{\varepsilon}(k')\big)\end{aligned}$$

$$\int_{H^\nu_{j\min}(z_i)}^{H^\nu_{j\max}(z_i)} \mathrm{d}H \int_{z_{i-1/2}}^{z_{i+1/2}} \mathrm{d}z\, \delta\big(H - H(z,k')\big)\delta(z-z')$$
$$= \left(\int_{H^\nu_{j\min}(z_i)}^{H^\nu_{j\max}(z_i)} Z^\nu(z_{i'},H)\,\mathrm{d}H\right)\delta_{i,i'}\delta_{j,j'}\Delta z_i\,. \tag{2.139}$$

Next, we want to project and box integrate the second set of coordinates of the Green's functions. The operators in (2.133) do not depend on the primed coordinates and thus we can simply evaluate the transformations on the Green's functions directly. The Green's function of the electron distribution function is transformed to:

$$\begin{aligned}
G^{f,\mathrm{BE}}_{ij} &= \mathrm{T}^{\mathrm{BE}}_{i'j'}\left(\bar{G}^{f,\mathrm{BE}}(\cdot;z',k')\right)\\
&= \int_{H^\nu_{j'\min}(z_i)}^{H^\nu_{j'\max}(z_i)}\mathrm{d}H'\int_{z_{i'-1/2}}^{z_{i'+1/2}}\mathrm{d}z'\iint \mathrm{d}\varepsilon'\,\frac{\mathrm{d}k'}{2\pi}\delta\big(\varepsilon'-(H'-\varepsilon_{\mathrm{sub}}(z'))\big)\delta\big(\varepsilon'-\tilde{\varepsilon}(k')\big)G^{f,\mathrm{BE}}(\cdot;z',k')\\
&= \int_{H^\nu_{j'\min}(z_i)}^{H^\nu_{j'\max}(z_i)}\mathrm{d}H'\int_{z_{i'-1/2}}^{z_{i'+1/2}}\mathrm{d}z'\int \mathrm{d}\varepsilon'\,\delta\big(\varepsilon'-(H'-\varepsilon_{\mathrm{sub}}(z'))\big)Z(z',H')G^{f,\mathrm{BE}}(\cdot;z',\varepsilon')\\
&= \left(\int_{H^\nu_{j\min}(z_i)}^{H^\nu_{j\max}(z_i)} Z(z_i,H)\,\mathrm{d}H\right)G^{f,\mathrm{BE}}(\cdot;z_{i'},H_{j'})\Delta z_i\,.
\end{aligned} \tag{2.140}$$

With the same derivation applied to $\bar{G}^{\varphi,\mathrm{BE}}_{ij}$, the Langevin BE for $\alpha,\gamma\in\tilde{\mathrm{D}}_{\mathrm{BE}}$ is transformed to:

$$\sum_{b\in\tilde{\mathrm{D}}_{\mathrm{PE}}}\frac{\partial \mathrm{F}_{\mathrm{BE}}(\alpha)}{\partial\varphi(b)}G^{\varphi,\mathrm{BE}}(b;\gamma)+\sum_{\beta\in\tilde{\mathrm{D}}_{\mathrm{BE}}}\frac{\partial \mathrm{F}_{\mathrm{BE}}(\alpha)}{\partial f(\beta)}G^{f,\mathrm{BE}}(\beta;\gamma)=\delta_{\alpha,\gamma} \tag{2.141}$$

the left-hand side of which is exactly the discretization used in small signal calculations, and the right-hand side (ignoring the BCs) is an identity matrix. Since the PE is coupled to the BE through the electron density, eq. (2.141) and the PE of Green's functions where there is no Langevin source in the PE form a complete set of equations which can be solved for $G^{\varphi,\mathrm{BE}}$ and $G^{f,\mathrm{BE}}$. As for the PE, we can write down the equation with a Langevin source term

$$\nabla_{\boldsymbol{r}}\cdot\big(\epsilon(\boldsymbol{r})\nabla_{\boldsymbol{r}}\delta\varphi(\boldsymbol{r},t)\big)-q\int\frac{\mathrm{d}k}{2\pi}\delta f(z,k,t)|\psi(\boldsymbol{r})|^2=\xi^{\mathrm{PE}}(\boldsymbol{r},t) \tag{2.142}$$

with the corresponding equation for the Green's functions,

$$\nabla_{\boldsymbol{r}}\cdot\big(\epsilon(\boldsymbol{r})\nabla_{\boldsymbol{r}}G^{\varphi,\mathrm{PE}}(\boldsymbol{r},t;\boldsymbol{r}')\big)-q\int\frac{\mathrm{d}k}{2\pi}G^{f,\mathrm{PE}}(z,k,t;\boldsymbol{r}')|\psi(\boldsymbol{r})|^2=\delta(\boldsymbol{r}-\boldsymbol{r}') \tag{2.143}$$

The transformation applied to the PE consists of a simple box integration over $\boldsymbol{r}=(x_i,y_j,z_k)$:

$$\mathrm{T}^{\mathrm{PE}}_{ijk}=\int_{x_{i-1/2}}^{x_{i+1/2}}\mathrm{d}x\int_{y_{j-1/2}}^{y_{j+1/2}}\mathrm{d}y\int_{z_{k-1/2}}^{z_{k+1/2}}\mathrm{d}z \tag{2.144}$$

Now, considering the small-signal linearization of (2.142):

$$\sum_{b\in\tilde{\mathrm{D}}_{\mathrm{PE}}}\frac{\partial \mathrm{F}_{\mathrm{PE}}(a)}{\partial\varphi(b)}\bar{G}^{\varphi,\mathrm{PE}}(b;\boldsymbol{r}')+\sum_{\beta\in\tilde{\mathrm{D}}_{\mathrm{BE}}}\frac{\partial \mathrm{F}_{\mathrm{PE}}(a)}{\partial f(\beta)}\bar{G}^{f,\mathrm{PE}}(\beta;\boldsymbol{r}')=\bar{E}^{\mathrm{PE}}(a;\boldsymbol{r}') \tag{2.145}$$

we can use the same approach to transform the Green's functions and the right-hand side of this equation. Transformation of the Langevin source reads:

$$\bar{E}^{\mathrm{PE}}_{ijk}=\mathrm{T}^{\mathrm{PE}}_{ijk}\Big(\delta(\boldsymbol{r}-\boldsymbol{r}')\Big)=\int_{x_{i-1/2}}^{x_{i+1/2}}\mathrm{d}x\int_{y_{j-1/2}}^{y_{j+1/2}}\mathrm{d}y\int_{z_{k-1/2}}^{z_{k+1/2}}\mathrm{d}z\,\delta(\boldsymbol{r}-\boldsymbol{r}') \tag{2.146}$$

and applying a second transformation on the primed coordinates yields:

$$\begin{aligned}E^{\mathrm{PE}}_{ijk}&=\mathrm{T}^{\mathrm{PE}}_{i'j'k'}\left(\int_{x_{i-1/2}}^{x_{i+1/2}}\mathrm{d}x\int_{y_{j-1/2}}^{y_{j+1/2}}\mathrm{d}y\int_{z_{k-1/2}}^{z_{k+1/2}}\mathrm{d}z\,\delta(\boldsymbol{r}-\boldsymbol{r}')\right)\\&=\delta_{ii'}\delta_{jj'}\delta_{kk'}\Delta x_i\Delta y_j\Delta z_i\end{aligned} \tag{2.147}$$

Transformation of the Green's functions follows the same way:

$$\mathrm{T}^{\mathrm{PE}}_{i'j'k'}\bar{G}^{f/\varphi,\mathrm{PE}}(a;\boldsymbol{r}')=G^{f/\varphi,\mathrm{PE}}(a;a')\Delta x_i\Delta y_j\Delta z_i \tag{2.148}$$

Therefore, we can divide everything by the box volume and get:

$$\sum_{b\in\tilde{\mathrm{D}}_{\mathrm{PE}}}\frac{\partial \mathrm{F}_{\mathrm{PE}}(a)}{\partial\varphi(b)}G^{\varphi,\mathrm{PE}}(b;a')+\sum_{\beta\in\tilde{\mathrm{D}}_{\mathrm{BE}}}\frac{\partial \mathrm{F}_{\mathrm{PE}}(a)}{\partial f(\beta)}G^{f,\mathrm{PE}}(\beta;a')=\delta_{aa'}\,. \tag{2.149}$$

The left-hand side is, analogous to the Langevin BE, the same discretization used in small signal calculations, whereas the right-hand side (ignoring the BCs) is an identity matrix. As for the BCs, we assume that there is no noise source for coordinates where Dirichlet BCs apply in the stationary case. That is, the right-hand side in (2.141) and (2.149) vanishes on the contacts.

Following a formal approach, Eqs. (2.141) and (2.149) can be cast into matrix form. It is important to note that in the case of spin and valley degeneracies, we cannot simply multiply the equations by the degeneracy factor and proceed analogous to the stationary case, because the Langevin sources in (2.132) represent the fluctuations in just *one* state. For example, we could have a Langevin source term for the state $(z,k,\nu,\uparrow)$ with $\uparrow$ representing the spin up particle, while there is no Langevin source for $(z,k,\nu,\downarrow)$ where $\downarrow$ stands for a spin down particle. In order to avoid double-counting of the degenerate states, we can start with the non-degenerate equation system and simplify the equations. The general matrix

equation for Green's functions looks like[15]:

$$\begin{pmatrix} \frac{\partial \hat{\mathrm{F}}_{\mathrm{PE}}}{\partial \boldsymbol{\varphi}} & \frac{\partial \hat{\mathrm{F}}_{\mathrm{PE}}}{\partial \boldsymbol{f}_\uparrow} & \frac{\partial \hat{\mathrm{F}}_{\mathrm{PE}}}{\partial \boldsymbol{f}_\downarrow} \\ \frac{\partial \hat{\mathrm{F}}_{\mathrm{BE}\uparrow}}{\partial \boldsymbol{\varphi}} & \frac{\partial \hat{\mathrm{F}}_{\mathrm{BE}\uparrow}}{\partial \boldsymbol{f}_\uparrow} & \frac{\partial \hat{\mathrm{F}}_{\mathrm{BE}\uparrow}}{\partial \boldsymbol{f}_\downarrow} \\ \frac{\partial \hat{\mathrm{F}}_{\mathrm{BE}\downarrow}}{\partial \boldsymbol{\varphi}} & \frac{\partial \hat{\mathrm{F}}_{\mathrm{BE}\downarrow}}{\partial \boldsymbol{f}_\uparrow} & \frac{\partial \hat{\mathrm{F}}_{\mathrm{BE}\downarrow}}{\partial \boldsymbol{f}_\downarrow} \end{pmatrix} \begin{pmatrix} \hat{G}^{\varphi,\mathrm{PE}} & \hat{G}^{\varphi,\mathrm{BE}\uparrow} & \hat{G}^{\varphi,\mathrm{BE}\downarrow} \\ \hat{G}^{f\uparrow,\mathrm{PE}} & \hat{G}^{f\uparrow,\mathrm{BE}\uparrow} & \hat{G}^{f\uparrow,\mathrm{BE}\downarrow} \\ \hat{G}^{f\downarrow,\mathrm{PE}} & \hat{G}^{f\downarrow,\mathrm{BE}\uparrow} & \hat{G}^{f\downarrow,\mathrm{BE}\downarrow} \end{pmatrix} = \begin{pmatrix} \hat{E}^{\mathrm{PE}} & \hat{0} & \hat{0} \\ \hat{0} & \hat{E}^{\mathrm{BE}}_{\uparrow} & \hat{0} \\ \hat{0} & \hat{0} & \hat{E}^{\mathrm{BE}}_{\downarrow} \end{pmatrix} \tag{2.150}$$

where all quantities are understood to be matrices, and the derivatives are evaluated in the noiseless steady state. Also, $\mathrm{F}_{\mathrm{BE}\uparrow}$ and $\mathrm{F}_{\mathrm{BE}\downarrow}$ denote the single-particle BE for spin up and spin down electrons, respectively. Since we want to find out how this system of equations relates to the spin-degenerate case, we can make a few simplifications. First of all, there is no transition between spin up and spin down particles, i.e.

$$\frac{\partial \hat{\mathrm{F}}_{\mathrm{BE}\uparrow}}{\partial \boldsymbol{f}_\downarrow} = \frac{\partial \hat{\mathrm{F}}_{\mathrm{BE}\downarrow}}{\partial \boldsymbol{f}_\uparrow} = \hat{0}\,. \tag{2.151}$$

Then, there is no difference to the charge density in the PE if we consider a spin up or a spin down particle, i.e.

$$\frac{\partial \hat{\mathrm{F}}_{\mathrm{PE}}}{\partial \boldsymbol{f}_\uparrow} = \frac{\partial \hat{\mathrm{F}}_{\mathrm{PE}}}{\partial \boldsymbol{f}_\downarrow} = \frac{\partial \hat{\mathrm{F}}_{\mathrm{PE}}}{\partial \boldsymbol{f}}\,. \tag{2.152}$$

Finally, the BE for spin up and spin down particles are the same,

$$\frac{\partial \hat{\mathrm{F}}_{\mathrm{BE}\uparrow}}{\partial \boldsymbol{f}_\uparrow} = \frac{\partial \hat{\mathrm{F}}_{\mathrm{BE}\downarrow}}{\partial \boldsymbol{f}_\downarrow} = \frac{\partial \hat{\mathrm{F}}_{\mathrm{BE}}}{\partial \boldsymbol{f}}\,, \tag{2.153}$$

$$\frac{\partial \hat{\mathrm{F}}_{\mathrm{BE}\uparrow}}{\partial \varphi} = \frac{\partial \hat{\mathrm{F}}_{\mathrm{BE}\downarrow}}{\partial \varphi} = \frac{\partial \hat{\mathrm{F}}_{\mathrm{BE}}}{\partial \varphi}\,. \tag{2.154}$$

Then we can write (2.150) as:

$$\begin{pmatrix} \frac{\partial \hat{\mathrm{F}}_{\mathrm{PE}}}{\partial \boldsymbol{\varphi}} & \frac{\partial \hat{\mathrm{F}}_{\mathrm{PE}}}{\partial \boldsymbol{f}} & \frac{\partial \hat{\mathrm{F}}_{\mathrm{PE}}}{\partial \boldsymbol{f}} \\ \frac{\partial \hat{\mathrm{F}}_{\mathrm{BE}}}{\partial \boldsymbol{\varphi}} & \frac{\partial \hat{\mathrm{F}}_{\mathrm{BE}}}{\partial \boldsymbol{f}} & \hat{0} \\ \frac{\partial \hat{\mathrm{F}}_{\mathrm{BE}}}{\partial \boldsymbol{\varphi}} & \hat{0} & \frac{\partial \hat{\mathrm{F}}_{\mathrm{BE}}}{\partial \boldsymbol{f}} \end{pmatrix} \begin{pmatrix} \hat{G}^{\varphi,\mathrm{PE}} & \hat{G}^{\varphi,\mathrm{BE}\uparrow} & \hat{G}^{\varphi,\mathrm{BE}\downarrow} \\ \hat{G}^{f\uparrow,\mathrm{PE}} & \hat{G}^{f\uparrow,\mathrm{BE}\uparrow} & \hat{G}^{f\uparrow,\mathrm{BE}\downarrow} \\ \hat{G}^{f\downarrow,\mathrm{PE}} & \hat{G}^{f\downarrow,\mathrm{BE}\uparrow} & \hat{G}^{f\downarrow,\mathrm{BE}\downarrow} \end{pmatrix} = \begin{pmatrix} \hat{E}_{\mathrm{PE}} & \hat{0} & \hat{0} \\ \hat{0} & \hat{E}_{\mathrm{BE}} & \hat{0} \\ \hat{0} & \hat{0} & \hat{E}_{\mathrm{BE}} \end{pmatrix} \tag{2.155}$$

[15] In order to better understand the meaning behind these matrix elements, note that the matrix element $\hat{G}(k,j)$ has something to do with the change in small signal distribution function $\delta f(k)$ when there is some change in j. Similarly, the matrix element $[\partial \hat{\mathrm{F}}_{\mathrm{BE}}/\partial \boldsymbol{f}](i,k)$ describes how much the Boltzmann operators (e.g. the free streaming term, the scattering term) change as a result of $\delta f(k)$.

What remains is to establish a relation between spin-independent Green's functions and the spin-dependent Green's functions in (2.155). Defining symmetric and antisymmetric linear combinations of the first spin arguments as:

$$\hat{G}^{f+,[\cdot]} = \hat{G}^{f\uparrow,[\cdot]} + \hat{G}^{f\downarrow,[\cdot]} \tag{2.156}$$

$$\hat{G}^{f-,[\cdot]} = \hat{G}^{f\uparrow,[\cdot]} - \hat{G}^{f\downarrow,[\cdot]} \tag{2.157}$$

where $[\cdot]$ is a placeholder for either PE, BE$\uparrow$ or BE$\downarrow$, we can obtain a matrix equation which does not depend on $G^{f-,[\cdot]}$:

$$\begin{pmatrix} \dfrac{\partial \hat{\mathrm{F}}_{\mathrm{PE}}}{\partial \boldsymbol{\varphi}} & \dfrac{\partial \hat{\mathrm{F}}_{\mathrm{PE}}}{\partial \boldsymbol{f}} \\ 2\dfrac{\partial \hat{\mathrm{F}}_{\mathrm{BE}}}{\partial \boldsymbol{\varphi}} & \dfrac{\partial \hat{\mathrm{F}}_{\mathrm{BE}}}{\partial \boldsymbol{f}} \end{pmatrix} \begin{pmatrix} \hat{G}^{\varphi,\mathrm{PE}} & \hat{G}^{\varphi,\mathrm{BE}\uparrow} \\ \hat{G}^{f+,\mathrm{PE}} & \hat{G}^{f+,\mathrm{BE}\uparrow} \end{pmatrix} = \begin{pmatrix} \hat{E}_{\mathrm{PE}} & \hat{0} \\ \hat{0} & \hat{E}_{\mathrm{BE}} \end{pmatrix} \tag{2.158}$$

Now, if the Green's function for the arbitrary observable x (e.g. the terminal current) is expressed as

$$\left(\hat{G}^{x}\right)^{\mathrm{T}} = \begin{pmatrix} \hat{P}_{\varphi}^{\mathrm{T}} & \hat{P}_{f\uparrow}^{\mathrm{T}} & \hat{P}_{f\downarrow}^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} \hat{G}^{\varphi,\mathrm{PE}} & \hat{G}^{\varphi,\mathrm{BE}\uparrow} & \hat{G}^{\varphi,\mathrm{BE}\downarrow} \\ \hat{G}^{f\uparrow,\mathrm{PE}} & \hat{G}^{f\uparrow,\mathrm{BE}\uparrow} & \hat{G}^{f\uparrow,\mathrm{BE}\downarrow} \\ \hat{G}^{f\downarrow,\mathrm{PE}} & \hat{G}^{f\downarrow,\mathrm{BE}\uparrow} & \hat{G}^{f\downarrow,\mathrm{BE}\downarrow} \end{pmatrix}, \tag{2.159}$$

the transformation to symmetric and antisymmetric Green's functions in (2.157) yields:

$$\begin{aligned} \left(\hat{G}^{x}\right)^{\mathrm{T}} &= \begin{pmatrix} \hat{P}_{\varphi}^{\mathrm{T}} & \hat{P}_{f\uparrow}^{\mathrm{T}} & \hat{P}_{f\downarrow}^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} \hat{I} & \hat{0} & \hat{0} \\ \hat{0} & \frac{1}{2}\hat{I} & \frac{-1}{2}\hat{I} \\ \hat{0} & \frac{1}{2}\hat{I} & \frac{1}{2}\hat{I} \end{pmatrix} \begin{pmatrix} \hat{G}^{\varphi,\mathrm{PE}} & \hat{G}^{\varphi,\mathrm{BE}\uparrow} & \hat{G}^{\varphi,\mathrm{BE}\downarrow} \\ \hat{G}^{f+,\mathrm{PE}} & \hat{G}^{f+,\mathrm{BE}\uparrow} & \hat{G}^{f+,\mathrm{BE}\downarrow} \\ \hat{G}^{f-,\mathrm{PE}} & \hat{G}^{f-,\mathrm{BE}\uparrow} & \hat{G}^{f-,\mathrm{BE}\downarrow} \end{pmatrix} \\ &= \begin{pmatrix} \hat{P}_{\varphi}^{\mathrm{T}} & \frac{1}{2}(\hat{P}_{f\uparrow}^{\mathrm{T}} + \hat{P}_{f\downarrow}^{\mathrm{T}}) & \frac{1}{2}(\hat{P}_{f\uparrow}^{\mathrm{T}} - \hat{P}_{f\downarrow}^{\mathrm{T}}) \end{pmatrix} \begin{pmatrix} \hat{G}^{\varphi,\mathrm{PE}} & \hat{G}^{\varphi,\mathrm{BE}\uparrow} & \hat{G}^{\varphi,\mathrm{BE}\downarrow} \\ \hat{G}^{f+,\mathrm{PE}} & \hat{G}^{f+,\mathrm{BE}\uparrow} & \hat{G}^{f+,\mathrm{BE}\downarrow} \\ \hat{G}^{f-,\mathrm{PE}} & \hat{G}^{f-,\mathrm{BE}\uparrow} & \hat{G}^{f-,\mathrm{BE}\downarrow} \end{pmatrix} \\ &= \begin{pmatrix} \hat{P}_{\varphi}^{\mathrm{T}} & \hat{P}_{f\uparrow}^{\mathrm{T}} & \hat{0} \end{pmatrix} \begin{pmatrix} \hat{G}^{\varphi,\mathrm{PE}} & \hat{G}^{\varphi,\mathrm{BE}\uparrow} & \hat{G}^{\varphi,\mathrm{BE}\downarrow} \\ \hat{G}^{f+,\mathrm{PE}} & \hat{G}^{f+,\mathrm{BE}\uparrow} & \hat{G}^{f+,\mathrm{BE}\downarrow} \\ \hat{G}^{f-,\mathrm{PE}} & \hat{G}^{f-,\mathrm{BE}\uparrow} & \hat{G}^{f-,\mathrm{BE}\downarrow} \end{pmatrix} \\ &= \begin{pmatrix} \hat{P}_{\varphi}^{\mathrm{T}} & \hat{P}_{f\uparrow}^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} \hat{G}^{\varphi,\mathrm{PE}} & \hat{G}^{\varphi,\mathrm{BE}\uparrow} & \hat{G}^{\varphi,\mathrm{BE}\uparrow} \\ \hat{G}^{f+,\mathrm{PE}} & \hat{G}^{f+,\mathrm{BE}\uparrow} & \hat{G}^{f+,\mathrm{BE}\uparrow} \end{pmatrix} \end{aligned} \tag{2.160}$$

where we have used the assumption $\hat{\mathrm{P}}_{f\uparrow} = \hat{\mathrm{P}}_{f\downarrow}$. For a detailed treatment of the degenerate states, see [69, 171].

2.7.2 Ramo-Shockley Theorem

In order to derive the Green's functions of the terminal current G^{I_k} from the Green's functions of distribution function and potential, we use the same method as in 2.6.1. Although in most cases the Green's functions of arbitrary variables can be obtained simply by replacing the distribution function and potential by their corresponding Green's functions, special care must be taken with the terminal currents. It must be noted that in 2.6.1 we relied on the continuity equations to derive an expression for the terminal currents. However, the continuity equation for the Green's functions contains an additional term due to the Langevin source appearing on the right-hand side as a generation rate. Let's start as we did in the previous section,

$$\frac{\partial}{\partial z}G_\gamma^{j^\nu}(z)+\mathrm{i}\omega G_\gamma^{n^\nu}(z)-G_\gamma^{S_\mathrm{I}^\nu}(z)-G_\gamma^{\Gamma_\mathrm{I}^\nu}(z)=\delta(z-z')\delta_{\nu\nu'} \tag{2.161}$$

where the subscript γ abbreviates the second set of coordinates. The Green's function of the terminal current for the $k-$th contact is:

$$G_\gamma^{I_k}=-\int_{\partial\mathrm{D}_k}\Big(\sum_\nu G_\gamma^{J^\nu}(\boldsymbol{r})+\mathrm{i}\omega\epsilon(\boldsymbol{r})\nabla G_\gamma^{\varphi}(\boldsymbol{r})\Big)\cdot\mathrm{d}\boldsymbol{A} \tag{2.162}$$

and multiply the integrand with $\tau_k(\boldsymbol{r})$,

$$\begin{aligned}
G_\gamma^{I_k}&=-\int_{\partial\mathrm{D}_k}\Big(\sum_\nu G_\gamma^{J^\nu}(\boldsymbol{r})+\mathrm{i}\omega\epsilon(\boldsymbol{r})\nabla G_\gamma^{\varphi}(\boldsymbol{r})\Big)\cdot\mathrm{d}\boldsymbol{A}\\
&=-\oint_{\partial\mathrm{D}}\tau_k(\boldsymbol{r})\Big(\sum_\nu G_\gamma^{J^\nu}(\boldsymbol{r})+\mathrm{i}\omega\epsilon(\boldsymbol{r})\nabla G_\gamma^{\varphi}(\boldsymbol{r})\Big)\cdot\mathrm{d}\boldsymbol{A}\\
&=\int_\mathrm{D}\nabla_{\boldsymbol{r}}\cdot\Big[\tau_k(\boldsymbol{r})\Big(\sum_\nu G_\gamma^{J^\nu}(\boldsymbol{r})+\mathrm{i}\omega\epsilon(\boldsymbol{r})\nabla G_\gamma^{\varphi}(\boldsymbol{r})\Big)\Big]\,\mathrm{d}V\\
&=\int_\mathrm{D}\tau_k(\boldsymbol{r})\nabla_{\boldsymbol{r}}\cdot\Big(\sum_\nu G_\gamma^{J^\nu}(\boldsymbol{r})+i\omega\epsilon(\boldsymbol{r})\nabla G_\gamma^{\varphi}(\boldsymbol{r})\Big)\,\mathrm{d}V\\
&\quad+\int_\mathrm{D}\Big(\sum_\nu G_\gamma^{J^\nu}(\boldsymbol{r})+\mathrm{i}\omega\epsilon(\boldsymbol{r})\nabla G_\gamma^{\varphi}(\boldsymbol{r})\Big)\nabla_{\boldsymbol{r}}\tau_k(\boldsymbol{r})\,\mathrm{d}V=G_\gamma^{I_k,(1)}+G_\gamma^{I_k,(2)}
\end{aligned}$$

For calculating $G_\gamma^{I_k,(1)}$, we use the Langevin BE continuity equation:

$$\begin{aligned}
G_\gamma^{I_k,(1)}&=\int_\mathrm{D}\tau_k(\boldsymbol{r})\nabla_{\boldsymbol{r}}\cdot\Big(\sum_\nu G_\gamma^{J^\nu}(\boldsymbol{r})+\mathrm{i}\omega\epsilon(\boldsymbol{r})\nabla G_\gamma^{\varphi}(\boldsymbol{r})\Big)\,\mathrm{d}V\\
&=\int_\mathrm{D}\tau_k(\boldsymbol{r})\sum_\nu\Big(-\mathrm{i}\omega G_\gamma^{n^\nu}(z)|\psi(\boldsymbol{r})|^2-2\mathrm{i}\omega n^\nu(z)\psi^\nu(\boldsymbol{r})G_\gamma^{\psi^\nu}(\boldsymbol{r})\\
&\quad+\big(G_\gamma^{S_\mathrm{I}^\nu}(z)+G_\gamma^{\Gamma_\mathrm{I}^\nu}(z)+\xi\delta_{\alpha;\gamma}\big)|\psi^\nu(\boldsymbol{r})|^2+\big(S_\mathrm{I}^\nu(z)+\Gamma_\mathrm{I}^\nu(z)\big)\psi^\nu(\boldsymbol{r})G_\gamma^{\psi^\nu}(\boldsymbol{r})
\end{aligned}$$

$$
\begin{aligned}
&+ \mathrm{i}\omega G_\gamma^{n^\nu}(z)|\psi^\nu(\boldsymbol{r})|^2 + 2\mathrm{i}\omega n^\nu(z)\psi^\nu(\boldsymbol{r})G_\gamma^{\psi^\nu}(\boldsymbol{r})\Big)\,\mathrm{d}V \\
&= \int_\mathrm{D} \mathrm{d}V\,\tau_k(\boldsymbol{r}) \sum_\nu \Big(\big(G_\gamma^{S_\mathrm{I}^\nu}(z) + G_\gamma^{\Gamma_\mathrm{I}^\nu}(z) + \xi\delta_{\alpha,\gamma}\big)|\psi^\nu(\boldsymbol{r})|^2 \\
&\quad + \big(S_\mathrm{I}^\nu(z) + \Gamma_\mathrm{I}^\nu(z)\big)\psi^\nu(\boldsymbol{r})G_\gamma^{\psi^\nu}(\boldsymbol{r})\Big)
\end{aligned} \tag{2.163}
$$

and for calculating $G_\gamma^{I_k,(2)}$,

$$
\begin{aligned}
G_\gamma^{I_k,(2)} &= \int_\mathrm{D} \Big(\sum_\nu G_\gamma^{J^\nu}(\boldsymbol{r}) + \mathrm{i}\omega\epsilon(\boldsymbol{r})\nabla G_\gamma^{\varphi}(\boldsymbol{r})\Big) \cdot \nabla_{\boldsymbol{r}}\tau_k(\boldsymbol{r})\,\mathrm{d}V \\
&= \int_\mathrm{D} \sum_\nu G_\gamma^{J^\nu}(\boldsymbol{r}) \cdot \nabla_{\boldsymbol{r}}\tau_k(\boldsymbol{r})\,\mathrm{d}V + \mathrm{i}\omega \int_\mathrm{D} \epsilon(\boldsymbol{r})\nabla G_\gamma^{\varphi}(\boldsymbol{r}) \cdot \nabla_{\boldsymbol{r}}\tau_k(\boldsymbol{r})\,\mathrm{d}V
\end{aligned} \tag{2.164}
$$

We note that in the small signal case, the last term was reformulated by breaking the potential up into a space charge and a contact bias part. The former vanished, leaving only capacitance terms proportional to the boundary conditions. However, since we are evaluating Green's functions at the coordinates γ, there are no Langevin sources within the Poisson part and hence the Green's functions of the potential $G_\gamma^{\varphi}(\boldsymbol{r})$, vanish on all boundaries. That means that the last term can be dropped. Therefore, what remains is to calculate the first integral in (2.164). This term can be rearranged just as in the small signal case. Thus, the Green's function of the terminal current reads:

$$
\begin{aligned}
G_\gamma^{I_k} = \sum_\nu \int_\mathrm{D} \mathrm{d}V\,\tau_k(\boldsymbol{r}) \Big[&2\mathrm{i}\omega n^\nu(z)\psi^\nu(\boldsymbol{r})G_\gamma^{\psi^\nu}(\boldsymbol{r}) \\
&+ \Big(G_\gamma^{S_\mathrm{I}^\nu}(z) - \frac{\partial}{\partial z} G_\gamma^{j^\nu}(z) + \xi\delta_{\alpha,\gamma}\left(1 - \delta(z' - z_\mathrm{S}) - \delta(z' - z_\mathrm{D})\right)\Big)|\psi^\nu(\boldsymbol{r})|^2 \Big].
\end{aligned} \tag{2.165}
$$

Calculating the matrix of all Green's functions $\hat{G}$ is computationally quite intensive. Fortunately, this is not necessary since we are often interested in the PSD of a certain observable, and the generalized adjoint approach can be exploited to greatly reduce the computational burden. For example, if we express the Green's functions of terminal current as $(\boldsymbol{G}^{I_k})^\mathrm{T} = \boldsymbol{a}^\mathrm{T}\hat{G}$, with $\boldsymbol{G}^{I_k}$ and $\boldsymbol{a}$ being $1 \times n_\mathrm{G}$ vectors and $\hat{G}$ being the $n_\mathrm{G} \times n_\mathrm{G}$ Green's functions matrix, we can just solve for $\vec{G}^{I_k}$ as follows,

$$
\hat{L}\hat{G} = \hat{E} \Rightarrow (\boldsymbol{G}^{I_k})^\mathrm{T} = \boldsymbol{a}^\mathrm{T}\hat{L}^{-1}\hat{E} \tag{2.166}
$$

Thus, we can calculate $\boldsymbol{G}^{I_k}$ by first finding $\boldsymbol{b}^\mathrm{T} = \boldsymbol{a}^\mathrm{T}\hat{L}^{-1}$ and then computing $\boldsymbol{G}^{I_k} = \hat{E}\boldsymbol{b}$. Finding $\vec{b}$ amounts to solving $\hat{L}\boldsymbol{b} = \boldsymbol{a}$ using a standard linear-equation solver and is much faster than the computation of $\hat{G}$.

2.7.3 PSD of the Noise Sources

Finally, we need to identify various contributions to noise in our device and calculate their PSD. The first source of stochastic fluctuations is the scattering term in (2.22). The one-sided PSD of fluctuations for the instantaneous scattering integral (2.42) can be expressed with the white and local noise of the transition rate [149]

$$P_{\xi\xi} = 2L_{\text{sys}}\delta(z-z')\sum_{\eta}\left(1-f^{\nu}(z,k)\right)W^{\nu\nu'}(z;k|k')f^{\nu'}(z,k') \tag{2.167}$$

where the factor of two is due to the scattering noise being a Poisson process [149]. Note, that the above description of the scattering noise sources is consistent with the transport parameters and no additional model parameter beyond the BE needs to be introduced. Using the Wiener-Lee theorem [172], the PSD of other quantities can be obtained with their corresponding transfer functions. In the case of scattering noise sources in the Langevin BE, the expression for the PSD of arbitrary variables $x(\boldsymbol{s},t)$ and $y(\boldsymbol{s}',t')$ is derived as

$$\begin{aligned} P_{xy}(\boldsymbol{s};\boldsymbol{s}';\omega) = &\sum_{\nu}\sum_{\nu'}\iint \mathrm{d}z\,\mathrm{d}z'\iint\frac{\mathrm{d}k}{2\pi}\frac{\mathrm{d}k'}{2\pi}P_{\xi\xi}^{\nu\nu'}(z,k;z',k') \\ &\times\Big((G^{x})^{\nu}(\boldsymbol{s};z,k;\omega)-(G^{x})^{\nu'}(\boldsymbol{s};z',k';\omega)\Big) \\ &\times\Big((G^{y})^{\nu}(\boldsymbol{s}';z,k;\omega)-(G^{y})^{\nu'}(\boldsymbol{s}';z',k';\omega)\Big)\,, \end{aligned} \tag{2.168}$$

because scattering events can be interpreted as perfectly correlated annihilation of electrons at the initial states and subsequent creation of electrons at the final states. Hence, the response of the arbitrary variable $x(\boldsymbol{s},t)$ to an scattering event from the initial state (z',k',ν') to the final state (z,k,ν) is

$$(G^{x})^{\nu}(\boldsymbol{s};z,k;\omega)-(G^{x})^{\nu'}(\boldsymbol{s};z',k';\omega) \tag{2.169}$$

where different signs of the Green's functions correspond to responses to the annihilation and creation of electrons. Having calculated the Green's function of the terminal current G^{I_k} in (2.165)

$$\delta I_k(\omega) = \sum_{\nu'}\int \mathrm{d}z'\int\frac{\mathrm{d}k'}{2\pi}(G^{I_k})^{\nu'}(z',k';\omega)\xi_{\text{BE}}^{\nu'}(z',k';\omega)\,, \tag{2.170}$$

the PSD of terminal currents $P_{kk'}(\omega)$ is obtained using (2.168). The position dependent PSD of the terminal currents is

$$K_{kk'}^{(\text{PH})}(z;\omega) = \sum_{\eta}\sum_{\nu,\nu'}\iint\frac{\mathrm{d}k}{2\pi}\frac{\mathrm{d}k'}{2\pi}c_{\eta}^{\nu\nu'}(z;k,k')f_{\text{DC}}^{\nu'}(z,k')\big(1-f_{\text{DC}}^{\nu}(z,k)\big)$$

$$\times \Big((G^{I_k})^{\nu}(z,k;\omega) - (G^{I_k})^{\nu'}(z,k';\omega) \Big)$$
$$\times \Big((G^{I_{k'}})^{\nu}(z,k;\omega) - (G^{I_{k'}})^{\nu'}(z,k';\omega) \Big)^{*} . \quad (2.171)$$

where $c_{\eta}^{\nu\nu'}(z;k,k')$ is defined in (2.46).

