Exploitation of the Maximum Entropy Principle in Mathematical Modeling of Charge Transport in Semiconductors

Giovanni Mascali [1,*,†] and Vittorio Romano [2,†]

[1] Dipartimento di Matematica ed Informatica, Università della Calabria e INFN-Gruppo c. Cosenza, Rende 87036, Italy

[2] Dipartimento di Matematica ed Informatica, Università di Catania, Catania 95125, Italy; romano@dmi.unict.it

* Correspondence: giovanni.mascali@unical.it; Tel.: +39-0984-496-418

† These authors contributed equally to this work.

Abstract: In the last two decades, the Maximum Entropy Principle (MEP) has been successfully employed to construct macroscopic models able to describe the charge and heat transport in semiconductor devices. These models are obtained, starting from the Boltzmann transport equations, for the charge and the phonon distribution functions, by taking—as macroscopic variables—suitable moments of the distributions and exploiting MEP in order to close the evolution equations for the chosen moments. Important results have also been obtained for the description of charge transport in devices made both of elemental and compound semiconductors, in cases where charge confinement is present and the carrier flow is two- or one-dimensional.

1. Introduction

In the design of new generation electronic devices, the modern micro- and nanoelectronics industry has an increasing need for mathematical models to simulate devices before they are realized in the laboratory. In particular, the accurate modeling of energy transport in semiconductors is necessary in order to describe high-field phenomena, such as hot electron propagation, impact ionization and heat generation in the bulk material. Semiclassically, the most effective way to describe these phenomena makes use of the semiclassical Boltzmann Transport Equation (BTE) [1]. However, solving the BTE is a daunting computational task, both using an indirect stochastic approach by Monte Carlo methods and direct numerical schemes based on discontinuous Galerkin methods or finite differences [2]. This is the reason why, in many cases, it is desirable to have simpler macroscopical models, known as hydrodynamical-like models, which are highly useful for computer aided design (CAD) purposes. These models are obtained from the infinite set of moment equations of the BTE by a suitable truncation procedure. It is well-known that a closure assumption is required in order to have a closed system of evolution equations. In the past, various closure assumptions have been made for the semiconductor transport moment systems, leading to various classes of hydrodynamical models, see e.g., [3–5]. However, these various closure assumptions are, at best, only phenomenological and often a consistent physical and mathematical justification is lacking. Lately, a closure assumption based on the Maximum Entropy Principle of extended thermodynamics [6,7] has been successfully applied, both in the parabolic and non-parabolic band approximation, to various types of semiconductors [8–13]. The resulting models, which differ for the choice of the moments to assume as field variables, are, in fact, able to describe charge transport due both to electrons and holes and also heat transport

due to phonons. All the main scattering mechanisms between carriers and phonons and among phonons themselves are taken into account. The models also have nice mathematical properties, being symmetric hyperbolic. In particular, this assures the well-posedness of the Cauchy problems and the finite velocity of the propagation of disturbances [14].

Due to the increasing shrinking of modern device dimensions, quantum effects are beginning to play a relevant role in charge transport. In the framework of the moment method and of Maximum Entropy Principle (MEP), a strategy to take into account these effects has been proposed in [15] and consists of using the moment system arising from the Wigner equation. The starting point is to expand the Wigner function and the relative transport equation with respect to the squared reduced Planck constant $\hbar^2$. The zero-order part of the collision operator is supposed to be the same as the semiclassical one, while the first-order contribution is supposed to act only on the $\hbar^2$ correction of the Wigner function and is modeled in a relaxation form. Therefore, at zero order, the Wigner function is given by the solution of the semiclassical Boltzmann equation, which is approximated with the standard maximum entropy method, while the $\hbar^2$ order correction is obtained with a Chapman–Enskog expansion in the high field scaling.

In the description of charge transport in some devices, such as double gate metal oxide semiconductor field effect transistors (DG-MOSFETs), where quantum effects are relevant only along one direction, called the confinement or transversal direction, another strategy can be used. One can adopt a quasi-static description along the confining direction based on a coupled Schrödinger–Poisson system which leads to a subband decomposition of the electron energy levels, while the transport along the longitudinal directions can be described by a semiclassical Boltzmann equation for each subband [16]. Therefore, a complete description can be done in terms of a coupled Schrödinger–Poisson–Boltzmann system. However, the numerical integration of the transport part, which has been performed by employing Monte Carlo or deterministic methods [16,17], is also in this case very expensive, from a computational point of view, for CAD purposes. Consequently, it can be convenient to substitute the Boltzmann equations with macroscopic models again obtained by using the moment method and closing the moment equations with MEP [18–20].

In this paper, we will give a review of all the above-mentioned models according to the following plan. In Section 2, the 3D semiclassical macroscopic models are presented, showing how they are closed by using MEP and in Section 3 their quantum correction is discussed. In Section 4, the case of quantum confinement is described. Eventually in Section 5, some numerical simulations are sketched.

2. The 3D Semiclassical Macroscopic Models

In the semiclassical description, roughly speaking, the charge and heat transport in semiconductors can be described by modeling them as consisting of a mixture of gases of electrons and phonons. As regards electrons, those which mainly contribute to the current flow occupy states near to the lowest minima of the lowest conductions bands and to the highest maxima of the highest valence bands. The latter contribution can be conveniently treated in terms of pseudo-particles called holes (missing electrons), which have electric charge, crystal momentum and energy opposite to those of the missing electrons in the considered valence band [1,21,22]. An electron/hole population for each neighborhood of the lowest/highest minima/maxima of the lowest/highest conduction/valence bands is taken into account. The neighborhoods of the minima/maxima are usually called valleys. The electron/hole state is characterized by the index of the band it occupies, by its wave vector and spin state.

The dependence of the carrier energy on the wave vector in each valley can be analytically approximated by non-parabolic dispersion relations of the following form [1,9]

$$\mathcal{E}_A(\mathbf{k}_A) = \frac{\hbar^2 |\mathbf{k}_A|^2}{2\,m_e}\,\gamma_A(\mathcal{E}_A)\,\psi_A^{-1}(\mathbf{n}_A), \tag{1}$$

where $\mathcal{E}_A$ is the carrier energy in the A-th valley measured from the bottom of the valley, the index A running over the considered valleys, $\mathbf{k}_A$ is the carrier quasi-wave vector, referred, for each valley, to the minimum or maximum of the valley [1], $\mathbf{n}_A := \frac{\mathbf{k}_A}{|\mathbf{k}_A|}$, γ_A is the non-parabolicity factor, and m_e is the free electron mass.

For electrons, if one uses the ellipsoidal approximation, the dependence of ψ_A on $\mathbf{n}_A$ is given by

$$\psi_A^{-1} = \frac{((\mathbf{n}_A)_1)^2}{(m_A^*)_1} + \frac{((\mathbf{n}_A)_2)^2}{(m_A^*)_2} + \frac{((\mathbf{n}_A)_3)^2}{(m_A^*)_3},$$

where $(m_A^*)_i^{-1}$, $i = 1, 2, 3$, are the diagonal elements (eigenvalues) of the inverse effective mass tensor of the A-th valley, multiplied by m_e, referred to an orthonormal basis of the tensor.

Analogously, for holes, one has

$$\psi_A^{-1} = |A_A| \mp \sqrt{B_A^2 + \frac{C_A^2}{4}\left(\sin^4\vartheta\sin^2 2\varphi + \sin^2 2\vartheta\right)},$$

in the case of warped bands, with φ and θ the azimuthal and polar angle of the wave vector respectively, and A_A, B_A and C_A inverse valence band parameters.

In the analytic approximations, for each valley, the domain of the wave-vector is extended to all $\mathbb{R}^3$ and the volume element in the $\mathbf{k}$-space can be written as (henceforth we omit the valley index unless there is a possibility of confusion)

$$d^3\mathbf{k} = \frac{m_e\sqrt{2m_e}}{\hbar^3}\sqrt{\frac{\mathcal{E}}{\gamma^5(\mathcal{E})}}\left(\gamma(\mathcal{E}) - \mathcal{E}\gamma'(\mathcal{E})\right)\psi^{\frac{3}{2}}(\varphi, \theta)d\mathcal{E}d\Omega,$$

where the prime denotes a derivative with respect to the argument of the function, and $d\Omega$ is the solid angle element. The charge carrier velocity, given by $\mathbf{v} = \frac{1}{\hbar}\nabla_{\mathbf{k}}\mathcal{E}$, has the following expression in terms of the energy and the angular variables

$$\begin{aligned}
v_i &= \frac{\hbar}{2\,m_e}\frac{\gamma^2(\mathcal{E})}{\gamma(\mathcal{E}) - \mathcal{E}\gamma'(\mathcal{E})}l_i, \quad l_i := \frac{\partial}{\partial k_i}(|\mathbf{k}|^2\psi^{-1}) = g(\mathcal{E})\eta_i(\varphi, \theta)n_i, \\
g &:= \frac{2\sqrt{2\,m_e\,\mathcal{E}}}{\hbar\,\sqrt{\gamma(\mathcal{E})}}, \quad \eta_i := \frac{\sqrt{\psi}}{m_i^*}.
\end{aligned}$$

Similarly, for phonons that are used to describe the crystal vibrations and strongly influence charge transport, one population for each branch has to be considered [23]. The number of branches is equal to 3ν, ν being the total number of atoms per unit cell of the crystal lattice, and a is the lattice constant. Among the most used analytic approximations for the phonon dispersion relation, we report the one which is isotropic and quadratic in the phonon wave vector $\mathbf{q}$

$$\epsilon_p = \epsilon_0^p + \hbar v_s^p\,|\mathbf{q}| + \hbar c^p\,|\mathbf{q}|^2\,,\,|\mathbf{q}| \in \left[0, \frac{2\pi}{a}\right], \tag{2}$$

where ϵ_p is the phonon energy, v_s^p and c^p, $p = 1, \ldots, 3\nu$, are suitable constants depending on the material under consideration, and p varies over the 3ν phonon branches. The Einstein and the Debye approximations are particular cases of (2), respectively obtained for $v_s^p = c^p = 0$ and $c^p = 0, v_s^p \neq 0$.