The second source for fluctuations is the stochastic generation and recombination of carriers at the source and drain contacts, represented by Γ in (2.22). This process can be interpreted as a scattering term between the subbands of the semiconductor and a fictional thermal bath which provides the equilibrium distribution function. Hence, similar to the scattering integral, the white noise of the generation/recombination rate is defined as

$$P_{1\,\xi\xi}^{\nu\nu'}(z,k;z',k') = 4\pi\delta_{\nu\nu'}\delta(z-z')\delta(k-k')\left[\delta(z-z_{\mathrm{S}})+\delta(z-z_{\mathrm{D}})\right] \times v_{\mathrm{GR}}\big(1-f^{\nu}(z,k)\big) f_{\mathrm{eq}}^{\nu'}(z',k') , \quad (2.172)$$

$$P_{2\,\xi\xi}^{\nu\nu'}(z,k;z',k') = 4\pi\delta_{\nu\nu'}\delta(z-z')\delta(k-k')\left[\delta(z-z_{\mathrm{S}})+\delta(z-z_{\mathrm{D}})\right] \times v_{\mathrm{GR}} f^{\nu}(z,k)\big(1-f_{\mathrm{eq}}^{\nu'}(z',k') . \quad (2.173)$$

which gives:

$$K_{kk'}^{(\mathrm{GR})}(z;\omega) = \quad (2.174)$$
$$\sum_{\nu,\nu'} \int \mathrm{d}z' \iint \frac{\mathrm{d}k}{2\pi}\frac{\mathrm{d}k'}{2\pi} (G^{I_k})^{\nu}(z,k;\omega) P_{1\,\xi\xi}^{\nu\nu'}(z,k;z',k') \left((G^{I_{k'}})^{\nu}(z,k;\omega)\right)^{*}$$
$$+\sum_{\nu,\nu'} \int \mathrm{d}z' \iint \frac{\mathrm{d}k}{2\pi}\frac{\mathrm{d}k'}{2\pi} \big(-(G^{I_k})^{\nu'}(z',k';\omega)\big) P_{2\,\xi\xi}^{\nu\nu'}(z,k;z',k') \big(-(G^{I_{k'}})^{\nu'}(z',k';\omega)\big)^{*}$$
$$= 2\sum_{\nu} \int \frac{\mathrm{d}k}{2\pi} (G^{I_k})^{\nu}(z,k;\omega) \left((G^{I_{k'}})^{\nu}(z,k;\omega)\right)^{*}$$
$$\times v_{\mathrm{GR}} \left[f^{\nu}(z,k) + f_{\mathrm{eq}}^{\nu}(z,k) - 2 f^{\nu}(z,k) f_{\mathrm{eq}}^{\nu}(z,k) \right] . \quad (2.175)$$

The total PSD is:

$$P_{kk'}(\omega) = \int \mathrm{d}z \left[K_{kk'}^{(\mathrm{PH})}(z;\omega) + K_{kk'}^{(\mathrm{GR})}(z;\omega) \right] . \quad (2.176)$$

Chapter 3

Moments Equations

3.1 Introduction

The macroscopic carrier transport equations can be obtained from the moments of the BE. The idea of projection of the BE onto Hermite polynomial bases for the case of Cartesian geometries was originally proposed by Grad in 1949 [173], and developed further by Shan, Chen, and co-workers [174,175]. This expansion maps the continuous k–space onto the space of expansion coefficients, yielding an open-ended hierarchy of first order nonlinear PDEs for the corresponding kinetic moments. The reason Hermite polynomials are best suited as the expansion basis rather than, say, Chebyshev or Legendre polynomials, is that in the case of Hermite polynomials, the expansion coefficients correspond exactly to the macroscopic transport variables up to the chosen degree [175].

The Hermite polynomials[1] are a sequence of classical orthogonal polynomials over the interval $(-\infty, \infty)$ with the kernel $\omega(x) = \exp\left(-x^2\right)$. They can be defined by means of the Rodrigues formula [176]

$$\mathcal{H}_n(x) = \frac{(-1)^n}{\omega(x)} \frac{\mathrm{d}^n \omega(x)}{\mathrm{d}x^n} \qquad n = 0, 1, 2, \ldots \tag{3.1}$$

and their scalar product satisfies the relation

$$\int_{-\infty}^{+\infty} \omega(x) \mathcal{H}_n(x) \mathcal{H}_m(x)\, \mathrm{d}x = 2^n n! \sqrt{\pi} \delta_{nm}\,. \tag{3.2}$$

The generalized relation

$$f(x) = \sum_{n=0}^{\infty} \frac{a_n \mathcal{H}_n(x)}{2^n n! \sqrt{\pi}}\,, \quad a_n = \int_{-\infty}^{\infty} f(x) \mathcal{H}_n(x) \omega(x)\, \mathrm{d}x \tag{3.3}$$

[1] In this work, the so-called physicists' Hermite polynomials are used.

projects the function $f(x)$ onto the Hermite bases, resulting in a set of expansion coefficients a_n. This is provided that $f(x)$ is piecewise smooth in every finite interval, and

$$\int_{-\infty}^{\infty} \exp\bigl(-x^2\bigr)|f(x)|^2\,\mathrm{d}x < \infty\,. \tag{3.4}$$

Because of the leapfrog properties of the coefficients in the expansion of $\mathcal{H}_n$, the expansion (3.3) splits into a part which is even in x and a part which is odd. There are useful relations among the polynomials and their derivatives, which can be found in handbooks such as [176]. Some of the useful recurrence relations are as follows:

$$\frac{\mathrm{d}\mathcal{H}_n(x)}{\mathrm{d}x} = 2n\mathcal{H}_{n-1}(x) \tag{3.5}$$

$$2x\mathcal{H}_n(x) = 2n\mathcal{H}_{n-1}(x) + \mathcal{H}_{n+1}(x) \tag{3.6}$$

$$\frac{\mathrm{d}\mathcal{H}_n(x)}{\mathrm{d}x} = 2x\mathcal{H}_n(x) - \mathcal{H}_{n+1}(x) \tag{3.7}$$

In this chapter, projection of the multi–subband BE onto the Hermite polynomials is presented and an approach based on the matrix exponentials is employed for the stabilization of discretized equations and improving the accuracy of the solution. Sections 3.2 and 3.3 treat the stationary and small signal problems, respectively.

3.2 Steady State Equations

3.2.1 Main Equation (RTA)

Starting from the multi-subband BE under the assumptions of parabolic band structure and RTA for the scattering term, which reads:

$$\frac{\partial f_k^\nu}{\partial t} + \frac{1}{\hbar}F^\nu(z,t)\frac{\partial f_k^\nu}{\partial k} + \frac{\hbar k}{m_{zz}^\nu}\frac{\partial f_k^\nu}{\partial z} = -\frac{f_k^\nu - f_{k\,\mathrm{eq}}^\nu}{\tau}\,, \tag{3.8}$$

we use the transformation[2]

$$f_k^\nu(z,k,t) = g^\nu\left(z, \frac{\hbar k}{\sqrt{2m_{zz}^\nu k_\mathrm{B}T}}, t\right)\exp\left(\frac{-\hbar^2k^2}{2m_{zz}^\nu k_\mathrm{B}T}\right) = g^\nu(z,u,t)\exp\left(-u^2\right) \tag{3.9}$$

[2]Based on this transformation, the Maxwellian will be expressed by only one nonzero coefficient a_0. Moreover, the zeroth and first expansion coefficients provide a natural description of charge and current density, respectively.

and after factoring out the weight function $\omega(u)$, obtain the following PDE for the new function $g^\nu(z,u,t)$:

$$\frac{\partial g^\nu}{\partial t} + \frac{F^\nu(z,t)}{\sqrt{2m^\nu_{zz}k_\mathrm{B}T}}\left(\frac{\partial g^\nu}{\partial u} - 2ug^\nu\right) + \sqrt{\frac{2k_\mathrm{B}T}{m^\nu_{zz}}}u\frac{\partial g^\nu}{\partial z} = -\frac{g^\nu - a^\nu_0/\sqrt{\pi}}{\tau}\,. \tag{3.10}$$

Projection of (3.10) onto the Hermite polynomial $\mathcal{H}_n(u)$ and using the recurrence relations (3.5)–(3.7) yields:

$$\frac{\partial a^\nu_n}{\partial t} + \frac{F^\nu(z,t)}{\sqrt{2m^\nu_{zz}k_\mathrm{B}T}}2na^\nu_{n-1} + \sqrt{\frac{2k_\mathrm{B}T}{m^\nu_{zz}}}\left(\frac{1}{2}\frac{\partial a^\nu_{n+1}}{\partial z} + n\frac{\partial a^\nu_{n-1}}{\partial z}\right) = -\frac{a^\nu_n - a^\nu_0\delta_{n,0}}{\tau}\,. \tag{3.11}$$

The BE, therefore, is rewritten as an infinite set of coupled PDEs for the expansion coefficients. In order to make this "infinite moment system" numerically solvable, (3.3) is approximated by

$$f(x) \approx \sum_{n=0}^{n_\mathrm{HP}} \frac{a_n\mathcal{H}_n(x)}{2^n n!\sqrt{\pi}}\,, \tag{3.12}$$

for the positive integer $n_\mathrm{HP} \geq 2$. This results in n_HP equations each containing information about the next one and describing the conservation of electron density, momentum, energy, and higher–order quantities. The coefficients for $n \geq n_\mathrm{HP}$ can be simply set to zero, or truncations based on various closure relations [173, 177–179] or extraction of the n_HPth moment from MC simulations [178] can be employed. The former choice has the advantage that the equations remain linear and is therefore used in this work, but it can lead to stability problems for large electric fields.

Defining $\vec{a}$ as the vector of n_HP expansion coefficients, (3.11) is cast into the matrix equation

$$\frac{\partial \vec{a}}{\partial t} + \hat{A}\frac{\partial \vec{a}}{\partial z} + \hat{B}(z,t)\vec{a} = \vec{0}\,. \tag{3.13}$$

As an example, for $n_\mathrm{HP} = 4$ the matrices $\hat{A}$ and $\hat{B}$ are given as:

$$\hat{A} = \sqrt{\frac{2k_\mathrm{B}T}{m^\nu_{zz}}}\begin{pmatrix} 0 & \frac{1}{2} & 0 & 0 \\ 1 & 0 & \frac{1}{2} & 0 \\ 0 & 2 & 0 & \frac{1}{2} \\ 0 & 0 & 3 & 0 \end{pmatrix}\,, \qquad \hat{B} = \frac{2F(z,t)}{\sqrt{2m^\nu_{zz}k_\mathrm{B}T}}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \end{pmatrix} + \frac{1}{\tau}\hat{I}'_{4\times 4}\,. \tag{3.14}$$

where $\hat{I}'$ is the identity matrix except for $\hat{I}'(1,1) = 0$. Under the assumption of position independent electric field in each box $z \in [z_i, z_{i+1}]$ and steady state conditions, (3.13) reduces to a linear system of differential equations of the first order for which the solution is simply given by

$$\vec{a}(z) = \exp\left(-\hat{D}_{i+1/2}(z)\right)\vec{a}(z_{i+1/2})\,. \tag{3.15}$$

In (3.15) we have used[3]

$$\hat{D}_{i+1/2}(z) = \hat{A}^{-1}\hat{B}(z_{i+1/2})(z - z_{i+1/2})\,. \tag{3.16}$$

If we split the vector $\vec{a}$ into its even and odd components with the non-square projection matrices of size $n_{\mathrm{HP}}/2 \times n_{\mathrm{HP}}$ (i.e. $\vec{a}_{\mathrm{e/o}} = \hat{\mathrm{P}}_{\mathrm{e/o}}\vec{a}$), we have:

$$\begin{pmatrix} \vec{a}_{\mathrm{e}}(z_i) \\ \vec{a}_{\mathrm{e}}(z_{i+1}) \end{pmatrix} = \begin{pmatrix} \hat{\mathrm{P}}_{\mathrm{e}} \exp\left(-\hat{D}_{i+1/2}(z_i)\right) \\ \hat{\mathrm{P}}_{\mathrm{e}} \exp\left(-\hat{D}_{i+1/2}(z_{i+1})\right) \end{pmatrix} \vec{a}(z_{i+1/2})\,, \tag{3.17}$$

$$\begin{pmatrix} \vec{a}_{\mathrm{o}}(z_i) \\ \vec{a}_{\mathrm{o}}(z_{i+1}) \end{pmatrix} = \begin{pmatrix} \hat{\mathrm{P}}_{\mathrm{o}} \exp\left(-\hat{D}_{i+1/2}(z_i)\right) \\ \hat{\mathrm{P}}_{\mathrm{o}} \exp\left(-\hat{D}_{i+1/2}(z_{i+1})\right) \end{pmatrix} \vec{a}(z_{i+1/2})\,. \tag{3.18}$$

Substitution of (3.17) into (3.18) gives

$$\begin{pmatrix} \vec{a}_{\mathrm{o}}(z_i) \\ \vec{a}_{\mathrm{o}}(z_{i+1}) \end{pmatrix} = \begin{pmatrix} \hat{\mathrm{P}}_{\mathrm{o}} \exp\left(-\hat{D}_{i+1/2}(z_i)\right) \\ \hat{\mathrm{P}}_{\mathrm{o}} \exp\left(-\hat{D}_{i+1/2}(z_{i+1})\right) \end{pmatrix} \begin{pmatrix} \hat{\mathrm{P}}_{\mathrm{e}} \exp\left(-\hat{D}_{i+1/2}(z_i)\right) \\ \hat{\mathrm{P}}_{\mathrm{e}} \exp\left(-\hat{D}_{i+1/2}(z_{i+1})\right) \end{pmatrix}^{-1} \begin{pmatrix} \vec{a}_{\mathrm{e}}(z_i) \\ \vec{a}_{\mathrm{e}}(z_{i+1}) \end{pmatrix}. \tag{3.19}$$

Using the continuity of flux on the grid nodes, (3.19) provides an even/odd coupling for the coefficients which links $\vec{a}_{\mathrm{e}}(z_{i-1})$, $\vec{a}_{\mathrm{e}}(z_i)$ and $\vec{a}_{\mathrm{e}}(z_{i+1})$. Hence, we have a well–defined system of equations provided that suitable boundary conditions are imposed on the first and last grid nodes. That is, if we take on the notation

$$\begin{pmatrix} \hat{\mathrm{P}}_{\mathrm{o}} \exp\left(-\hat{D}_{i+1/2}(z_i)\right) \\ \hat{\mathrm{P}}_{\mathrm{o}} \exp\left(-\hat{D}_{i+1/2}(z_{i+1})\right) \end{pmatrix} \begin{pmatrix} \hat{\mathrm{P}}_{\mathrm{e}} \exp\left(-\hat{D}_{i+1/2}(z_i)\right) \\ \hat{\mathrm{P}}_{\mathrm{e}} \exp\left(-\hat{D}_{i+1/2}(z_{i+1})\right) \end{pmatrix}^{-1} = \begin{pmatrix} \hat{\mathcal{B}}^{(11)}_{i,i+1} & \hat{\mathcal{B}}^{(12)}_{i,i+1} \\ \hat{\mathcal{B}}^{(21)}_{i,i+1} & \hat{\mathcal{B}}^{(22)}_{i,i+1} \end{pmatrix} \tag{3.20}$$

where submatrices $\hat{\mathcal{B}}$ are of size $n_{\mathrm{HP}}/2 \times n_{\mathrm{HP}}/2$, the left-hand side matrix equation for $\vec{a}_{\mathrm{e}}$ then looks like:

$$\begin{pmatrix} \times & \times & & & \\ \hat{\mathcal{B}}^{(21)}_{1,2} & \hat{\mathcal{B}}^{(22)}_{1,2} - \hat{\mathcal{B}}^{(11)}_{2,3} & -\hat{\mathcal{B}}^{(12)}_{2,3} & & \\ 0 & \hat{\mathcal{B}}^{(21)}_{2,3} & \hat{\mathcal{B}}^{(22)}_{2,3} - \hat{\mathcal{B}}^{(11)}_{3,4} & -\hat{\mathcal{B}}^{(12)}_{3,4} & \\ & & & \ddots & \\ & & & \times & \times \end{pmatrix} \begin{pmatrix} \vec{a}_{\mathrm{e}}(z_1) \\ \vec{a}_{\mathrm{e}}(z_2) \\ \vec{a}_{\mathrm{e}}(z_3) \\ \vdots \\ \vec{a}_{\mathrm{e}}(z_{\mathrm{N}}) \end{pmatrix} = \begin{pmatrix} \times \\ \vec{0} \\ \vec{0} \\ \vdots \\ \times \end{pmatrix}, \tag{3.21}$$

where $\times$ stands for the modifications due to the boundary conditions. After solving (3.21) for the even coefficients, the odd ones are recovered by using (3.19). The electron density $n(z)$ and current density $j(z)$ along the transport direction are then calculated as follows:

$$n(z) = 2\sum_{\nu} \int_{-\infty}^{\infty} f^{\nu}_k(z,k) \frac{\mathrm{d}k}{2\pi} = \sum_{\nu} \frac{\sqrt{2m^{\nu}_{zz} k_{\mathrm{B}} T}}{\pi\hbar} \int_{-\infty}^{\infty} g^{\nu}(z,u)\omega(u)\,\mathrm{d}u$$

[3] The matrix $\hat{A}$ is always invertible.

$$= \sum_\nu \frac{\sqrt{2m^\nu_{zz}k_\mathrm{B}T}}{\pi\hbar}\int_{-\infty}^{\infty} g^\nu(z,u)\omega(u)\mathcal{H}_0(u)\,\mathrm{d}u = \sum_\nu \frac{\sqrt{2m^\nu_{zz}k_\mathrm{B}T}}{\pi\hbar}a^\nu_0(z) \tag{3.22}$$

$$\begin{aligned} j(z) &= 2\sum_\nu \int_{-\infty}^{\infty} v^\nu_z(k) f^\nu_k(z,k)\frac{\mathrm{d}k}{2\pi} = \frac{2k_\mathrm{B}T}{\pi\hbar}\sum_\nu \int_{-\infty}^{\infty} u \sum_{m=0}^{n_\mathrm{HP}} \frac{a^\nu_m(z)\mathcal{H}_m(u)}{2^m m!\sqrt{\pi}}\omega(u)\,\mathrm{d}u \\ &= \frac{2k_\mathrm{B}T}{\pi\hbar}\sum_\nu \int_{-\infty}^{\infty} u \sum_{m=0}^{n_\mathrm{HP}} \frac{a^\nu_m(z)\mathcal{H}_m(u)}{2^m m!\sqrt{\pi}}\omega(u)\mathcal{H}_0(u)\,\mathrm{d}u = \frac{k_\mathrm{B}T}{\pi\hbar}\sum_\nu a^\nu_1(z)\,. \end{aligned} \tag{3.23}$$

For the calculation of the energy density, the recurrence relations (3.6) is used:

$$\begin{aligned} E(z) &= 2\sum_\nu \int \frac{\hbar^2k^2}{2m^\nu_{zz}} f^\nu_k(z,k)\frac{\mathrm{d}k}{2\pi} = \sum_\nu \frac{(2m^\nu_{zz}k_\mathrm{B}T)^{\frac{3}{2}}}{2m^\nu_{zz}\pi\hbar}\int u^2 \sum_{m=0}^{n_\mathrm{HP}} \frac{a^\nu_m\mathcal{H}_m(u)}{2^m m!\sqrt{\pi}}\omega(u)\,\mathrm{d}u \\ &= \sum_\nu \frac{(2m^\nu_{zz}k_\mathrm{B}T)^{\frac{3}{2}}}{2m^\nu_{zz}\pi\hbar}\int \sum_{m=0}^{n_\mathrm{HP}} \frac{a^\nu_m\left[m(m-1)\mathcal{H}_{m-2} + \frac{2m+1}{2}\mathcal{H}_m + \frac{1}{4}\mathcal{H}_{m+2}\right]}{2^m m!\sqrt{\pi}}\omega(u)\,\mathrm{d}u \\ &= \sum_\nu \frac{(2m^\nu_{zz}k_\mathrm{B}T)^{\frac{3}{2}}}{2m^\nu_{zz}\pi\hbar}\int \sum_{m=0}^{n_\mathrm{HP}} \frac{a^\nu_m\left[m(m-1)\mathcal{H}_{m-2}(u) + \frac{2m+1}{2}\mathcal{H}_m(u)\right]}{2^m m!\sqrt{\pi}}\omega(u)\mathcal{H}_0(u)\,\mathrm{d}u \\ &= \sum_\nu \frac{(2m^\nu_{zz}k_\mathrm{B}T)^{\frac{3}{2}}}{2m^\nu_{zz}\pi\hbar}\left[\frac{1}{2}a^\nu_0(z) + \frac{1}{4}a^\nu_2(z)\right]. \end{aligned} \tag{3.24}$$

3.2.2 Projection of the Scattering Integral

Neglecting the Pauli exclusion principle, the general form of the scattering integral is written as

$$S^\nu\{f_k\} = \frac{L_\mathrm{sys}}{2\pi}\sum_{\nu'}\sum_{\eta,\sigma}\int \mathrm{d}k'\left[f^{\nu'}_k(z,k')W^{\nu\nu'}_{\eta\sigma}(z;k|k') - f^\nu_k(z,k)W^{\nu'\nu}_{\eta\sigma}(z;k'|k)\right],$$

where in the case of electron–phonon scattering, the physical parameters for various intra– and inter–valley transitions are lumped into a general coefficient $c^{\nu\nu'}_{\eta\sigma}(z)$, as explained in section 2.4. The in–scattering term is

$$\begin{aligned} S^\nu_\mathrm{in}\{f_k\} &= \frac{1}{2\pi}\sum_{\nu'}\sum_{\eta,\sigma}\int \mathrm{d}k'\, f^{\nu'}_k(z,k')c^{\nu\nu'}_{\eta\sigma}(z)\delta\big(\varepsilon^\nu_\mathrm{tot}(z,k) - \varepsilon^{\nu'}_\mathrm{tot}(z,k') + \sigma\hbar\omega_\eta\big) \\ &= \frac{1}{2\pi}\sum_{\nu'}\sum_{\eta,\sigma}\left(\frac{2m^{\nu'}_{zz}}{\hbar^2}\right)c^{\nu\nu'}_{\eta\sigma}(z)\int \mathrm{d}k'\, f^{\nu'}_k(z,k')\delta\Big(k'^2 - \tfrac{m^{\nu'}_{zz}}{m^\nu_{zz}}k^2 - \alpha\Big) \\ &= \sum_{\nu'}\sum_{\eta,\sigma}\left(\frac{m^{\nu'}_{zz}}{\pi\hbar^2}\right)\frac{c^{\nu\nu'}_{\eta\sigma}(z)}{\sqrt{\frac{m^{\nu'}_{zz}}{m^\nu_{zz}}k^2+\alpha}} f^{\nu'}_\mathrm{ke}\left(z,\sqrt{\tfrac{m^{\nu'}_{zz}}{m^\nu_{zz}}k^2+\alpha}\right)\Theta\left(\tfrac{m^{\nu'}_{zz}}{m^\nu_{zz}}k^2+\alpha\right) \end{aligned}$$

with $\alpha = 2m_{zz}^{\nu'}\big(\varepsilon_{\rm sub}^{\nu}(z) - \varepsilon_{\rm sub}^{\nu'}(z) + \sigma\hbar\omega_\eta\big)/\hbar^2$. The subscript e of $f_{\rm e}$ denotes the even part of the distribution function. Similarly, the out–scattering term is simplified as:

$$\begin{aligned} S_{\rm out}^{\nu}\{f_k\} &= \frac{1}{2\pi}\sum_{\nu'}\sum_{\eta,\sigma}\int \mathrm{d}k'\, f_k^{\nu}(z,k) c_{\eta\sigma}^{\nu'\nu}(z)\delta\big(\varepsilon_{\rm tot}^{\nu}(z,k) - \varepsilon_{\rm tot}^{\nu'}(z,k') - \sigma\hbar\omega_\eta\big) \\ &= \frac{1}{2\pi}\sum_{\nu'}\sum_{\eta,\sigma}\left(\frac{2m_{zz}^{\nu'}}{\hbar^2}\right) c_{\eta\sigma}^{\nu'\nu}(z) f_k^{\nu}(z,k)\int \mathrm{d}k'\,\delta\Big(k'^2 - \tfrac{m_{zz}^{\nu'}}{m_{zz}^{\nu}}k^2 - \beta\Big) \\ &= \sum_{\nu'}\sum_{\eta,\sigma}\left(\frac{m_{zz}^{\nu'}}{\pi\hbar^2}\right)\frac{c_{\eta\sigma}^{\nu'\nu}(z)}{\sqrt{\frac{m_{zz}^{\nu'}}{m_{zz}^{\nu}}k^2+\beta}} f_k^{\nu}(z,k)\,\Theta\Big(\tfrac{m_{zz}^{\nu'}}{m_{zz}^{\nu}}k^2+\beta\Big) \end{aligned}$$

with $\beta = 2m_{zz}^{\nu'}\big(\varepsilon_{\rm sub}^{\nu}(z) - \varepsilon_{\rm sub}^{\nu'}(z) - \sigma\hbar\omega_\eta\big)/\hbar^2$. Applying the transformation of (3.9) to the in– and out–scattering terms results in

$$\begin{aligned} S_{\rm in}^{\nu}\{g\} &= \frac{1}{\pi\hbar}\sum_{\nu'}\sum_{\eta,\sigma}\frac{m_{zz}^{\nu'}}{\sqrt{2m_{zz}^{\nu}k_{\rm B}T}}\frac{c_{\eta\sigma}^{\nu\nu'}(z)e^{-\alpha'}}{\sqrt{\frac{m_{zz}^{\nu'}}{m_{zz}^{\nu}}u^2+\alpha'}}\Theta\Big(\tfrac{m_{zz}^{\nu'}}{m_{zz}^{\nu}}u^2+\alpha'\Big) \\ &\qquad g_{\rm e}^{\nu'}\left(z,\sqrt{\tfrac{m_{zz}^{\nu'}}{m_{zz}^{\nu}}u^2+\alpha'}\right)\exp\Big(\tfrac{-m_{zz}^{\nu'}}{m_{zz}^{\nu}}u^2\Big) \\ S_{\rm out}^{\nu}\{g\} &= \frac{1}{\pi\hbar}\sum_{\nu'}\sum_{\eta,\sigma}\frac{m_{zz}^{\nu'}}{\sqrt{2m_{zz}^{\nu}k_{\rm B}T}}\frac{c_{\eta\sigma}^{\nu'\nu}(z)\exp\big(-u^2\big)}{\sqrt{\frac{m_{zz}^{\nu'}}{m_{zz}^{\nu}}u^2+\beta'}}g^{\nu}(z,u)\,\Theta\Big(\frac{m_{zz}^{\nu'}}{m_{zz}^{\nu}}u^2+\beta'\Big) \end{aligned}$$

where

$$\alpha' = \frac{\hbar^2\alpha}{2m_{zz}^{\nu}k_{\rm B}T}\,,\quad \beta' = \frac{\hbar^2\beta}{2m_{zz}^{\nu}k_{\rm B}T}\,. \tag{3.25}$$

For the case $\nu = \nu'$, we have $\alpha' = -\beta' = \sigma\hbar\omega_\eta/k_{\rm B}T$. Projection of $S_{\rm in}^{\nu}\{g\}$ and $S_{\rm out}^{\nu}\{g\}$ onto the Hermite polynomials gives:

$$\begin{aligned} &\int S_{\rm in}^{\nu}\{g\}\mathcal{H}_n(u)\,\mathrm{d}u = \sum_{\nu'}\sum_{\eta,\sigma}\frac{m_{zz}^{\nu'}}{\pi\hbar\sqrt{2m_{zz}^{\nu}k_{\rm B}T}} \\ &\int \frac{c_{\eta\sigma}^{\nu\nu'}e^{-\alpha'}\Theta\Big(\frac{m_{zz}^{\nu'}}{m_{zz}^{\nu}}u^2+\alpha'\Big)}{\sqrt{\frac{m_{zz}^{\nu'}}{m_{zz}^{\nu}}u^2+\alpha'}}g_{\rm e}^{\nu'}\left(z,\sqrt{\tfrac{m_{zz}^{\nu'}}{m_{zz}^{\nu}}u^2+\alpha'}\right)\exp\Big(\tfrac{-m_{zz}^{\nu'}}{m_{zz}^{\nu}}u^2\Big)\mathcal{H}_n(u)\,\mathrm{d}u\,. \end{aligned} \tag{3.26}$$

Hence, we have for even m:

$$\int S_{\rm in}^{\nu}\{g\}\mathcal{H}_n(u)\,\mathrm{d}u = \sum_{\nu'}\sum_{\eta,\sigma}\frac{m_{zz}^{\nu'}}{\pi\hbar\sqrt{2m_{zz}^{\nu}k_{\rm B}T}}$$

$$\int \frac{c_{\eta\sigma}^{\nu\nu'} e^{-\alpha'} \Theta\left(\frac{m_{zz}^{\nu'}}{m_{zz}^{\nu}} u^2 + \alpha'\right)}{\sqrt{\frac{m_{zz}^{\nu'}}{m_{zz}^{\nu}} u^2 + \alpha'}} \sum_{m=0}^{n_{\mathrm{HP}}} \frac{a_m^{\nu'} \mathcal{H}_m\left(\sqrt{\frac{m_{zz}^{\nu'}}{m_{zz}^{\nu}} u^2 + \alpha'}\right)}{2^m m! \sqrt{\pi}} \exp\left(\frac{-m_{zz}^{\nu'}}{m_{zz}^{\nu}} u^2\right) \mathcal{H}_n(u)\,\mathrm{d}u$$

$$= \sum_{m=0}^{n_{\mathrm{HP}}} \left(\sum_{\nu'} \sum_{\eta,\sigma} \Gamma_m^{\nu\nu'} \int \frac{c_{\eta\sigma}^{\nu\nu'} e^{-\alpha'} \Theta\left(\frac{m_{zz}^{\nu'}}{m_{zz}^{\nu}} u^2 + \alpha'\right)}{\sqrt{\frac{m_{zz}^{\nu'}}{m_{zz}^{\nu}} u^2 + \alpha'}} \exp\left(\frac{-m_{zz}^{\nu'}}{m_{zz}^{\nu}} u^2\right) \mathcal{H}_m\left(\sqrt{\frac{m_{zz}^{\nu'}}{m_{zz}^{\nu}} u^2 + \alpha'}\right) \mathcal{H}_n(u)\,\mathrm{d}u \right) a_m^{\nu'} \tag{3.27}$$

which is calculated by numerical integration for all n and m values. Similarly, for the out–scattering term we have:

$$\int S_{\mathrm{out}}^{\nu}\{g\} \mathcal{H}_n(u)\,\mathrm{d}u$$

$$= \sum_{\nu'} \sum_{\eta,\sigma} \frac{m_{zz}^{\nu'}}{\pi\hbar\sqrt{2 m_{zz}^{\nu} k_{\mathrm{B}} T}} \int \frac{c_{\eta\sigma}^{\nu'\nu} \Theta\left(\frac{m_{zz}^{\nu'}}{m_{zz}^{\nu}} u^2 + \beta'\right) \mathrm{e}^{-u^2}}{\sqrt{\frac{m_{zz}^{\nu'}}{m_{zz}^{\nu}} u^2 + \beta'}} \sum_{m=0}^{n_{\mathrm{HP}}} \frac{a_m^{\nu} \mathcal{H}_m(u)}{2^m m! \sqrt{\pi}} \mathcal{H}_n(u)\,\mathrm{d}u$$

$$= \sum_{m=0}^{n_{\mathrm{HP}}} \left(\sum_{\nu'} \sum_{\eta,\sigma} \Gamma_m^{\nu\nu'} \int \frac{c_{\eta\sigma}^{\nu'\nu} \Theta\left(\frac{m_{zz}^{\nu'}}{m_{zz}^{\nu}} u^2 + \beta'\right) \mathrm{e}^{-u^2}}{\sqrt{\frac{m_{zz}^{\nu'}}{m_{zz}^{\nu}} u^2 + \beta'}} \mathcal{H}_m(u) \mathcal{H}_n(u)\,\mathrm{d}u \right) a_m^{\nu} \tag{3.28}$$

In Eqs. (3.27) and (3.28), the spatial arguments are omitted for better readability, and the pre–factor Γ_m is defined as

$$\Gamma_m^{\nu\nu'} = \left(\frac{m_{zz}^{\nu'}}{\pi\hbar\sqrt{2 m_{zz}^{\nu} k_{\mathrm{B}} T}} \right) \frac{1}{2^m m! \sqrt{\pi}}\,. \tag{3.29}$$

As an example, the phonon scattering matrix for $n_{\mathrm{HP}} = 4$ looks like,

$$\hat{S}_{\mathrm{PH}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \times & 0 & \times \\ \times & 0 & \times & 0 \\ 0 & \times & 0 & \times \end{pmatrix}. \tag{3.30}$$

It is noteworthy, that all of the elements of the first row of $\hat{S}_{\mathrm{PH}}$ are zeros, which implies that charge conservation is guaranteed for the inter– and intra–valley transitions. Moreover, Eqs. (3.27) and (3.28) clearly show that for different parities of n and m, coefficients for both a_m^{ν} and $a_m^{\nu'}$ are zero.

3.2.3 Boundary Conditions

The moments equations need to be supplemented with suitable BCs which describe the contacts. In this section, various BCs are projected onto the Hermite polynomial bases and their performance in describing the contacts will be compared later.

Fixed Boundary Conditions

Using the fixed BCs, the distribution function is specified on the source and drain boundaries. Given the equilibrium distribution function,

$$f^{\nu}_{\mathrm{eq}}(z,k) = \exp\left(-\frac{\hbar^2 k^2}{2m^{\nu}_{zz}} - \varepsilon^{\nu}_{\mathrm{sub}}(z) + \eta\right) = \gamma^{\nu}(z)\exp\left(\frac{-\hbar^2 k^2}{2m^{\nu}_{zz}}\right), \tag{3.31}$$

the fixed boundary condition requires that

$$f^{\nu}_{k}(z_{\mathrm{S/D}},k) = f^{\nu}_{\mathrm{eq}}(z_{\mathrm{S/D}},k)\,. \tag{3.32}$$

Projection of (3.32) onto Hermite polynomials results in:

$$a^{\nu}_{n}(z_{\mathrm{S/D}}) = \gamma^{\nu}_{\mathrm{S/D}}\sqrt{\pi}\delta_{n,0}\,. \tag{3.33}$$

In other words, in the matrix equation (3.21) the left-hand side is an identity matrix for the source and drain grid nodes and the right-hand side vector is zero everywhere except for the zeroth coefficient on the contacts.

Generation/Recombination Velocity

Projection of the generation/recombination BCs of (2.55) onto Hermite polynomials is straightforward, and they are added to the first and last blocks of the matrix equation. In the limit $v_{\mathrm{GR}} \to \infty$, fixed BCs are recovered.