At the kinetic level, the state of the electrons, holes and phonons can be described by their one-particle distribution functions f_{A_β}, $\beta = e$ (electron), h (hole), A_β indicating the valleys occupied by β-carriers ($\{A_e\} \cup \{A_h\} = \{A\}$), and g_p, with $p = 1, \ldots, 3\nu$, whose time evolution is determined by the Boltzmann–Bloch–Peierls (BBP) system (see [24] and references therein)

$$\frac{\partial f_{A_\beta}}{\partial t} + \mathbf{v}_{A_\beta} \cdot \nabla_{\mathbf{x}}f_{A_\beta} + \frac{q_\beta}{\hbar}\mathbf{E} \cdot \nabla_{\mathbf{k}}f_{A_\beta} = \mathcal{C}_{im}(f_{A_\beta}) + \sum_{A'_{\beta'},p}\mathcal{C}_p^{A_\beta A'_{\beta'}}(f_{A_\beta}, f_{A'_{\beta'}}, g_p) + \sum_{A'_{\bar\beta},p}\mathcal{I}_p^{A_\beta A'_{\bar\beta}}[f_{A_\beta}, f_{A'_{\bar\beta}}, g_p], \tag{3A}$$

$$\frac{\partial g_p}{\partial t} + \mathbf{v}_p \cdot \nabla_{\mathbf{x}} g_p = \sum_{A_\beta, A'_{\beta'}} \mathcal{C}^p_{A_\beta, A'_{\beta'}}(g_p, f_{A_\beta}, f_{A'_{\beta'}}) + \sum_\eta \mathcal{C}^p_\eta(g_p) + \sum_{A_\beta, A'_{\bar\beta}} \mathcal{I}^p_{A_\beta A'_{\bar\beta}}[g_p, f_{A_\beta}, f_{A'_{\bar\beta}}], \quad (3B)$$

$$\Delta(\varepsilon_s V) = q\big(n(\mathbf{x}, t) - p(\mathbf{x}, t) + N_a(\mathbf{x}) - N_d(\mathbf{x})\big), \quad (3C)$$

where q_β is the charge of the particles, $q_{e,h} = \mp q$ with q absolute value of the elementary charge, $\bar\beta$ is the complement of β with respect to the set $\{e, h\}$ $\mathbf{v}_p = \frac{1}{\hbar}\nabla_{\mathbf{q}}\epsilon_p$ is the phonon group velocity, ε_s is the material permittivity, V and $\mathbf{E}$ are the electric potential and field, η labels the type of scattering among phonons themselves (see below), N_a and N_d are the acceptor and donor concentrations, while n is the total electron density

$$n = \sum_{A_e} f_{A_e}, \quad (4)$$

and p the total hole density

$$p = \sum_{A_h} f_{A_h}. \quad (5)$$

The Boltzmann Equation (3A) is coupled among them through the Poisson Equation (3C) and some of the collision operators which appear at the right-hand side of (3A) and (3B).

The operator $\mathcal{C}_{im}$ [1] takes into account scatterings between carriers and impurities and is elastic and intravalley, which means that the initial and the final state of the carrier lie in the same valley. Its form is

$$\mathcal{C}_{im}(f_{A_\beta}) = \int_{\mathbb{R}^3} \big[w_{im}(\mathbf{k}', \mathbf{k})f_{A_\beta}(\mathbf{k}') - w_{im}(\mathbf{k}, \mathbf{k}')f_{A_\beta}(\mathbf{k})\big]d\mathbf{k}',$$

the impurity scattering transition rate being given by

$$w_{im}(\mathbf{k}, \mathbf{k}') = \mathcal{K}^{(im)}\frac{1}{\big[|\mathbf{k} - \mathbf{k}'|^2 + \lambda_D^2\big]^2}\delta(\mathcal{E}(\mathbf{k}') - \mathcal{E}(\mathbf{k})),$$

with $\lambda_D = \sqrt{\frac{(N_{a/d})q^2}{\epsilon_s k_B T_L}}$ the inverse Debye length and $\mathcal{K}^{(im)} = \frac{Z_{a/d}^2 N_{a/d} q^4}{4\pi\hbar\epsilon_s^2}$, where Z is the impurity charge number, T_L the lattice temperature and k_B the Boltzmann constant. According to whether interactions with donors or acceptors are considered, N_d or N_a, and Z_d or Z_a have to be taken.

The collision operators $\mathcal{C}_p^{AA'}(f_{A, f_{A'}, g_p})$ describe interactions between carriers of the same type and phonons. These scatterings can be intravalley (intraband) ($A = A'$), or intervalley (intraband or interband) ($A \neq A'$) [1], which means that the initial and final state belong to different valleys. These operators read [24]

$$\mathcal{C}_p^{A\,A'}(f_A, f_{A'}, g_p) = \int_{S_{\frac{2\pi}{a}}} \int_{\mathbb{R}^3} \big[w_p^{AA'+}(\mathbf{k}, \mathbf{k}', \mathbf{q})\kappa_1^{A\,A'}(g_p, f_A, f_{A'})$$

$$+ w_p^{A_\beta A'-}(\mathbf{k}, \mathbf{k}', \mathbf{q})\kappa_2^{A\,A'}(g_p, f_A, f_{A'})\big]d\mathbf{k}'d\mathbf{q},$$

where $S_{\frac{2\pi}{a}}$ is the sphere of radius $\frac{2\pi}{a}$, which approximates the first Brillouin zone $\mathcal{B}$ in the case of isotropic approximations for the phonon dispersion relations, and, for example, the integrands are given by

$$\kappa_1^{A_e A'_e}(g_p, f_{A_e}, f_{A'_e}) := \left(\frac{g_p}{y_p} + 1\right)f_{A'_e}(\mathbf{k}') - \frac{g_p}{y_p}f_{A_e}(\mathbf{k}),$$

$$\kappa_2^{A_e A'_e}(g_p, f_{A_e}, f_{A'_e}) := \frac{g_p}{y_p}f_{A'_e}(\mathbf{k}') - \left(\frac{g_p}{y_p} + 1\right)f_{A_e(\mathbf{k})},$$

$$w_p^{A_e A'_e \pm}(\mathbf{k}, \mathbf{k}', \mathbf{q}) := s_p^{A_e A'_e}(\mathbf{q})\delta(\mathcal{E}_{A'_e}(\mathbf{k}') - \mathcal{E}_{A_e}(\mathbf{k}) \mp \epsilon_p)\delta(\mathbf{k}' - \mathbf{k} \mp \mathbf{q} + \mathbf{G}),$$

y_p being the p-phonon density of states, δ the Dirac delta function, and $\mathbf{G}$ a vector of the Brillouin zone, whose presence is due to the fact that the total wave vector is conserved up to it. The scattering functions $s_p^{AA'}$ are characteristic of the type of interaction of carriers, for example with acoustic and non-polar optical phonons for elemental semiconductors such as Si and Ge, and also with polar optical phonons for compound semiconductors, such as GaAs and SiC. The previous expressions of the scattering rates w are valid in the case when only conduction electrons are involved; for all the other cases as well as for the expressions of the scattering functions, we refer the interested reader to [1,8,9].

The generation-recombination collision operators $\mathcal{I}_p{}^{A_\beta A'_{\tilde{\beta}}}[f_{A_\beta}, f_{A'_{\tilde{\beta}}}, g_p]$ include several mechanisms, among which the most important are the Auger and the Schockley–Read–Hall (SRH) processes. In a relaxation time approximation and in the simpler case of a single conduction and a single valence band, the operators read [25]

$$\mathcal{I}[f_A, f_{\tilde{A}}] = -\left(\Gamma_n\, n + \Gamma_h\, p\right)\left(\tilde{n} f_A - n_i{}^2 \mathcal{M}_A\right) - \frac{\tilde{n} f_A - n_i{}^2 \mathcal{M}_A}{\tau_h(n + n_i) + \tau_n(p + n_i)}, \tag{6}$$

where $\tilde{n} = n$ or p according to whether the A-th valley is populated by electrons or holes, the $\mathcal{M}_A$'s are the Maxwellians normalized to unit, Γ_n and Γ_h are the Auger electron and hole constant coefficients, τ_n and τ_h are the carrier lifetimes, and n_i is the intrinsic concentration [1].

The collision operators which appear in the Boltzmann–Peierls equations for phonons, relative to their interactions with carriers, read [24]

$$\mathcal{C}^p_{A_\beta A'_{\beta'}}(g_p, f_{A_\beta}, f_{A'_{\beta'}}) = \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} w_p^{A_\beta A'_{\beta'}, +}(\mathbf{k}, \mathbf{k}', \mathbf{q}) \kappa_1^{A_\beta A'_{\beta'}}(g_p, f_{A_\beta}, f_{A'_{\beta'}})\, d\mathbf{k}'\, d\mathbf{k}.$$

The other interactions of phonons can be distinguished into intrinsic ones, arising from the anharmonicity of the interatomic forces, and extrinsic ones, due to phonon scatterings at the boundaries of the crystal and at various types of crystal defects and imperfections. In their turn, the anharmonic scatterings can be normal processes (*N*-processes), in which the phonon total momentum after the interaction is conserved, and umklapp processes (*U*-processes) for which the total momentum changes by a reciprocal lattice vector multiplied by $\hbar$ after a collision. On the other hand, all extrinsic processes do not conserve the total momentum after a collision and, together with the umklapp ones, are usually called resistive processes. For the expressions of the corresponding collision operators, we refer the interested readers to [23] and references therein.

Macroscopic models can be constructed starting from the BBP system by taking a suitable number of moments of the distribution functions [14]. The most common choice is that of considering the following functions of the carrier and phonon wave vectors $\psi_{A_\beta}(\mathbf{k}) := \left(1, \mathbf{v}_{A_\beta}, \mathcal{E}_{A_\beta}, \mathcal{E}_{A_\beta} \mathbf{v}_{A_\beta}\right)$ and $(\psi_p(\mathbf{q})) := (\epsilon_p, \epsilon_p \mathbf{v}_p)$, to which the following macroscopic state variables correspond:

$$\begin{pmatrix} n_{A_\beta} \\ W_{A_\beta} \end{pmatrix} = \int_{\mathbb{R}^3} \begin{pmatrix} 1 \\ \frac{\mathcal{E}_{A_\beta}}{n_{A_\beta}} \end{pmatrix} f_{A_\beta}\, d\mathbf{k}, \quad \begin{pmatrix} \mathbf{V}_{A_\beta} \\ \mathbf{S}_{A_\beta} \end{pmatrix} = \frac{1}{n_{A_\beta}} \int_{\mathbb{R}^3} \begin{pmatrix} \mathbf{v}_{A_\beta} \\ \mathcal{E}_{A_\beta} \mathbf{v}_{A_\beta} \end{pmatrix} f_{A_\beta}\, d\mathbf{k}, \tag{7}$$

$$\begin{pmatrix} W_p \\ \mathbf{Q}_p \end{pmatrix} = \int_{S_{\frac{2\pi}{a}}} \begin{pmatrix} \epsilon_p \\ \epsilon_p \mathbf{v}_p \end{pmatrix} g_p\, d\mathbf{q}, \tag{8}$$

which are the carrier number densities, the average energies, velocities and energy-fluxes per carrier, and the phonon average energies and energy-fluxes, respectively. The above choice involves the minimal number of moments necessary for describing the thermal energy transport, but this number, if required by the physical problem under study, can be easily extended to cover, for example, an arbitrary number of scalar and vector moments both for carriers and phonons, by taking into account, e.g., higher microscopic energy powers in the functions ψ [7,26].

The evolution equations for the state variables (7) and (8) can be obtained directly from the Boltzmann equations by integration:

$$\frac{\partial}{\partial t}\begin{pmatrix} n_{A_\beta} \\ n_{A_\beta} W_{A_\beta} \end{pmatrix} + \nabla_{\mathbf{x}} \cdot \left(n_{A_\beta} \begin{pmatrix} \mathbf{V}_{A_\beta} \\ \mathbf{S}_{A_\beta} \end{pmatrix} \right) - n_{A_\beta} q_\beta \begin{pmatrix} 0 \\ \mathbf{E} \cdot \mathbf{V}_{A_\beta} \end{pmatrix} = n_{A_\beta} \begin{pmatrix} C_{n_{A_\beta}} \\ C_{W_{A_\beta}} \end{pmatrix}, \tag{9}$$

$$\frac{\partial}{\partial t} \left(n_{A_\beta} \begin{pmatrix} \mathbf{V}_{A_\beta} \\ \mathbf{S}_{A_\beta} \end{pmatrix} \right) + \nabla_{\mathbf{x}} \left(n_{A_\beta} \begin{pmatrix} F^{(0)}_{A_\beta} \\ F^{(1)}_{A_\beta} \end{pmatrix} \right)$$
$$- n_{A_\beta} q_\beta \begin{pmatrix} G^{(0)}_{A_\beta} \\ G^{(1)}_{A_\beta} \end{pmatrix} \mathbf{E} = n_{A_\beta} \begin{pmatrix} C_{\mathbf{V}_{A_\beta}} \\ C_{\mathbf{S}_{A_\beta}} \end{pmatrix} \tag{10}$$

$$\frac{\partial}{\partial t} \begin{pmatrix} W_p \\ \mathbf{Q}_p \end{pmatrix} + \nabla_{\mathbf{x}} \cdot \begin{pmatrix} \mathbf{Q}_p \\ T_p \end{pmatrix} = \begin{pmatrix} C_{W_p} \\ C_{\mathbf{Q}_p} \end{pmatrix}. \tag{11}$$

In the above equations, the extra-fluxes and production terms relative to charges respectively read

$$\begin{pmatrix} F^{(0)}_{A_\beta} \\ F^{(1)}_{A_\beta} \end{pmatrix} = \frac{1}{n_{A_\beta}} \int \begin{pmatrix} 1 \\ \mathcal{E}_{A_\beta} \end{pmatrix} \mathbf{v}_{A_\beta} \otimes \mathbf{v}_{A_\beta} f_{A_\beta} d\mathbf{k}, \quad \begin{pmatrix} \text{velocity flux,} \\ \text{flux of the energy-flux} \end{pmatrix}$$

$$\begin{pmatrix} G^{(0)}_{A_\beta} \\ G^{(1)}_{A_\beta} \end{pmatrix} = \frac{1}{\hbar\, n_{A_\beta}} \int \begin{pmatrix} \nabla_{\mathbf{k}} \mathbf{v}_{A_\beta} \\ \nabla_{\mathbf{k}} \left(\mathcal{E}_{A_\beta} \mathbf{v}_{A_\beta} \right) \end{pmatrix} f_{A_\beta} d\mathbf{k},$$

$$n_{A_\beta} C_{M_{\psi_{A_\beta}}} = \int \psi_{A_\beta}(\mathbf{k}) \Big[\mathcal{C}_{im}(f_{A_\beta}) + \sum_{A'_{\beta'},p} \mathcal{C}^{A_\beta A'_{\beta'}}_p (f_{A_\beta}, f_{A'_{\beta'}}, g_p)$$
$$+ \sum_{A'_{\tilde\beta},p} \mathcal{I}^{A_\beta A'_{\tilde\beta}}_p [f_{A_\beta}, f_{A'_{\tilde\beta}, g_p}] \Big] d\mathbf{k}, \quad M_{\psi_{A_\beta}}\text{-production,}$$

with $M_{\psi_{A_\beta}} := \left(n_{A_\beta}, \mathbf{V}_{A_\beta}, W_{A_\beta}, \mathbf{S}_{A_\beta} \right)$, while for the phonons, we have

$$T_p = \int \epsilon_p \mathbf{v}_p \otimes \mathbf{v}_p\, g_p(\mathbf{q}) d\mathbf{q}, \qquad \text{flux of the energy-flux,}$$

$$\begin{pmatrix} C_{W_p} \\ C_{\mathbf{Q}_p} \end{pmatrix} = \int \begin{pmatrix} \epsilon_p \\ \epsilon_p \mathbf{v}_p \end{pmatrix} \sum_{A_\beta, A'_{\beta'}} \mathcal{C}^p_{A_\beta, A'_{\beta'}} (g_p, f_{A_\beta}, f_{A'_{\beta'}}) + \sum_\eta \mathcal{C}^p_\eta \tag{12}$$

$$+ \sum_{A_\beta, A'_{\tilde\beta}} \mathcal{I}^p_{A_\beta A'_{\tilde\beta}} [g_p, f_{A_\beta}, f_{A'_{\tilde\beta}}] d\mathbf{q}, \quad \begin{pmatrix} \text{energy production} \\ \text{energy-flux production} \end{pmatrix}.$$