Thermal Bath Boundary Conditions

In the case of thermal bath BCs, the source and drain contacts are treated as reservoirs in thermodynamic equilibrium, and only the fluxes injected into the active region are specified. These BCs are better suited for handling highly non–equilibrium distributions. However, their projection onto the Hermite polynomials is more involved since the Heaviside functions limit the integrations to only positive or negative wave numbers. Starting with

$$\Gamma^{\nu}_{\mathrm{S}}\{f\} = \left[f^{\nu}_{k}(z_{\mathrm{S}},k)\Theta(-k) + f_{\mathrm{eq}}(z_{\mathrm{S}},k)\Theta(k)\right]v(k) \tag{3.34}$$

$$\Gamma^{\nu}_{\mathrm{D}}\{f\} = -\left[f^{\nu}_{k}(z_{\mathrm{D}},k)\Theta(k) + f_{\mathrm{eq}}(z_{\mathrm{D}},k)\Theta(-k)\right]v(k)\,, \tag{3.35}$$

transformation of the distribution functions and subsequent expansion with the nth Hermite polynomial gives:

$$\int \Gamma^{\nu}_{\mathrm{S}}\{g\}\mathcal{H}_n(u)\,\mathrm{d}u$$

$$= \sqrt{\frac{2k_{\mathrm{B}}T}{m^{\nu}_{zz}}}\int_{-\infty}^{0} u g^{\nu}(z_{\mathrm{S}},u)\mathcal{H}_n(u)\omega(u)\,\mathrm{d}u + \sqrt{\frac{2k_{\mathrm{B}}T}{m^{\nu}_{zz}}}\gamma^{\nu}_{\mathrm{S}}\int_{0}^{\infty} u\mathcal{H}_n(u)\omega(u)\,\mathrm{d}u$$

$$= \sqrt{\frac{2k_\mathrm{B}T}{m^\nu_{zz}}} \left[\sum_{m=0}^{n_\mathrm{HP}} \frac{a^\nu_m(z_\mathrm{S})}{2^m m! \sqrt{\pi}} \int_{-\infty}^{0} u\mathcal{H}_m(u)\mathcal{H}_n(u)\omega(u)\,\mathrm{d}u + \gamma^\nu_\mathrm{S} \int_0^\infty u\mathcal{H}_n(u)\omega(u)\,\mathrm{d}u \right]$$
$$= \sqrt{\frac{2k_\mathrm{B}T}{m^\nu_{zz}}} \left[\sum_{m=0}^{n_\mathrm{HP}} \mathcal{I}^-_{n,m} a^\nu_m(z_\mathrm{S}) + \gamma^\nu_\mathrm{S}\sqrt{\pi}\mathcal{I}^+_{n,0} \right] \tag{3.36}$$

$$\int \Gamma^\nu_\mathrm{D}\{g\}\mathcal{H}_n(u)\,\mathrm{d}u$$
$$= -\sqrt{\frac{2k_\mathrm{B}T}{m^\nu_{zz}}} \int_0^\infty u g^\nu(z_\mathrm{D},u)\mathcal{H}_n(u)\omega(u)\,\mathrm{d}u + \sqrt{\frac{2k_\mathrm{B}T}{m^\nu_{zz}}}\gamma^\nu_\mathrm{D} \int_{-\infty}^{0} u\mathcal{H}_n(u)\omega(u)\,\mathrm{d}u$$
$$= -\sqrt{\frac{2k_\mathrm{B}T}{m^\nu_{zz}}} \left[\sum_{m=0}^{n_\mathrm{HP}} \frac{a^\nu_m(z_\mathrm{D})}{2^m m! \sqrt{\pi}} \int_0^\infty u\mathcal{H}_m(u)\mathcal{H}_n(u)\omega(u)\,\mathrm{d}u + \gamma^\nu_\mathrm{D} \int_{-\infty}^{0} u\mathcal{H}_n(u)\omega(u)\,\mathrm{d}u \right]$$
$$= -\sqrt{\frac{2k_\mathrm{B}T}{m^\nu_{zz}}} \left[\sum_{m=0}^{n_\mathrm{HP}} \mathcal{I}^+_{n,m} a^\nu_m(z_\mathrm{D}) + \gamma^\nu_\mathrm{D}\sqrt{\pi}\mathcal{I}^-_{n,0} \right]. \tag{3.37}$$

The above definite integrals can be calculated analytically for different expansion orders. Using the recursive formula

$$\begin{aligned}
\mathcal{I}'^+_{n,m} &= \int_0^\infty \mathcal{H}_n(u)\mathcal{H}_m(u)\omega(u)\,\mathrm{d}u \\
&= \int_0^\infty \left(2u\mathcal{H}_{n-1}(u) - 2(n-1)\mathcal{H}_{n-2}\right)\mathcal{H}_m(u)\omega(u)\,\mathrm{d}u \\
&= -2(n-1)\mathcal{I}'^+_{n-2,m} - \mathcal{H}_{n-1}(0)\mathcal{H}_m(0) + \int_0^\infty \frac{\partial}{\partial u}\big(\mathcal{H}_{n-1}(u)\mathcal{H}_m(u)\big)\omega(u)\,\mathrm{d}u \\
&= -2(n-1)\mathcal{I}'^+_{n-2,m} - \mathcal{H}_{n-1}(0)\mathcal{H}_m(0) + 2(n-1)\mathcal{I}'^+_{n-2,m} + 2m\mathcal{I}'^+_{n-1,m-1} \\
&= 2m\mathcal{I}'^+_{n-1,m-1} - \mathcal{H}_{n-1}(0)\mathcal{H}_m(0)\,,
\end{aligned} \tag{3.38}$$

and the relation $2u\mathcal{H}_n(u) = 2n\mathcal{H}_{n-1}(u) + \mathcal{H}_{n+1}(u)$, they can be obtained for arbitrary values of n and m. Some of the results for $\mathcal{I}'^+_{n,m}$ are listed in Table 3.1. As an example, for $n_\mathrm{HP} = 3$ we have:

$$\int \Gamma^\nu_\mathrm{S}\{g\}\mathcal{H}_0(u)\,\mathrm{d}u = \sqrt{\frac{2k_\mathrm{B}T}{m^\nu_{zz}}} \left[\frac{a^\nu_0(z_\mathrm{S})}{2\sqrt{\pi}} + \frac{a^\nu_1(z_\mathrm{S})}{4} + \frac{a^\nu_2(z_\mathrm{S})}{8\sqrt{\pi}} - \frac{\gamma^\nu_\mathrm{S}}{2} \right]$$
$$\int \Gamma^\nu_\mathrm{S}\{g\}\mathcal{H}_1(u)\,\mathrm{d}u = \sqrt{\frac{2k_\mathrm{B}T}{m^\nu_{zz}}} \left[\frac{a^\nu_0(z_\mathrm{S})}{2} + \frac{a^\nu_1(z_\mathrm{S})}{\sqrt{\pi}} + \frac{a^\nu_2(z_\mathrm{S})}{4} + \frac{a^\nu_3(z_\mathrm{S})}{12\sqrt{\pi}} - \frac{\gamma^\nu_\mathrm{S}\sqrt{\pi}}{2} \right]$$
$$\int \Gamma^\nu_\mathrm{S}\{g\}\mathcal{H}_2(u)\,\mathrm{d}u = \sqrt{\frac{2k_\mathrm{B}T}{m^\nu_{zz}}} \left[\frac{a^\nu_0(z_\mathrm{S})}{\sqrt{\pi}} + a^\nu_1(z_\mathrm{S}) + \frac{5a^\nu_2(z_\mathrm{S})}{4\sqrt{\pi}} + \frac{a^\nu_3(z_\mathrm{S})}{4} - \gamma^\nu_\mathrm{S} \right]$$

$$\int \Gamma_{\mathrm{S}}^{\nu}\{g\}\mathcal{H}_3(u)\,\mathrm{d}u = \sqrt{\frac{2k_{\mathrm{B}}T}{m_{zz}^{\nu}}}\left[\frac{2a_1^{\nu}(z_{\mathrm{S}})}{\sqrt{\pi}} + \frac{3a_2^{\nu}(z_{\mathrm{S}})}{2} + \frac{3a_3^{\nu}(z_{\mathrm{S}})}{2\sqrt{\pi}}\right]. \tag{3.39}$$

Using the even/odd coupling in the box $z \in [z_1, z_2]$ (see (3.19)), we can express the odd coefficients $\vec{a}_{\mathrm{o}}(z_1)$ in terms of the even ones $\vec{a}_{\mathrm{e}}(z_1)$, $\vec{a}_{\mathrm{e}}(z_2)$ and then add the resulting BCs to the first block equations of (3.21). The coefficient terms for the drain contact are obtained similarly and added to the last block equations of (3.21).

(n,m)	0	1	2	3	4	5
0	$\frac{1}{2}$	$\frac{1}{2\sqrt{\pi}}$	0	$\frac{-1}{24\sqrt{\pi}}$	0	$\frac{1}{320\sqrt{\pi}}$
1	$\frac{1}{\sqrt{\pi}}$	$\frac{1}{2}$	$\frac{1}{4\sqrt{\pi}}$	0	$\frac{-1}{96\sqrt{\pi}}$	0
2	0	$\frac{1}{\sqrt{\pi}}$	$\frac{1}{2}$	$\frac{1}{4\sqrt{\pi}}$	0	$\frac{-1}{96\sqrt{\pi}}$
3	$\frac{-2}{\sqrt{\pi}}$	0	$\frac{3}{2\sqrt{\pi}}$	$\frac{1}{2}$	$\frac{3}{16\sqrt{\pi}}$	0
4	0	$\frac{-2}{\sqrt{\pi}}$	0	$\frac{3}{2\sqrt{\pi}}$	$\frac{1}{2}$	$\frac{3}{16\sqrt{\pi}}$
5	$\frac{12}{\sqrt{\pi}}$	0	$\frac{-5}{\sqrt{\pi}}$	0	$\frac{15}{8\sqrt{\pi}}$	$\frac{1}{2}$

Table 3.1: The definite integral $\mathcal{I}'^{+}_{n,m} = \frac{1}{2^m m!\sqrt{\pi}}\int_0^{\infty} \mathcal{H}_m(u)\mathcal{H}_n(u)\exp\bigl(-u^2\bigr)\,\mathrm{d}u$.

3.3 Small Signal Analysis

Small signal analysis for the SSSC is straightforward. If we express the vector of expansion coefficients as $\vec{a}(z,t) = \vec{a}_{\mathrm{DC}}(z) + \mathrm{Re}\{\underline{\vec{a}}(z)\mathrm{e}^{\mathrm{i}\omega t}\}$, from (3.13) we get an expression for the complex phasors

$$\mathrm{i}\omega\underline{\vec{a}} + \hat{A}\frac{\partial\underline{\vec{a}}}{\partial z} + \hat{B}_{\mathrm{DC}}\underline{\vec{a}} = -\frac{\partial\hat{B}}{\partial\varphi}\underline{\varphi}\,\vec{a}_{\mathrm{DC}}\,, \tag{3.40}$$

which is an inhomogeneous differential equation due to $\hat{B}$ depending on the small signal potential through the electric field $F_{i+½}$ and the overlap integral in the electron–phonon scattering term. Defining

$$\hat{D}'_{i+½}(z) = \hat{A}^{-1}(\hat{B}_{\mathrm{DC}} + \mathrm{i}\omega\hat{I})(z - z_{i+½})\,, \tag{3.41}$$

the solution to (3.40) is given by

$$\underline{\vec{a}}(z) = \exp\Bigl(-\hat{D}'_{i+½}(z)\Bigr)\underline{\vec{a}}(z_{i+½}) + \underline{\vec{a}}_{\mathrm{P}}(z)\,. \tag{3.42}$$

The second term in (3.42) is the particular solution for the inhomogeneous equation, obtained by numerically calculating the following expression:

$$\begin{aligned}
\underline{\vec{a}}_{\mathrm{p}}(z) &= \exp\Big(-\hat{D}'_{i+1/2}(z)\Big)\int_{z_{i+1/2}}^{z}\exp\Big(\hat{D}'_{i+1/2}(z')\Big)\frac{\partial(-\hat{A}^{-1}\hat{B})}{\partial\varphi}\vec{a}_{\mathrm{DC}}(z')\underline{\varphi}\,\mathrm{d}z' \\
&= \int_{z_{i+1/2}}^{z}\exp\Big(\hat{A}^{-1}(\hat{B}_{\mathrm{DC}}+\mathrm{i}\omega\hat{I})(z'-z)\Big)\frac{\partial(-\hat{A}^{-1}\hat{B})}{\partial\varphi}\vec{a}_{\mathrm{DC}}(z')\underline{\varphi}\,\mathrm{d}z' \\
&= \left[\int_0^1\exp\Big(\hat{D}'_{i+1/2}(z)(1-\alpha)\Big)\frac{\partial\hat{D}}{\partial\varphi}\exp\Big(\hat{D}_{i+1/2}(z)\alpha\Big)\,\mathrm{d}\alpha\right]\vec{a}_{\mathrm{DC}}(z_{i+1/2})\underline{\varphi} \\
&=: \hat{H}(z)\vec{a}_{\mathrm{DC}}(z_{i+1/2})\underline{\varphi}\,.
\end{aligned}$$

Now if we evaluate the solution at the grid points,

$$\begin{aligned}
\begin{pmatrix}\underline{\vec{a}}_{\mathrm{e}}(z_i)\\ \underline{\vec{a}}_{\mathrm{e}}(z_{i+1})\end{pmatrix} &= \begin{pmatrix}\hat{\mathrm{P}}_{\mathrm{e}}\exp\big(-\hat{D}'_{i+1/2}(z_i)\big)\\ \hat{\mathrm{P}}_{\mathrm{e}}\exp\big(-\hat{D}'_{i+1/2}(z_{i+1})\big)\end{pmatrix}\underline{\vec{a}}(z_{i+1/2}) + \begin{pmatrix}\hat{\mathrm{P}}_{\mathrm{e}}\hat{H}(z_i)\\ \hat{\mathrm{P}}_{\mathrm{e}}\hat{H}(z_{i+1})\end{pmatrix}\vec{a}_{\mathrm{DC}}(z_{i+1/2})\underline{\varphi} \\
&=: \hat{M}_1\,\underline{\vec{a}}(z_{i+1/2}) + \hat{M}_2\,\vec{a}_{\mathrm{DC}}(z_{i+1/2})\underline{\varphi}\,. \qquad (3.43)
\end{aligned}$$

the small signal quantity $\vec{a}(z_{i+1/2})$ can be expressed in terms of the small signal even coefficients on the grid nodes, small signal potential and the stationary solution:

$$\underline{\vec{a}}(z_{i+1/2}) = \hat{M}_1^{-1}\begin{pmatrix}\underline{\vec{a}}_{\mathrm{e}}(z_i)\\ \underline{\vec{a}}_{\mathrm{e}}(z_{i+1})\end{pmatrix} - \hat{M}_1^{-1}\hat{M}_2\,\vec{a}_{\mathrm{DC}}(z_{i+1/2})\underline{\varphi}\,, \qquad (3.44)$$

Hence, the final expression for the complete solution in terms of the even components on grid nodes z_i and z_{i+1} is derived:

$$\begin{aligned}
\underline{\vec{a}}(z) &= \exp\big(-\hat{D}'_{i+1/2}(z)\big)\hat{M}_1^{-1}\begin{pmatrix}\underline{\vec{a}}_{\mathrm{e}}(z_i)\\ \underline{\vec{a}}_{\mathrm{e}}(z_{i+1})\end{pmatrix} \\
&+ \Big[\hat{H}(z) - \exp\big(-\hat{D}'_{i+1/2}(z)\big)\hat{M}_1^{-1}\hat{M}_2\Big]\begin{pmatrix}\hat{\mathrm{P}}_{\mathrm{e}}\exp\big(-\hat{D}_{i+1/2}(z_i)\big)\\ \hat{\mathrm{P}}_{\mathrm{e}}\exp\big(-\hat{D}_{i+1/2}(z_{i+1})\big)\end{pmatrix}^{-1}\begin{pmatrix}\vec{a}_{\mathrm{e,DC}}(z_i)\\ \vec{a}_{\mathrm{e,DC}}(z_{i+1})\end{pmatrix}\underline{\varphi}. \qquad (3.45)
\end{aligned}$$

Equation (3.45) can be used to build a linear system of equations for the even components of the small signal solution similar to the DC case.

Chapter 4

Discretization in Phase Space

Transformation of the BE from wave number to energy causes several numerical issues, especially in the case of a 1D k–space. Firstly, since the 1D density of states goes to infinity at zero energy, any quantity plotted against gate or drain biases will exhibit spurious discontinuities. This is due to the fact that for a static grid (i.e. an H–grid which remains constant throughout the whole computation of the stationary and small signal solutions) and any function h which depends on the kinetic energy, the derivative of the box-integrated function with respect to the potential vanishes for all boxes expect the lowest one on which it is proportional to $h(H_l)$, assuming l is the index of the lowest H–box truncated by the subband energy. With an infinitesimal change in any $\varepsilon^{\nu}_{\mathrm{sub}}(z_i)$ from one H–box to the next, this derivative exhibits a discontinuity from $h(H_l)$ to $h(H_{l+1})$ or $h(H_{l-1})$. This is a direct consequence of the H–transformation and degrades the numerical quality of the small signal results. The problem is not pronounced for the 3D k-space where both the density of states and group velocity are zero at zero kinetic energy. In the case of 2D k–space with a constant nonzero density of states, the discontinuities become smaller as the energy space decreases (see e.g. [69, 141]). However, the problem is fatal for the 1D density of states and will be addressed in Chapter 5. Furthermore, as we move towards the ballistic limit, the distribution function can get strongly asymmetric and discontinuous with respect to the wave number. This leads to further problems with the usual stabilization schemes based on the maximum entropy dissipation scheme and the H–transformation. Although these schemes are useful in evaluating the performance of conventional devices and mobilities, they fail at numerical analysis of quasi-ballistic phenomena such as plasma resonances and a numerically robust solver for the BE is required for handling the ballistic limit.

In order to solve the above mentioned problems, in this chapter the BE is discretized directly in the phase space and the ballistic limit of the presented

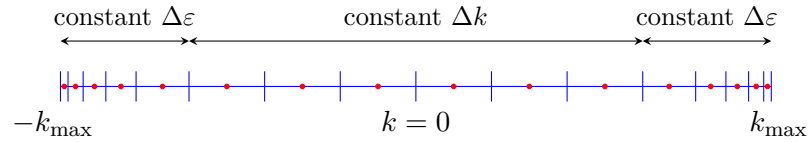

Figure 4.1: The non-equidistant phase space grid. The grid nodes are shown by red dots.

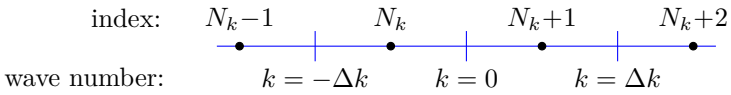

Figure 4.2: The phase space grid around $k=0$.

N^+NN^+ nanowire FET is investigated. The particular focus is on demonstrating a stabilization method, which is based on the method of characteristic curves and matrix exponentials.

4.1 Discretization

4.1.1 Steady State Equation

The BE is discretized on a grid, which is the Cartesian product of the position along the channel axis z and the axial wave vector k. An equidistant spatial grid

$$z_i = (i-1)\Delta z \qquad \text{for } i \in \{1, \cdots, n_z\} \tag{4.1}$$

is introduced. In the k–space a nonuniform grid with constant Δk near $k=0$ and a Δk at higher energies corresponding to a constant step in energy are chosen. The k–grid has $2n_\mathrm{k}$ grid nodes, which are symmetrically distributed around zero and the value zero itself is excluded from the grid. The jth box is given by $k \in [k_{j-1/2}, k_{j+1/2}]$ with its boundaries defined as

$$\begin{aligned} k_{j+1/2} &= \frac{k_j + k_{j+1}}{2} \quad \text{for } j \in \{1, \ldots, 2n_\mathrm{k}-1\}\,, \\ k_{j-1/2} &= \frac{k_j + k_{j-1}}{2} \quad \text{for } j \in \{2, \ldots, 2n_\mathrm{k}\}\,. \end{aligned} \tag{4.2}$$

We start our formulation by considering the ballistic limit ($W^{\nu\nu'}(z;k|k')=0$), because this is the most critical case with respect to the stability. Assuming a linear approximation for the subband energy $\varepsilon^\nu_\mathrm{sub}(z)$ between the adjacent grid nodes z_i and z_{i+1}, the force

$$F^\nu(z_{i+1/2}) = -\frac{\varepsilon^\nu_\mathrm{sub}(z_{i+1}) - \varepsilon^\nu_\mathrm{sub}(z_i)}{z_{i+1} - z_i} \tag{4.3}$$

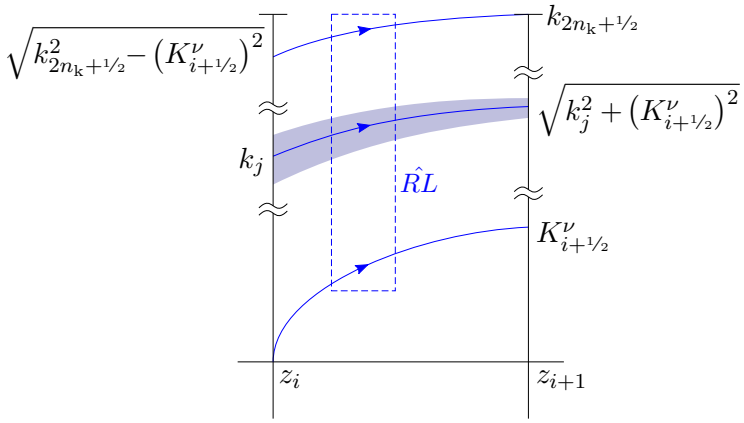

Figure 4.3: Schematic representation of the characteristic curves for constant electric force in $[z_i, z_{i+1}]$ and the case $\varepsilon_{\text{sub}}^{\nu}(z_i) > \varepsilon_{\text{sub}}^{\nu}(z_{i+1})$ and $k(z_i) > 0$.

is constant in the box $[z_i, z_{i+1}]$ and the characteristic curves for the distribution function (ballistic electron path in the phase space) can be obtained. The characteristics curves are defined as the functions $z = z(k)$ in the z–k plane along which the hyperbolic first order PDE

$$v_z^{\nu}(k)\frac{\partial}{\partial z}f_k^{\nu}(z,k) + \frac{1}{\hbar}F^{\nu}(z_{i+1/2})\frac{\partial}{\partial k}f_k^{\nu}(z,k) = 0 \tag{4.4}$$

becomes an ODE. For $f = f(z(k), k)$, the rate of change of z along $z = z(k)$ is given by the chain rule,

$$\frac{\mathrm{d}f}{\mathrm{d}z} = \frac{\partial f}{\partial z} + \frac{\mathrm{d}k}{\mathrm{d}z}\frac{\partial f}{\partial k}\,. \tag{4.5}$$

If the characteristic curve satisfies the ODE

$$\frac{\mathrm{d}k}{\mathrm{d}z} = \frac{-m_{zz}^{\nu}}{\hbar^2 k}F^{\nu}(z_{i+1/2})\,, \tag{4.6}$$

then the rate of change of distribution function along the characteristic curve $z = z(k)$ is zero. Equation (4.6) yields:

$$z = \frac{\hbar^2}{2m_{zz}^{\nu}F^{\nu}(z_{i+1/2})}k^2 + z_0\,, \tag{4.7}$$

which is the *exact* solution to (4.4). Equation (4.7) shows that the particle trajectories do not intersect, and if we follow the ballistic movement of electrons, all of the electrons within the volume element $\mathrm{d}k$ at position z will end up on the corresponding volume element $\mathrm{d}k'$ at position $z' = z + \mathrm{d}z$. In the discretized problem, this suggests simply moving the distribution $f^{\nu}(z_i, k_j)$ to the neighboring grid nodes $(z_{i'}, k_{j'})$ and does not require computation but just rearrangement

of the allocated memory. In other words, z_i and k_j can be interpreted as the "initial" position and wave number of an electron. It is noteworthy that (4.7) is the equation of parabolas due to the force term being constant inside the box. If $F^\nu(z_{i+1/2})$ is positive, the electrons are accelerated in the positive z-direction and for $k > 0$ they move from z_i to z_{i+1} (see Fig. 4.3). For an electron in state ν starting at z_i with $k(z_i)$, the wave number at position z_{i+1} is calculated as

$$k(z_{i+1}) = \sqrt{k(z_i)^2 + \left(K^\nu_{i+1/2}\right)^2}\,, \tag{4.8}$$

where we have:

$$K^\nu_{i+1/2} = \sqrt{\frac{2m^\nu_{zz}\left|\varepsilon^\nu_{\mathrm{sub}}(z_{i+1}) - \varepsilon^\nu_{\mathrm{sub}}(z_i)\right|}{\hbar^2}}\,. \tag{4.9}$$

Thus, $k(z_{i+1}) \geq K^\nu_{i+1/2}$ holds. For electrons starting at z_{i+1} and moving into the negative direction we get

$$k(z_i) = -\sqrt{k(z_{i+1})^2 - \left(K^\nu_{i+1/2}\right)^2}\,. \tag{4.10}$$

The electrons can pass the barrier only, if their kinetic energy is larger than $\varepsilon^\nu_{\mathrm{sub}}(z_i) - \varepsilon^\nu_{\mathrm{sub}}(z_{i+1})$ and thus $k(z_{i+1}) \leq -K^\nu_{i+1/2}$ (see the lower shaded area in Fig. 4.4). If this is not the case (i.e. $-K^\nu_{i+1/2} \leq k(z_{i+1}) \leq 0$) the electron is reflected and returns to position z_{i+1} with its wave number reversed sign as shown by the red line in Fig. 4.4.

The change in the wave number depends only on the difference in the subband energy and not on the exact force profile between both grid nodes. We can therefore assume that the subband energy is piecewise constant and changes abruptly at position $z_{i+1/2} = (z_{i+1}+z_i)/2$, as shown in Fig. 4.5. Due to the Liouville theorem [180] the electron flux at position $z_{i+1/2}$ on the left-hand side of the step must equal the one on the right-hand side. Using the subscripts L and R to indicate quantities on the left-hand side and right-hand side of the step, respectively, for $k_\mathrm{L} > 0$ we have:

$$v^\nu_z(k_\mathrm{R}) f^\nu_\mathrm{R}(z_{i+1/2}, k_\mathrm{R})\,\mathrm{d}k_\mathrm{R} = v^\nu_z(k_\mathrm{L}) f^\nu_\mathrm{L}(z_{i+1/2}, k_\mathrm{L})\,\mathrm{d}k_\mathrm{L}\,, \tag{4.11}$$

where the wave numbers k_L and k_R are related using (4.8):

$$k_\mathrm{R} = \sqrt{k_\mathrm{L}^2 + \left(K^\nu_{i+1/2}\right)^2}\,. \tag{4.12}$$

Thus, the distribution function is no longer continuous and its values on the left and right of the step have to be distinguished. For $k_\mathrm{R} < -K^\nu_{i+1/2}$ the electrons can pass the barrier from right to left with

$$k_\mathrm{L} = -\sqrt{k_\mathrm{R}^2 - (K^\nu_{i+1/2})^2} \tag{4.13}$$

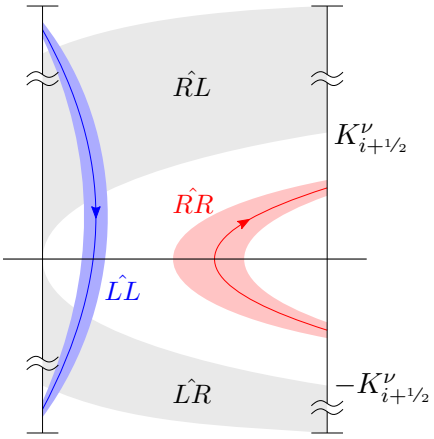

Figure 4.4: Schematic representation of the characteristic curves for constant electric force in $[z_i, z_{i+1}]$ and the case $\varepsilon^{\nu}_{\rm sub}(z_i) > \varepsilon^{\nu}_{\rm sub}(z_{i+1})$. The electrons at z_{i+1} with $k(z_{i+1}) \leq -K^{\nu}_{i+1/2}$ cannot pass the barrier and are reflected back to z_{i+1} with their wave numbers reversed sign (red curve). Moreover, the electrons that exceed the maximum wave number in moving from z_i to z_{i+1} are reflected back to z_i (blue curves). Although transition from z_i to z_{i+1} and vice versa changes the phase space volume element for nonzero electric field (gray shaded areas), reflection does not change the volume element.

and for $-K^{\nu}_{i+1/2} < k_{\rm R} < 0$, the electrons are reflected:

$$f^{\nu}_{\rm R}(z_{i+1/2}, k_{\rm R}) = f^{\nu}_{\rm R}(z_{i+1/2}, -k_{\rm R})\,. \tag{4.14}$$

In order to obtain a discrete system of equations, the above relations are integrated on the sub–intervals of the phase space. For better legibility, integration over $[k_{j-1/2}, k_{j+1/2}]$ is shown by $\int_{\Delta k_j}$. We also use the selection function Ω_j for the jth box, which is one inside the jth box and zero outside:

$$\Omega_j(k) = \begin{cases} 1 & k \in [k_{j-1/2}, k_{j+1/2}] \\ 0 & \text{otherwise} \end{cases} \tag{4.15}$$

The distribution functions are assumed to be piecewise constant with the value $f^{\nu,j}(z)$ for the jth box

$$f^{\nu}(z,k) \approx \sum_{j=1}^{2n_{\rm k}} f^{\nu,j}(z)\Omega_j(k)\,. \tag{4.16}$$

Moreover, we distinguish the distribution function on the left and right side of the steps in subband energy using the following notation:

$$f^{\nu,j}_{{\rm L},i+1/2} = f^{\nu,j}_{\rm L}(z_{i+1/2})\,, \quad f^{\nu,j}_{{\rm R},i+1/2} = f^{\nu,j}_{\rm R}(z_{i+1/2})\,. \tag{4.17}$$

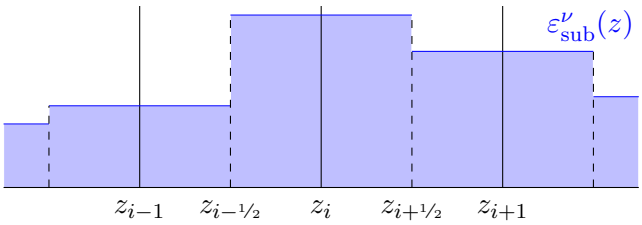

Figure 4.5: The subband energy is assumed to be piecewise constant, and changes at the intermediary $z-$grid points.

Now, we can map the particle flux conservation onto the discrete $k-$grid using box integration. This allows to discretize (4.11) together with (4.14) for $j > n_{\mathrm{k}}$ (i.e. positive wave numbers):

$$f^{\nu,j}_{\mathrm{R},i+1/2} \int_{\Delta k_j} v^{\nu}_z(k_{\mathrm{R}})\,\mathrm{d}k_{\mathrm{R}} = \sum_{j'=n_{\mathrm{k}}+1}^{2n_{\mathrm{k}}} f^{\nu,j'}_{\mathrm{L},i+1/2} \int_{\mathrm{K}^{\nu,j}_{\min}}^{\mathrm{K}^{\nu,j}_{\max}} v^{\nu}_z(k_{\mathrm{L}})\Omega_{j'}(k_{\mathrm{L}})\,\mathrm{d}k_{\mathrm{L}} + f^{\nu,2n_{\mathrm{k}}-j+1}_{\mathrm{R},i+1/2} \int_{\Delta k_j} v^{\nu}_z(k_{\mathrm{R}})\,\Theta\big(K^{\nu}_{i+1/2}-k_{\mathrm{R}}\big)\,\mathrm{d}k_{\mathrm{R}}\ . \tag{4.18}$$

with

$$\mathrm{K}^{\nu,j}_{\min} = \sqrt{\max\left\{k^2_{j-1/2} - (K^{\nu}_{i+1/2})^2, 0\right\}}\,, \tag{4.19}$$

$$\mathrm{K}^{\nu,j}_{\max} = \sqrt{\max\left\{k^2_{j+1/2} - (K^{\nu}_{i+1/2})^2, 0\right\}}\,. \tag{4.20}$$

The first term on the right–hand side of (4.18) describes the flux of electrons from the left to the right and the second term accounts for the electrons reflected at the step. If we assemble the n_{k} values of the distribution function for positive wave numbers into a vector on the left side of the step $\vec{f}^{\nu}_{\mathrm{Lp},i+1/2}$, on the right side $\vec{f}^{\nu}_{\mathrm{Rp},i+1/2}$ and for negative wave numbers $\vec{f}^{\nu}_{\mathrm{Ln},i+1/2}$, $\vec{f}^{\nu}_{\mathrm{Rn},i+1/2}$, we can formulate (4.18) using matrices

$$\hat{D}^{\nu}_{\mathrm{p}}\vec{f}^{\nu}_{\mathrm{Rp},i+1/2} = \hat{RL}^{\nu}_{i+1/2}\vec{f}^{\nu}_{\mathrm{Lp},i+1/2} + \hat{RR}^{\nu}_{i+1/2}\vec{f}^{\nu}_{\mathrm{Rn},i+1/2}\ , \tag{4.21}$$

where the diagonal $n_{\mathrm{k}}\times n_{\mathrm{k}}$ matrix $\hat{D}^{\nu}_{\mathrm{p}}$ is invertible and does not depend on the position in real space

$$\left[\hat{D}^{\nu}_{\mathrm{p}}\right]_{jj'} = \int_{\Delta k_{j+n_{\mathrm{k}}}} v^{\nu}_z(k)\,\mathrm{d}k\,\delta_{j,j'} \qquad \text{for}\ \ j,j' \in \{1,\ldots,n_{\mathrm{k}}\}\,. \tag{4.22}$$

The other two matrices are

$$\left[\hat{RL}^{\nu}_{i+1/2}\right]_{jj'} = \int_{\mathrm{K}^{\nu,j+n_{\mathrm{k}}}_{\min}}^{\mathrm{K}^{\nu,j+n_{\mathrm{k}}}_{\max}} v^{\nu}_z(k)\,\Omega_{j'+n_{\mathrm{k}}}(k)\,\mathrm{d}k\ , \tag{4.23}$$

$$\left[\hat{RR}^{\nu}_{i+1/2}\right]_{jj'} = \int_{\Delta k_{j+n_\mathrm{k}}} v^{\nu}_{z}(k)\Theta\big(K^{\nu}_{i+1/2} - k\big)\,\mathrm{d}k\,\delta_{j',n_\mathrm{k}-j+1}\ . \tag{4.24}$$

Flux conservation for particles moving from right to left results with (4.11) and (4.13) in

$$\hat{D}^{\nu}_{\mathrm{n}}\vec{f}^{\nu}_{\mathrm{Ln,i+1/2}} = \hat{LR}^{\nu}_{i+1/2}\vec{f}^{\nu}_{\mathrm{Rn,i+1/2}} + \hat{LL}^{\nu}_{i+1/2}\vec{f}^{\nu}_{\mathrm{Lp,i+1/2}}\ . \tag{4.25}$$

The second term on the right–hand side is due to the assumption of a maximal positive wave number $k_{2n_\mathrm{k}+1/2}$. All of the particles that exceed this value while moving from left to right, i.e.

$$\sqrt{k_\mathrm{L}^2 + (K^{\nu}_{i+1/2})^2} > k_{2n_\mathrm{k}+1/2}\,, \tag{4.26}$$

are instead reflected to ensure flux conservation (see the blue curve in Fig. 4.4). A similar set of equations can be formulated for a positive step in energy.

Under stationary conditions the ballistic distribution function does not depend on position for a constant subband energy and within a box of the real space grid $z \in [z_{i-1/2}, z_{i+1/2}]$ we get

$$\vec{f}^{\nu}_{\mathrm{Lp}}(z_{i+1/2}) = \vec{f}^{\nu}_{\mathrm{Rp}}(z_{i-1/2}) =: \vec{f}^{\nu}_{\mathrm{p},i} \tag{4.27}$$

$$\vec{f}^{\nu}_{\mathrm{Ln}}(z_{i+1/2}) = \vec{f}^{\nu}_{\mathrm{Rn}}(z_{i-1/2}) =: \vec{f}^{\nu}_{\mathrm{n},i}\,. \tag{4.28}$$

Due to the thermal bath BCs the distribution function of the incoming particles is given by the equilibrium distribution [11]

$$\vec{f}^{\nu}_{\mathrm{p},1} = \vec{f}^{\nu}_{\mathrm{p,eq}}\,, \quad \vec{f}^{\nu}_{\mathrm{n},n_z} = \vec{f}^{\nu}_{\mathrm{n,eq}}\,. \tag{4.29}$$

This closes the system of equations and a unique solution can be obtained for $\vec{f}^{\nu}_{\mathrm{p},i}$ and $\vec{f}^{\nu}_{\mathrm{n},i}$ as long as the subband energy does not have a minimum within the device. Since the Pauli exclusion principle occurs in the BE only in the scattering integral, it could be included in the ballistic case by using Fermi–Dirac distribution functions instead of Boltzmann distributions in the BCs.

The solution becomes unique for arbitrary subband profiles (i.e. with several minima or maxima) if scattering is taken into account, because scattering couples all states for a given position. Since scattering couples different subbands and valleys, the index ν is dropped and the vector of the distribution function contains from now on the distribution functions for all values of ν. Due to the assumption of a piecewise constant subband energy the force is zero within a box of the real space. Integration of the BE over a box in k–space yields

$$\underbrace{\begin{pmatrix}\hat{D}_\mathrm{p} & \hat{0}\\ \hat{0} & \hat{D}_\mathrm{n}\end{pmatrix}}_{=\hat{D}_v}\frac{\partial}{\partial z}\underbrace{\begin{pmatrix}\vec{f}_\mathrm{p}(z)\\ \vec{f}_\mathrm{n}(z)\end{pmatrix}}_{=\vec{f}(z)} + \underbrace{\begin{pmatrix}\hat{S}_\mathrm{pp} & \hat{S}_\mathrm{pn}\\ \hat{S}_\mathrm{np} & \hat{S}_\mathrm{nn}\end{pmatrix}}_{=\hat{S}(z)}\begin{pmatrix}\vec{f}_\mathrm{p}(z)\\ \vec{f}_\mathrm{n}(z)\end{pmatrix} = \vec{0} \tag{4.30}$$

with the scattering integral

$$\left[\hat{S}(z)\right]_{jj'}^{\nu\nu'} = \frac{L_{\mathrm{sys}}}{2\pi}\int_{\Delta k_j}\int_{\Delta k_{j'}} W^{\nu\nu'}(z;k|k')\,\mathrm{d}k'\,\mathrm{d}k - \frac{L_{\mathrm{sys}}}{2\pi}\sum_{\nu''}\int_{\Delta k_j}\int_{-\infty}^{\infty} W^{\nu''\nu}(z;k''|k)\,\mathrm{d}k''\,\mathrm{d}k\;\delta_{\nu,\nu'}\delta_{j,j'}\,. \tag{4.31}$$

Due to scattering the solution depends on the z–coordinate within a box of the real space, even if we assume that the scattering integral itself does not depend on position within the box. The solution within the ith box of the real space grid is calculated by a matrix exponential [181]:

$$\vec{f}(z) = \exp\left((z_i - z)\hat{D}_v^{-1}\hat{S}(z_i)\right)\vec{f}(z_i)\,. \tag{4.32}$$

With the abbreviation $\hat{C}_i(z) = (z_i - z)\hat{D}_v^{-1}\hat{S}(z_i)$ the solution at the boundaries of the ith box in real space is

$$\vec{f}_{\mathrm{L},i+1/2} = \exp\left(\hat{C}_i(z_{i+1/2})\right)\vec{f}(z_i) \tag{4.33}$$

$$\vec{f}_{\mathrm{R},i-1/2} = \exp\left(\hat{C}_i(z_{i-1/2})\right)\vec{f}(z_i)\,. \tag{4.34}$$

These equations replace (4.27), (4.28) of the ballistic case, for which they reproduce the previous case. Thus, it is possible to calculate the distribution function on the left and right-hand side of the steps based on the distribution function on the ith grid node and a complete set of equations can again be assembled. The matrix exponential can be calculated with the methods described in Ref. [181], but for large matrices the numerical accuracy might not be sufficient. In this case Expokit together with an iterative solver, which requires only multiplications of the matrix exponential with a vector, yields results with a higher precision and is more CPU efficient [182]. If the Pauli exclusion principle were included in the scattering integral, (4.30) would become nonlinear and could in principle be integrated by a sophisticated ODE–solver.

4.1.2 Small Signal Analysis

Once the DC solution has been evaluated, small signal analysis can be carried out through linearization of the BE. For small signal analysis we assume the SSSC and the time dependence of the solution quantities is given by $x(t) = x_{\mathrm{DC}} + \mathrm{Re}\{\underline{x}\mathrm{e}^{\mathrm{i}\omega t}\}$, where x_{DC} is the large signal stationary solution and $\underline{x}$ the complex phasor of the small signal solution. The time derivative in the BE can therefore be replaced by $\mathrm{i}\omega$ and within a box of the spatial grid, the linearized BE has two additional terms compared to the stationary case (4.30)

$$\mathrm{i}\omega\hat{D}_k\underline{\vec{f}}(z) + \hat{D}_v\frac{\partial\underline{\vec{f}}(z)}{\partial z} + \hat{S}\underline{\vec{f}}(z) = -\underline{\hat{S}}\vec{f}_{\mathrm{DC}}(z)\,. \tag{4.35}$$

$\hat{D}_k$ is a diagonal matrix, which is due to box integration in k–space and contains the k–space box volumes

$$\left[\hat{D}_k\right]_{jj'}^{\nu\nu'} = (k_{j+1/2} - k_{j-1/2})\delta_{\nu,\nu'}\delta_{j,j'} \,. \tag{4.36}$$

Since the scattering integral depends on the subband energies and wave functions (e.g. through the overlap integral), linearization of the scattering integral yields the term on the right-hand side of (4.35). The corresponding derivatives are calculated by the chain rule and stationary perturbation theory, as explained in section 2.3. With the abbreviation $\hat{C}_i'(z) = (z_i - z)\hat{D}_v^{-1}(\hat{S}(z_i) + \mathrm{i}\omega\hat{D}_k)$ the solution of the small signal BE is obtained

$$\underline{\vec{f}}(z) = \exp\left(\hat{C}_i'(z)\right)\underline{\vec{f}}(z_i) + \int_0^1 \exp\left(\hat{C}_i'(z)(1-\alpha)\right)\underline{\hat{C}}_i(z)\exp\left(\hat{C}_i(z)\alpha\right)\mathrm{d}\alpha\,\vec{f}(z_i) \,. \tag{4.37}$$

With this result the values of the small signal distribution function at the boundaries of the real–space box can be calculated by numerical means. Linearization of (4.21) and (4.25) yields

$$\begin{aligned}\hat{D}_{\mathrm{p}}^{\nu}\underline{\vec{f}}_{\mathrm{Rp,i+1/2}}^{\nu} =& \hat{R}L_{i+1/2}^{\nu}\underline{\vec{f}}_{\mathrm{Lp,i+1/2}}^{\nu} + \underline{\hat{R}L}_{i+1/2}^{\nu}\vec{f}_{\mathrm{Lp,i+1/2}}^{\nu} \\ &+ \hat{R}R_{i+1/2}^{\nu}\underline{\vec{f}}_{\mathrm{Rn,i+1/2}}^{\nu} + \underline{\hat{R}R}_{i+1/2}^{\nu}\vec{f}_{\mathrm{Rn,i+1/2}}^{\nu} \,, \end{aligned} \tag{4.38}$$

$$\begin{aligned}\hat{D}_{\mathrm{n}}^{\nu}\underline{\vec{f}}_{\mathrm{Ln,i+1/2}}^{\nu} =& \hat{L}R_{i+1/2}^{\nu}\underline{\vec{f}}_{\mathrm{Rn,i+1/2}}^{\nu} + \underline{\hat{L}R}_{i+1/2}^{\nu}\vec{f}_{\mathrm{Rn,i+1/2}}^{\nu} \\ &+ \hat{L}L_{i+1/2}^{\nu}\underline{\vec{f}}_{\mathrm{Lp,i+1/2}}^{\nu} + \underline{\hat{L}L}_{i+1/2}^{\nu}\vec{f}_{\mathrm{Lp,i+1/2}}^{\nu} \,. \end{aligned} \tag{4.39}$$

These equations together with the small signal PE and the small signal version of the boundary conditions (4.29)

$$\underline{\vec{f}}_{\mathrm{p},1}^{\nu} = \vec{0} \,, \quad \underline{\vec{f}}_{\mathrm{n},n_z}^{\nu} = \vec{0} \tag{4.40}$$

can be used to assemble a linear system of equations. For the calculation of the small signal terminal currents and admittance parameters, a formulation of the Ramo–Shockly theorem is used that is consistent with the presented numerical framework. For further details the reader is referred to [81, 183].

Chapter 5

Results

In this chapter, the developed deterministic PE/SE/BE solver is applied to the GAA nanowire nFET of Fig. 2.1, and the simulation results are presented. Section 5.1 introduces some of the physical and discretization parameters that are used throughout this chapter. In Sect. 5.2 the stationary solutions are presented and the DC performance of the device is evaluated in terms of its transfer characteristics, subthreshold behavior and the well–known output characteristics. These curves paint an almost complete picture of the device's static behavior. Section 5.3 gives the small signal characterization of the transistor at various operating DC biases. The focus is on the calculation of admittance parameters, from which many important figures of merit such as the cut–off frequency, maximum oscillation frequency, and stability factor are obtained and discussed. The RF noise behavior of the device in terms of the power spectral densities of terminal currents is investigated in Sect. 5.4. The usual figures of merit consisting of the drain and gate excess noise factors, the cross–correlation as well as noise suppression factors are presented. The results of Sects. 5.2 through 5.4 are obtained using the discretization of the BE in H–space. In Sect. 5.5 the ballistic limits of the device are investigated using the proposed discretization in k–space, with comparisons to the results H–space discretization and moments equations.

5.1 Simulation Parameters

The nanowire N^+NN^+ FET of Fig. 2.1 is simulated with different gate lengths of $L_G = 10\,\text{nm}$, $22\,\text{nm}$, and $100\,\text{nm}$. The source and drain regions are $10\,\text{nm}$ long each, and a rectangular cross–section of $5\,\text{nm} \times 4\,\text{nm}$ is assumed[1]. The gate oxide thickness is $t_{ox} = 1\,\text{nm}$, and the gate contact covers the complete channel

[1] The cross–section of the nanowire needs to be smaller than 6 nm in each direction so that the transverse electric field penetrates the entire volume and effectively depletes the channel in the OFF–state [131, 184].

length. Hence, the expressions "gate length" and "channel length" may be used interchangably. The shades of color in the silicon region of Fig. 2.1 indicate the donor doping concentration, with $N_{\mathrm{D}}^{+} = 1 \times 10^{20}\,\mathrm{cm}^{-3}$ for the source/drain extensions and $N_{\mathrm{D}} = 1 \times 10^{18}\,\mathrm{cm}^{-3}$ for the channel region. The high values of doping (well above the Mott criteria) in the source and drain regions are necessary to ensure high values of I_{ON}, whilst the GAA architecture aims for a low I_{OFF}. The channel $\mathrm{N^{+}N}$ junctions are abrupt and no gate to source/drain overlap is considered. A silicon crystal with a $\langle 100 \rangle$ orientation is used in this work.

In order to correctly calculate the first–order perturbations in constructing the Newton-Raphson system, we need to solve the SE for all of the subbands. However, for a nanowire cross-section of 5 nm×4 nm, we can safely limit ourselves to the 5 lowest subbands in the BE. Regarding the H-grid, a grid spacing of $\Delta \mathrm{H} = 2.6\,\mathrm{meV}$ is chosen in order to resolve also the phonons with lower energies.. In order to speed up the simulations, the electrostatic potential obtained from the solution of the drift–diffusion equation is used as an initial condition for the BE solver. The Gummel iterations converge to the error of $1 \times 10^{-3}\,\mathrm{V}$ within 3–8 iterations (depending on the bias), and then the Newton–Raphson approach converges to $1 \times 10^{-12}\,\mathrm{V}$ in 4–5 steps.

5.2 Stationary Results

We shall first look at the steady state results, which are obtained for time invariant boundary conditions. The voltages V_{GS} and V_{DS}, defined relative to the $V_{\mathrm{S}} = 0\,\mathrm{V}$ on the source contact, are applied on the gate and drain contacts, respectively. Figure 5.1 demonstrates the I_{DS}–V_{GS} characteristics for devices with different gate lengths biased at $V_{\mathrm{DS}} = 0.5\,\mathrm{V}$. The voltage at gate terminal is varied from 0 to 0.85 V, and the results are plotted on both linear and logarithmic scale. The threshold voltage V_{TH} is defined by the constant–current method at the reference current of $I_{\mathrm{DS}} = 1\,\mu\mathrm{A}$. We have $V_{\mathrm{TH}} = 0.575\,\mathrm{V}, 0.627\,\mathrm{V}$, and $0.699\,\mathrm{V}$ for $L_{\mathrm{G}} = 10\,\mathrm{nm}, 22\,\mathrm{nm}$, and 100 nm, respectively. The transconductance of the device can be estimated by taking the ratio of the change in drain current with respect to the change in V_{GS}. For the 10 nm device, the maximum transconductance of 89.5 µS is obtained at $V_{\mathrm{GS}} = 0.81\,\mathrm{V}$ (see Fig. 5.1).

Another important metric is the subthreshold slope (SS), defined as the inverse of the slope of $\log_{10}(I_{\mathrm{DS}})$ versus V_{GS} below the threshold voltage and represents the sharpness of the ON–OFF switching of a transistor. While the typical MOS transistor has a subthreshold slope of about 80 mV/dec and the best trigate SOI devices have $\mathrm{SS} \approx 63\,\mathrm{mV/dec}$, the GAA nanowire structure gets even closer to the theoretical best value of $\mathrm{SS} = (k_{\mathrm{B}}T/q)\ln(10) = 60\,\mathrm{mV/dec}$ at room temperature. The near–ideal subthreshold slope was also experimentally observed in e.g. [129, 185, 186]. Even for the 10 nm device, the GAA structure combined

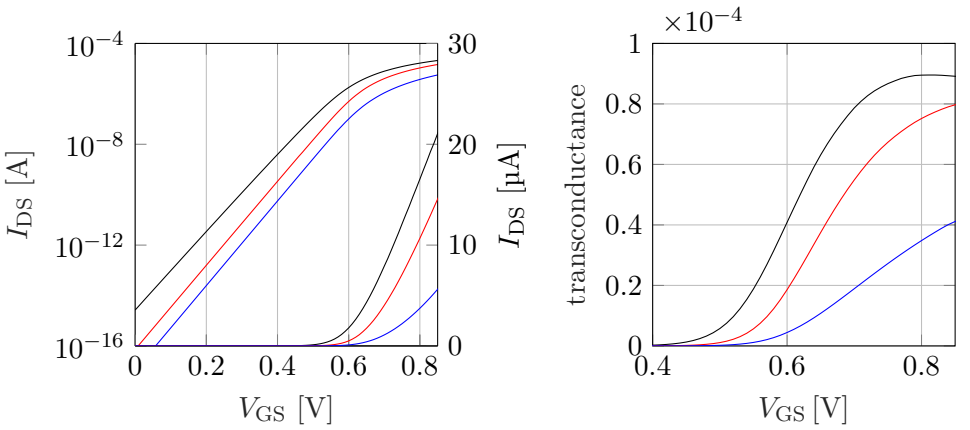

Figure 5.1: Variation of the drain current I_{DS} (left) and transconductance (right) as a function of the gate voltage V_{GS} at $V_{DS} = 0.5\,V$. The simulations are performed for $L_G = 10\,nm$ (black), $L_G = 22\,nm$ (red), and $L_G = 100\,nm$ (blue).

with the ultrathin cross–section still shows acceptable gate controllability and SS $\approx 68.3\,mV/dec$ is obtained. Defining the OFF–current as the drain current for $V_{GS} = V_{TH} - 0.5\,V$ and $V_{DS} = 0.5\,V$, the ON/OFF current ratio changes from 4.05×10^7 to 2.08×10^7 as the channel length is decreased from 100 nm to 10 nm. It is noteworthy that the I_{DS}–V_{GS} characteristics of Fig. 5.1 span over 10 decades of magnitude and the developed solver is able to accurately simulate the deep–subthreshold regime. Such simulations are practically impossible with the stochastic methods because the CPU time of an MC simulation is inversely proportional to current density, whereas in the presented solver the CPU time is almost independent of the current.

In Fig. 5.2 (left) the electron concentration under the gated region is plotted along the x–direction ($y = 3\,nm$). The results are shown for different values of V_{GS} ranging from pinch–off (e.g. $V_{GS} = 0.5\,V$) to flatband conditions[2] ($V_{GS} \approx 0.7\,V$). As expected, the electron density increases by increasing V_{GS} and the channel goes from volume inversion into partial depletion and subsequently into a flatband condition. It is important to note that the conduction path is located near the center of the nanowire and not at the silicon/oxide interface. This is due to the small cross–section, which effectively pushes the carriers away from the boundaries of the potential well. The depleted part of the semiconductor acts as a capacitance in silicon connected to the gate oxide capacitance, degrading the gate–channel coupling. For $V_{GS} > V_{TH}$, the current flow is distributed throughout the entire

[2] In junctionless nanowire FETs, the device is said to be in flatband condition when the electric field in the direction perpendicular to the current flow is zero at the center of the cross–sectional area [129].

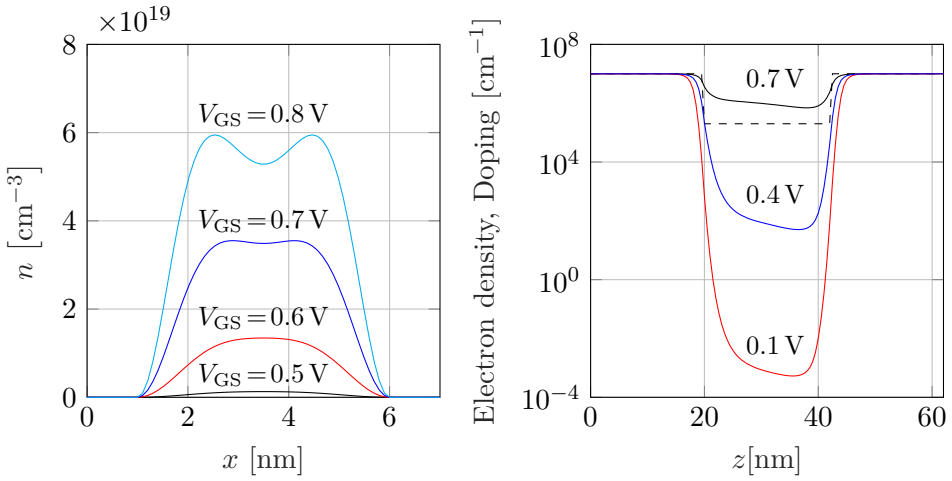

Figure 5.2: Electron density under the gated area of the 22 nm device along the $x-$axis for different V_{GS} values (left). Doping profile and the electron density at $V_{DS} = 0.5\,V$ and different V_{GS} values (right).

volume of the active region. Increasing V_{GS} beyond the flatband voltage creates conduction channels in the close proximity of silicon/oxide interface and the device is said to be operating in the surface conduction regime. Fig. 5.2 confirms that our simulation framework is able to capture the underlying conduction mechanism in junctionless nanowire devices, as discussed in [128, 185, 187]. The N^+NN^+ FETs are normally designed to operate in flatband condition in the ON–state [129]. This is because in the surface conduction regime, electrons are susceptible to interactions with the interface nonidealities, traps, etc. Therefore, throughout this section and Sects. 5.3 and 5.4 we have chosen $V_{GS} = 0.7\,V$ and $V_{DS} = 0.5\,V$ as the ON–state of the simulated devices[3].

Figure 5.2 (right) shows the stationary electron density and the doping profile integrated on the confinement cross–section along the transport direction. The results are presented for $V_{DS} = 0.5\,V$ and different V_{GS} values. The numerical approach is clearly robust and produces very smooth results for arbitrarily sharp gradients of the potential profile. Although the charge neutrality is satisfied in the N^+ source/drain extension regions, the electron density near the source and drain contacts deviates from its equilibrium value due to the finite generation/recombination velocity and nonzero current flowing through the device. This small change is not discernible in Fig. 5.2.

Figure 5.3 shows the calculated I_{DS}–V_{DS} characteristics for different V_{GS}

[3]The bulk conduction mechanism also turns out to improve device's long–term reliability. It was shown in [188, 189] that N^+NN^+ trigate transistors with SiO_2/HfO_2 dielectric stacks have a lower hot carrier degradation compared to their inversion–mode counterparts.

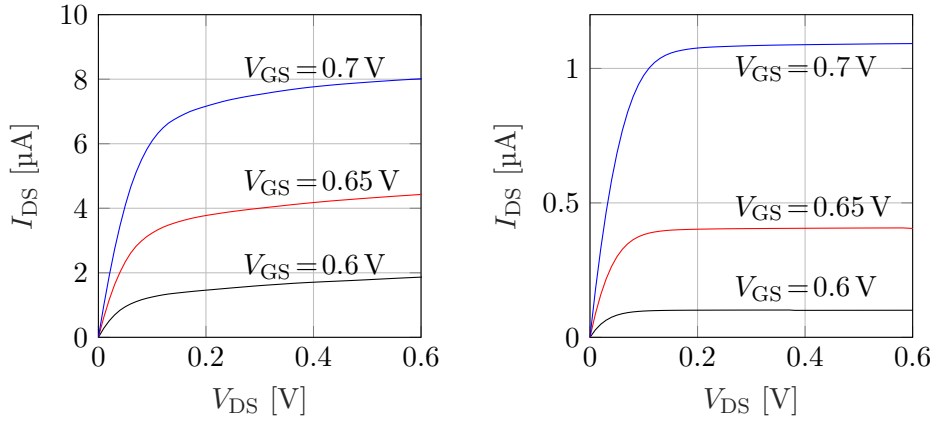

Figure 5.3: Stationary drain current I_{DS} vs. drain bias V_{DS} at different V_{GS} values. Simulations are performed for the 10 nm device (left) and the 100 nm device (right).

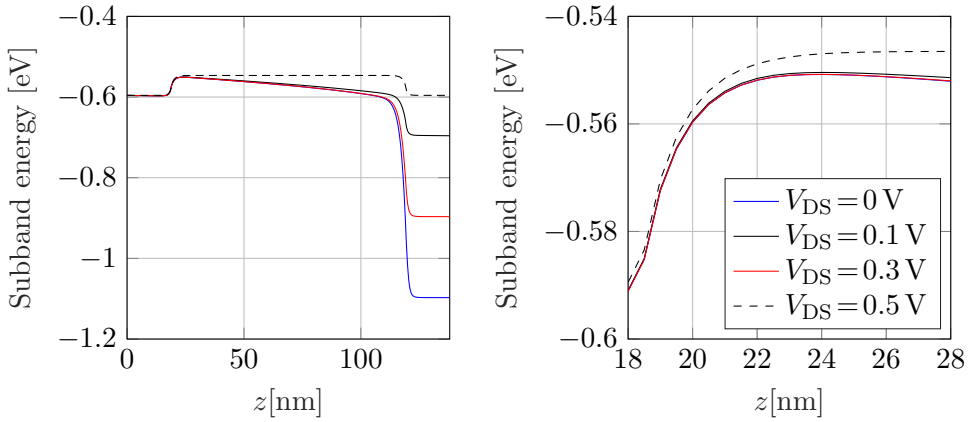

Figure 5.4: Subband energy profile for the energetically lowest subband of the 100 nm device. The results are for $V_{GS} = 0.7\,V$ and different V_{DS} values. The height of the energy barrier does not change by changing the drain bias (right).

values. The characteristic curves of the simulated N^+NN^+ structure are remarkably similar to those of a regular MOSFET although the latter uses an inversion layer and reverse–biased junctions for turning the channel on and off. Hence, the simulated device can provide full CMOS functionality. While the output current is larger for shorter devices due to the proportionality of channel resistance and channel length, the 10 nm device does not saturate at higher drain bias. The reason is shown in Figs. 5.4 and 5.5, where the subband energy of the energetically lowest state is plotted for $V_{GS} = 0.7\,V$ and different drain biases. As the drain bias increases, the subband energy on the drain side decreases which

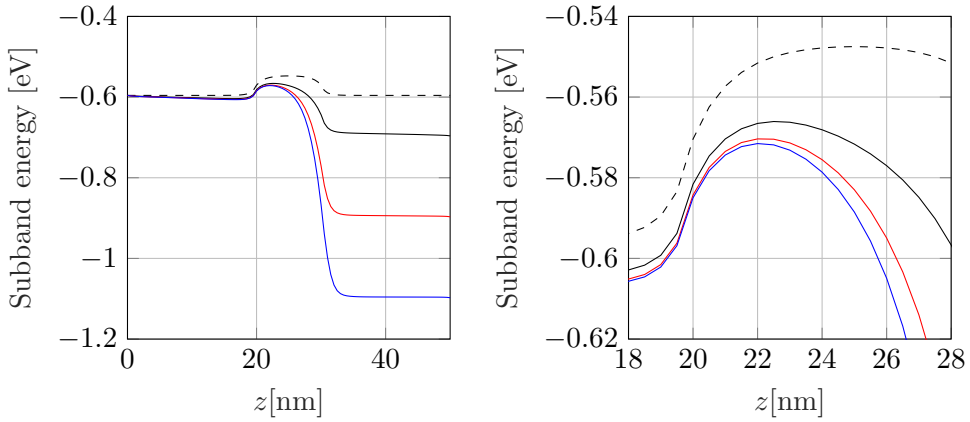

Figure 5.5: Subband energy profile for the energetically lowest subband of the 10 nm device. The results are for $V_{GS} = 0.7\,V$ and different drain biases, $V_{DS} = 0\,V$ (dashed), $V_{DS} = 0.1\,V$ (solid black), $V_{DS} = 0.3\,V$ (red), and $V_{DS} = 0.5\,V$ (blue). DIBL is more pronounced compared to the 100 nm device (right).

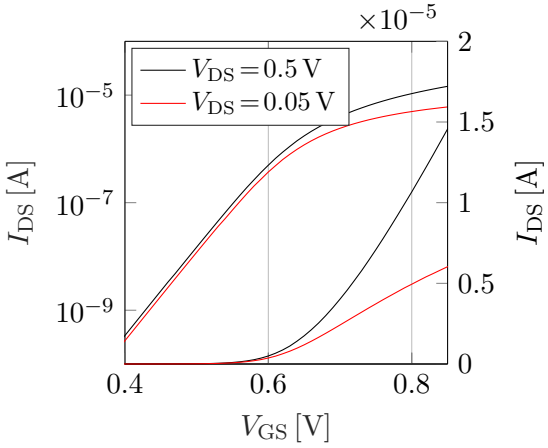

Figure 5.6: Drain current I_{DS} vs. gate voltage V_{GS} for the 22 nm device at $V_{DS} = 0.05\,V$ (red) and $V_{DS} = 0.5\,V$ (black).

can pull down the potential barrier at the source/channel homojunction, leading to an increase in I_{DS}. The 10 nm device shows to be strongly affected by this drain–induced barrier lowering (DIBL) effect, whereas the 100 nm device exhibits much better gate control because the drain side of the channel is too far from the top of the potential barrier and cannot influence the barrier's height. DIBL is usually quantified as the ratio of the variation of V_{TH} to the variation of the

drain voltage (see Fig. 5.6):

$$\mathrm{DIBL} = \frac{V_{\mathrm{TH}}\big|_{V_{\mathrm{DS}}=V_1} - V_{\mathrm{TH}}\big|_{V_{\mathrm{DS}}=V_2}}{V_2 - V_1}\,. \tag{5.1}$$

For $V_1 = 0.05\,\mathrm{V}$ and $V_2 = 0.5\,\mathrm{V}$, we get DIBL values of $60\,\mathrm{mV\,V^{-1}}$ and $13.3\,\mathrm{mV\,V^{-1}}$ for the 10 nm and 100 nm device, respectively.

It is also useful to look at the microscopic distribution of charge carriers inside the device. Fig. 5.7 shows the distribution function for $\nu = (1, 1)$ at the source and drain contacts when the transistor is under high bias $V_{\mathrm{GS}} = 0.7\,\mathrm{V}$ and $V_{\mathrm{DS}} = 0.5\,\mathrm{V}$. Near the source contact, the distribution function follows the expected Fermi–Dirac distribution whereas a second peak at higher energies is added to the Fermi–Dirac distribution of the drain contact. This peak is due to the incoming electrons from the source side, which are subject to phonon cascade due to inelastic scattering mechanisms. The high energy peak highlights the ballistic component of the transport, and its magnitude decreases by increasing L_{G}.

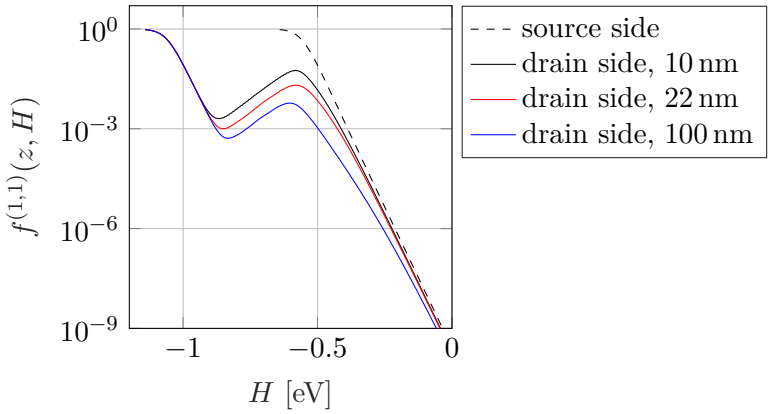

Figure 5.7: Distribution function $f^{(1,1)}(z, H)$ at the source and drain ends of the simulated devices.

5.3 Small Signal Results

In order to evaluate the performance of junctionless nanowire FETs in circuits, investigation of their dynamic response is also necessary. Although several works [190–194] have developed small signal models for various multigate N^+NN^+ devices, the proposed models are usually validated by TCAD simulations based on the drift–diffusion transport mechanism. Small signal analysis based on the BE can provide higher accuracy for the future models. In this section, the results of our small signal simulations are presented and several ac figures of merit are compared for the simulated devices.

Figures 5.8 and 5.9 show the calculated admittance parameters of the 22 nm device for a common–source configuration biased at $V_{GS} = 0.7\,V$ and $V_{DS} = 0.5\,V$. The results are smooth over frequency, and can be verified at $f = 0$ using the steady state characteristics of Sect. 5.2. For $f < 1\,THz$, the imaginary parts of the admittance parameters represent the capacitive effects and change linearly with the frequency, whereas the real part of the gate self–admittance changes quadratically with the frequency since the gate current only has a displacement component and oxide tunneling is not included. The real part of the drain self–admittance is positive for all frequencies and nearly constant for frequencies up to 100 GHz. At $f = 27.5\,THz$, the drain self–admittance peaks due to a plasma resonance caused by the inductive effect of electrons' finite mass and the capacitance of the lowly doped region [159]. Increasing the electron density in the channel shifts this resonance peak to lower frequencies. The sharp extremum in Fig. 5.8 is due to $\mathrm{Re}\{\underline{Y}_{DG}\}$ changing sign at 1.73 THz. Note that while the admittance parameters can be easily computed for the entire frequency range,

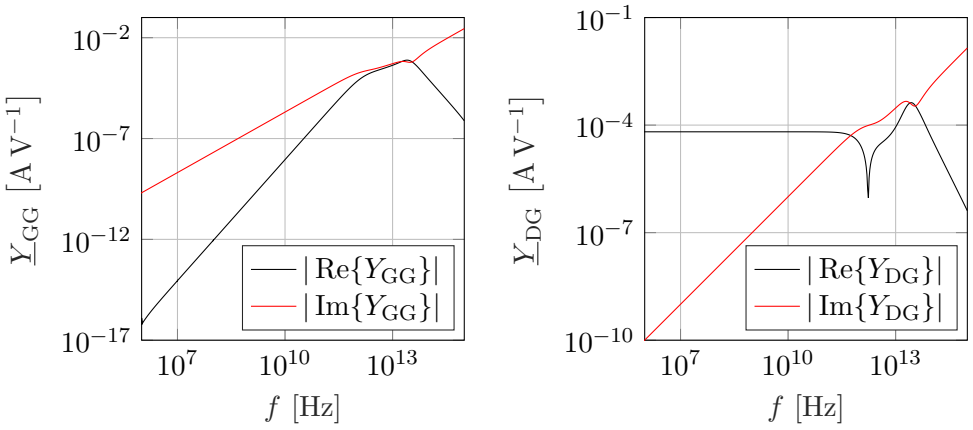

Figure 5.8: Real and imaginary parts of $\underline{Y}_{GG}$ (left) and $\underline{Y}_{DG}$ (right) vs. frequency for the 22 nm device biased at $V_{GS} = 0.7\,V$ and $V_{DS} = 0.5\,V$.

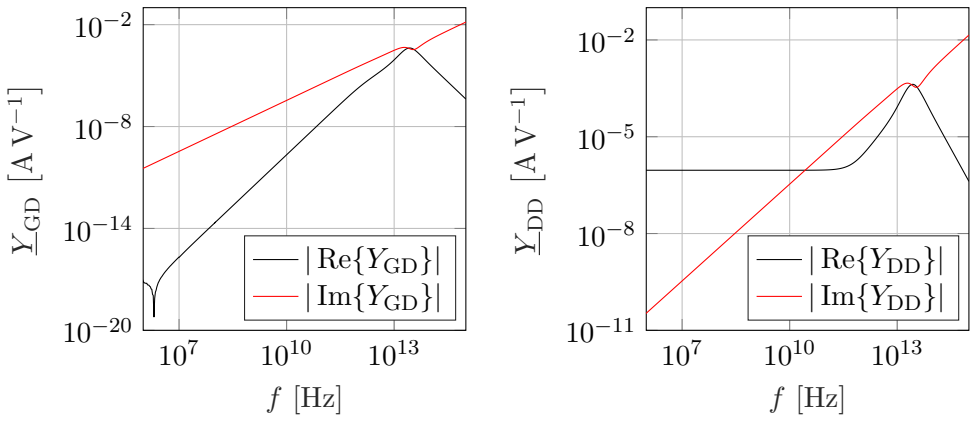

Figure 5.9: Real and imaginary parts of $\underline{Y}_{GD}$ (left) and $\underline{Y}_{DD}$ (right) vs. frequency for the 22 nm device biased at $V_{GS} = 0.7\,\mathrm{V}$ and $V_{DS} = 0.5\,\mathrm{V}$.

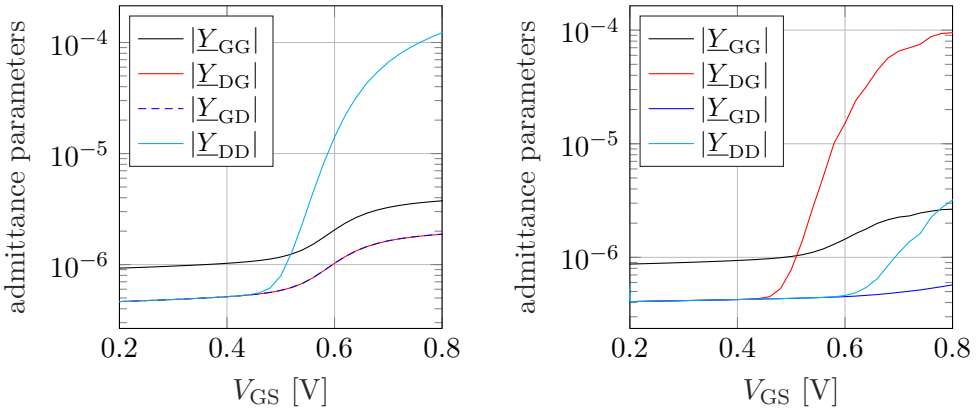

Figure 5.10: Absolute value of the admittance parameters of the 22 nm device versus the gate bias at equilibrium (left), and at $V_{DS} = 0.5\,\mathrm{V}$ (right). Simulations are performed at $f = 10\,\mathrm{THz}$.

the results are not predictive of the device's behavior for $f > 10\,\mathrm{THz}$, since the transistor is then operating in the UV/visible frequency range and the wave energies are high enough to excite carriers into higher subbands. The physics of such phenomena is not included in this work.

In Fig. 5.10 the admittance parameters are calculated for different gate voltages and $V_{DS} = 0.5\,\mathrm{V}$. As mentioned before, the calculated small signal results suffer from numerical artifacts due to H−transformation and show discontinuities whenever the subband profile crosses from one energy box to another leading to severe changes in the derivatives with respect to the potential. This deficiency

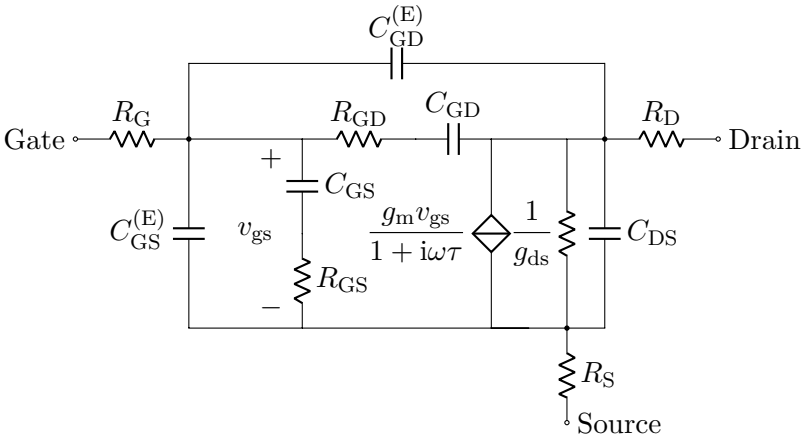

Figure 5.11: Small signal model of the N^+NN^+ nanowire FET [191].

is more evident in the non–equilibrium results for $V_{GS} > V_{TH}$ (see Fig. 5.10), but generally shows when plotting any small signal quantity (e.g. capacitances, cut–off frequency, stability factor) versus the contact biases. In this work, the effect is reduced by choosing a finer H–grid and slightly changing the density of states so that it does not diverge at the subband minimum. Discretization in phase space does not suffer from such discontinuities. At $V_{GS} = 0.7\,\mathrm{V}$, the magnitude of $\underline{Y}_{DD}$ decreases from $66.4\,\mu\mathrm{A\,V^{-1}}$ to $1.09\,\mu\mathrm{A\,V^{-1}}$ when the drain bias is increases to $V_{DS} = 0.5\,\mathrm{V}$ which shows the saturation of the current. Also, the magnitude of $\underline{Y}_{GD}$ decreases from $1.63\,\mu\mathrm{A\,V^{-1}}$ to $0.49\,\mu\mathrm{A\,V^{-1}}$ due to the increased lateral field and velocity overshoot, which reduces the transit time of electrons in the gated area.