In the evolution equations, the number of the unknowns is greater than that of the equations, therefore constitutive equations are needed for the extra-variables $F^{A(0)}_{ij}, G^{A(0)}_{ij}, F^{A(1)}_{ij}, G^{A(1)}_{ij}, T^p_{ij}, C_{n_A}, C_{W_A}, C_{\mathbf{V}_A}, C_{\mathbf{S}_A}, C_{W_p}, C_{\mathbf{Q}_p}$. A systematic way to find these relations is founded on a universal physical principle: the maximum entropy principle [6,14,27,28]. It states that, if a certain number of moments is known, then the least biased distribution functions, which can be used for evaluating the unknown moments, are those maximizing the total entropy functional under the constraint that they reproduce the known moments. In the case under consideration, neglecting the mutual interactions

among the subsystems and degeneration effects (the degeneracy case can be tackled in the same way; we present the non-degenerate case only for the sake of simplicity), the total entropy is

$$
\mathcal{S} = -k_B \left\{ \sum_{A_\beta} \int_{\mathbb{R}^3} \left(f_{A_\beta} \ln \frac{f_{A_\beta}}{y_{A_\beta}} - f_{A_\beta} \right) d\mathbf{k} + \sum_p \int_{S_{\frac{2\pi}{a}}} \left[g_p \ln \frac{g_p}{y_p} \right. \right.
$$
$$
\left. \left. - y_p \left(1 + \frac{g_p}{y_p} \right) \ln \left(1 + \frac{g_p}{y_p} \right) \right] d\mathbf{q} \right\},
$$

with $y_{A_\beta} = \frac{2}{(2\pi)^3}$ the charge density of states, while the constraints are given by (7) and (8). Let us introduce the functional spaces

$$
\mathcal{F}_{\psi_{A_\beta}} = \left\{ h : \mathbb{R}^3 \mapsto \mathbb{R} \quad s.t. \quad \int_{\mathbb{R}^3} |h(\mathbf{k}) \psi_j(\mathbf{k})| \, d\mathbf{k} < +\infty, \quad \forall \psi_j \in \{\psi_{A_\beta}\} \right\},
$$

$$
\mathcal{F}_{\psi_p} = \left\{ h : \mathbb{R}^3 \mapsto \mathbb{R} \quad s.t. \quad \int_{S_{\frac{2\pi}{a}}} |h(\mathbf{q}) \psi_j(\mathbf{q})| \, d\mathbf{q} < +\infty, \quad \forall \psi_j \in \{\psi_p\} \right\}.
$$

Given the values of the moments $M_{\psi_{A_\beta}}$ and $W_p, \mathbf{Q}_p$, MEP amounts to solve the following optimization problem:

$$
\max_{f_{A_\beta} \in \mathcal{F}_{\psi_{A_\beta}}, g_p \in \mathcal{F}_{\psi_p}} \mathcal{S}
$$

under the constraints

$$
\int_{\mathbb{R}^3} \psi_{A_\beta}(\mathbf{k}) f_{A_\beta} \, d\mathbf{k} = M_{\psi_{A_\beta}}, \quad \int_{S_{\frac{2\pi}{a}}} g_p \left(\begin{array}{c} \epsilon_p \\ \epsilon_p \mathbf{v}_p \end{array} \right) d\mathbf{q} = \left(\begin{array}{c} W_P \\ \mathbf{Q}_p \end{array} \right)
$$

where time and position are considered fixed and omitted for the sake of simplifying the notation. The solution of this maximization problem is given by [29]

$$
f_{A_\beta}^{ME} = \exp \left(-\Lambda_{A_\beta} - \Lambda_{W_{A_\beta}} \mathcal{E}_{A_\beta} - \mathbf{v}_{A_\beta} \cdot (\Lambda_{\mathbf{V}_{A_\beta}} + \mathcal{E}_{A_\beta} \Lambda_{\mathbf{S}_{A_\beta}}) \right),
$$
$$
g_p^{ME} = \frac{y_p}{\exp \left(\epsilon_p \Lambda_{W_p + \epsilon_p \mathbf{v}_p \cdot \Lambda_{\mathbf{Q}_p}} \right) - 1},
$$

which, linearized with respect to the vector variables (small anisotropy assumption) [8,9,27], becomes

$$
f_{A_\beta}^{ME} = \exp \left(-\Lambda_{A_\beta} - \Lambda_{W_{A_\beta}} \mathcal{E}_{A_\beta} \right) \left(1 - \mathbf{v}_{A_\beta} \cdot (\Lambda_{\mathbf{V}_{A_\beta}} + \mathcal{E}_{A_\beta} \Lambda_{\mathbf{S}_{A_\beta}}) \right),
$$
$$
g_p^{ME} = \frac{y_p}{\exp \left(\epsilon_p \Lambda_{W_p} \right) - 1} - \frac{y_p \epsilon_p \exp \left(\epsilon_p \Lambda_{W_p} \right)}{(\exp \left(\epsilon_p \Lambda_{W_p} \right) - 1)^2} \mathbf{v}_p \cdot \Lambda_{\mathbf{Q}_p},
$$

where the Λ's are Lagrange multipliers, related to the state variables by means of the constraint relations (7) and (8). Linearization is made in order to simplify the inversion of the constraints, otherwise it has to be done numerically [29]. After inversion, the dependence of the distribution functions on $(\mathbf{x}, t)$ will only be through the state variables, and substituting the distributions into the integrals defining the extra-variables, the needed closure relations can be obtained. It can be proven that the balance equations with the closure relations given by MEP form a quasilinear hyperbolic system in the relevant physical region of the field variables.

Following this procedure, mathematical models for the description of charge transport in unipolar and bipolar silicon devices have been constructed, see for example [10,30], while results relative to

Si thermal properties can be found in [23,31]. The procedure has also been applied to compound semiconductors such as GaAs, GaN and SiC [8,9].

3. Quantum Corrections to the Semiclassical Models

Based on the previous considerations, a natural way to get a quantum macroscopic model is to use MEP in a quantum framework to close the moment system arising from the Wigner equations. The general guidelines can be found in [32,33]. This approach has been followed in [34] (see also [35] and references therein). However, one has to deal with operatorial equations which are very complex and therefore hard to be solved numerically. Moreover, drastic simplifications of the collision terms are introduced in order to make the problem tractable.

Another strategy can be adopted to close the moment system arising from the Wigner equation. According to this strategy, the Wigner function and the relative transport equation are expanded with respect to $\hbar^2$. The zero-order part of the collision operator is supposed to be the same as the semiclassical one, while the first-order contribution is supposed to act only on the $\hbar^2$ correction of the Wigner function and is modeled in a relaxation form. Therefore, at the zero order, the Wigner function is given by the solution of the semiclassical Boltzmann equation, which is approximated with the standard maximum entropy method, while the $\hbar^2$ order correction is obtained with a Chapman–Enskog expansion in the high field scaling.

The typical physical situation which can be described in this way is that when the main contribution to the charge transport is semiclassical, the quantum effects enter as small perturbations. For example, this is reasonable for devices such as MOSFETs of characteristic length of about ten nanometers subjected to strong electric fields.

The main assumption is that there is a balance between the $\hbar^2$ drift and collision terms. This can be motivated by the fact that the collision frequency of the semiclassical scatterings tends to increase as the energy rises. Therefore, quantum effects are expected to be relevant at high fields; in such conditions, there should be high energies with consequently high collision frequencies.

The starting point for our derivation of the quantum corrections to the semiclassical model is the single particle Wigner–Poisson system, which represents the quantum analogue of the semiclassical Boltzmann–Poisson system (for the sake of simplicity, only the case of a single valley in the conduction band is considered),

$$\frac{\partial w}{\partial t} + S[\mathcal{E}]w + \frac{q}{m^*}\Theta[V]w = \mathcal{C}[w], \tag{13}$$

$$\mathrm{div}\,(\epsilon_s \nabla_{\mathbf{x}} V_S) = -q(N_d - N_a - n). \tag{14}$$

where m^* is the electron effective mass, the potential V usually is the sum of a self-consistent part V_S, solution of the Poisson equation, and an additional part which models the potential barriers in heterostructures. The unknown function $w(\mathbf{x}, \mathbf{p}, t)$, depending on the position $\mathbf{x}$, crystal momentum $\mathbf{p} = \hbar\mathbf{k}$ and time t, is the Wigner quasi-distribution, defined as

$$\begin{aligned} w(\mathbf{x}, \mathbf{p}, t) = \quad & \mathcal{F}^{-1}[\rho(\mathbf{x} + \frac{\hbar}{2m^*}\boldsymbol{\eta}, \mathbf{x} - \frac{\hbar}{2m^*}\boldsymbol{\eta}, t)](\mathbf{p}) \\ = \quad & \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} \rho(\mathbf{x} + \frac{\hbar}{2m^*}\boldsymbol{\eta}, \mathbf{x} - \frac{\hbar}{2m^*}\boldsymbol{\eta}, t)\, e^{i\mathbf{p}\cdot\boldsymbol{\eta}}\, d\boldsymbol{\eta}. \end{aligned} \tag{15}$$

Here, $\rho(\mathbf{x}, \mathbf{y}, t)$ is the density matrix, which is related to the wave function $\psi(\mathbf{x}, t)$ by

$$\rho(\mathbf{x}, \mathbf{y}, t) = \overline{\psi(\mathbf{x}, t)}\, \psi(\mathbf{y}, t). \tag{16}$$

The operator $\mathcal{F}$ denotes the Fourier transform, given for a function $g(\mathbf{p}) \in L^1(\mathbb{R}^3)$ by

$$\mathcal{F}[g(\mathbf{p})](\boldsymbol{\eta}) = \int_{\mathbb{R}^3_{\mathbf{p}}} g(\mathbf{p})\, e^{-i\mathbf{p}\cdot\boldsymbol{\eta}}\, d\,\mathbf{p},$$

with i the imaginary unit, and $\mathcal{F}^{-1}$ denotes the inverse Fourier transform

$$\mathcal{F}^{-1}[h(\boldsymbol{\eta})] = \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3_{\boldsymbol{\eta}}} h(\boldsymbol{\eta})\, e^{i\mathbf{p}\cdot\boldsymbol{\eta}}\, d\,\boldsymbol{\eta}.$$