Before continuing further with the small signal and noise analysis of the devices, we have to make sure that the numerical scheme satisfies reciprocity under equilibrium conditions for the entire frequency range because our small signal results cannot be more accurate than the error we find in the reciprocity. Moreover, since the calculation of noise involves the same transfer functions, the accuracy of the admittance parameters is expected to set an upper limit for the accuracy of the noise results. The relative error of reciprocity, defined as

$$\mathrm{E_{rec}} = \frac{|\underline{Y}_{GD} - \underline{Y}_{DG}|}{|\underline{Y}_{GD}| + |\underline{Y}_{DG}|} \tag{5.2}$$

is calculated for different frequencies and V_{GS} values under equilibrium condition. In all the simulated cases, this error is less than 10^{-3} and the techniques in Sect. 2.6 reproduce the expected symmetry of small signal admittance parameters at equilibrium. We also observe that changing the spatial and energy grid spacing does not affect the reciprocity of the device. In addition, by checking the positive

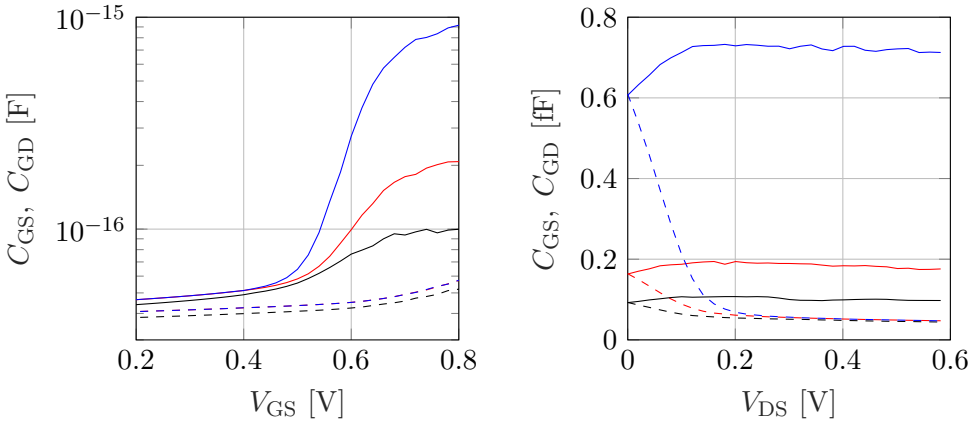

Figure 5.12: C_{GD} (dashed) and C_{GS} (solid) capacitances as a function of V_{GS} (left) and V_{DS} (right) calculated at $f = 10\,\text{GHz}$. The simulations are performed for $L_G = 10\,\text{nm}$ (black), $L_G = 22\,\text{nm}$ (red), and $L_G = 100\,\text{nm}$ (blue).

definiteness of $\text{Im}\{\underline{\hat{Y}}\}$ for all frequencies and simulation parameters we can confirm that the simulated device is passive under equilibrium conditions. Here, the admittance matrix $\underline{\hat{Y}}$ is defined as

$$\underline{\hat{Y}} = \begin{pmatrix} \underline{Y}_{GG} & \underline{Y}_{GD} \\ \underline{Y}_{DG} & \underline{Y}_{DD} \end{pmatrix} . \tag{5.3}$$

A simplified small signal equivalent circuit for the N^+NN^+ nanowire FETs is shown in Fig. 5.11, which helps in developing a crude intuition into the important device parameters and can be used to model the device in a complete electronic circuit. In this equivalent circuit, R_S and R_D are the source and drain resistances, respectively. $C_{GS}^{(E)}$ and $C_{GD}^{(E)}$ are extrinsic gate–source and gate–drain capacitances which are closely linked to the fabrication process, whereas g_{ds} and g_m are intrinsic source–drain conductance and transconductance. R_{GS} and R_{GD} are distributed channel resistances, and C_{GS} and C_{GD} are intrinsic gate–source and gate–drain capacitances, respectively. Under the assumptions $\omega^2 R_{GS}^2 C_{GS}^2 \ll 1$, $\omega^2 R_{GD}^2 C_{GD}^2 \ll 1$, and $\omega^2 \tau^2 \ll 1$, the admittance parameters can be approximated as

$$\underline{Y}_{GG} \approx \omega^2 (R_{GS} C_{GS}^2 + R_{GD} C_{GD}^2) + \mathrm{i}\omega (C_{GS} + C_{GD}) \tag{5.4}$$

$$\underline{Y}_{GD} \approx -\omega^2 R_{GD} C_{GD}^2 - \mathrm{i}\omega C_{GD} \tag{5.5}$$

$$\underline{Y}_{DG} \approx g_m (1 + \tau) - \omega^2 R_{GD} C_{GD}^2 - \mathrm{i}\omega C_{GD} \tag{5.6}$$

$$\underline{Y}_{DD} \approx g_{ds} + \omega^2 R_{GD} C_{GD}^2 + \mathrm{i}\omega (C_{SD} + C_{GD}) \tag{5.7}$$

after the extrinsic capacitances and resistances (parasitics) are de-embedded. Hence, various elements in the equivalent circuit of Fig. 5.11 can be extracted

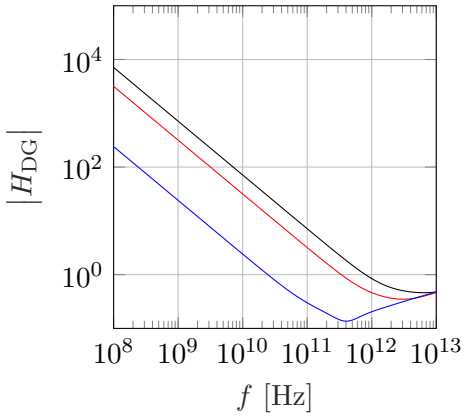

Figure 5.13: The small signal current gain vs. frequency at $V_{GS} = 0.7\,V$ and $V_{DS} = 0.5\,V$. The simulations are performed for $L_G = 10\,nm$ (black), $L_G = 22\,nm$ (red), and $L_G = 100\,nm$ (blue).

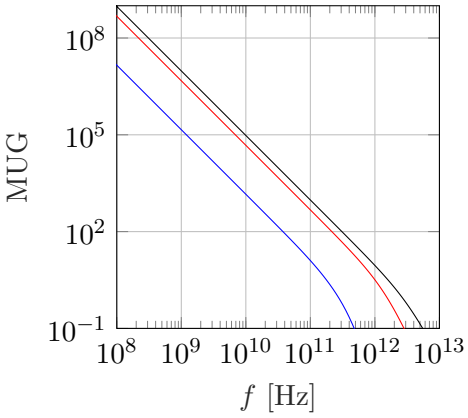

Figure 5.14: The maximum unilateral power gain vs. frequency at $V_{GS} = 0.7\,V$ and $V_{DS} = 0.5\,V$. The simulations are performed for $L_G = 10\,nm$ (black), $L_G = 22\,nm$ (red), and $L_G = 100\,nm$ (blue).

from the real and imaginary parts of the admittance parameters. As an example, the calculated gate–source and gate–drain capacitances are extracted as

$$C_{GD} = -\frac{\mathrm{Im}\{\underline{Y}_{GD}\}}{\omega}\,, \quad C_{GS} = \frac{\mathrm{Im}\{\underline{Y}_{GD} + \underline{Y}_{GG}\}}{\omega} \tag{5.8}$$

and plotted in Fig. 5.12. The results are shown as a function of V_{GS} (for $V_{DS} = 0.5\,V$), and as a function of V_{DS} (for $V_{GS} = 0.7\,V$). For $V_{GS} < V_{TH}$, the depletion region in the channel extends beyond the edges of the gate to some distance inside the source and the drain (see Fig. 5.2). The extension of the depletion region is larger in the drain side of the channel and C_{GD} is strongly reduced for large $V_{DS} > 0$ (i.e., when the device is operating in the saturation regime). This makes C_{GS} the main capacitive component for the determination of dynamic figures of merit. At $V_{GS} > V_{TH}$, this capacitance begins to saturate to the oxide capacitance due to surface potential pinning in the gated area. As for C_{GD}, the curves tend to constant values for $V_{DS} > V_{DS,sat}$ (about 0.2 V for the simulated device) because the channel is depleted on the drain side and V_{GD} does not make it conducting. Since our simulator considers the 3D electrostatics and quantum capacitances due to quantization of energy subbands, it can be effectively used for predicting various model parameters for devices with different cross–sectional dimensions, oxide thickness, and doping concentrations.

Considering the target of high frequency electronics, the cut–off frequency (f_T) and the maximum oscillation frequency (f_{max}) are the most widely used performance indicators. As the signal frequency applied to the gated N^+NN^+ structure increases, the charging of gate capacitance by the drain current gets

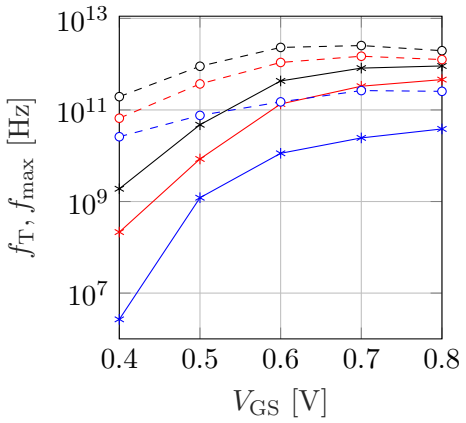

Figure 5.15: Cut-off frequency (solid) and maximum oscillation frequency (dashed) at $V_{GS} = 0.7\,V$ and $V_{DS} = 0.5\,V$. The simulations are performed for $L_G = 10\,nm$ (black), $L_G = 22\,nm$ (red), and $L_G = 100\,nm$ (blue).

suppressed and the device loses its current gain due to the corresponding RC delay. The cut–off frequency is defined as the frequency for which the magnitude of the small signal current gain ($\underline{H}_{DG}$) is reduced to unity, i.e.

$$|\underline{H}_{DG}(f_T)| = \left|\frac{\underline{Y}_{DG}}{\underline{Y}_{GG}}\right| = 1\,. \tag{5.9}$$

On the other hand, the maximum oscillation frequency is defined as the frequency for which the magnitude of the maximum unilateral power gain,

$$\mathrm{MUG} = \frac{|\underline{Y}_{DG} - \underline{Y}_{GD}|^2}{4\,(\mathrm{Re}\{\underline{Y}_{GG}\}\,\mathrm{Re}\{\underline{Y}_{DD}\} - \mathrm{Re}\{\underline{Y}_{GD}\}\,\mathrm{Re}\{\underline{Y}_{DG}\})}\,, \tag{5.10}$$

becomes unity. The values of f_T and f_{max} can be obtained by extrapolating the small signal current gain and the unilateral power gain to the unity–gain points (i.e. intersection with the 0 dB line) considering they decrease with $-10\,dB/dec$ and $-20\,dB/dec$, respectively. That is, for the arbitrary frequency $f_0 \ll f_T$, we have[4]:

$$f_T \approx |\underline{H}_{DG}(f_0)|f_0\,, \quad f_{max} \approx \sqrt{U(f_0)}f_0\,. \tag{5.11}$$

As shown in Figs. 5.13 and 5.14, the curves of $\underline{H}_{DG}$ and MUG exhibit the theoretically expected slope, and (5.11) gives a sufficiently accurate approximation. In Fig. 5.15, the quantities f_T and f_{max} are plotted versus V_{GS} for devices of different gate lengths. They initially increase in the subthreshold regime, and

[4]The frequency f_0 needs to be chosen so that the admittance parameters are above the numerical precision. Here, $f_0 = 1\,MHz$ is considered for the extrapolations.

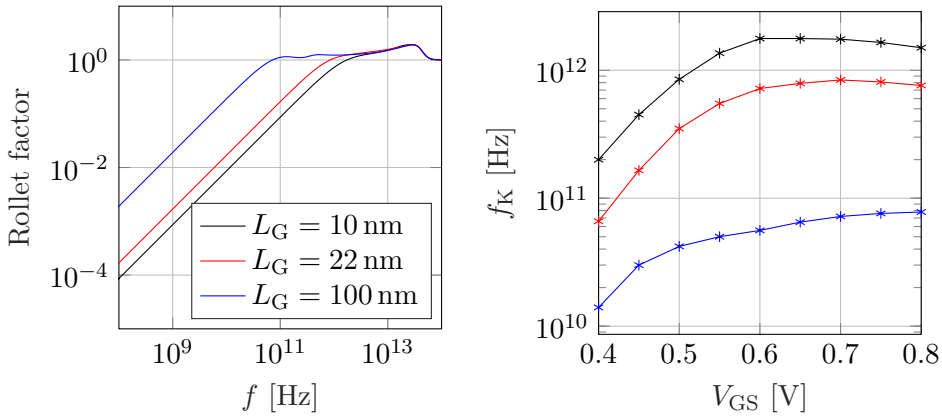

Figure 5.16: Rollet factor K versus frequency for devices of different L_G (left), f_K versus V_{GS} for devices of different L_G (right).

after V_{GS} reaches the threshold voltage, they start to saturate. The initial increase of f_{max} is driven by increasing I_{DS}, and then, it slightly decreases due to larger gate–drain capacitance at higher gate bias which masks the improvement brought by the increased g_m. At $V_{GS} = 0.7\,\text{V}$ and $V_{DS} = 0.5\,\text{V}$, we get $f_T = 820\,\text{GHz}$, $f_T = 330\,\text{GHz}$, and $f_T = 24.5\,\text{GHz}$ for gate lengths of 10 nm, 22 nm, and 100 nm, respectively. It must be noted that these values correspond to the de-embedded device, and extrinsic elements limit the aforementioned f_T and f_{max} values significantly, especially when the gate length is very small.

Our simulations show good f_{max}/f_T ratios (higher than 3 for the 10 nm device and at least one order of magnitude for longer channels) in a large V_{GS} range. Since f_T and f_{max} are related to the speed and the power gain of a transistor, respectively, it is important that this ratio is as high as possible in order to realize high-power operations at high frequencies. Another important aspect in the overall analysis of a transistor is to identify its potential stability. This may be achieved by calculating the Rollet stability factor K, which is computed from the obtained admittance parameters and specifies whether the device is unconditionally stable or conditionally stable [195]. Assuming $\text{Re}\{\underline{Y}_{GG}\} > 0$ and $\text{Re}\{\underline{Y}_{DD}\} > 0$ which is satisfied across the whole frequency range (see Fig. 5.8), the transistor is unconditionally stable if $K > 1$ for all frequencies. Otherwise, it is only conditionally stable and might lead to oscillations for some combination of generator and load admittances. The Rollet stability factor is given by [195]

$$K = \frac{2\,\text{Re}\{Y_{GG}\}\,\text{Re}\{Y_{DD}\} - \text{Re}\{Y_{GD}Y_{DG}\}}{|Y_{GD}Y_{DG}|}\,, \tag{5.12}$$

and is plotted in Fig. 5.16 for different gate lengths. The results show that $K < 1$ for frequencies below some f_K (shown as a function of V_{GS} in Fig. 5.16), and the

simulated devices are not unconditionally stable.

5.4 Noise Results

While a fair amount of the published literature has been dedicated to the static and dynamic performance of junctionless nanowire structures, little work has been published related to their RF noise behavior. In the past decades, several publications reported dramatic enhancement of the thermal noise in submicron devices compared to the long–channel theory [145, 196, 197], and the presented results call for more advanced models of noise in short channels and novel architectures. This is especially the case for multigate designs, in which the carrier transport happens through the bulk and approximations based on surface conduction are inaccurate for modeling the RF noise. In this section, the results of our noise simulations are presented and the impact of downscaling on the noise performance of junctionless nanowire FETs is discussed. The results are expressed mostly in terms of the PSDs of terminal currents, since the fluctuations of terminal currents are the quantities of interest for circuit level analyses.

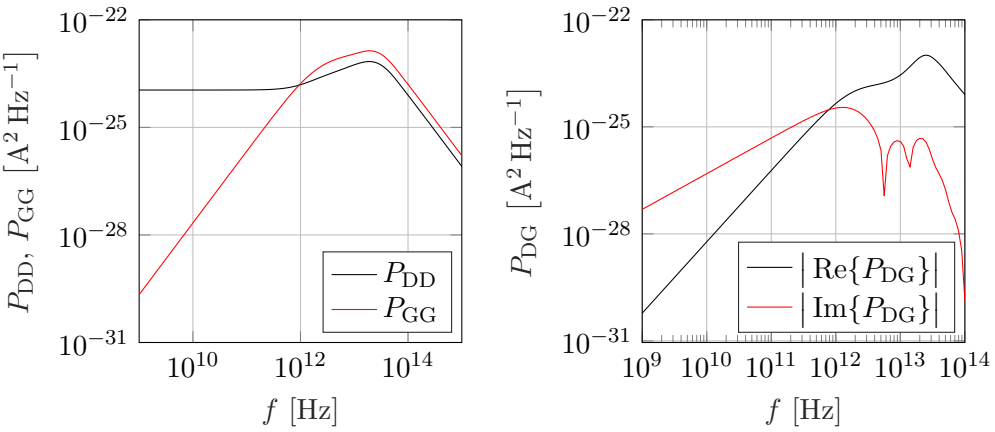

Figure 5.17: Power spectral densities for the 22 nm device vs. frequency, calculated at $V_{GS} = 0.7\,V$ and $V_{DS} = 0.5\,V$.

First, we show the calculated PSDs for the terminal currents in Fig. 5.17 at $V_{GS} = 0.7\,V$ and $V_{DS} = 0.5\,V$. In RF applications, random potential fluctuations resulting from the channel noise are strongly coupled to the gate terminal through the gate capacitance, causing an induced gate noise. At low frequencies, this noise is proportional to ω^2 because it is generated through a capacitor coupling with the noise current $\propto \omega$ flowing from the channel into the gate terminal[5]. The

[5] As explained in chapter 2, our simulation framework assumes ideal gate dielectrics and the gate leakage current due to quantum–mechanical tunneling [57, 198] is neglected. A leaky gate dielectric is expected to add a shot noise component $P_{GG} = 2qI_G$ to the calculated results. It was shown in [57], that for $f > 1\,GHz$ the shot noise of gate leakage is completely dominated by

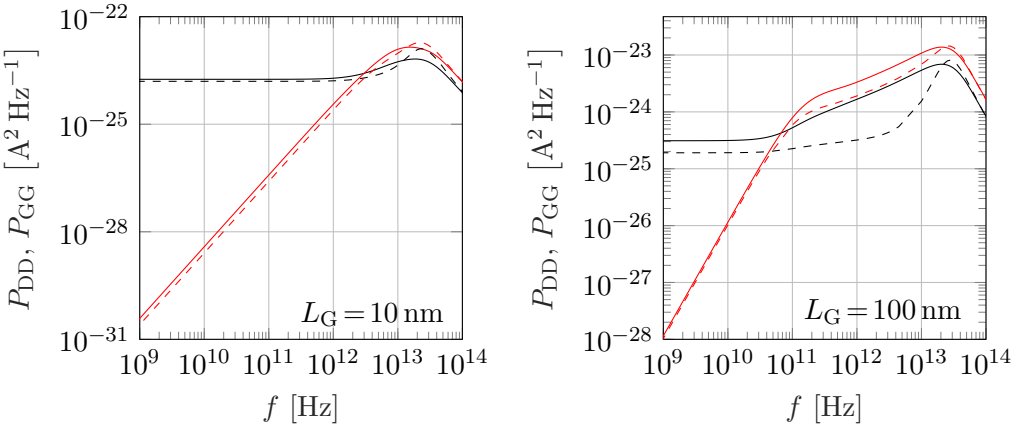

Figure 5.18: Power spectral densities P_{DD} (black) and P_{GG} (red) vs. frequency at $V_{\mathrm{GS}} = 0.7\,\mathrm{V}$. The results are calculated for $V_{\mathrm{DS}} = 0\,\mathrm{V}$ (solid lines) and $V_{\mathrm{DS}} = 0.5\,\mathrm{V}$ (dashed lines).

PSD of the drain current, on the other hand, is given by white noise at lower frequencies and decreases with ω^{-2} at frequencies above 500 THz. The values of P_{DD} at lower frequencies decrease as the device is moved away from equilibrium. Since the cross PSDs are complex conjugate pairs (i.e. $P_{\mathrm{DG}} = P^{*}_{\mathrm{GD}}$), only P_{DG} is shown in Fig. 5.17, which is dominated by its imaginary part at lower frequencies. For frequencies above 1 THz, $\mathrm{Re}\{P_{\mathrm{DG}}\}$ becomes larger compared to its imaginary part.

In Fig. 5.18 the drain and gate PSDs are compared for the 10 nm and 100 nm device. At lower frequencies and under equilibrium conditions, P_{DD} of the 10 nm device is almost six times larger than that of the 100 nm device, while its P_{GG} is 2–3 orders of magnitude smaller. As we move to the saturation regime ($V_{\mathrm{DS}} = 0.5\,\mathrm{V}$), the difference of P_{DD} becomes larger, which is explained when we calculate the drain excess noise for the simulated devices. Figure 5.18 also shows that the equilibrium and non–equilibrium curves of the 10 nm device are rather similar for frequencies above 100 GHz, whereas in the 100 nm device they differ significantly.

The validity of our noise calculations is first checked by showing that the Nyquist theorem[6] holds under equilibrium conditions. The relative error of the PSD of the drain current compared to the Johnson–Nyquist noise, defined as

$$\mathrm{E}_{\mathrm{nyq}} = \left| \frac{P_{\mathrm{DD}} - 4k_{\mathrm{B}}T\,\mathrm{Re}\{\underline{Y}_{\mathrm{DD}}\}}{P_{\mathrm{DD}}} \right| \tag{5.13}$$

induced gate noise which has an ω^2 dependence.

[6] which is a special case of the fluctuation–dissipation theorem. The reader can refer to [198] for more information.

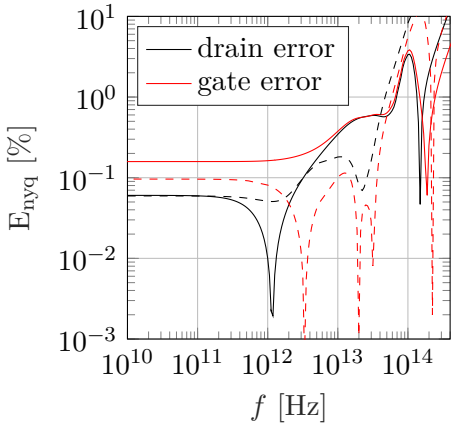

Figure 5.19: Relative deviation of numerical results from the Nyquist theorem. The simulations are performed with $\Delta z = 1\,\mathrm{nm}$, $\Delta \mathrm{H} = 5.2\,\mathrm{eV}$ (solid) and $\Delta z = 0.5\,\mathrm{nm}$, $\Delta \mathrm{H} = 2.6\,\mathrm{eV}$ (dashed).

is plotted versus frequency in Fig. 5.19. For frequencies below 2 THz the maximum deviation is about 0.06 %, which shows excellent agreement considering the very different methods used for calculation of the drain self–admittance and P_{DD}. For $v_{\mathrm{GR}} \approx 1.23 \times 10^{7}\,\mathrm{cm\,s^{-1}}$ at the source and drain boundaries, it is very important to include the contact generation/recombination noise (with its nonlinear terms) in the total PSD of the drain current since its contribution to the fulfillment of the Nyquist theorem is not negligible. Moreover, the high accuracy in reproduction of the Nyquist theorem is due to careful treatment of the reciprocity in [69] and correct expressions for the PSD of scattering integral and generation/recombination fluctuations. Without such considerations, the resulting error is at least one order of magnitude larger and the simulations cannot be trusted to give a quantitative description of the noise. However, beyond a frequency of 4 THz the error in the Nyquist theorem grows and reaches about 3.5 % at a frequency of 100 THz, calling into question the reliability of our simulator at $f > 10\,\mathrm{THz}$. As mentioned before, these frequencies reach beyond the physical phenomena included in our simulation approach and it is not important whether or not the Johnson–Nyquist noise is reproduced anyway.

The drain and gate noise in MOSFETs are often characterized by the respective excess noise factors. The drain excess noise factor γ is defined by

$$\gamma = \frac{P_{\mathrm{DD}}}{4 k_{\mathrm{B}} T g_{\mathrm{D0}}}\,, \tag{5.14}$$

where g_{D0} is the drain output conductance under zero drain bias and zero frequency. In the long channel limit we have $2/3 \leq \gamma \leq 1$, where $\gamma = 2/3$ holds in the saturation regime and $\gamma = 1$ holds at equilibrium [198]. Analogously, the

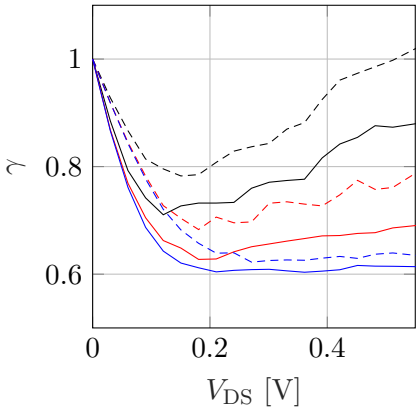

Figure 5.20: Drain excess noise factor γ vs. drain bias V_{DS} at $V_{GS} = 0.7\,V$ (solid) and $V_{GS} = 0.8\,V$ (dashed). Simulations are performed at $f = 1\,GHz$, and for gate lengths of 10 nm (black), 22 nm (red), and 100 nm (blue).

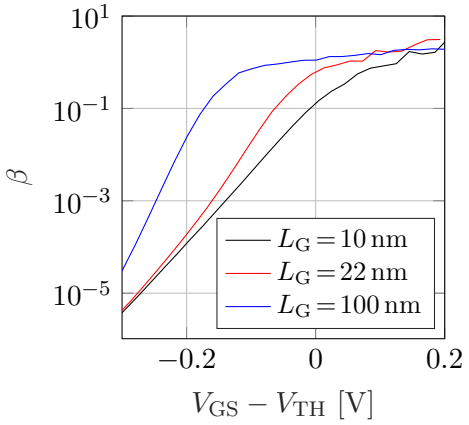

Figure 5.21: Gate excess noise factor β vs. gate bias at $V_{DS} = 0.5\,V$. Simulations are performed at $f = 1\,GHz$, and for gate lengths of 10 nm (black), 22 nm (red), and 100 nm (blue).

gate excess noise factor is defined by [198, 199]

$$\beta = \frac{5 g_{D0} P_{GG}}{4 k_B T (\omega C_{GS})^2} \tag{5.15}$$

where C_{GS} is the gate–source capacitance at the given bias and frequency. The theoretical long–channel value of β is $\beta = 4/3$ for planar MOSFETs.

In Fig. 5.20 the drain excess noise is plotted versus V_{DS} for the ON–state. The results at equilibrium show that our noise simulations conform with the Nyquist theorem, i.e., P_{DD} is equal to the corresponding Johnson–Nyquist noise. The drain excess noise increases for decreasing channel length, and in the long–channel limit, the drain excess noise approaches the theoretical van der Ziel limit of $\gamma = 2/3$ for $V_{DS} > V_{DS,sat}$. For $L_G = 10\,nm$, the PSD of drain current noise increases steadily with the increase in drain bias when $V_{DS} > 0.15\,V$, whereas it is almost constant and weakly dependent on V_{DS} for the 100 nm device. The increase of γ with V_{DS} for the 10 nm channel length is contrary to most of the noise models which predict no dependence of noise on V_{DS} at high drain bias, and can be attributed to DIBL. Other investigations, such as [58, 144] have also found increased values of γ in deep submicron nMOSFETs. Our results show that the enhancement of γ for short channels is not as large as the values given in [145, 196].

Figure 5.21 shows the gate excess noise factor versus the gate bias for various channel lengths. The results show that unlike γ, the gate excess noise for shorter

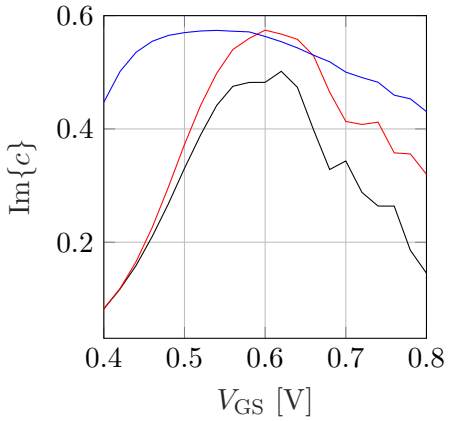

Figure 5.22: Imaginary part of the correlation factor vs. gate bias V_{GS} for $V_{DS} = 0.5\,V$. Simulations are performed at the frequency $f = 1\,GHz$, and for gate lengths of 10 nm (black), 22 nm (red), and 100 nm (blue).

devices is not always larger than for the longer devices. Indeed, for $V_{GS} < V_{TH}$ we can observe that the gate excess noise for the 100 nm device exceeds the 10 nm device. The drain and gate excess noise factors show the same bias dependence as in the conventional inversion–mode MOSFETs, and the results are in agreement with the HD model of [58].

Since P_{DD} and the induced gate noise P_{GG} are generated from the same source of noise, it is possible to evaluate the correlation between them. This is expressed in terms of a correlation coefficient defined by

$$c = \frac{P_{DG}}{\sqrt{P_{DD} P_{GG}}}\,. \tag{5.16}$$

In the long gate limit and for channel induced noise, the correlation coefficient is about 0.395i for planar MOSFETs [198]. In Fig. 5.22 this quantity is shown for different devices. The real part of the correlation factor is negligible (less than 10^{-5}) and not shown in Fig. 5.22. But $\mathrm{Im}\{c\}$ is a strong function of V_{GS}. By increasing the gate voltage in subthreshold regime, the correlation factor rises up to approximately 0.57i for devices with longer channels and then it decreases. While the 100 nm device retains the gate–drain noise correlation for a larger V_{GS} range, the gate–drain correlation of the 10 nm device drops more rapidly and its value peaks at $c = 0.5i$ for $V_{GS} = 0.62\,V$. These results are quite different from the long–channel limit for planar MOSFETs, and can be a subject of further study.

The shot noise suppression factor (Fano factor) is plotted in Fig. 5.23 versus the gate bias. This quantity is defined as

$$\alpha_{Fano} = \frac{P_{DD}}{2qI_{DC}}\,, \tag{5.17}$$

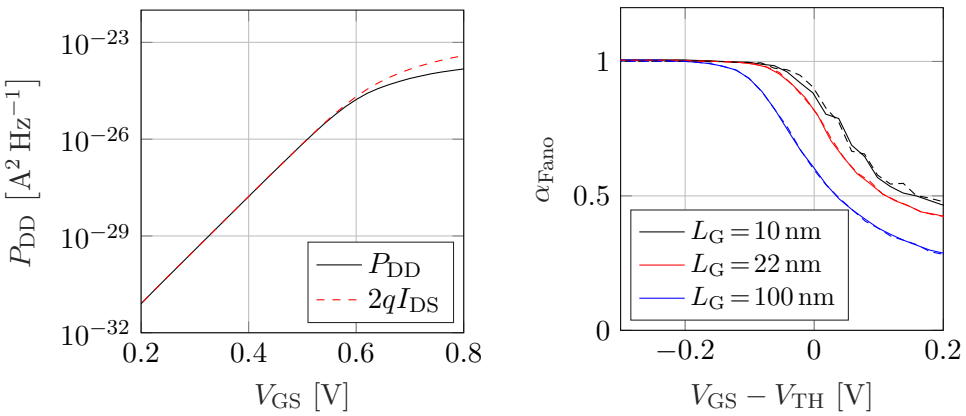

Figure 5.23: Left: PSD of drain current fluctuations (black) vs. V_{GS} for $V_{DS} = 0.5$ V and $f = 1$ GHz of the 22 nm device, together with the equivalent shot noise $2qI_D$ (red dashed line). Right: Fano factor F for different devices biased at $V_{GS} = 0.7$ V and $V_{DS} = 0.5$ V.

with $\alpha_{Fano} = 1$ implying that the injection of electrons over the potential barrier is a Poisson process and the drain current noise is purely shot noise, while $\alpha_{Fano} < 1$ represents the suppression of shot noise [198]. The Fano factor is nearly unity in the subthreshold regime due to the low density of carriers and then decreases for $V_{GS} > V_{TH}$ when the carriers in the channel become abundant and inevitably interact with the electrons injected into the channel from the source side[7] For the 10 nm device, the limited number of scattering events in the channel results in reduced suppression of the non–equilibrium noise component, whereas in the case of the 100 nm device the correlation of electrons in the channel and consequently the noise suppression becomes more pronounced [200]. The small dependency of α_{Fano} on the drain bias is mainly due to the DIBL.

In Fig. 5.24 the local contribution to the noise is shown for equilibrium and non–equilibrium conditions. Such insights can only be obtained by solving the Langevin–BE. The results show that under equilibrium conditions, the drain current noise originates mostly from the low density region while the high density N^+–regions contribute mostly to the gate current noise. In the strong non–equilibrium conditions, the noise originates from the left side of the potential barrier (source/channel region), which is consistent with the shot noise [198]. Subsequent heating of the carriers in the channel/drain region has very little impact on the drain current noise, mainly because both the electron density and Green's function of the drain current noise are very small in the region of hottest electrons. These results are consistent with the findings for double–gate N^+NN^+

[7] This is due to the self–consistent solution of the PE and BE, as well as the Pauli exclusion principle and should not be confused with the electron–electron scattering mechanism.

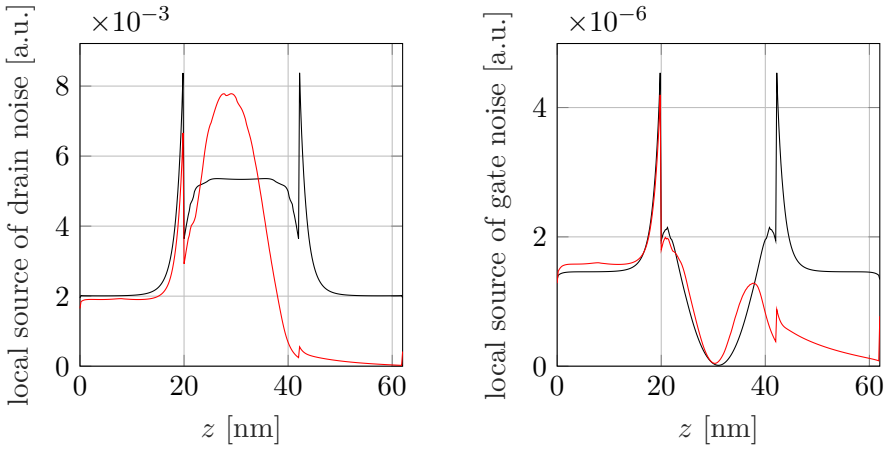

Figure 5.24: Spatial distribution of the terminal current noise at equilibrium (black) and $V_{\mathrm{DS}} = 0.5\,\mathrm{V}$ (red). Simulations are performed for the 22 nm device at $V_{\mathrm{GS}} = 0.7\,\mathrm{V}$ and $f = 10\,\mathrm{GHz}$.

structures [69] and MOSFETs [201, 202].