The terms $S[\mathcal{E}]$ and $\Theta[V]$ represent the pseudo-differential operators

$$\begin{aligned}
S[\mathcal{E}]w(\mathbf{x},\mathbf{p},t) &= \frac{i}{\hbar(2\pi)^3} \int_{\mathbb{R}^3_{\mathbf{x}'}\times\mathbb{R}^3_{\boldsymbol{\nu}}} \left[\mathcal{E}\left(\mathbf{p}+\frac{\hbar}{2}\boldsymbol{\nu},t\right) - \mathcal{E}\left(\mathbf{p}+\frac{\hbar}{2}\boldsymbol{\nu},t\right) \right] \\
&\quad w(\mathbf{x}',\mathbf{p},t)\, e^{-i(\mathbf{x}'-\mathbf{x})\cdot\boldsymbol{\nu}}\, d\,\mathbf{x}'\, d\,\boldsymbol{\nu} \\
\Theta[V]w(\mathbf{x},\mathbf{p},t) &= \frac{i\, m^*}{\hbar(2\pi)^3} \int_{\mathbb{R}^3_{\mathbf{p}'}\times\mathbb{R}^3_{\boldsymbol{\eta}}} \left[V\left(\mathbf{x}+\frac{\hbar}{2m^*}\boldsymbol{\eta},t\right) - V\left(\mathbf{x}-\frac{\hbar}{2m^*}\boldsymbol{\eta},t\right) \right] \\
&\quad w(\mathbf{x},\mathbf{p}',t)\, e^{-i(\mathbf{p}'-\mathbf{p})\cdot\boldsymbol{\eta}}\, d\,\mathbf{p}'\, d\,\boldsymbol{\eta}.
\end{aligned}$$

As is well known, $w(\mathbf{x},\mathbf{p},t)$ is not in general positive definite. However, it is possible to calculate the macroscopic quantities of interest as expectation values (moments) of $w(\mathbf{x},\mathbf{p},t)$ in the same way as in the semiclassical case, e.g.,

$$n(\mathbf{x},t) = \int_{\mathbb{R}^3} w(\mathbf{x},\mathbf{p},t)\, d\,\mathbf{p} \quad \text{(density)}, \tag{17}$$

$$\mathbf{V}(\mathbf{x},t) = \frac{1}{n(\mathbf{x},t)} \int_{\mathbb{R}^3} \mathbf{v}\, w(\mathbf{x},\mathbf{p},t)\, d\,\mathbf{p} \quad \text{(velocity)}, \tag{18}$$

$$W(\mathbf{x},t) = \frac{1}{n(\mathbf{x},t)} \int_{\mathbb{R}^3} \mathcal{E}(\mathbf{p})\, w(\mathbf{x},\mathbf{p},t)\, d\,\mathbf{p} \quad \text{(energy)}, \tag{19}$$

$$\mathbf{S}(\mathbf{x},t) = \frac{1}{n(\mathbf{x},t)} \int_{\mathbb{R}^3} \mathcal{E}(\mathbf{p})\, \mathbf{v}\, w(\mathbf{x},\mathbf{p},t)\, d\,\mathbf{p} \quad \text{(energy-flux)}. \tag{20}$$

The operator $\mathcal{C}[w]$ represents the quantum collision term. Its formulation is itself an open problem. Some attempts can be found in [36,37], but a derivation suitable for applications in electron devices is still lacking. In [15], an expression has been proposed which is a perturbation of the semiclassical collision term, useful for the formulation of macroscopic models.

As a general guideline, $\mathcal{C}[w]$ should drive the system towards the equilibrium. Let us consider the electrons in a thermal bath at the lattice temperature $T_L = 1/k_B\beta$. The equilibrium Wigner function w_{eq} for an arbitrary energy band, up to $\hbar^2$ terms, reads [38,39]

$$w_{eq}(\mathbf{x},\mathbf{p},\beta) = \frac{n(\mathbf{x},t)e^{-\beta\mathcal{E}}}{A_0(\beta,m^*)} \left\{ 1 + \hbar^2 \left[\left(\frac{q\beta^2}{8}\frac{\partial^2 V}{\partial x_i\partial x_j} + \frac{q^2\beta^3}{24}\frac{\partial V}{\partial x_i}\frac{\partial V}{\partial x_j} \right) \right.\right.$$
$$\left.\left. \left(\frac{\partial^2\mathcal{E}}{\partial p_i\partial p_j} - \frac{A_{ij}(\beta,m^*)}{A_0(\beta,m^*)} \right) - \frac{q\beta^3}{24}\frac{\partial^2 V}{\partial x_i\partial x_j}\left(v_iv_j - \frac{B_{ij}(\beta,m^*)}{A_0(\beta,m^*)} \right) \right] \right\} + \mathcal{O}(\hbar^4),$$

where

$$\begin{aligned}
A_0(\beta,m^*) &= \int_{\mathbb{R}^3} e^{-\beta\mathcal{E}}\, d\,\mathbf{p}, \quad A_{ij}(\beta,m^*) = \int_{\mathbb{R}^3} e^{-\beta\mathcal{E}}\frac{\partial^2\mathcal{E}}{\partial p_i\partial p_j}\, d\,\mathbf{p}, \\
B_{ij}(\beta,m^*) &= \int_{\mathbb{R}^3} e^{-\beta\mathcal{E}} v_iv_j d\,\mathbf{p}.
\end{aligned}$$

We suppose that the expansion

$$w = w^{(0)} + \hbar^2\, w^{(1)} + \mathcal{O}(\hbar^4) \tag{21}$$

holds. By proceeding in a formal way, as $\hbar \to 0$ the Wigner equation gives the semiclassical Boltzmann equation

$$\frac{\partial w^{(0)}}{\partial t} + \mathbf{v}\cdot\nabla_{\mathbf{x}}w^{(0)} + \frac{q}{m^*}\nabla_{\mathbf{x}}V\cdot\nabla_{\mathbf{v}}w^{(0)} = \mathcal{C}^{(0)}[w^{(0)}] \tag{22}$$

While, at the first order in $\hbar^2$, we have (Einstein's convention is used: summation with respect to repeated dummy indices is understood.)

$$\begin{aligned}
\frac{\partial w^{(1)}}{\partial t} &+ \mathbf{v}\cdot\nabla_{\mathbf{x}}w^{(1)} - \frac{1}{24}\frac{\partial^3 \mathcal{E}}{\partial p_i\partial p_j\partial p_k}\frac{\partial^3 w^{(1)}}{\partial x_i\partial x_j\partial x_k} \\
&+ \frac{q}{m^*}\nabla_{\mathbf{x}}V\cdot\nabla_{\mathbf{v}}w^{(1)} - \frac{q}{24 m^{*3}}\frac{\partial^3 V}{\partial x_i\,\partial x_j\,\partial x_k}\frac{\partial^3 w^{(0)}}{\partial v_i\,\partial v_j\,\partial v_k} = \mathcal{C}^{(1)}
\end{aligned} \tag{23}$$

with $\mathcal{C}^{(1)}$ to be modeled. We make the following

Assumption 1.

$$\begin{aligned}
\mathcal{C}[w] &= \mathcal{C}^{(0)}[w^{(0)}] + \hbar^2\,\mathcal{C}^{(1)}[w^{(1)}] = \mathcal{C}_C[w^{(0)}] - \hbar^2\nu\left(w^{(1)} - w^{(1)}_{eq}\right) + \mathcal{O}(\hbar^4), \\
&\quad \text{with} \quad \mathcal{C}_C[w^{(0)}] \quad \text{the semiclassical collision operator} \\
&\quad \text{and} \quad \nu > 0 \quad \text{quantum collision frequency}
\end{aligned} \tag{24}$$

Remark 1. *At variance with other approaches, only the $\hbar^2$ correction to the collision term has a relaxation form. This assures that as $\hbar \to 0$ one gets the semiclassical scattering of electrons with phonons and impurities. We note that $w^{(0)} > 0$ and therefore the semiclassical expression of the collision term makes sense.*

The value of the quantum collision frequency ν is a fitting parameter that can be determined by comparing the results with the experimental data.

We require that $\mathcal{C}[w]$ conserves the electron density, that is we make the following

Basic Assumption 2.

$$\int_{\mathbb{R}^3}\mathcal{C}[w]\,d\mathbf{p} = 0. \tag{25}$$

Proposition 1. *The collision operator $\mathcal{C}[w]$ of the form (24) satisfies up to terms $\mathcal{O}(\hbar^4)$, the following properties:*

1. *Ker $(\mathcal{C}[w])$ is given by the quantum Maxwellian*

$$w_{(eq)} = w^{(0)}_{eq} + \hbar^2 w^{(1)}_{eq},$$

with $w^{(0)}_{eq}$ the classical Maxwellian.

2.

$$-k_B\int_{\mathbb{R}^3}\mathcal{C}^{(0)}[w^{(0)}]\ln\frac{w^{(0)}}{\exp(-\frac{\beta\,m^*|\mathbf{v}|^2}{2})}\,d\mathbf{p} = -k_B\int_{\mathbb{R}^3}\left[\ln w^{(0)} + \frac{\beta m^*|\mathbf{v}|^2}{2}\right]\mathcal{C}^{(0)}\,d\mathbf{p} \geq 0,$$

3.

$$-\frac{1}{2}\mathcal{C}^{(1)}[w^{(1)}]\left(w^{(1)} - w^{(1)}_{eq}\right) \geq 0$$

Moreover, the equality holds if and only if w is the quantum Maxwellian, defined above.