5.5 Quasi-Ballistic Simulations

As the semiconductor devices are scaled down to nanometer scales, carrier transport approaches the ballistic limit [203–206] and investigation of ballistic transport becomes an important aspect of device simulation[8]. The distribution function in the quasi–ballistic limit of a submicron N^+NN^+ structure was first evaluated in [211] using the 3D MC simulations. Other works, such as [71] have investigated the ballistic limits of electron transport over the potential barriers and present the distinctive features of ballistic distribution function (e.g. asymmetry at the top of the energy barrier, and the development of a ballistic peak along the channel) using MC simulations or analytical models based on the difference of source and drain injection fluxes. Previous studies are, however, not adequate for accurate characterization and understanding the details of quasi-ballistic transport in nanoscale FETs because they do not consider vertical quantum confinement and are limited to the steady state behavior of the quasi–ballistic devices. In this section, transport in the quasi–ballistic regime for the junctionless device of Fig. 2.1 is discussed. Since the qualitative results do not depend on the gate length, most of the simulations are performed for the 22 nm device. We have used $m_{xx} = m_{zz} = 0.19m_e$ and $m_{yy} = 0.98m_e$, and the doping concentrations for the N^+– and N–region are set to $5\times10^{19}\,\mathrm{cm}^{-3}$ and $1\times10^{18}\,\mathrm{cm}^{-3}$, respectively. For the sake of CPU efficiency, the comparisons

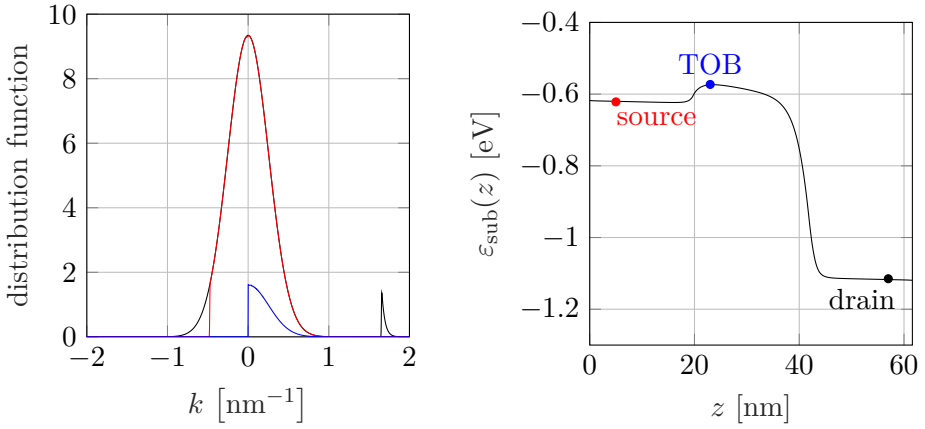

Figure 5.25: Distribution functions at the drain contact (black), source contact (red), and top of the potential barrier (blue) for the ballistic case at $V_{GS} = 0.5\,\mathrm{V}$, $V_{DS} = 0.5\,\mathrm{V}$.

[8] In fact, several structures other than the short–channel N^+NN^+ FETs have the conditions which result in ballistic transport. Some examples are the graded bandgap structures and heterojunction transistors formed of $Al_xGa_{1-x}As$ and GaAs [207–209], as well as planar doped barrier structures [210].

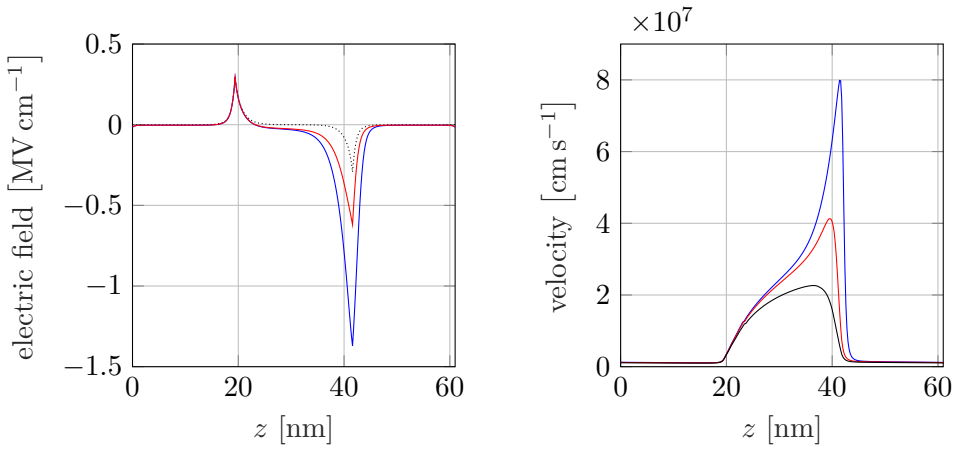

Figure 5.26: The electric field profile (left) and average velocity profile (right) of the 22 nm device for $V_{GS} = 0.5\,V$ and different V_{DS} values using the single band RTA model. The solid curves correspond to $V_{DS} = 0.1\,V$ (black), $V_{DS} = 0.2\,V$ (red), and $V_{DS} = 0.5\,V$ (blue). The dashed curve shows the electric field under equilibrium condition.

are made for single subband and single valley transport, with the RTA for the scattering integral. These assumptions are justified since for high mobilities the impact of scattering is anyway weak and the error due to the RTA is small. However, the results of the phase space discretization for the full band structure of silicon and various electron–phonon scattering modes are verified by comparing the I_{DS}–V_{DS} characteristics to the results of H-transformation.

First, simulation of the ballistic case (i.e. $S\{f\} = 0$) is demonstrated, because this is the case for which the conventional transformations to energy fail[9]. In Fig. 5.25, the distribution functions at three different locations of the 22 nm device (near the source side, near the drain side, and at the top of the potential barrier as marked on the subband energy profile of Fig. 5.25) are plotted for $V_{GS} = 0.5\,V$ and $V_{DS} = 0.5\,V$. The results are for a k–grid spacing of $\Delta k = 2.2 \times 10^{-3}\,nm^{-1}$ at low energies, and k–boxes corresponding to $\Delta\varepsilon = 2.5\,meV$ at higher energies. The maximum wave number gives $\varepsilon_{max} = 1.01\,eV$. The distribution function is positive in the whole phase space and shows no spurious oscillations. The obtained results depend on the height of the source–channel energy barrier ε_B, which is about 43 meV in this example. In the source region, electrons injected from the source contact with $\varepsilon < \varepsilon_B$ get reflected back by the barrier, filling their negative counterpart while electrons with $\varepsilon > \varepsilon_B$ get transmitted across the barrier. For $qV_{DS} > 4k_BT$ there are almost no carriers injected from the drain contact. Hence,

[9] In the case of H-transformation, the left-hand side matrix becomes singular in the ballistic limit.

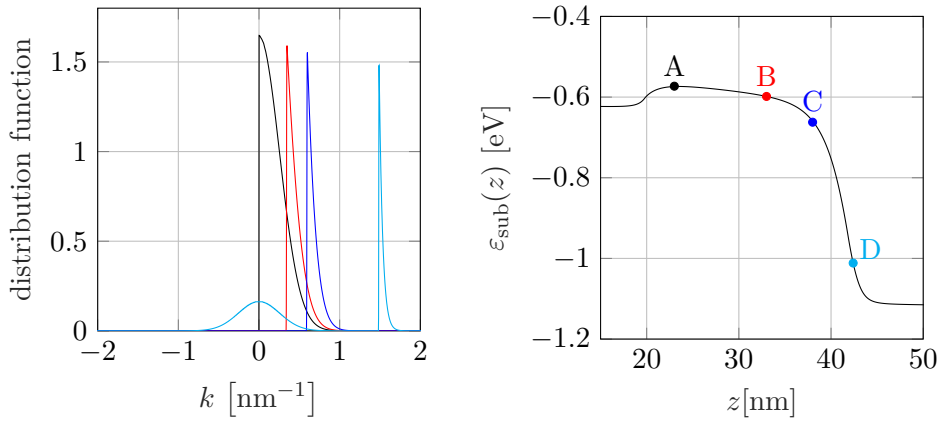

Figure 5.27: The distribution function at different positions of the gated N-region for $V_{GS} = 0.5\,V$ and $V_{DS} = 0.5\,V$.

the distribution at the source side is effectively zero for $k < -0.478\,nm^{-1}$. At the top of the potential barrier, the distribution function becomes a hemi–Maxwellian for $k > 0$ and the semiclassical transmission shows a discontinuous change from 0 to 1. At the drain contact, the thermal equilibrium injection from the drain side with $\varepsilon < \varepsilon_B + qV_{DS}$ (nearly the entire hemi–Maxwellian) gets reflected by the barrier, and builds up a symmetric thermal distribution. While the thermal portion remains in equilibrium, the electrons injected from the source side with $\varepsilon > \varepsilon_B$ add a non-symmetric contribution which peaks at approximately $1.65\,nm^{-1}$ (corresponding to qV_{DS}), thereby providing a ballistic limit to the current density.

It is important to be able to evaluate the ballistic distributions throughout the entire channel and not just at the top of the source–channel barrier. In order to better understand the ballistic transport of electrons from top of the barrier to the drain end, in Fig. 5.26 the electric field profile of the 22 nm device for the ballistic case is plotted under different bias conditions. At equilibrium, the sharp gradients of potential near the NN^+ doping steps result in very large built–in electric fields ($\approx 0.3\,MV\,cm^{-1}$). Applying a nonzero drain bias increases the electric field in the high–resistivity N–region and at the drain side of the channel, while the electric field at the source–channel barrier remains nearly unchanged. This leads to the $k > 0$ electrons at the top of the source–channel barrier being accelerated by the high field in the N–region, creating a prominent ballistic peak in the positive half of the distribution function. As the carriers are accelerated along the channel, the location of this peak moves toward higher velocities. Moreover, the peak gets narrower as the carrier density decreases which preserves the current density (see Fig. 5.27). In the highly non–equilibrium

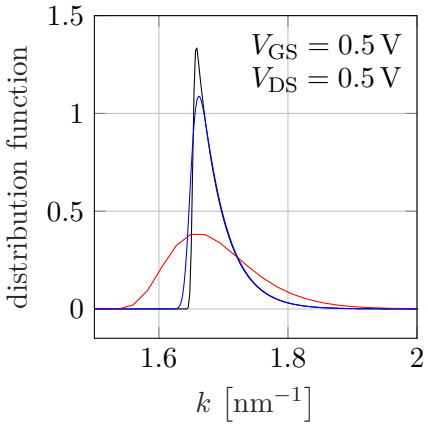

Figure 5.28: Ballistic peak of the distribution function for the 22 nm device. The results are for $\Delta k = 2.2\,\mu m^{-1}$, $\Delta z = 2\,nm$ (black), $\Delta k = 22\,\mu m^{-1}$, $\Delta z = 2\,nm$ (red), and $\Delta k = 2.2\,\mu m^{-1}$, $\Delta z = 0.2\,nm$ (blue).

case, the average velocity of the electrons in the channel region increases and exceeds the thermal velocity ($v_{th} \approx 2\times10^7\,cm\,s^{-1}$) by a substantial amount. For $V_{DS} = 0.5\,V$, the velocity peaks at approximately $8\times10^7\,cm\,s^{-1}$ at the drain side of the NN^+ interface (see Fig. 5.26).

Although the cutoff at $k = 0$ by the source–channel barrier is abrupt, the left flank of the ballistic peak at the drain side is not as abrupt as at the top of the barrier. This is due to the finite k–space grid which leads to artificial broadening of the distribution function in the k–space, i.e. for a nonzero step of the subband energy at $z_{i+½}$, any k–space box on z_i is connected to multiple boxes on z_{i+1} and vice versa (see Sect. 4.1.1). This is a fundamental problem of tensor–product grids in phase space and occurs regardless of the treatment of the derivative with respect to the wave number. In principle, a finer k–grid could reduce the diffusion, but a finer k–space grid requires a finer real space grid to prevent numerical problems and a certain level of diffusion cannot be avoided. In Figs. 5.25 and 5.27, the numerical diffusion of the ballistic peak is not discernible because a very fine k–grid with a very coarse z–grid ($\Delta z = 2\,nm$) are chosen in order to minimize the numerical diffusion and show the ballistic effects better at the expense of the quantitative accuracy. Refining the real space grid (which is necessary for accurately resolving the sharp gradients especially in the deep–subthreshold regime) increases the numerical diffusion. In Fig. 5.28 the ballistic peak of the distribution function at the drain side is calculated for various Δz and Δk values. We observe that both a coarser k–grid (red curve) and a finer z–grid (blue curve) increase the artificial broadening.

Moreover, the results in Figs. 5.25 and 5.27 show that the height of the ballistic peak decreases as the electrons move toward the drain contact. This

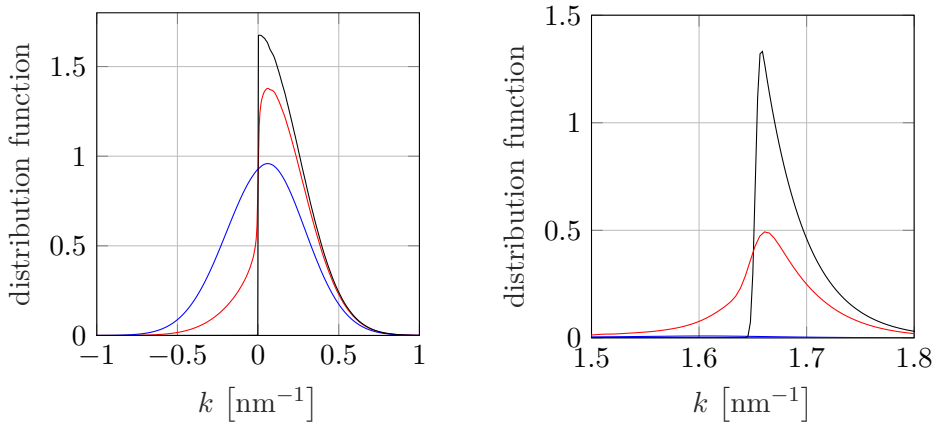

Figure 5.29: Distribution function at the top of the barrier (left), and the ballistic peak of the distribution function at the drain side (right) for $V_{GS} = 0.5\,V$ and $V_{DS} = 0.5\,V$. The results are for the mobilities $\mu = 10^6\,cm^2/Vs$ (black), $\mu = 1000\,cm^2/Vs$ (red), and $\mu = 100\,cm^2/Vs$ (blue).

effect, which is another consequence of the numerical diffusion, suggests that our discretization in phase space does not conserve the total electron energy. However, the impact of the artificial diffusion is weak, and the phase space solver based on the tensor–product grid does not show spurious oscillations or negative values of distribution function in the ballistic case.

The observed features of the carrier distribution function in the ballistic regime, such as the discontinuous asymmtery and development of the sharp ballistic peak leading to velocity overshoot, cannot be captured by any lower order polynomial approximation as employed in conventional macroscopic models and needs to be treated in the k–space. In addition, thermal injection of the electrons from the source side is a current limiting constraint that needs to be incorporated in the BCs of any quasi–ballistic description of the device. While the thermal bath BCs are inherent in the first principles such as NEGF formalism and MC simulations, they are usually neglected when macroscopic models are used for investigation of quasi-ballistic phenomena.

Figure 5.29 shows the effect of scattering on the non–equilibrium distribution function. Here, the focus is on the distinctive features of the shape of distribution function in ballistic limit. Therefore, the discontinuity of $f(z,k)$ at the top of the barrier and the ballistic peak at the drain side are plotted for $V_{GS} = 0.5\,V$ and $V_{DS} = 0.5\,V$ and various mobilities. It is evident that an increase in the mobility sharpens the distribution at the top of the barrier, bringing it closer to the thermoionic emission limit (a hemi–Maxwellian at $k = 0$). Moving to the drain side, the ballistic peak spreads for lower mobilities. At $\mu = 100\,cm^2/Vs$

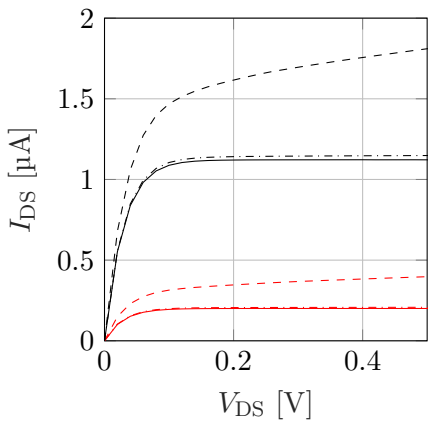

Figure 5.30: I_{DS} versus V_{DS} for ballistic simulations. Simulations are performed for $V_{GS} = 0.5\,V$ (red) and $V_{GS} = 0.6\,V$ (black). The simulated gate lengths are $L_G = 100\,nm$ (solid), $L_G = 22\,nm$ (dash-dotted), and $L_G = 10\,nm$ (dashed).

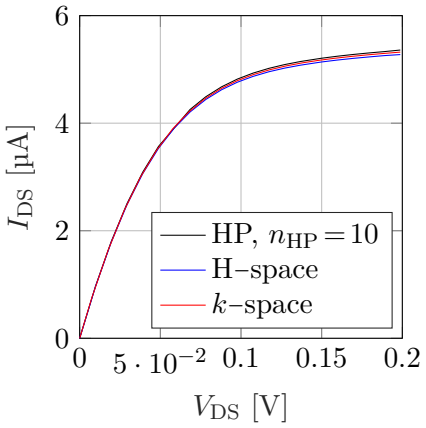

Figure 5.31: I_{DS}–V_{DS} curves of the 22 nm silicon junctionless transistor for $V_{GS} = 0.6\,V$. The results are obtained from simulation of the BE under H-transformation, simulation of the BE in phase space, and the moments model with $n_{HP} = 10$.

this ballistic peak has nearly vanished and the distribution function can be safely approximated with a drifted Maxwellian, which fits into the macroscopic description based on moments equations.

Having calculated the distribution function for several sets of grid parameters, we can now discuss other quantities of interest, such as the steady state current for various bias conditions. In equilibrium, the distribution function is completely symmetric and $I_{DS} = 0$. The symmetry of the distribution function is due to the balanced injections from the source and drain contacts instead of detailed balance properties of the scattering terms in the diffusive regime which relax the distribution function to the symmetrical Fermi–Dirac or Maxwell–Boltzmann functions. As V_{DS} increases, the negative half of the distribution function becomes suppressed leaving the positive half (source injection) unbalanced. The device operates in linear regime for $V_{DS} < 0.15\,V$, and at the total suppression of the drain injection (at $V_{DS} = 0.15\,V$) I_{DS} saturates to its ballistic limit. The output characteristics in Fig. 5.30 show that for the 22 nm and 100 nm devices, the saturation current is nearly independent of the channel length. However, the 10 nm device gives a considerably larger current (nearly twice the long channel values at $V_{GS} = 0.5\,V$ and $V_{DS} = 0.5\,V$) due to DIBL.

The artificial diffusion observed in k–space simulations is especially important in the subthreshold region, as it translates to artificial heating of the electrons and increases the subthreshold slope of the device. This is shown in Fig. 5.32 for $\Delta k = 0.055\,nm^{-1}$ and the spatial grid spacings of $\Delta z = 2\,nm$ and $\Delta z = 0.2\,nm$. The

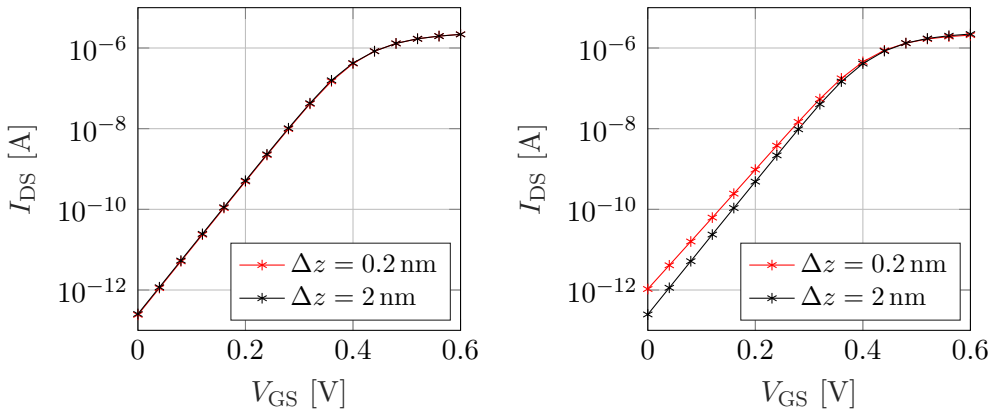

Figure 5.32: Calculated $I_{DS} - V_{GS}$ curves of the silicon N^+NN^+ transistor, obtained from simulation of the BE under H-transformation (left) and simulation of the BE in phase space (right).

effect is not significant and even for this rather coarse k–grid, the subthreshold slope changes from SS = 60.5 mV/dec to SS = 66.5 mV/dec as we move from $\Delta z = 2$ nm to $\Delta z = 0.2$ nm. As expected, refining the phase space grid alleviates this problem and we have SS = 62.5 mV/dec for $\Delta k = 0.0225\,\text{nm}^{-1}$ (see Fig. 5.33). Simulations with the H–transformation do not show this problem, as shown in Fig. 5.32.

Before moving on to investigation of the ballistic limits of small signal parameters, we need to verify that the numerical methods developed in Chapter 4 also work in the diffusive regime. Hence, the 22 nm device is simulated using the conventional model for the band structure of silicon (six ellipsoidal and parabolic valleys) with five subbands per valley. The calculated I_{DS}–V_{DS} curves are presented in Fig. 5.31 (red curves), and compared to consistent simulations by other means. We observe that the different methods produce nearly identical steady state results. At $V_{DS} = 0.2$ V, the calculations by the H–transformed BE and moments equations differ by $\approx 1.5\%$ for $V_{GS} = 0.6$ V and $\approx 0.8\%$ for $V_{GS} = 0.5$ V. Our results confirm the validity of all three approaches for the simulation of conventional devices with relatively low mobilities.

Next, the accuracy of the phase space BE solver for small signal analysis is checked by refining the simulation grid and calculating the ballistic small signal drain self–admittance $\underline{Y}_{DD}$ at $f = 10$ GHz. In Fig. 5.34 the relative error (i.e. error with respect to the finest grid) is plotted versus the grid refinement factor. Our simulations show that the refinements of the z– and k–grids have contrary impacts on the numerical stability of the discretized equations and we have a Courant–Friedrich–Lewy–like condition for Δz and Δk. While we choose a fine k–grid for capturing the ballistic discontinuities of the distribution function, the

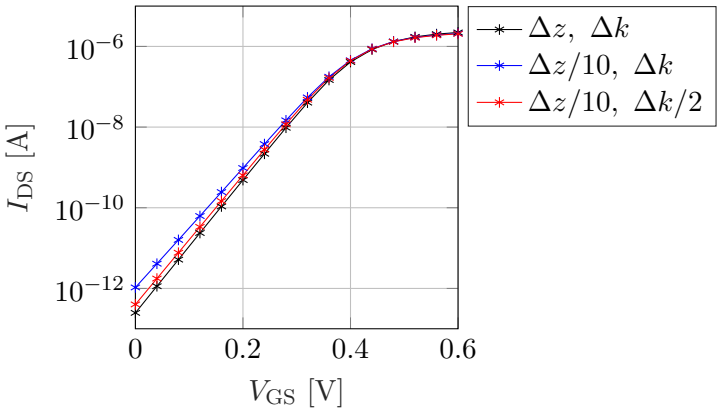

Figure 5.33: Calculated $I_{\mathrm{DS}} - V_{\mathrm{GS}}$ curves of the silicon $\mathrm{N^+NN^+}$ transistor, obtained from simulation of the BE in phase space. $\Delta z = 2\,\mathrm{nm}$ and $\Delta k = 0.055\,\mathrm{nm}^{-1}$ were chosen for these simulations.

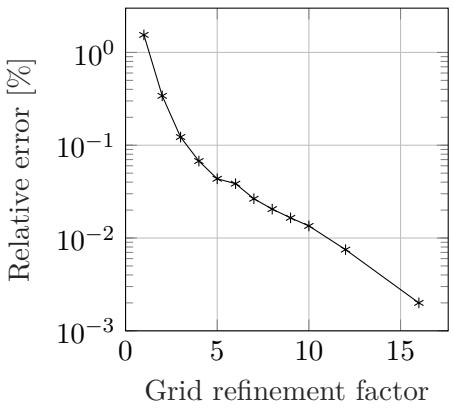

Figure 5.34: Convergence of $\mathrm{Re}\{\underline{Y}_{\mathrm{DD}}\}$ at $f = 10\,\mathrm{GHz}$. The relative error is calculated with respect to a grid refinement of 20 with $\Delta z = 0.1\,\mathrm{nm}$ and $\Delta k = 0.0045\,\mathrm{nm}^{-1}$, and ballistic simulations are performed for the 22 nm device at $V_{\mathrm{GS}} = 0.5\,\mathrm{V}$ and $V_{\mathrm{DS}} = 0\,\mathrm{V}$.

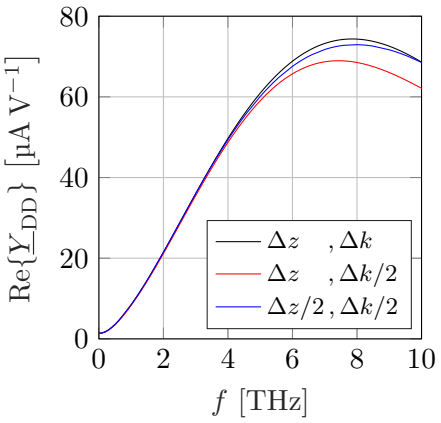

Figure 5.35: $\mathrm{Re}\{\underline{Y}_{\mathrm{DD}}\}$ versus frequency for the 22 nm device at $V_{\mathrm{GS}} = 0.5\,\mathrm{V}$ and $V_{\mathrm{DS}} = 0\,\mathrm{V}$. Simulations are performed for the mobility of $\mu = 100\,\mathrm{cm^2/Vs}$, and different grid spacings.

z–grid must be refined accordingly in order to prevent numerical instabilities. It is also observed that decreasing Δz and Δk changes the simulation results in opposite directions (see Fig. 5.35). Hence, in Fig. 5.34 we have calculated the relative error with simultaneous refinement of both z– and k–grids. The error decreases to less than 0.01 % for finer grids and the method seems to converge.

Regarding the small signal analysis, the drain self–admittance is chosen for

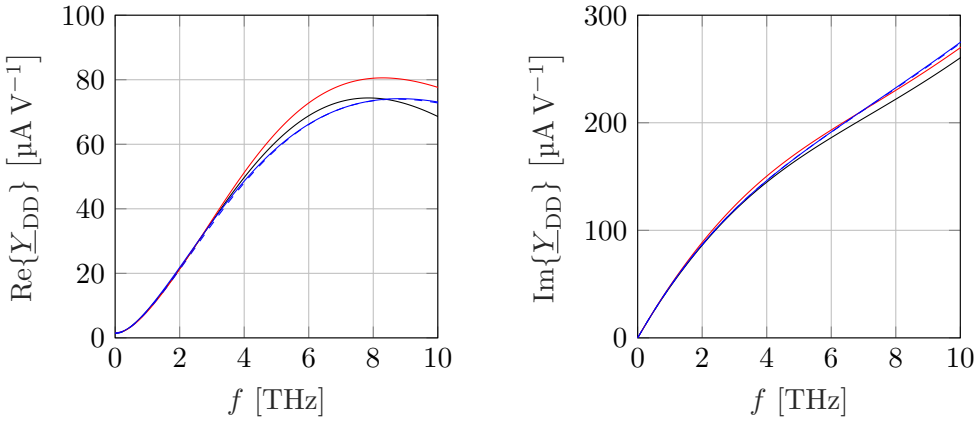

Figure 5.36: Real and imaginary parts of the drain self-admittance versus frequency at $V_{\mathrm{GS}} = 0.5\,\mathrm{V}$ and $V_{\mathrm{DS}} = 0\,\mathrm{V}$. Simulations are performed for the mobility of $\mu = 100\,\mathrm{cm}^2/\mathrm{Vs}$ using the BE (black), DD model with fixed BCs (red), DD model with thermal bath BCs (solid blue), and TM with thermal bath BCs (dashed blue).

comparing the results of the BE solver to those of the macroscopic models. The drain self–admittance is an important quantity for determining the impact of plasma waves on the device's behavior. For example, a negative real part of the drain self–admittance corresponds to an instability and could enable the generation of THz waves [199,212]. In order to obtain a plasma instability specific bias conditions have to be applied to the contacts of the device. The gate/source port should be short–circuited and the drain/source port open (i.e. the drain self–admittance should be zero)[10]. In the following discussions, models based on projection of the BE onto the Hermite polynomials with $n_{\mathrm{HP}} = 2$ and $n_{\mathrm{HP}} = 10$ are abbreviated as the drift–diffusion (DD) and the 10th–moment (TM) model, respectively[11].

In Fig. 5.36 the real and imaginary parts of the drain self–admittance are shown for $\mu = 100\,\mathrm{cm}^2/\mathrm{Vs}$ (i.e. strong damping of the plasma waves)[12]. The results of the phase space BE solver are compared to the values obtained from the moments equations with thermal bath and Dirichlet BCs. In the diffusive transport, good agreement is observed over the entire frequency range and even the DD model with Dirichlet BCs gives reasonably accurate predictions for

[10] These bias conditions should not be confused with the boundary conditions of the transport models at the contacts inside the device, which are determined by the contact model.

[11] In this work, we use the term "drift–diffusion" model in the sense of the TCAD community, although the $n_{\mathrm{HP}} = 2$ model contains the time derivative of the current density. These terms are usually neglected in TCAD.

[12] Since the electron mobility is a more familiar parameter compared to the microscopic relaxation time, in our simulations $\mu = q\tau_{\mathrm{RTA}}/m^{\nu}_{zz}$ is used as an adjustable parameter.

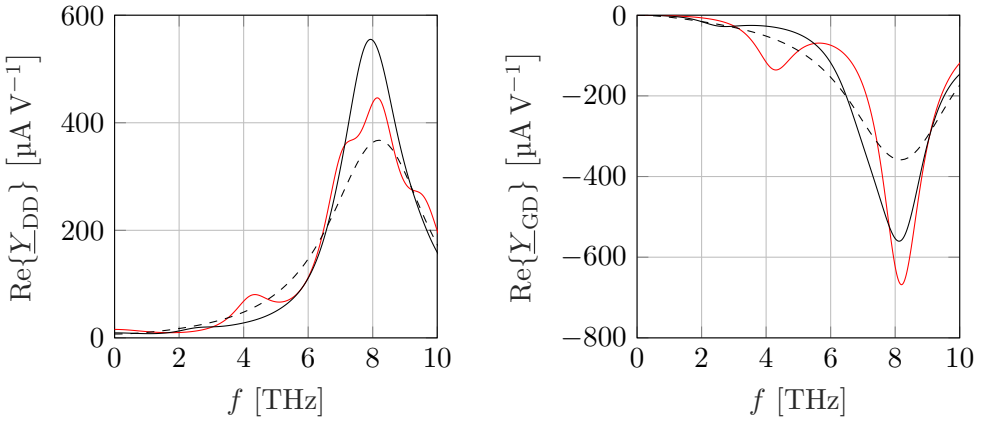

Figure 5.37: $\mathrm{Re}\{\underline{Y}_{\mathrm{DD}}\}$ (left) and $\mathrm{Re}\{\underline{Y}_{\mathrm{GD}}\}$ (right) of the 22 nm device evaluated using the phase space BE solver (dashed) and the moments equations with $n_{\mathrm{HP}} = 2$ (red), $n_{\mathrm{HP}} = 10$ (solid black). Simulations are performed for $\mu = 10^3\,\mathrm{cm}^2/\mathrm{Vs}$ and Dirichlet BCs are imposed.

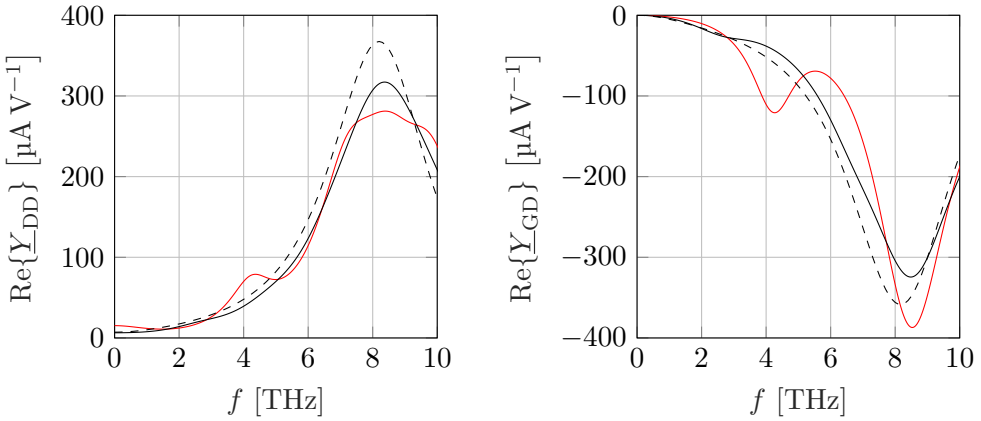

Figure 5.38: $\mathrm{Re}\{\underline{Y}_{\mathrm{DD}}\}$ (left) and $\mathrm{Re}\{\underline{Y}_{\mathrm{GD}}\}$ (right) of the 22 nm device evaluated using the phase space BE solver (dashed) and the moments equations with $n_{\mathrm{HP}} = 2$ (red), $n_{\mathrm{HP}} = 10$ (solid black). Simulations are performed for $\mu = 10^3\,\mathrm{cm}^2/\mathrm{Vs}$ and thermal bath BCs are imposed.

$f < 3\,\mathrm{THz}$. Although the DD model deviates from the results of the BE solver at higher frequencies (e.g. $\mathrm{Re}\{\underline{Y}_{\mathrm{DD}}\}$ has an error of $\approx 8.2\%$ at $f = 8\,\mathrm{THz}$), it performs well in predicting the overall behavior of the admittance parameters. Truncation of the Hermite polynomial expansion at higher orders improves the accuracy of the results, but the results are almost the same for Dirichlet and

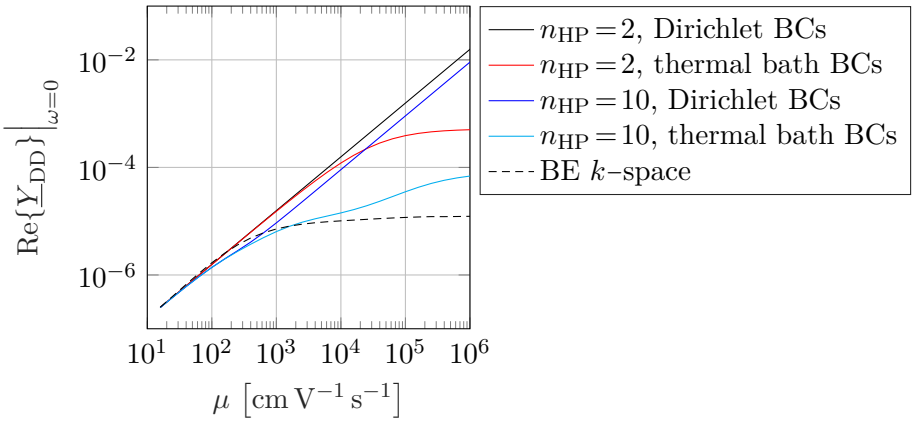

Figure 5.39: Real part of the drain self-admittance for $V_{\mathrm{GS}}=0.5\,\mathrm{V}$ and $V_{\mathrm{DS}}=0\,\mathrm{V}$ at zero frequency.

thermal bath BCs.

As the mobility is increased, the moments equations begin to lose their validity in describing the transport. In order to get a quantitative measure of this discrepancy, Fig. 5.37 compares the real parts of $\underline{Y}_{\mathrm{DD}}$ and $\underline{Y}_{\mathrm{GD}}$ obtained from the BE solver (dashed lines) to the results of moments equations with Dirichlet BCs. Simulations are performed for $\mu=10^3\,\mathrm{cm^2/Vs}$. The zero frequency drain self–admittance of the DD model is twice the value obtained from the BE. Moreover, the results of the DD model show several distinct additional peaks for $\mathrm{Re}\{\underline{Y}_{\mathrm{DD}}\}$ (e.g. at 4.3 THz and 7.4 THz) and are qualitatively different from those of the BE solver. Although truncation of the moments equations at higher orders dampens the additional peaks and reduces the steady state conductance, even the TM model overestimates the zero frequency drain self–admittance by 30 % and predicts a much larger peak at $f=7.9\,\mathrm{THz}$. In Fig. 5.38 the BCs of the moments equations are changed to thermal bath BCs. Although the results of the DD model do not change at zero frequency, the relative error of the zero frequency drain self–admittance for the TM model decreases to below 10 %. Moreover, thermal bath BCs dampen the additional peaks of the DD model and improve the accuracy of the TM model at higher frequencies considerably. The impact of BCs on the small signal results is evident even for the mobility of $\mu=10^3\,\mathrm{cm^2/Vs}$.