Properties 1 and 3 are straightforward. Property 2 is based on the proof in [40–42] valid in the classical case.

3.1. Quantum Corrections in the High Field Approximation

In the case of high electric fields and effective mass approximation for the dispersion relation

$$\mathcal{E} = \frac{|\mathbf{p}|^2}{2m^*}, \quad \mathbf{v} = \frac{\mathbf{p}}{m^*},$$

it is possible to find an approximation for $w^{(1)}$ by a suitable Chapman–Enskog expansion. Let us introduce the dimensionless variables

$$\tilde{x} = \frac{|\mathbf{x}|}{l_0}, \quad \tilde{t} = \frac{t}{t_0}, \quad \tilde{v} = \frac{|\mathbf{v}|}{v_0},$$

with l_0, t_0 and $v_0 = l_0/t_0$ the typical length, time and velocity. Let l_V be the characteristic length of the electrical potential and $1/t_C$ the characteristic collision frequency. After scaling the collision frequency according to $\tilde{\nu} = t_C \, \nu$, Equation (23) can be rewritten as

$$\frac{1}{t_0}\frac{\partial w^{(1)}}{\partial t} + \frac{v_0}{l_0}\mathbf{v}\cdot\nabla_{\mathbf{x}}w^{(1)} \; + \; \frac{v_0}{l_V}\left[\frac{q}{m^*}\nabla_{\mathbf{x}}V\cdot\nabla_{\mathbf{v}}w^{(1)} - \frac{q}{24m^{*3}}\frac{\partial^3 V}{\partial x_i\,\partial x_j\,\partial x_k}\frac{\partial^3 w^{(0)}}{\partial v_i\,\partial v_j\,\partial v_k}\right]$$
$$= \; -\frac{1}{t_C}\nu\left(w^{(1)} - w^{(1)}_{eq}\right).$$

We will continue to denote the scaled variables as the unscaled ones in order to simplify the notation.

Let us introduce the characteristic length associated with the quantum correction of the collision term (a kind of mean free path in a semiclassical context)

$$l_C = v_0 \, t_C.$$

Let us assume that the quantum effects occur in the high field and collision dominated regime, where drift and collision mechanisms have the same characteristic length. Therefore, we set formally

$$\frac{l_C}{l_V} = 1$$

and observe that, in the high frequency regime, the Knudsen number

$$\alpha = \frac{l_C}{l_0}$$

is a small parameter. Substituting it in the previous equation, we get

$$\alpha\frac{\partial w^{(1)}}{\partial t} + \alpha\mathbf{v}\cdot\nabla_{\mathbf{x}}w^{(1)} + \frac{q}{m^*}\nabla_{\mathbf{x}}V\cdot\nabla_{\mathbf{v}}w^{(1)} \; - \; \frac{q}{24m^{*3}}\frac{\partial^3 V}{\partial x_i\,\partial x_j\,\partial x_k}\frac{\partial^3 w^{(0)}}{\partial v_i\,\partial v_j\,\partial v_k} \qquad (26)$$
$$= \; -\nu\left(w^{(1)} - w^{(1)}_{eq}\right).$$

The zero order in α gives

$$\frac{q}{m^*}\nabla_{\mathbf{x}}V\cdot\nabla_{\mathbf{v}}w^{(1)} - \frac{q}{24m^{*3}}\frac{\partial^3 V}{\partial x_i\,\partial x_j\,\partial x_k}\frac{\partial^3 w^{(0)}}{\partial v_i\partial v_j\,\partial v_k} = -\nu\left(w^{(1)} - w^{(1)}_{eq}\right), \qquad (27)$$

and by Fourier transforming, one has

$$w^{(1)}(\mathbf{x}, \mathbf{v}, t) = \mathcal{F}^{-1} \left\{ \frac{1}{\nu + \frac{iq}{m^*} \boldsymbol{\eta} \cdot \nabla_{\mathbf{x}} V} \left[-\frac{iq}{24m^{*3}} \frac{\partial^3 V}{\partial x_i \partial x_j \partial x_k} \eta_i \eta_j \eta_k \mathcal{F} w^{(0)}(\boldsymbol{\eta}) \right. \right.$$

$$\left. \left. + \nu \mathcal{F} w_{eq}^{(1)}(\boldsymbol{\eta}) \right] \right\} (\mathbf{x}, \mathbf{v}, t). \tag{28}$$

Approximating $w^{(0)}$ with MEP distribution function, we obtain

$$w(\mathbf{x}, \mathbf{v}, t) \approx f_{ME}(\mathbf{x}, \mathbf{v}, t) + \hbar^2 w^{(1)}(\mathbf{x}, \mathbf{v}, t). \tag{29}$$

For example, in the 8-moment case, using the f_{ME} found in [10], the following explicit approximation for the Wigner function is obtained

$$w(\mathbf{x}, \mathbf{v}, t) \approx \frac{n \exp\left(-\frac{3m^*|\mathbf{v}|^2}{4W}\right)}{\left(\frac{4}{3}\pi m^* W\right)^{3/2}} \left\{ 1 - \left[\left(-\frac{21m^*}{4W} + \mathcal{E} \frac{9m^*}{4W^2} \right) \mathbf{V}^{(0)} + \left(\frac{9m^*}{4W^2} - \mathcal{E} \frac{27m^*}{20W^3} \right) \mathbf{S}^{(0)} \right] \cdot \mathbf{v} \right\}$$

$$+ \hbar^2 \mathcal{F}^{-1} \left\{ \frac{1}{\nu + \frac{iq}{m^*} \boldsymbol{\eta} \cdot \nabla_{\mathbf{x}} V} \left[-\frac{iq}{24m^{*3}} \frac{\partial^3 V}{\partial x_i \partial x_j \partial x_k} \eta_i \eta_j \eta_k \mathcal{F} w^{(0)}(\boldsymbol{\eta}) + \nu \mathcal{F} w_{eq}^{(1)}(\boldsymbol{\eta}) \right] \right\} (\mathbf{x}, \mathbf{v}, t) \tag{30}$$

which can be used for evaluating the unknown quantities in the moment system, associated with the Wigner equation. The vectors $\mathbf{V}^{(0)}$ and $\mathbf{S}^{(0)}$ represent the velocity and the energy-flux at zero order in $\hbar^2$.

3.2. The Quantum Moment Equations

In analogy with the semiclassical case, multiplying (13) by suitable weight functions ψ, depending, in the physically relevant cases, on the velocity $\mathbf{v}$, and integrating over the velocity space, one has the balance equations for the macroscopic quantities of interest

$$\frac{\partial}{\partial t} \int_{\mathbb{R}^3} w(\mathbf{x}, \mathbf{v}, t) \, \psi(\mathbf{v}) \, d\mathbf{v} + \nabla_{\mathbf{x}} \cdot \int_{\mathbb{R}^3} \psi(\mathbf{v}) \mathbf{v} w \, d\mathbf{v} + \frac{q}{m^*} \int_{\mathbb{R}^3} \psi(\mathbf{v}) \, \Theta[V] w \, d\mathbf{v}$$

$$= \int_{\mathbb{R}^3} \psi(\mathbf{v}) \, \mathcal{C}[w] \, d\mathbf{v}. \tag{31}$$

In the 8-moment model, the basic variables are the macroscopic density, velocity, energy and energy-flux, that are the moments relative to the weight functions 1, $\mathbf{v}$, $\frac{1}{2} m^* |\mathbf{v}|^2$, $\frac{1}{2} m^* |\mathbf{v}|^2 \mathbf{v}$.

By evaluating (31) for $\psi = 1$, under the assumption that the necessary moments of $w^{(1)}(\mathbf{x}, \mathbf{v}, t)$ and $\frac{\partial^3 w^{(0)}}{\partial v_i v_j v_k}$ with respect to $\mathbf{v}$ exist, one has

$$\frac{q}{m^*} \int_{\mathbb{R}^3} \Theta[V] w \, d\mathbf{v} = \frac{q}{m^*} \nabla_{\mathbf{x}} \cdot \int_{\mathbb{R}^3} \nabla_{\mathbf{v}} w^{(0)} \, d\mathbf{v} + \hbar^2 \left[\frac{q}{m^*} \nabla_{\mathbf{x}} V \cdot \int_{\mathbb{R}^3} \nabla_{\mathbf{v}} w^{(1)} \, d\mathbf{v} \right.$$

$$\left. - \frac{q}{24m^3} \frac{\partial^3 V}{\partial x_i \partial x_j \partial x_k} \int_{\mathbb{R}^3} \frac{\partial^3 w^{(0)}}{\partial v_i \partial v_j \partial v_k} \, d\mathbf{v} \right] = 0,$$

obtaining the continuity equation

$$\frac{\partial}{\partial t} n + \frac{\partial(n V_i)}{\partial x^i} = 0. \tag{32}$$

In order to get the other moment equations, we observe that from (27) it follows that

$$\frac{q}{m^*}\nabla_{\mathbf{x}}V \cdot \int_{\mathbb{R}^3} \psi(\mathbf{v})\,\nabla_{\mathbf{v}}w^{(1)}\,d\,\mathbf{v} - \frac{q}{24m^3}\frac{\partial^3 V}{\partial x_i\,\partial x_j\,\partial x_k}\int_{\mathbb{R}^3}\psi(\mathbf{v})\frac{\partial^3 w^{(0)}}{\partial v_i\,\partial v_j\,\partial v_k}\,d\,\mathbf{v}$$
$$+\nu\int_{\mathbb{R}^3}\psi(\mathbf{v})\left(w^{(1)}-w^{(1)}_{eq}\right)d\,\mathbf{v}=0, \tag{33}$$

for any weight function $\psi(\mathbf{v})$ such that the integrals exist.

By taking into account (33), multiplying Equation (13) by the weight functions $\mathbf{v}$, $\frac{1}{2}m^*\,|\mathbf{v}|^2$, $\frac{1}{2}m^*\,|\mathbf{v}|^2\mathbf{v}$, and after integration, one finds the balance equations for velocity, energy and energy-flux

$$\frac{\partial}{\partial t}(nV_i) + \frac{\partial(nU_{ij})}{\partial x_j} + n\,q\,E_i = nC_{V_i}\left(W^{(0)}, V_i^{(0)}, S_i^{(0)}\right), \tag{34}$$

$$\frac{\partial}{\partial t}(nW) + \frac{\partial(nS_j)}{\partial x_j} + nqV_k^{(0)}E^k = nC_W(W^{(0)}), \tag{35}$$

$$\frac{\partial}{\partial t}(nS_i) + \frac{\partial(nF_{ij})}{\partial x_j} + \frac{5}{3}n\frac{q}{m^*}E_iW^{(0)} = nC_{S_i}\left(W^{(0)}, V_i^{(0)}, S_i^{(0)}\right). \tag{36}$$

Here, $V_i^{(0)}$, $W^{(0)}$ and $P_i^{(0)}$ are the zero-order components of the average velocity, energy and energy-flux. The components of the flux of the velocity and the flux of the energy-flux are defined as

$$U_{ij} \;=\; \frac{1}{n(\mathbf{x},t)}\int_{\mathbb{R}^3}v_i\,v_j\,w(\mathbf{x},\mathbf{v},t)\,d\,\mathbf{v}, \tag{37}$$

$$F_{ij} \;=\; \frac{1}{n(\mathbf{x},t)}\int_{\mathbb{R}^3}\frac{1}{2}m^*v_i\,v_j\,|\mathbf{v}|^2\,w(\mathbf{x},\mathbf{v},t)\,d\,\mathbf{v}. \tag{38}$$

The production terms are defined as

$$n\,C_{V_i} \;=\; \int_{\mathbb{R}^3}v_i\,\mathcal{C}[w]\,d\,\mathbf{v}, \tag{39}$$

$$n\,C_W \;=\; \int_{\mathbb{R}^3}\frac{1}{2}m^*v^2\,\mathcal{C}[w]\,d\,\mathbf{v}, \tag{40}$$

$$n\,C_{S_i} \;=\; \int_{\mathbb{R}^3}\frac{1}{2}m^*|\mathbf{v}|^2\,v_i\,\mathcal{C}[w]\,d\,\mathbf{v}. \tag{41}$$

Remark 2. *The quantum corrections affect only the free streaming part, while the drift and production terms appear only at the zero order.*

Therefore, $C_W(W^{(0)})$, $C_{V_i}\left(W^{(0)}, V_i^{(0)}, S_i^{(0)}\right)$ and $C_{S_i}\left(W^{(0)}, V_i^{(0)}, S_i^{(0)}\right)$ are the same as in the semiclassical case.

In the system (32), (34)–(36) are not closed for the presence of the unknown quantities U_{ij}, F_{ij}, C_{V_i}, C_W and C_{S_i}. We solve the closure problem with the approximation (29), assuming a collision dominated high field regime for the quantum effects.

In order to evaluate the unknown quantities present in the moment equations, the following formal lemmas are useful

Lemma 1.

$$\int_{\mathbb{R}^3}v_{i_1}\cdots v_{i_k}\,w^{(1)}\,d\,\mathbf{v} = \left[i^k\frac{\partial}{\partial \eta_{i_1}}\cdots\frac{\partial}{\partial \eta_{i_k}}\mathcal{F}w^{(1)}(\eta)\right]_{\eta=0}. \tag{42}$$

Lemma 2.

$$\mathcal{F}w_{eq}^{(1)}(0) = \frac{\partial}{\partial\eta_{i_k}}\mathcal{F}w_{eq}^{(1)}(0) = \frac{\partial}{\partial\eta_i}\frac{\partial}{\partial\eta_j}\frac{\partial}{\partial\eta_k}\mathcal{F}w_{eq}^{(1)}(0) = 0 \tag{43}$$

$$\frac{\partial}{\partial\eta_i}\frac{\partial}{\partial\eta_j}\mathcal{F}w_{eq}^{(1)}(0) = n(\mathbf{x},t)\,\frac{\hbar^2\beta q}{12m^{*2}}\,\frac{\partial^2 V}{\partial x_i\partial x_j} \tag{44}$$

$$\frac{\partial}{\partial\eta_i}\frac{\partial}{\partial\eta_j}\Delta_\eta\,\mathcal{F}w_{eq}^{(1)}(0) = -n(\mathbf{x},t)\,\frac{\hbar^2 q}{12m^{*3}}\left[\Delta V\,\delta_{ij} + 5\frac{\partial^2 V}{\partial x_i\partial x_j}\right]. \tag{45}$$

The proof follows by a simple computation. With the aid of these lemmas, we get the following closure relations:

Proposition 2.

$$J_i = nV_i = nV_i^{(0)} + \mathcal{O}(\hbar^4), \tag{46}$$

$$W = W^{(0)} - \frac{\hbar^2\beta q}{24m^*}\,\Delta V + \mathcal{O}(\hbar^4), \tag{47}$$

$$U_{ij} = \frac{2}{3}\frac{W^{(0)}}{m^*}\delta_{ij} - \frac{\hbar^2\beta q}{12(m^*)^2}\frac{\partial^2 V}{\partial x_i\partial x_j} + \mathcal{O}(\hbar^4), \tag{48}$$

$$S_i = S_i^{(0)} - \frac{\hbar^2\beta^2 q^2}{24m^{*2}v}\left(2\frac{\partial^2 V}{\partial x_i\partial x_r}\frac{\partial V}{\partial x_r} + \frac{\partial V}{\partial x_i}\Delta V\right) - \frac{\hbar^2 q}{8m^{*2}v}\frac{\partial}{\partial x_i}\Delta V + \mathcal{O}(\hbar^4), \tag{49}$$

$$\begin{aligned}
F_{ij} = {} & \frac{10}{9m^*}(W^{(0)})^2 - \frac{\hbar^2\beta q^3}{3m^{*3}v^2}\frac{\partial V}{\partial x_{(i}}\frac{\partial^2 V}{\partial x_{j)}\partial x_r}\frac{\partial V}{\partial x_r} \\
& -\frac{\hbar^2 q^2}{4m^{*3}v^2}\frac{\partial^3 V}{\partial x_i\partial x_j\partial x_r}\frac{\partial V}{\partial x_r} - \frac{\hbar^2\beta q^3}{12m^{*3}v^2}\left(\frac{\partial V}{\partial x_i}\frac{\partial V}{\partial x_j}\Delta V + |\nabla V|^2\frac{\partial^2 V}{\partial x_i\partial x_j}\right) \\
& -\frac{\hbar^2 q^2}{4m^{*3}v^2}\frac{\partial\Delta V}{\partial x_{(i}}\frac{\partial V}{\partial x_{j)}} - \frac{\hbar^2 q}{24m^{*2}}\left(\Delta V\,\delta_{ij} + 5\frac{\partial^2 V}{\partial x_i\partial x_j}\right) \\
& -\frac{\hbar^2 q}{4m^{*2}v}\left(\frac{\partial\Delta V}{\partial x_{(i}}V_{j)} + \frac{\partial^3 V}{\partial x_{(i}\partial x_j\partial x_{k)}}V_k\right) + \mathcal{O}(\hbar^4).
\end{aligned} \tag{50}$$

In the previous relationships, parentheses indicate symmetrization, e.g.,

$$\begin{aligned}
A_{i(jk)} &= \frac{1}{2}\left(A_{ijk} + A_{ikj}\right), \\
A_{(ijk)} &= \frac{1}{3!}\left(A_{ijk} + A_{ikj} + A_{jik} + A_{jki} + A_{kij} + A_{kji}\right).
\end{aligned}$$

For the proof, see [15].

Remark 3. *The zero order in $\hbar^2$ is the same as that obtained in [10,30].*

Remark 4. *Since*

$$\int_{\mathbb{R}^3} w^{(1)}\,d\mathbf{v} = \int_{\mathbb{R}^3} w_{eq}^{(1)}\,d\mathbf{v} = 0,$$

we have

$$\int_{\mathbb{R}^3}\mathcal{C}[w]\,d\mathbf{v} = 0,$$

that is the assumption (25) is fulfilled.