This is better demonstrated in Fig. 5.39, where $\underline{Y}_{\mathrm{DD}}$ of the 22 nm device at $f=0$ is plotted as the mobility is varied in the range 20–$10^6\,\mathrm{cm^2/Vs}$. The result depends strongly on the choice of the transport model and the details of BCs. Although the results of the BE saturate at $1.23\,\mu\mathrm{A\,V^{-1}}$, Dirichlet BCs offer no current limiting mechanism and lead to arbitrarily large values of $\underline{Y}_{\mathrm{DD}}$ at $f=0$, $V_{\mathrm{DS}}=0$. Thermal bath BCs improve the results of the moments equations

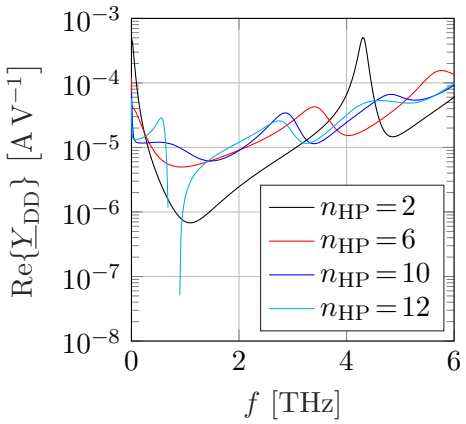

Figure 5.40: $\mathrm{Re}\{\underline{Y}_{\mathrm{DD}}\}$ versus frequency for the 22 nm device at $V_{\mathrm{GS}} = 0.5\,\mathrm{V}$ and $V_{\mathrm{DS}} = 0\,\mathrm{V}$. Simulations are performed for a mobility of $\mu = 10^6\,\mathrm{cm}^2/\mathrm{Vs}$ and thermal bath BCs are applied.

considerably and limit the zero frequency $\underline{Y}_{\mathrm{DD}}$ in the quasi–ballistic regime even for the DD model.

When the mobility is increased to $\mu = 10^6\,\mathrm{cm}^2/\mathrm{Vs}$, moments of the BE fail to even provide a qualitative description of the device's small signal behavior (Fig. 5.41). Although the TM model does a better job than the DD model, the overall error below 5 THz is rather large. The DD model not only shows an artificial peak at about 5 THz, it also yields a far too large admittance at zero frequency [4,7]. Moreover, the results of the moments equations indicate that the higher order equations up to $n_{\mathrm{HP}} = 10$ do not converge to a solution. For $n_{\mathrm{HP}} = 12$, the calculated drain self–admittance becomes negative for $0.68\,\mathrm{THz} < f < 0.9\,\mathrm{THz}$ which is indicative of instability. The situation becomes worse, if we change the BCs of the moments–based models from thermal bath to Dirichlet, as were used by Dyakonov and Shur (Fig. 5.42). The unrealistically large values of $\mathrm{Re}\{\underline{Y}_{\mathrm{DD}}\}$ at zero frequency increase by another 2–3 orders of magnitude. In addition, the peaks of the real part of the admittance become much sharper and higher and the minima smaller. The latter result is important, because the Dyakonov–Shur instability corresponds to a zero of the drain self–admittance (pole of the drain self–impedance for a short circuited input). In Fig. 5.44 the absolute value of the drain self-admittance is shown for the DD model with different BCs, where the maxima and minima correspond to the poles and zeros, respectively. For Dirichlet BCs the poles and zeros are lined up at $\sigma = -1/2\tau_{\mathrm{RTA}}$ [213], which is very small for $\mu = 10^6\,\mathrm{cm}^2/\mathrm{Vs}$ and results in sharp peaks on the imaginary axis. On the other hand, thermal bath BCs shift these poles (and zeros) to the left and the plasma resonances are strongly damped. This reduces their impact on the self-admittance at $\sigma = 0$ which is why the results in Fig. 5.41 are much smoother

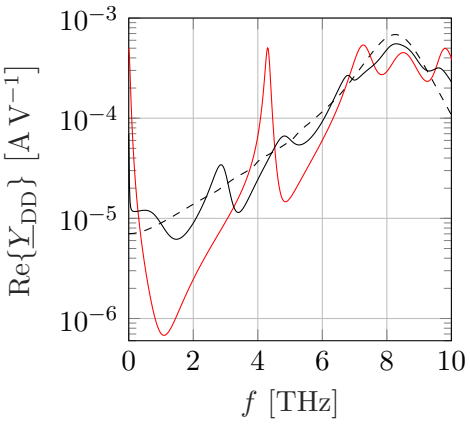

Figure 5.41: $\mathrm{Re}\{\underline{Y}_{\mathrm{DD}}\}$ of the 22 nm device evaluated using the phase space BE solver (dashed) and the moments equations with $n_{\mathrm{HP}} = 2$ (red), $n_{\mathrm{HP}} = 10$ (solid black). Simulations are performed for $\mu = 10^6\,\mathrm{cm}^2/\mathrm{Vs}$ and thermal bath BCs are imposed.

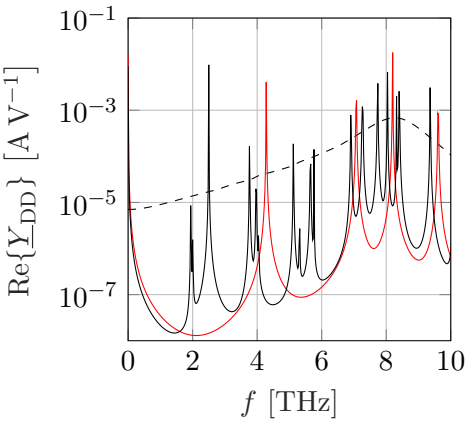

Figure 5.42: $\mathrm{Re}\{\underline{Y}_{\mathrm{DD}}\}$ of the 22 nm device evaluated using the phase space BE solver (dashed) and the moments equations with $n_{\mathrm{HP}} = 2$ (red), $n_{\mathrm{HP}} = 10$ (solid black). Simulations are performed for $\mu = 10^6\,\mathrm{cm}^2/\mathrm{Vs}$ and Dirichlet BCs are imposed.

than in Fig. 5.42. For a plasma instability poles or zeros with $\sigma > 0$ are required. Unfortunately, the more realistic thermal bath BCs move the poles and zeros far to the left making plasma instabilities even more improbable. Figure 5.45 shows the absolute values of the drain self–admittance calculated by the BE. The poles and zeros of the BE also occur at large negative real parts of the complex frequency and even a nonzero V_{DS} does not result in a significant shift to the right.

In Fig. 5.43 the non–equilibrium behavior of $\mathrm{Re}\{\underline{Y}_{\mathrm{DD}}\}$ is shown for various V_{DS} values as a function of frequency. The BE results are compared to the TM model. As it is evident at zero frequency, the transistor operates in the saturation regime for $V_{\mathrm{DS}} > 0.2\,\mathrm{V}$. Increasing the drain bias shifts the peak at $f = 8.2\,\mathrm{THz}$ to higher frequencies (see Fig. 5.45). The moments-based model follows this behavior up to $V_{\mathrm{DS}} = 0.1\,\mathrm{V}$, whereas for higher biases it becomes instable and cannot produce meaningful results. The BE, on the other hand, can be solved without problems for larger voltages and its results do not show any active behavior (i.e. $\mathrm{Re}\{\underline{Y}_{\mathrm{DD}}\} < 0$) for drain voltages larger than zero. The above results clearly show that the internal contact BCs of the transport model play an important role and that the more realistic thermal bath BCs strongly damp plasma resonances.

Since the damping of the plasma resonances depends on the type of the BCs, it would be interesting to investigate the impact of the Dirichlet BCs onto the

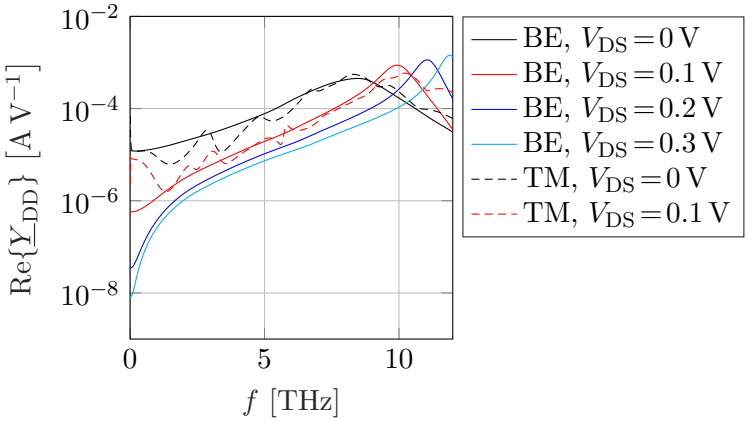

Figure 5.43: $\mathrm{Re}\{\underline{Y}_{\mathrm{DD}}\}$ versus frequency for the 22 nm device at $V_{\mathrm{GS}} = 0.5$ V using the BE and tenth-order moments-based model. Simulations are performed for a mobility of $\mu = 10^6\,\mathrm{cm}^2/\mathrm{Vs}$ and thermal bath BCs are applied.

BE results. The Dirichlet BCs are applied in the case of the moments–based models to the even moments (density, energy density (temperature) etc.). In the case of the BE this would correspond to the application of Dirichlet BCs to the even parts of the distribution functions. Without scattering such a BC violates the Liouville theorem, because the ratio of the even part of the distribution function at the source and drain at energies above the top of the barrier is fixed by transport for an inversion symmetric bandstructure and can not be imposed by the BCs. Thus, Dirichlet BCs for the even part of the distribution function and the ballistic BE are fundamentally incompatible.

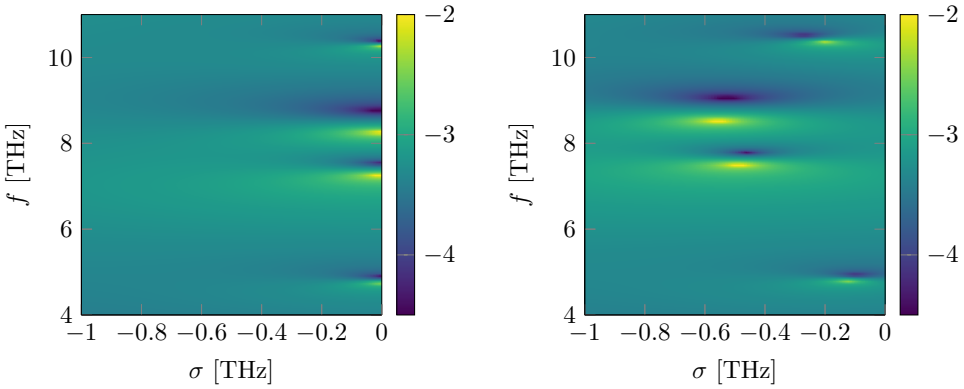

Figure 5.44: Logarithm of the absolute value of the drain self-admittance for $V_{\mathrm{GS}} = 0.5\,\mathrm{V}$, $V_{\mathrm{DS}} = 0\,\mathrm{V}$, and $\mu = 10^6\,\mathrm{cm}^2/\mathrm{Vs}$. The results are calculated by the DD model with Dirichlet BCs (left) and thermal bath BCs (right).

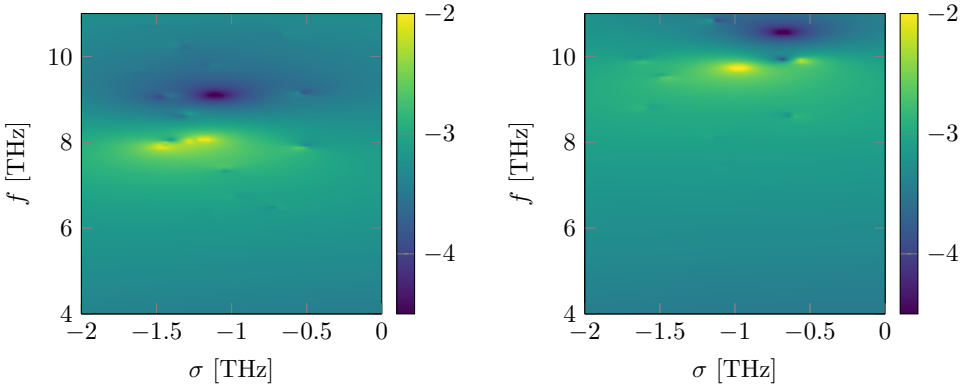

Figure 5.45: Logarithm of the absolute value of the drain self-admittance for $V_{\mathrm{GS}} = 0.5\,\mathrm{V}$ and the ballistic case. The results are calculated by the BE for $V_{\mathrm{DS}} = 0\,\mathrm{V}$ (left) and $V_{\mathrm{DS}} = 0.1\,\mathrm{V}$ (right).

Chapter 6

Concluding Remarks

6.1 Summary

The main objective of this dissertation was to develop a fully self–consistent and deterministic solver for the system of Poisson, Schrödinger, and Boltzmann equations in a gate–all–around junctionless nanowire FET. The simulation framework was built upon the existing works on deterministic BE solvers [69, 84] and employed various numerical techniques such as H–transformation and even/odd decomposition of the distribution function on a staggered spatial grid for stabilization of the BE. We solved the system of equations using the Newton–Raphson approach, with careful implementation of the necessary derivatives and the SE being included through the first order perturbation theory.

The small signal system of equations was set up in Sect. 2.6, and ideas developed in [69] were employed to properly discretize the time derivative term and conserve the necessary symmetries under equilibrium condition. In order to calculate the small signal terminal currents, we used a formulation of the Ramo–Shockley theorem consistent with the dimensionality of the problem, i.e. 1D BE along the transport direction and 2D SE on the transverse planes. It was shown that the in–house simulator is robust and allows for computations with unprecedented precision in the complete frequency range and for a wide range of bias conditions.

In Sect. 2.7 the Langevin–source approach was exploited for self–consistent calculation of the noise, resulting in the first deterministic solver for noise analysis of nanowire FETs. The Langevin–BE makes it possible to calculate the spatial distribution of noise in the simulated device, which is an absolute advantage over the Monte Carlo method.

In Chapter 5 the simulation results were put forward. Section 5.2 showcased the capabilities of the developed simulator for DC characterization of the device and providing a microscopic description of its inner working mechanisms. In

Sect. 5.3 the small signal results and high frequency figures of merit were calculated, which were perfectly smooth when plotted against frequency and fulfilled vital requirements such as reciprocity and passivity at equilibrium. Finally, the RF noise behavior of the junctionless nanowire FET was studied in Sect. 5.4 and important quantities such as the drain and gate excess noise factors, correlation coefficient, and noise suppression factor were calculated. For the sake of comparison, devices with different gate lengths were simulated and the effect of gate length downscaling on the DC, AC, and noise performance of nanowire nFET was investigated.

The second part of this work was dedicated to development of an alternative approach to discretization of the BE, which was motivated by the failure of energy–based solvers in quasi–ballistic regime and several numerical deficiencies of H–transformation for 1D phase space. In Chapter 4 the BE was discretized in k–space, and a stabilization method based on the characteristic curves and matrix exponentials was presented. The developed phase space solver is able to capture the complicated nature of ballistic distribution function and show its discontinuous behavior.

In Chapter 3 the BE was projected onto Hermite polynomials, so that the solutions of the phase space solver could be compared to the results of the moments equations. The comparisons, presented in Sect. 5.5, showed that the limitations of moments equations in describing the quasi–ballistic transport cannot be addressed through models obtained by higher moments as the ballistic transport is fundamentally different from diffusion or drift of high energy carriers. It was also observed that the boundary conditions of the transport model have a significant impact on the stationary and small signal results. These observations have important implications about the possibility of terahertz wave generation in high mobility 1D devices and suggest that the predictions of the Dyakonov–Shur model are too optimistic.

This thesis summarizes many years of research into various numerical approaches, and its results have been reported in several publications. In Ref. [79] we presented our first implementation of the Newton–Raphson solver for the stationary solution of the system of constituent equations. Refs. [80,81] presented our in–house simulator capable of small signal and noise analysis of the junctionless nanowire FETs, and in Ref. [82] we used the developed solver for assessing the effect of downscaling on the RF and noise performance of these devices. It was reported that due to the improved electrostatic control and better immunity to short–channel effects, the drain and gate excess noise factors and correlation coefficient demonstrate classical long–channel behavior for gate lengths as small as 16 nm.

6.2 Outlook

Moving forward, the developed simulator can be used for investigation of the device's behavior with respect to various design parameters. For example, the impact of high–κ spacers, graded doping profiles, asymmetric designs and additional gates on the RF and noise characteristics could be studied in detail. Another interesting area would be the variability of ultrashort transistors. While the sensitivity of longer devices to random variations in the gate length, oxide thickness, or the doping concentration was negligible, the impact of such changes on the characteristics of a 10 nm device is significant and can be predicted by numerical means. Regarding the numbers of published papers on variability, it is obvious that dealing with fluctuations of characteristics in ultrashort devices has become a major critical point for the next technology nodes.

Regarding the numerical aspects, there are many areas where more detailed models can be added. The methods in this work can be extended to III–V materials and more complex band structures. Inclusion of additional scattering mechanisms, e.g. impurity scattering and electron–electron interactions and the penetration of wave functions into the dielectric material will improve the accuracy of the simulation results. Moreover, there is enormous potential to speed up the computations and reduce memory requirements by using unstructured grids or applying suitable parallel algorithms. For example, the methods in [214, 215] have resulted in performance gains of up to one order of magnitude compared to single–threaded executions and lower memory requirements.

Finally, we need to mention that the transient simulation, which is a fundamental capability, is still lacking for the case of H–transformation [216]. This is because the H–transformation requires interpolation of the previous solution variables under the present band profile, which inevitably results in interpolation errors. The situation is in contrast to the MC method, which is inherently transient and allows for simulations in time domain [217,218]. Although some reports on transient simulation results can be found in [219] and [220], those works adopt the explicit time marching scheme and the maximum time steps are restricted due to the stability issues. Developing a new stabilization scheme, which allows the transient simulation while taking advantages of the H–transformation, would be a formidable task.

Nomenclature

This section presents an alphabetic list of acronyms and symbols used in the manuscript.

Acronym	**Expansion**
BC	boundary condition
BE	Boltzmann equation
CMOS	complementary metal–oxide–semiconductor
D	drain
DD	drift–diffusion
DIBL	drain-induced barrier lowering
EMA	effective mass approximation
FET	field–effect transistor
G	gate
GAA	gate–all–around
GR	generation and recombination
HP	Hermite polynomials
HD	hydrodynamic
MC	Monte Carlo
MOSFET	metal-oxide-semiconductor field-effect transistor
ODE	ordinary differential equation
PDE	partial differential equation
PE	Poisson equation
PSD	power spectral density
RF	radio frequency
RTA	relaxation–time approximation
S	source
SE	Schrödinger equation
SOI	silicon–on–insulator
SR	surface roughness
SS	subthreshold slope
SSSC	sinusoidal steady state condition
TCAD	technology computer-aided design

Symbol	Description
a	aggregate index for the Poisson equation.
α	aggregate index for the Boltzmann equation
b	aggregate index for the Poisson equation
β	aggregate index for the Boltzmann equation
$\boldsymbol{D}$	electric displacement current
$\mathrm{D_{PE}}$	domain of definition for the PE
$\mathrm{D_{BE}}$	domain of definition for the BE
Δx	box in x–direction surrounding the grid node
Δy	box in y–direction surrounding the grid node
Δz	box in z–direction surrounding the grid node
$\epsilon(\boldsymbol{r})$	dielectric constant
$\boldsymbol{E}$	electric field
$\varepsilon_{\mathrm{sub}}^{\nu}$	subband energy in the valley and subband ν
f	frequency
F	force acting on the electrons
$\mathrm{F}_{\nu\nu'}$	overlap integral for the states ν and ν'
$\mathrm{F_{PE}}$	stationary PE
$\underline{\mathrm{F}}_{\mathrm{PE}}$	small signal phasor of the PE
$\mathrm{F_{BE}}$	stationary BE
$\underline{\mathrm{F}}_{\mathrm{BE}}$	small signal phasor of the BE
$f_k^{\nu}(z,k)$	distribution function of the stationary BE in k–space
$f_{\varepsilon}^{\nu}(z,\varepsilon)$	distribution function of the stationary BE in ε–space
$f_{\mathrm{H}}^{\nu}(z,H)$	distribution function of the stationary BE in H–space
$f_k^{\nu}(z,k,t)$	distribution function of the time–dependent BE in k–space
$f_{\varepsilon}^{\nu}(z,\varepsilon,t)$	distribution function of the time–dependent BE in ε–space
$f_{\mathrm{H}}^{\nu}(z,H,t)$	distribution function of the time–dependent BE in H–space
f_{eq}	equilibrium part of the distribution function
f_{ne}	non–equilibrium part of the distribution function
$\varphi(\boldsymbol{r})$	quasistatic electric potential
$\underline{\varphi}(\boldsymbol{r})$	small signal phasor of electric potential
φ_{MS}	metal–semiconductor work function difference
ϕ_n	quasi Fermi potential of electrons
$\Gamma_{\mathrm{GR}}\{f\}$	generation/recombination boundary condition in the BE
$\Gamma_{\mathrm{TH}}\{f\}$	thermal bath boundary conditions in the BE
$\hbar$	reduced Planck constant
k_{B}	Boltzmann constant
L_x	length of the semiconductor region in x–direction
L_y	length of the semiconductor region in y–direction
L_z	length of the semiconductor region in z–direction
$\hat{M}$	(diagonal) mass tensor

m_e	electron rest mass
m_{xx}^ν	effective mass in x–direction in the valley and subband ν
m_{yy}^ν	effective mass in y–direction in the valley and subband ν
m_{zz}^ν	effective mass in z–direction in the valley and subband ν
N_A	acceptor doping concentration
N_D	donor doping concentration
n_BE	dimension of the discretized Boltzmann equation system
n_PE	dimension of the discretized Poisson equation system
n_x	highest index of grid in x–direction
n_y	highest index of grid in y–direction
n_z	highest index of grid in z–direction
∇_r	nabla operator in real space
ν	combined valley and subband index $\nu = (v, s)$
n_i	intrinsic carrier density
n_3D	3D electron density
ω	angular frequency
p_3D	3D hole density
q	positive elementary charge
$\boldsymbol{r}$	2D spatial vector $\boldsymbol{r} = (x, y, z)$
ρ	mass density of the silicon crystal
$\rho(\boldsymbol{r})$	space charge density
s	subband index
$S_\mathrm{RTA}\{f\}$	scattering term in the BE under RTA
$S_\mathrm{PH}\{f\}$	electron–phonon scattering integral in the BE
$\psi^\nu(\boldsymbol{r})$	wavefunction of the Schrödinger equation
$\psi^\nu(\boldsymbol{r}, t)$	time dependent wavefunction of the Schrödinger equation
T	temperature
τ	relaxation time for scattering under RTA
$\Theta(\cdot)$	Heaviside step function
v	valley index of silicon band structure
v_λ	sound velocity of acoustic phonons
$v_z^\nu(k)$	electron group velocity in z–direction
v_GR	generation/recombination velocity
V_DS	drain–source voltage
V_GS	gate–source voltage
V_T	thermal voltage $V_\mathrm{T} = k_\mathrm{B}T/q$
$W_\eta^{\nu\nu'}(z; k\|k')$	transition rate of the scattering integral
z	transport equation
Z^ν	density of states function

Bibliography

[1] International Roadmap Committee, "The international technology roadmap for semiconductors," *http://www.itrs2.net/*, 2017.

[2] E. Sicard, "Introducing 10-nm FinFET technology in microwind," 2017.

[3] D. Bishop, "Nanotechnology and the end of moore's law?" *Bell Labs Technical Journal*, vol. 10, no. 3, pp. 23–28, 2005.

[4] M. Nekovee, B. J. Geurts, H. M. J. Boots, and M. F. H. Schuurmans, "Failure of extended-moment-equation approaches to describe ballistic transport in submicrometer structures," *Phys. Rev. B*, vol. 45, no. 12, pp. 6643–6651, 1992.

[5] M. V. Fischetti, S. E. Laux, and E. Crabbe, "Understanding hot-electron transport in silicon devices: Is there a shortcut?" *J. Appl. Phys.*, vol. 78, pp. 1058–1087, 1995.

[6] M. S. Shur, "Low ballistic mobility in submicron HEMTs," *IEEE Electron Device Lett.*, vol. 23, no. 9, pp. 511–513, 2002.

[7] C. Jungemann, T. Grasser, B. Neinhüs, and B. Meinerzhagen, "Failure of moments-based transport models in nanoscale devices near equilibrium," *Electron Devices, IEEE Transactions on*, vol. 52, no. 11, pp. 2404–2408, 2005.

[8] O. Madelung, *Introduction to Solid State Theory*. Berlin: Springer, 1978.

[9] C. Jacoboni and P. Lugli, *The Monte Carlo Method for Semiconductor Device Simulation*. Wien: Springer, 1989.

[10] P. J. Price, "Monte Carlo calculation of electron transport in solids," *Semiconductors and Semimetals*, vol. 14, pp. 249–309, 1979.

[11] M. Lundstrom, *Fundamentals of Carrier Transport*, ser. Modular Series on Solid State Devices. New York: Addison-Wesley, 1990, vol. 10.

[12] E. M. Conwell, *High field transport in semiconductors.* New York: Academic Press, 1967.

[13] D. L. Rode, *Semiconductors and Semimetals.* New York: Academic Press, 1972.

[14] J. A. Aubert, J. C. Vaissiere, and J. P. Nougier, "Matrix determination of the stationary solution of the Boltzmann equation for hot carriers in semiconductors," *J. Appl. Phys.*, vol. 56, no. 4, pp. 1128–1132, 1984.

[15] M. Saraniti and S. M. Goodnick, "Hybrid fullband cellular automaton/Monte Carlo approach for fast simulation of charge transport in semiconductors," *Electron Devices, IEEE Transactions on*, vol. 47, no. 10, pp. 1909–1916, 2000.

[16] J. M. Hammarsley and D. C. Handscomb, *Monte Carlo Methods.* London: Methuen/Chapman and Hall, 1964.

[17] S. M. Ermakow, *Die Monte-Carlo-Methode und verwandte Fragen.* München, Wien: R. Oldenbourg, 1975.

[18] R. Y. Rubinstein, *Simulation and the Monte Carlo method*, ser. Wiley series in probability and mathematical statistics. New York: John Wiley & Sons, 1981.

[19] T. Kurosawa, "Monte Carlo calculation of hot electron problems," *J. Phys. Soc. Jap.*, vol. 21, pp. 424–426, 1966.

[20] W. Fawcett, A. D. Boardman, and S. Swain, "Monte Carlo determination of electron transport properties in gallium arsenide," *J. Phys. Chem. Solids*, vol. 31, pp. 1963–1990, 1970.

[21] J. Y. Tang, H. Shichijo, K. Hess, and G. J. Iafrate, "Band-structure dependent impact ionization in silicon and gallium arsenide," *Journal de Physique*, vol. 42, pp. 63–69, 1981.

[22] M. V. Fischetti and S. E. Laux, "Monte Carlo analysis of electron transport in small semiconductor devices including band–structure and space–charge effects," *Phys. Rev. B*, vol. 38, pp. 9721–9745, 1988.

[23] H. J. Peifer, R. Thoma, A. Emunds, and W. L. Engl, "Hot carriers in a LDD–MOSFET investigated with a Monte Carlo simulator," in *VLSI Process/Device Modeling Workshop*, Tokyo, Aug. 1988.

[24] C. Moglestue, *Monte Carlo Simulation of Semiconductor Devices.* London: Chapman & Hall, 1993.

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

[26] P. D. Yoder and K. Hess, "First-principles Monte Carlo simulation of transport in Si ," *Semicond. Sci. Technol.*, vol. 9, pp. 852–854, 1994.

[27] K. Hess, Ed., *Monte Carlo Device Simulation: Full Band and Beyond*. Boston: Kluwer, 1991.

[28] M. Lundstrom, Z. Ren, and S. Datta, "Essential physics of carrrier transport in nanoscale MOSFETs," in *Proc. SISPAD*, 2000, pp. 1–5.

[29] P. W. Rambo and J. Denavit, "Time stability of Monte Carlo device simulation," *IEEE Trans. Computer–Aided Des.*, vol. 12, pp. 1734–1741, 1993.

[30] C. Jungemann, B. Neinhüs, S. Decker, and B. Meinerzhagen, "Hierarchical 2–D DD and HD noise simulations of Si and SiGe devices: Part II—Results," *Electron Devices, IEEE Transactions on*, vol. 49, no. 7, pp. 1258–1264, 2002.

[31] P. Lugli and D. Ferry, "Degeneracy in the Ensemble Monte Carlo Method for High-Field Transport in Semiconductors," *Electron Devices, IEEE Transactions on*, vol. 32, no. 11, pp. 2431–2334, 1985.

[32] L. Lucci, P. Palestri, D. Esseni, L. Bergagnini, and L. Selmi, "Multisubband Monte Carlo study of transport, quantization, and electron-gas degeneration in ultrathin SOI n-MOSFETs," *Electron Devices, IEEE Transactions on*, vol. 54, pp. 1156–1164, 2007.

[33] D. Ventura, A. Gnudi, G. Baccarani, and F. Odeh, "Multidimensional spherical harmonics expansion of Boltzmann equation for transport in semiconductors," *Appl. Math. Lett.*, vol. 5, p. 85, 1992.

[34] A. Gnudi, D. Ventura, G. Baccarani, and F. Odeh, "Two-dimensional MOSFET simulation by means of a multidimensional spherical harmonics expansion of the Boltzmann transport equation," *Solid–State Electron.*, vol. 36, no. 4, pp. 575 – 581, 1993.

[35] K. Rahmat, J. White, and D. A. Antoniadis, "Simulation of semiconductor devices using a Galerkin/spherical harmonic expansion approach to solving the coulped Poisson-Boltzmann system," *IEEE Trans. Computer–Aided Des.*, vol. 15, no. 10, pp. 1181–1196, 1996.

[36] C. Ringhofer, "Numerical methods for the semiconductor Boltzmann equation based on spherical harmonics expansions and entropy discretizations,"

Transport Theory and Statistical Physics, vol. 31, no. 4-6, pp. 431–452, 2002.

[37] C. Jungemann, A.-T. Pham, B. Meinerzhagen, C. Ringhofer, and M. Bollhöfer, "Stable discretization of the Boltzmann equation based on spherical harmonics, box integration, and a maximum entropy dissipation principle," *J. Appl. Phys.*, vol. 100, p. 024502, 2006.

[38] S. Smirnov and C. Jungemann, "A full band deterministic model for semiclassical carrier transport in semiconductors," *J. Appl. Phys.*, vol. 99, pp. 063 707–1–11, 2006.

[39] S.-M. Hong and C. Jungemann, "Simulation of magnetotransport in nanoscale devices," in *International Conference on Solid State and Integrated Circuits Technology*, 2008, pp. 377–380.

[40] H. Kamrani, T. Kochubey, D. Jabs, and C. Jungemann, "Electrothermal simulation of sige hbts and investigation of experimental extraction methods for junction temperature," in *2015 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, Sept 2015, pp. 108–111.

[41] H. Kamrani, D. Jabs, V. d'Alessandro, N. Rinaldi, and C. Jungemann, "Physics-based hot-carrier degradation model for sige hbts," in *2016 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD)*, Sept 2016, pp. 341–344.

[42] M. Ramonas and C. Jungemann, "Spherical harmonics solver for coupled hot-electron–hot-phonon system," in *Simulation of Semiconductor Processes and Devices (SISPAD), 2013 International Conference on*, Sept 2013, pp. 360–363.

[43] D. Jabs and C. Jungemann, "Avalanche breakdown of pn-junctions: Simulation by spherical harmonics expansion of the Boltzmann transport equation," in *Simulation of Semiconductor Processes and Devices (SISPAD), 2014 International Conference on*, Sept 2014, pp. 173–176.

[44] D. Jabs, C. Jungemann, and K. Bach, "A robust algorithm for microscopic simulation of avalanche breakdown in semiconductor devices," *Electron Devices, IEEE Transactions on*, vol. 62, no. 8, pp. 2614–2619, Aug 2015.

[45] Y. Huang, M. Chiang, S. Wang, and J. G. Fossum, "GAAFET versus pragmatic FinFET at the 5 nm si-based CMOS technology node," *IEEE J. Elec. Dev. Soc.*, vol. 5, no. 3, pp. 164–169, 2017.

[46] M. Luisier, A. Schenk, and W. Fichtner, "Three-dimensional fullband simulations of si nanowire transistors," *IEDM Tech. Dig.*, pp. 1–4, 2006.

[47] Y. Zheng, C. Rivas, R. Lake, K. Alam, T. B. Boykin, and G. Klimeck, "Electronic properties of silicon nanowires," *Electron Devices, IEEE Transactions on*, vol. 52, no. 6, pp. 1097–1103, 2005.

[48] J. Wang, A. Rahman, A. Gosh, G. Klimeck, and M. Lundstrom, "On the validity of the parabolic effective-mass approximation for the I-V calculation of silicon nanowire transistors," *Electron Devices, IEEE Transactions on*, vol. 52, no. 7, pp. 1589–1595, 2005.

[49] 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 MOSFETs: An atomistic study," *Solid State Electron.*, vol. 50, no. 4, pp. 716–721, 2006.

[50] J. Wang, E. Polizzi, and M. Lundstrom, "A three-dimensional quantum simulation of silicon nanowire transistors with the effectivemass approximation," *J. Appl. Phys.*, vol. 96, no. 4, pp. 2192–2203, 2004.

[51] E. Polizzi and N. B. Abdallah, "Subband decomposition approach for the simulation of quantum electron transport in nanostructures," *J. Compu. Phys.*, vol. 202, no. 1, pp. 150–180, 2005.

[52] S. Jin, Y. J. Park, and H. S. Min, "A three-dimensional simulation of quantum transport in silicon nanowire transistor in the presence of electron–phonon interactions," *J. Appl. Phys.*, vol. 99, no. 12, p. 123719, 2006.

[53] M. Luisier, A. Schenk, and W. Fichtner, "Atomistic treatment of interface roughness in si nanowire transistors with different channel orientations," *Appl. Phys. Lett.*, vol. 90, no. 10, p. 102103, 2007.

[54] 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," *Electron Devices, IEEE Transactions on*, vol. 54, no. 9, pp. 2213–2222, 2007.

[55] C. Jungemann, B. Neinhüs, S. Decker, and B. Meinerzhagen, "Hierarchical 2D RF noise simulation of Si and SiGe devices by Langevin-type DD and HD models based on MC generated noise parameters," in *IEEE Tech. Dig. IEDM*, Washington (USA), 2001, pp. 481–484.

[56] A. J. Scholten, L. F. Tiemeijer, R. van Langevelde, R. J. Havens, V. C. Venezia, A. T. A. Z. van Duijnhoven, B. Neinhüs, C. Jungemann, and D. B. M. Klaassen, "Compact modeling of drain and gate current noise for rf cmos," *IEEE Tech. Dig. IEDM*, pp. 129–132, 2002.

[57] A. J. Scholten, L. F. Tiemeijer, R. van Langevelde, R. J. Havens, A. T. A. Z. van Duijnhoven, and V. C. Venezia, "Noise modeling for RF CMOS circuit simulation," *Electron Devices, IEEE Transactions on*, vol. 50, no. 3, pp. 618–632, 2003.

[58] C. Jungemann, B. Neinhüs, C. D. Nguyen, A. J. Scholten, L. F. Tiemeijer, and B. Meinerzhagen, "Numerical modeling of RF noise in scaled MOS devices," *Solid–State Electron.*, vol. 50, pp. 10–17, 2006.

[59] Y. Chen and R. Xu, "Analysis of the rf and noise performance of junctionless mosfetsusing monte carlo simulation," *Int. J. Numer. Model.*, vol. 27, pp. 822–833, 2014.

[60] M. Fallahnejad, M. Vadizadeh, A. Salehi, A. Kashaniniya, and F. Razaghian, "Impact of channel doping engineering on the high-frequency noise performance of junctionless $in_{0.3}ga_{0.7}as$/gaas FET: A numerical simulation study," *Physica E*, vol. 115, p. 113715, 2020.

[61] P. Langevin, "Sur la theéorie du mouvement brownien," *C. R. Acad. Sci. Paris*, vol. 146, pp. 530–533, 1908.

[62] S. M. Kogan and A. Y. Shul'man, "Theory of fluctuations in a nonequilibrium electron gas," *Sov. Phys. JETP*, vol. 29, no. 3, pp. 467–474, 1969.

[63] H. S. Min and D. Ahn, "Langevin noise sources for the Boltzmann transport equations with the relaxation-time approximation in nondegenerate semiconductors," *J. Appl. Phys.*, vol. 58, pp. 2262–2265, 1985.

[64] C. Jungemann and B. Meinerzhagen, "Do hot electrons cause excess noise?" *Solid–State Electron.*, vol. 50, pp. 674–679, 2006.

[65] ——, "Numerical simulation of RF noise in Si devices," in *Proc. SISPAD*, 2006, pp. 87–94.

[66] ——, "Noise calculation in the semiclassical framework: A critical analysis of the monte carlo method and a numerical alternative," in *Large-Scale Scientific Computing*, ser. Lecture Notes in Computer Science, I. Lirkov, S. Margenov, and J. Wa niewski, Eds. Springer Berlin Heidelberg, 2006, vol. 3743, pp. 164–171. [Online]. Available: http://dx.doi.org/10.1007/11666806_17

[67] C. Jungemann, "A deterministic approach to RF noise in silicon devices based on the Langevin Boltzmann equation," *Electron Devices, IEEE Transactions on*, vol. 54, no. 5, pp. 1185–1192, 2007.

[68] M. Ramonas and C. Jungemann, "A deterministic approach to noise in a non-equilibrium electron-phonon system based on the Boltzmann equation," *Journal of Computational Electronics*, vol. 14, no. 1, pp. 43–50, 2015. [Online]. Available: http://dx.doi.org/10.1007/s10825-014-0627-3

[69] D. Ruić and C. Jungemann, "Numerical aspects of noise simulation in MOSFETs by a Langevin-Boltzmann solver," *Journal of Computational Electronics*, vol. 14, no. 1, pp. 21–36, 2015. [Online]. Available: http://dx.doi.org/10.1007/s10825-014-0642-4

[70] ——, "Microscopic noise simulation of long- and short-channel nMOSFETs by a deterministic approach," *Journal of Computational Electronics*, vol. 15, no. 3, p. 809–819, 2016.

[71] H. U. Baranger and J. W. Wilkins, "Ballistic structure in the electron distribution function of small semiconducting structures: General features and specific trends," *Phys. Rev. B*, vol. 36, no. 3, pp. 1487–1502, 1987.

[72] J.-H. Rhew, Z. Ren, and M. S. Lundstrom, "A numerical study of ballistic transport in a nanoscale MOSFET," *Solid–State Electron.*, vol. 46, no. 11, pp. 1899–1906, 2002.

[73] E. Fuchs, P. Dollfus, G. L. Carval, S. Barraud, D. Villanueva, F. Salvetti, H. Jaouen, and T. Skotnicki, "A new backscattering model giving a description of the quasi-ballistic transport in nano-MOSFET," *Electron Devices, IEEE Transactions on*, vol. 52, no. 10, pp. 2280–2289, 2005.

[74] L. P. Kadanoff and G. Baym, *Quantum Statistical Mechanics*. New York: Benjamin, 1962.

[75] S. Datta, Ed., *Quantum Transport: Atom to Transistor*. U.K.: Cambridge Univ. Press, 2005.

[76] E. Wigner, "On the quantum correction for thermodynamic equilibrium," *Phys. Rev.*, vol. 40, no. 5, pp. 749–759, 1932.

[77] M. V. Fischetti, "Theory of electron transport in small semiconductor devices using the Pauli master equation," *J. Appl. Phys.*, vol. 83 (1), pp. 270–291, 1998.

[78] ——, "Master-equation approach to the study of electronic transport in small semiconductor devices," *Phys. Rev. B*, vol. 59, no. 7, pp. 4901–4917, 1999.

[79] M. Noei and C. Jungemann, "Numerical investigation of junctionless nanowire transistors using a boltzmann/schrdinger/poisson full newton-raphson solver," *Proc. SISPAD*, 2016. [Online]. Available: http://ieeexplore.ieee.org/document/7605137/

[80] M. Noei, D. Ruic, and C. Jungemann, "Small-signal analysis of silicon nanowire transistors based on a poisson/schrödinger/boltzmann solver," *Proc. SISPAD*, pp. 65–68, 2017.

[81] M. Noei and C. Jungemann, "Microscopic simulation of RF noise in junctionless nanowire transistors," *J. Computational Electronics*, vol. 17, no. 5, 2018.

[82] ——, "RF analysis and noise characterization of junctionless nanowire FETs by a boltzmann transport equation solver," *J. Computational Electronics*, vol. 18, pp. 1347–1353, 2019.

[83] T. Ando, A. Fowler, and F. Stern, "Electronic properties of two-dimensional systems," *Rev. Mod. Phys.*, vol. 54, pp. 437–672, 1982.

[84] S.-M. Hong, A. T. Pham, and C. Jungemann, *Deterministic solvers for the Boltzmann transport equation*, ser. Computational Microelectronics, S. Selberherr, Ed. Wien, New York: Springer, 2011.

[85] S. Selberherr, *Analysis and Simulation of Semiconductor Devices.* Wien: Springer, 1984.

[86] R. D. .Cook, D. S. Malkus, and M. E. Plesha, "Concepts and applications of finite element analysis: 3rd edition," *New York: John Wiley*, 1989.

[87] E. A. W. Maunder, H. G. Zhong, and P. Beckers, "A posteriori error estimators related to equilibrium defaults of finite element solutions for elastostatic problems," *Finite Elements in Analysis and Design*, vol. 26, no. 3, pp. 171–192, 1997.

[88] J. M. Fiard and R. Herbin, "Comparison between finite volume and finite element methods for an elliptic system arising in electrochemical engineering," *Computer Methods in Appl. Mechanics and Engineering*, vol. 115, pp. 315–338, 1994.

[89] R. S. Varga, *Matrix Iterative Analysis*, ser. Series in Automatic Computation. Englewood Cliffs, New Jersey: Prentice-Hall, 1962.

[90] N. Taniguchi and T. Kobayashi, "Finite volume method on the unstructured grid system," *Computers & Fluids*, vol. 19, no. 3–4, pp. 287–295, 1991.

[91] C. Farhat and S. Lanteri, "Simulation of compressible viscous flows a variety of MPPS: computational algorithms for unstructured dynamic meshes and performance results," *Comp. Methods in Appl. Mechanics and Engineering*, vol. 119, no. 1–2, pp. 35–60, 1994.

[92] I. D. Mishev, "Finite volume methods on voronoi meshes," *Numerical Methods for Partial Differential Equations*, vol. 14, no. 2, pp. 193–212, 1998.

[93] A. Berman and R. J. Plemmons, *Nonnegative matrices in the mathematical sciences*, ser. Computer Science and Applied Mathematics. Academic Press, 1979.

[94] J.-P. Colinge and J. C. Greer, *Nanowire Transistors*. Cambridge University Press, 2016.

[95] R. Shankar, *Principles of quantum mechanics*. New York: Springer, 1994.

[96] N. Neophytou, A. Paul, M. S. Lundstrom, and G. Klimeck, "Bandstructure effects in silicon nanowire electron transport," *Electron Devices, IEEE Transactions on*, vol. 55, no. 6, pp. 1286–1297, 2008.

[97] D. Esseni and P. Palestri, "Linear combination of bulk bands method for investigating the low-dimensional electron gas in nanostructured devices," *Phys. Rev. B, Condens. Matter*, vol. 72, no. 16, p. 165342, 2005.

[98] E. Gnani, S. Reggiani, A. Gnudi, P. Parruccini, R. Colle, M. Rudan, and G. Baccarani, "Band-structure effects in ultrascaled silicon nanowires," *Electron Devices, IEEE Transactions on*, vol. 54, no. 9, pp. 2243–2254, 2007.

[99] S. Jin, M. V. Fischetti, and T.-W. Tang, "Modeling of electron mobility in gated silicon nanowires at room temperature: Surface roughness scattering, dielectric screening, and band nonparabolicity," *J. Appl. Phys.*, vol. 102, 2007.

[100] D. Esseni, P. Palestri, and L. Selmi, *Nanoscale MOS Transistors. Semi-Classical Transport and Applications*. Cambridge University Press, 2011.

[101] C. Jungemann, A. Emunds, and W. Engl, "Simulation of linear and nonlinear electron transport in homogeneous silicon inversion layers," *Solid-State Electronics*, vol. 36, no. 11, pp. 1529 – 1540, 1993. [Online]. Available: http://www.sciencedirect.com/science/article/pii/003811019390024K

[102] E. Polizzi, "Density-matrix-based algorithm for solving eigenvalue problems," *Phys. Rev. B*, vol. 79, p. 115112, Mar 2009. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevB.79.115112

[103] F. Balestra, S. Cristoloveanu, M. Benachir, J. Brini, and T. Elewa, "Double-gate silicon-on-insulator transistor with volume inversion: A new device with greatly enhanced performance," *IEEE Electron Device Lett.*, vol. 8, no. 9, 1987.

[104] E. G. Marin, F. J. G. Ruiz, I. M. Tienda-Luna, A. Godoy, and F. Gámiz, "Analytical gate capacitance modeling of III–V nanowire transistors," *Electron Devices, IEEE Transactions on*, vol. 60, no. 5, pp. 1590–1599, 2013.

[105] C. Buet, "A discrete velocity scheme for the Boltzmann operator of rarefied gas dynamics," *Transp. Theory Stat. Phys.*, vol. 25, pp. 33–60, 1996.

[106] C. Cercignani, R. Illner, and M. Pulvirenti, *The mathematical theory of dilute gases*. New York: Springer, 1994.

[107] L. Pareschi and B. Perthame, "A Fourier spectral method for homogeneous Boltzmann equations," *Transp. Theory Stat. Phys.*, vol. 25, pp. 369–383, 1996.

[108] E. Gabetta, L. Pareschi, and G. Toscani, "Relaxation schemes for nonlinear kinetic equations," *SIAM J. Numer. Anal.*, vol. 34, pp. 2168–2194, 1997.

[109] V. Aristov, *Direct methods for solving the Boltzmann equation and study of nonequilibrium flows*. Dordrecht: Kluwer, 2001.

[110] N. Bellomo, *Lecture notes on the discretization of Boltzmann equation*. Singapore: World Scientific, 2003.

[111] S.-M. Hong, C. Jungemann, and M. Bollhöfer, "A deterministic Boltzmann equation solver for two-dimensional semiconductor devices," in *Proc. SISPAD*, 2008, pp. 293–296.

[112] A.-T. Pham, B. Meinerzhagen, and C. Jungemann, "A fast $\vec{k} \cdot \vec{p}$ solver for hole inversion layers with an efficient 2d $\vec{k}$ -space discretization," *Journal of Computational Electronics*, vol. 7, no. 3, pp. 99–102, 2008. [Online]. Available: http://dx.doi.org/10.1007/s10825-007-0155-5

[113] N. Goldsman, L. Henrickson, and J. Frey, "A physics-based analytical/numerical solution to the Boltzmann transport equation for use in device simulation," *Solid–State Electron.*, vol. 34, pp. 389–396, 1991.

[114] J. A. Carrillo, I. M. Gamba, A. Majorana, and C.-W. Shu, "A direct solver for 2D non-stationary Boltzmann-Poisson systems for semiconductor devices: a MESFET simulation by WENO-Boltzmann schemes," *Journal of Computational Electronics*, vol. 2, pp. 375–379, 2003.

[115] W. Liang, N. Goldsman, I. Mayergoyz, and P. Oldiges, "2-D MOSFET modeling including surface effects and impact ionization by self-consistent solution of the Boltzmann, Poisson, and hole-continuity equations," *Electron Devices, IEEE Transactions on*, vol. 44, no. 2, pp. 257–267, 1997.

[116] M. V. Fischetti, Z. Ren, P. M. Solomon, M. Yang, and K. Rim, "Six-band $k \cdot p$ calculation of the hole mobility in silicon inversion layers: Dependence on surface orientation, strain, and silicon thickness," *J. Appl. Phys.*, vol. 94, pp. 1079–1095, 2003.

[117] C. Jacoboni and L. Reggiani, "The Monte Carlo method for the solution of charge transport in semiconductors with application to covalent materials," *Rev. Mod. Phys.*, vol. 55, pp. 645–705, 1983.

[118] R. Brunetti, C. Jacoboni, F. Nava, L. Reggiani, G. Bosman, and R. J. J. Zijlstra, "Diffusion coefficient of electrons in silicon," *J. Appl. Phys.*, vol. 52, pp. 6713–6722, 1981.

[119] F. M. Bufler, P. Graf, S. Keith, and B. Meinerzhagen, "Full band Monte Carlo investigation of electron transport in strained Si grown on $Si_{1-x}Ge_x$ substrates," *Appl. Phys. Lett.*, vol. 70, pp. 2144–2146, 1997.

[120] C. Jungemann, S. Keith, F. M. Bufler, and B. Meinerzhagen, "Effects of band structure and phonon models on hot electron transport in silicon," *Electrical Engineering*, vol. 79, pp. 99–101, 1996.

[121] W. A. Harrison, "Scattering of electrons by lattice vibrations in nonpolar crystals," *Phys. Rev.*, vol. 104, pp. 1281–1290, 1956.

[122] K. Moors, B. Sorée, and W. Magnus, "Modeling surface roughness scattering in metallic nanowires," *J. Appl. Phys.*, vol. 118, p. 124307, 2015.

[123] H. Tanaka, J. Suda, and T. Kimoto, "Modeling of surface roughness scattering in nanowires based on atomistic wave function: Application to hole mobility in rectangular germanium nanowires," *Phys. Rev. B*, vol. 93, p. 155303, 2016.

[124] I. M. Tienda-Luna, F. G. Ruiz, A. Godoy, B. Biel, and F. Gámiz, "Surface roughness scattering model for arbitrarily oriented silicon nanowires," *J. Appl. Phys.*, vol. 110, no. 8, p. 084514, 2011.

[125] R. E. Prange and T. W. Nee, "Quantum spectroscopy of the low-field oscillations in the surface impedance," *Phys. Rev.*, vol. 168, pp. 779–785, 1968.

[126] S. M. Goodnick, D. K. Ferry, C. W. Wilmsen, Z. Liliental, D. Fathy, and O. L. Krivanek, "Surface roughness at the Si(100)-SiO_2 interface," *Phys. Rev. B*, vol. 32, pp. 8171–8186, 1985.

[127] F. Gámiz, J. B. Roldán, J. A. López-Villanueva, P. Cartujo-Cassinello, and J. E. Carceller, "Surface roughness at the Si–SiO2 interfaces in fully depleted silicon-on-insulator inversion layers," *Journal of Applied Physics*, vol. 86, no. 12, pp. 6854–6863, 1999. [Online]. Available: http://dx.doi.org/10.1063/1.371763

[128] A. Kranti, R. Yan, C. W. Lee, I. Ferain, R. Yu, N. D. Akhavan, P. Razavi, and J. Colinge, "Junctionless nanowire transistor (JNT): Properties and design guidelines," *Solid-State Device Research Conference (ESSDERC), 2010 Proceedings of the European*, 2010. [Online]. Available: http://ieeexplore.ieee.org/document/5618216/

[129] J. P. Colinge, C.-W. Lee, A. Afzalian, N. D. Akhavan, R. Yan, I. Ferain, P. Razavi, B. O'Neill, A. Blake, M. White, A.-M. Kelleher, B. McCarthy, and R. Murphy, "Nanowire transistors without junctions," *Nature Nanotechnol.*, vol. 5, pp. 225–229, 2010.

[130] R. Rios, A. Cappellani, M. Armstrong, A. Budrevich, H. Gomez, R. Pai, N. Rahhal-Orabi, and K. Kuhn, "Comparison of junctionless and conventional trigate transistors with Lg down to 26nm," *IEEE Electron Device Lett.*, vol. 32, no. 9, pp. 1170–1172, 2011.

[131] S. Sahay and M. J. Kumar, "Realizing efficient volume depletion in SOI junctionless FETs," *IEEE J. Electron Devices Soc.*, vol. 4, no. 3, pp. 110–115, 2016.

[132] B. Sorée, W. Magnus, and W. Vandenberghe, "Low-field mobility in ultrathin silicon nanowire junctionless transistors," *Appl. Phys. Lett.*, vol. 99, 2011.

[133] C. Ringhofer, "Space - time discretization of series expansion methods for the Boltzmann transport equation," *SIAM J. Num. Ana.*, vol. 38, pp. 442–465, 2000.

[134] C. Ringhofer, C. Schmeiser, and A. Zwirchmayer, "Moment methods for the semiconductor Boltzmann equation in bounded position domains," *SIAM J. Num. Ana.*, vol. 39, pp. 1078–1095, 2001.

[135] C. Ringhofer, "A mixed spectral-difference method for the steady state boltzmann-poisson system," *SIAMNumAna*, vol. 41, no. 1, pp. 64 – 89, 2003.

[136] M. Bollhöfer and Y. Saad, "ILUPACK — preconditioning software package," release 2.2 available online at `www.icm.tu-bs.de/~bolle/ilupack/`, Dept. of Mathematics and Computer Science, November 2008.

[137] ——, "Multilevel preconditioners constructed from inverse–based ILUs," *SIAM J. Sci. Comput.*, vol. 27, no. 5, pp. 1627–1650, 2006.

[138] S. E. Laux, "Techniques for small-signal analysis of semiconductor devices," *Electron Devices, IEEE Transactions on*, vol. 32, no. 10, pp. 2028–2037, Oct. 1985.

[139] I. Getreu, *Modeling the Bipolar Transistor*. Amsterdam: Elsevier, 1978.

[140] T. Schulz and W. Stieler, "Unter der Haube: Optimierung von CMOS–Transistoren," *c't*, vol. 5, pp. 260–265, 2000.

[141] D. Ruić and C. Jungemann, "Small Signal and Microscopic Noise Simulation of an nMOSFET by a Self-Consistent, Semi-Classical and Deterministic Approach," in *Simulation of Semiconductor Processes and Devices (SISPAD), 2015 International Conference on*. IEEE, Sept 2015, pp. 20–23.

[142] S. Ramo, "Currents induced by electron motion," *Proc. IRE*, vol. 27, pp. 584–585, 1939.

[143] W. Shockley, "Currents to conductors induced by a moving point charge," *J. Appl. Phys.*, vol. 9, pp. 635–636, 1938.

[144] R. P. Jindal, "Hot-electron effects on channel thermal noise in fine-line NMOS field-effect transistors," *Electron Devices, IEEE Transactions on*, vol. ED-33, no. 9, pp. 1395 – 1397, 1986.

[145] A. A. Abidi, "High-frequency noise measurements on FET's with small dimensions," *Electron Devices, IEEE Transactions on*, vol. 33, no. 11, pp. 1801–1805, 1986.

[146] D. K. Shaeffer and T. H. Lee, "A 1.5–V, 1.5–GHz CMOS low noise amplifier," *IEEE J. Solid–State Circuits*, vol. 32, no. 5, pp. 745–759, 1997.

[147] A. van der Ziel, *Solid state physical electronics*, 2nd ed. Englewood Cliffs, New Jersey: Prentice-Hall, 1968.

[148] W. C. Lee, Y. C. King, T. J. King, and C. Hu, "Investigation of poly-$Si_{1-x}Ge_x$ for dual-gate CMOS technology," *IEEE Electron Device Lett.*, vol. 19, no. 7, pp. 247–249, 1998.

[149] S. Kogan, *Electronic Noise and Fluctuations in Solids*. Cambridge, New York, Melbourne: Cambridge University Press, 1996.

[150] W. Shockley, J. A. Copeland, and R. P. James, "The impedance field method of noise calculation in active semiconductor devices," in *Quantum theory of atoms, molecules and solid state*, P. O. Lowdin, Ed. Academic Press, 1966, pp. 537–563.

[151] K. M. van Vliet, "Markov approach to density fluctuations due to transport and scattering. I. Mathematical Formalism*," *J. Math. Phys.*, vol. 12, pp. 1981–1998, 1971.

[152] A. Cappy, F. Danneville, and G. Dambrine, "Noise modelling in linear and nonlinear devices," *IEICE Trans. on Electronics*, vol. E82–C, no. 6, pp. 900–907, 1999.

[153] F. Bonani, G. Ghione, M. R. Pinto, and R. K. Smith, "An efficient approach to noise analysis through multidimentional physics-based models," *Electron Devices, IEEE Transactions on*, vol. 45, no. 1, pp. 261–269, 1998.

[154] P. Shiktorov, E. Starikov, V. Gruzinskis, T. Gonzalez, J. Mateos, D. Pardo, L. Reggiani, L. Varani, and J. C. Vaissere, "Langevin forces and generalized transfer fields for noise modeling in deep submicron devices," *Electron Devices, IEEE Transactions on*, vol. 47, no. 10, pp. 1992–1998, 2000.

[155] C. Jungemann, B. Neinhüs, and B. Meinerzhagen, "Hierarchical 2–D DD and HD noise simulations of Si and SiGe devices: Part I—Theory," *Electron Devices, IEEE Transactions on*, vol. 49, no. 7, pp. 1250–1257, 2002.

[156] D. O. Martin, F.-C. Hou, J. E. Sanchez, and G. Bosman, "Impedance field noise simulations of silicon devices operating under DC and AC steady state conditions," in *Noise and Fluctuations Control in Electronic Devices*, A. A. Balandin, Ed. American Scientific Publishers, 2002, pp. 267–290.

[157] F. Bonani, G. Ghoine, M. R. Pinto, and R. K. Smith, "A novel implementation of noise analysis in general-purpose PDE-based semiconductor device simulators," in *IEEE Tech. Dig. IEDM*, 1995, pp. 777–780.

[158] S. Donati, M. A. Alam, K. S. Krisch, S. Martin, M. R. Pinto, and H. H. Vuong, "Physics-based RF noise modeling of submicron MOSFETs," in *IEEE Tech. Dig. IEDM*, 1998, pp. 81–84.

[159] L. Varani, L. Reggiani, T. Kuhn, T. González, and D. Pardo, "Microscopic simulation of electronic noise in semiconductor materials and devices," *Electron Devices, IEEE Transactions on*, vol. 41, no. 11, pp. 1916–1925, 1994.

[160] T. Gonzalez, J. Mateos, M. J. Martin-Martinez, S. Perez, R. Rengel, B. G. Vasallo, and D. Pardo, "Monte Carlo simulation of noise in electronic devices: limitations and perspectives," *Proceedings of the 3rd International Conference on Unsolved Problems of Noise*, pp. 496–503, 2003.

[161] F. N. Hooge, "$1/f$ noise sources," *Electron Devices, IEEE Transactions on*, vol. 41, no. 11, pp. 1926–1945, 1994.

[162] P. H. Handel, "Fundamental quantum $1/f$ noise in semiconductor devices," *Electron Devices, IEEE Transactions on*, vol. 41, no. 11, pp. 2023–2033, 1994.

[163] N. Wiener, "Generalized harmonic analysis," *Acta Mathematica*, vol. 55, no. 1, pp. 117–258, 1930. [Online]. Available: http://dx.doi.org/10.1007/BF02546511

[164] A. Khintchine, "Korrelationstheorie der stationären stochastischen Prozesse," *Mathematische Annalen*, vol. 109, no. 1, pp. 604–615, 1934. [Online]. Available: http://dx.doi.org/10.1007/BF01449156

[165] F. Bonani and G. Ghione, *Noise in Semiconductor Devices, Modeling and Simulation*, ser. Advanced Microelectronics. Berlin, Heidelberg, New York: Springer, 2001.

[166] C. Jungemann and B. Meinerzhagen, "A frequency domain spherical harmonics solver for the Langevin Boltzmann equation," in *International Conference on Noise in Physical Systems and 1/f Fluctuations, AIP Conf. Proc.*, vol. 780, 2005, pp. 777–782.

[167] J. E. Sanchez, G. Bosman, and M. E. Law, "Device simulation of generation-recombination noise under periodic large-signal conditions," in *IEEE Tech. Dig. IEDM*, Washington,DC(USA), 2001, pp. 477–480.

[168] F. Bonani, S. D. Guerrieri, and G. Ghione, "Physics-based simulation techniques for small-signal and large-signal device noise analysis in RF applications," *Electron Devices, IEEE Transactions on*, vol. 50, no. 3, pp. 633–644, 2003.

[169] C. M. van Vliet, "Macroscopic and microscopic methods for noise in devices," *Electron Devices, IEEE Transactions on*, vol. 41, no. 11, pp. 1902–1915, 1994.

[170] K. M. van Vliet, A. Friedmann, R. J. J. Zijlstra, A. Gisolf, and A. van der Ziel, "Noise in single injection diodes. I. A survey of methods," *J. Appl. Phys.*, vol. 46, no. 4, pp. 1804–1813, 1975.

[171] C. Jungemann, "A deterministic solver for the Langevin Boltzmann equation including the Pauli principle," in *SPIE: Fluctuations and Noise*, vol. 6600, 2007, pp. 660 007–1–660 007–12.

[172] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 4th ed. McGraw–Hill, 2001.

[173] H. Grad, "On the kinetic theory of rarefied gases," *Commun. Pure Appl. Maths*, vol. 2, pp. 331–407, 1949.

[174] X. W. Shan and X. Y. He, "Discretization of the velocity space in the solution of the Boltzmann equation," *Phys. Rev. Lett.*, vol. 80, no. 1, pp. 65–68, 1998.

[175] X. Shan, X. F. Yuan, and H. Chen, "Kinetic theory representation of hydrodynamics: a way beyond the Navier Stokes equation," *J Fluid Mech*, vol. 550, pp. 413–441, 2006.

[176] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York: Dover Publications, INC., 1972.

[177] C. Levermore, "Moment closure hierarchies for kinetic theories," *J. Stat. Phys.*, vol. 83, pp. 1021–1065, 1996.

[178] T. Grasser, R. Kosik, C. Jungemann, H. Kosina, and S. Selberherr, "Non-parabolic macroscopic transport models for device simualtion based on bulk Monte Carlo data," *J. Appl. Phys.*, vol. 97, pp. 093 710–1–12, 2005.

[179] M. R. A. Abdelmalik and E. H. van Brummelen, "Moment closure approximations of the boltzmann equation based on φ-divergences," *J. Stat. Phys.*, vol. 164, pp. 77–104, 2016. [Online]. Available: https://doi.org/10.1007/s10955-016-1529-5

[180] L. D. Landau, E. M. Lifshitz, and L. P. Pitaevskii, *Statistical Physics*. New York: Pergamon, 1969.

[181] C. Moler and C. Van Loan, "Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later," *SIAM review*, vol. 45, no. 1, pp. 3–49, 2003.

[182] R. B. Sidje, "Expokit: A software package for computing matrix exponentials," *ACM Transactions on Mathematical Software (TOMS)*, vol. 24, no. 1, pp. 130–156, 1998.

[183] H. Kim, H. S. Min, T. W. Tang, and Y. J. Park, "An extended proof of the Ramo-Shockley theorem," *Solid–State Electron.*, vol. 34, pp. 1251–1253, 1991.

[184] S. Gundapaneni, S. Ganguly, and A. Kottantharayil, "Bulk planar junctionless transistor (BPJLT): An attractive device alternative for scaling," *IEEE Electron Device Lett.*, vol. 32, no. 3, pp. 261–263, 2011.

[185] C.-W. Lee, A. Afzalian, N. D. Akhavan, R. Yan, I. Ferain, and J.-P. Colinge, "Junctionless multigate field-effect transistor," *Appl. Phys. Lett.*, vol. 94, no. 5, pp. 0 535 111–0 535 112, 2009.

[186] V. P. Georgiev, M. M. Mirza, A.-I. Dochioiu, F. Adamu-Lema, S. M. Amoroso, E. Towie, C. Riddet, D. A. MacLaren, A. Asenov, and D. J. Paul, "Experimental and simulation study of silicon nanowire transistors using heavily doped channels," *IEEE Transactions on Nanotechnology*, vol. 16, no. 5, pp. 727–735, 2017.

[187] S. C. Wagaj and Y. V. Chavan, "A review paper: A comprehensive study of junctionless MOSFETs," *American International Journal of Research in Science, Technology, Engineering & Mathematics*, pp. 40–47, 2015.

[188] M. C. et al., "On and off state hot carrier reliability in junctionless high-k mg gate-all-around nanowires," *IEDM Tech. Dig.*, vol. 14, pp. 366–369, 2015.

[189] S.-Y. K. et al., "A comparative study on hot-carrier injection in 5-story vertically integrated inversion-mode and junctionless-mode gateall-around MOSFETs," *IEEE Electron Device Lett.*, vol. 39, no. 1, pp. 4–7, 2018.

[190] R. T. Doria, M. A. Pavanello, R. Trevisoli, and M. de Souza, "Junctionless multiple-gate transistor for analog applications," *Electron Devices, IEEE Transactions on*, vol. 58, no. 8, pp. 2511–2519, 2011.

[191] S. Cho, K. R. Kim, B.-G. Park, and I. M. Kang, "RF performance and small-signal parameter extraction of junctionless silicon nanowire transistors," *Electron Devices, IEEE Transactions on*, vol. 58, no. 5, pp. 1388–1396, 2011.

[192] F. Jazaeri, L. Barbut, and J.-M. Sallese, "Trans-capacitance modeling in junctionless symmetric double-gate MOSFETs," *Electron Devices, IEEE Transactions on*, vol. 60, no. 12, pp. 4034–4040, 2013.

[193] O. Moldovan, F. Lime, and B. I. niguez, "A compact explicit model for long-channel gate-all-around junctionless MOSFETs. part II: Total charges and intrinsic capacitance characteristics," *Electron Devices, IEEE Transactions on*, vol. 61, no. 9, pp. 3042–3046, 2014.

[194] R. T. Doria, R. Trevisoli, M. de Souza, and M. A. Pavanello, "Physical insights on the dynamic response of junctionless nanowire transistors," *Symposium on Microelectronics Technology and Devices (SBMicro)*, 2016.

[195] J. Rollett, "Stability and Power-Gain Invariants of Linear Twoports," *IRE Transactions on Circuit Theory*, vol. 9, no. 1, pp. 29–32, March 1962.

[196] P. Klein, "An analytical thermal noise model of deep submicron MOSFET's for circuit simulations with emphasis on the BSIM3v3 SPICE model," in *Proc. ESSDERC*, 1998, pp. 460–463.

[197] L. M. Franca-Neto and J. S. Harris, "Excess noise in sub-micron silicon FET: Characterization, prediction and control," *29th European Solid-State Device Research Conference*, pp. 252–255, 1999.

[198] A. van der Ziel, *Noise in Solid State Devices and Circuits.* Canada: John Wiley and Sons, 1986.

[199] G. Gonzalez, *Microwave Transistor Amplifiers: Analysis and Design*, 2nd ed. Upper Saddle River, New Jersey: Prentice-Hall, 1997.

[200] R. Navid and R. W. Dutton, "The physical phenomena responsible for excess noise in short-channel MOS devices," *IEEE Tech. Dig. IEDM*, pp. 75–78, 2002.

[201] J.-S. Goo, C.-H. Choi, F. Danneville, E. Morifuji, H. S. Momose, Z. Yu, H. Iwai, T. H. Lee, and R. W. Dutton, "An accurate and efficient high frequency noise simulation technique for deep submicron MOSFETs," *Electron Devices, IEEE Transactions on*, vol. 47, no. 12, pp. 2410–2419, 2000.

[202] A. Schenk, B. Schmithüsen, A. Wettstein, A. Erlebach, S. Brugger, F. M. Bufler, T. Feudel, and W. Fichtner, "Simulation of RF noise in MOSFETs using different transport models," *IEICE Trans. on Electronics*, vol. E86-C, no. 3, pp. 481–489, 2003.

[203] P. A. Childs and D. W. Dyke, "Hot carrier quasi-ballistic transport in semiconductor devices," *Solid–State Electron.*, no. 48, pp. 765–772, 2004.

[204] P. Palestri, D. Esseni, S. Eminente, C. Fiegna, E. Sangiorgi, and L. Selmi, "Understanding quasi-ballistic transport in nano-MOSFETs: Part i-scattering in the channel and in the drain," *Electron Devices, IEEE Transactions on*, vol. 52, no. 12, pp. 2727–2735, 2005.

[205] C. Jungemann, "Ballisticity of the linear response transport in nanometric silicon devices," in *Semiconductor Device Research Symposium, 2007 International*, Dec 2007, pp. 1–2.

[206] K. Zhao, C. Jungemann, and X. Liu, "Quasi-ballisticity of the electron transport in a 16nm silicon double-gate nmosfet," in *Electron Devices and Solid-State Circuits (EDSSC), 2011 International Conference of*, nov. 2011, pp. 1 –2.

[207] C. L. Allyn, A. C. Gossard, and W. Wiegmann, "New rectifying semiconductor structure by molecular beam epitaxy," *Appl. Phys. Lett.*, vol. 36, no. 5, pp. 373–376, 1980.

[208] H. Kroemer, "Heterostructure bipolar transistors and integrated circuits," *Proceedings of the IEEE*, vol. 70, no. 1, pp. 13–25, 1982.

[209] D. Ankri and L. F. Eastman, "Gaalas-gaas ballistic hetero-junction bipolar transistor," *Electronics Letters*, vol. 18, no. 17, pp. 750–751, 1982.

[210] R. J. Malik, T. R. Aucoin, R. L. Ross, K. Board, C. E. C. Wood, and L. F. Eastman, "Planar-doped barriers in gaas by molecular beam epitaxy," *Electronics Letters*, vol. 16, no. 22, pp. 836–838, 1980.

[211] P. Hesto, J. F. Pone, R. Castagne, and J. L. Pelouard, *A dynamic Monte Carlo simulation of conduction in submicron GaAs devices at 77K. In: Grubin H.L., Hess K., Iafrate G.J., Ferry D.K. (eds) The Physics of Submicron Structures.* Boston, MA: Springer, 1984.

[212] M. Dyakonov and M. Shur, "Plasma wave electronics: novel terahertz devices using two dimensional electron fluid," *Electron Devices, IEEE Transactions on*, vol. 43, no. 10, pp. 1640–1645, Oct 1996.

[213] ——, "Shallow water analogy for a ballistic field effect transistor: New mechanism of plasma wave generation by dc current," *Phys. Rev. Lett.*, vol. 71, pp. 2465–2468, Oct 1993. [Online]. Available: http://link.aps.org/doi/10.1103/PhysRevLett.71.2465

[214] K. Rupp, T. Grasser, and A. Jüngel, "On the feasibility of spherical harmonics expansions of the Boltzmann transport equation for three-dimensional device geometries," in *IEEE International Electron Devices Meeting (IEDM)*, dec. 2011, pp. 34.1.1 –34.1.4.

[215] K. Rupp, A. Jüngel, and T. Grasser, "A gpu-accelerated parallel preconditioner for the solution of the boltzmann transport equation for semiconductors," *In: Keller R., Kramer D., Weiss JP. (eds) Facing the Multicore - Challenge II. Lecture Notes in Computer Science*, vol. 7174, pp. 147–157, 2012.

[216] K. Rupp, C. Jungemann, S.-M. Hong, M. Bina, T. Grasser, and A. Jünge, "A review of recent advances in the spherical harmonics expansion method for semiconductor device simulation," *J. Computational Electronics*, vol. 15, pp. 939–958, 2016.

[217] S. Tiwari, M. V. Fischetti, and S. E. Laux, "Overshoot in transient and steady-state in GaAs, InP, Ga.47In.53As and InAs bipolar transistors," *Proc. Int. Electron Device Meeting*, pp. 435–438, 1990.

[218] S. E. Laux and M. V. Fischetti, "Monte Carlo study of velocity overshoot in switching a 0.1-micron CMOS inverter," in *IEEE Tech. Dig. IEDM*, 1997, pp. 877–880.

[219] Y. Koseki, V. Rhzhii, T. Otsuji, V. V. Popov, and A. Satou, "Giant plasmon instability in a dual-grating-gate graphene field-effect transistor," *Phys. Rev. B, Condens. Matter*, vol. 93, no. 24, p. 245408, 2016.

[220] S. Di, K. Zhao, T. Lu, G. Du, and X. Liu, "Investigation of transient responses of nanoscale transistors by deterministic solution of the time-dependent bte," *J. Computational Electronics*, vol. 15, no. 3, pp. 770–777, 2016.