Remark 5. *In the limit of high frequency $\nu \to \infty$ one has the simplified model*

$$J_i = n V_i = n V_i^{(0)} + \mathcal{O}(\hbar^4), \tag{51}$$

$$W = W^{(0)} - \frac{\hbar^2 \beta q}{24 m^*} \Delta V + \mathcal{O}(\hbar^4) \tag{52}$$

$$U_{ij} = \frac{2}{3} \frac{W^{(0)}}{m^*} \delta_{ij} - \frac{\hbar^2 \beta q}{12(m^*)^2} \frac{\partial^2 V}{\partial x_i \partial x_j} + \mathcal{O}(\hbar^4) \tag{53}$$

$$S_i = S_i^{(0)} + \mathcal{O}(\hbar^4), \tag{54}$$

$$F_{ij} = \frac{10}{9m^*} (W^{(0)})^2 - \frac{\hbar^2 q}{24 m^{*2}} \left(\Delta V \delta_{ij} + 5 \frac{\partial^2 V}{\partial x_i \partial x_j} \right) \tag{55}$$

From Equation (28), one sees that in the limit $\nu \to \infty$, $w^{(1)}$ reduces to the quantum correction of the equilibrium Wigner function $w_{eq}^{(1)}$. The resulting quantum corrections to the tensor U_{ij} are the same as those obtained in [43] by using a shifted Wigner function, but with the semiclassical contribution which also contains a heat flux, not added ad hoc.

Remark 6. *Under the assumption (see [43]) that the equilibrium relation*

$$\nabla V = (q \beta)^{-1} \nabla \ln n + \mathcal{O}(\hbar^2) \tag{56}$$

is valid, one can recover a formulation of the quantum corrections of density gradient type.

The density gradient version is more familiar because the quantum corrections take a form similar to the Bohm potential arising in the Madelung model of quantum fluids in the zero temperature limit. Moreover, some numerical experiments [43] lead to it being considered more robust than the original formulation in terms of the derivatives of the electric potential. In the limit $\nu \to \infty$, the closure relations in the density gradient form explicitly read as

$$J_i = n V_i = n V_i^{(0)} + \mathcal{O}(\hbar^4), \tag{57}$$

$$W = W^{(0)} - \frac{\hbar^2}{24 m^*} \Delta \ln n + \mathcal{O}(\hbar^4), \tag{58}$$

$$U_{ij} = \frac{2}{3} \frac{W^{(0)}}{m^*} \delta_{ij} - \frac{\hbar^2}{12(m^*)^2} \frac{\partial^2 \ln n}{\partial x_i \partial x_j} + \mathcal{O}(\hbar^4), \tag{59}$$

$$S_i = S_i^{(0)} + \mathcal{O}(\hbar^4), \tag{60}$$

$$F_{ij} = \frac{10}{9m^*} n (W^{(0)})^2 - \frac{\hbar^2}{24 m^{*2} \beta} \left(\Delta \ln n \, \delta_{ij} + 5 \frac{\partial^2 \ln n}{\partial x_i \partial x_j} \right). \tag{61}$$

4. Two-Dimensional Electron Gases: The Case of Quantum Confinement

In this section, we consider the case of an ensemble of electrons (the treatment of holes would be analogous) confined along one dimension, called quasi 2D electron gas (2DEG) [16,18–20,44]. This situation arises when the length scale in one (the confined) space direction of the semiconductor device under study is of the order of de Broglie wavelength of electrons, while the nonconfined directions have a much bigger length scale. In other words, electrons are in a quantum regime in the confined direction and exhibit a classical behaviour in the nonconfined ones [16] as shown in Figure 1.
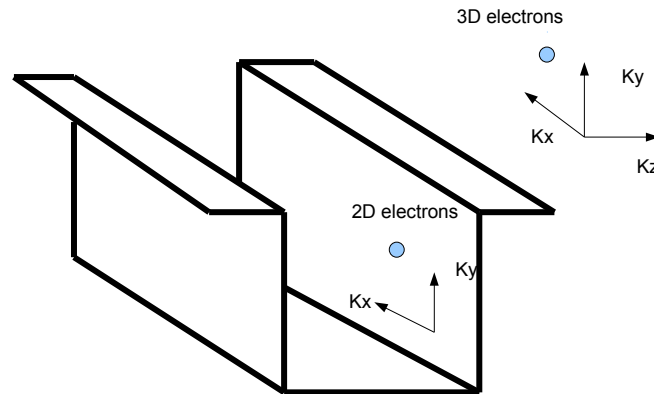
Figure 1. Schematic representation of electron confinement in one dimension due to a potential barrier.

Let us suppose that electrons are quantized along one direction, which we choose as the *z* direction, and free to move in the orthogonal *x*-*y* plane. Let $\Omega = [0, L_x] \times [0, L_y] \times [0, L_z] \subset \mathbb{R}^3$ be the spatial domain. Here, L_x, L_y, $L_z > 0$ are fixed. More general cases with a variable confining direction can be easily incorporated.

Let us introduce the parameter σ defined as the ratio between the transversal (perpendicular to the *x*-*y* plane) typical length scale L_t and the longitudinal typical length scale L_l

$$\sigma = \frac{L_t}{L_l}$$

and assume that $\sigma \ll 1$. Moreover, as is customary, let us assume the following ansatz about the electron wave function

$$\psi(\mathbf{k}, \mathbf{r}) = \psi(k_x, k_y, k_z, x, y, z) = \frac{1}{\sqrt{\mathcal{A}}} \, \varphi(\sigma \, \mathbf{r}_{||}, z) e^{i \mathbf{k}_{||} \cdot \sigma \mathbf{r}_{||}} \quad \text{in} \quad \Omega, \tag{62}$$

with $\mathbf{k}_{||} = (k_x, k_y)$ and $\mathbf{r}_{||} = (x, y)$ denoting the longitudinal components of the wave-vector $\mathbf{k}$ and the position vector $\mathbf{r}$, respectively, and $\mathcal{A}$ being the area of the *x*-*y* cross-section. The previous wave function is inserted into the stationary Schrödinger equation in the effective mass approximation

$$\left[-\frac{\hbar^2}{2m^*} \Delta + E_C(\mathbf{r}) \right] \psi = E \, \psi, \tag{63}$$

where m^* is the electron effective mass, E_C is the conduction band minimum, $E_C = -q(V_C + V)$, with $V_C(z)$ the confining potential and $V(\mathbf{r}_{||}, z)$ the self-consistent electrostatic potential.

Under the assumption that the confining potential is so high that it gives rise to a barrier which can be considered infinite, it enters into the equations only through the boundary conditions

$$\psi(\mathbf{r}_{||}, z) = 0 \quad \text{at} \quad z = 0 \quad \text{and} \quad z = L_z,$$

and $E_C = -q \, V$ is taken.

Therefore, introducing the slowly varying variable $\tilde{\mathbf{r}}_{||} = \sigma \mathbf{r}_{||}$ and inserting (62) into (63), one has

$$-\frac{\hbar^2}{2m^*} \sigma^2 \left[\Delta_{\mathbf{r}_{||}} \varphi(\tilde{\mathbf{r}}_{||}, z) + 2i\mathbf{k}_{||} \cdot \nabla_{\mathbf{r}_{||}} \varphi(\tilde{\mathbf{r}}_{||}, z) - \mathbf{k}_{||}^2 \varphi(\tilde{\mathbf{r}}_{||}, z) \right]$$
$$-\frac{\hbar^2}{2m^*} \frac{\partial}{\partial z^2} \varphi(\tilde{\mathbf{r}}_{||}, z) - qV(\tilde{\mathbf{r}}_{||}, z)\varphi(\tilde{\mathbf{r}}_{||}, z) = E\varphi(\tilde{\mathbf{r}}_{||}, z).$$

In the limit $\sigma \to 0^+$, one formally gets that the envelope function φ must solve the reduced Schrödinger equation

$$\left[-\frac{\hbar^2}{2m^*}\frac{d^2}{dz^2} - qV(\mathbf{r}_{||}, z) \right] \varphi(\sigma \mathbf{r}_{||}, z) = \varepsilon \varphi(\sigma \mathbf{r}_{||}, z), \quad 0 \le z \le L_z, \tag{64}$$

with boundary conditions

$$\varphi = 0 \quad \text{at} \quad z = 0 \quad \text{and} \quad z = L_z. \tag{65}$$

The solution of (64) and (65) depends only parametrically on $\mathbf{r}_{||}$ (and in a slow way), and more in general also on time t if a non-steady-state solution is considered. In fact, electrons as waves achieve equilibrium along the confined direction in a time which is much shorter than the typical transport time, so that one can adopt a quasi-static description along the z-direction. The couple (64) and (65) constitutes a one dimensional Sturm–Liouville problem, which admits a countable set of eigen-pairs (subbands) $(\varphi_\nu(\mathbf{r}_{||}, z), \varepsilon_\nu(\mathbf{r}_{||}))$, $\nu = 1, 2...$, whose eigenvalues are simple and do not cross. The potential V is obtained from the Poisson equation

$$\nabla \cdot (\epsilon_s \nabla V) = -q(N_d(\mathbf{r}) - n), \tag{66}$$

where the electron density n is given by

$$n(\mathbf{r}) = \sum_{\nu=1}^{+\infty} \rho_\nu(\mathbf{r}_{||}) |\varphi_\nu(\mathbf{r}_{||}, z,)|^2,$$

with ρ_ν the (areal) density of electrons of the νth subband and for simplicity an n–doped material has been considered. Of course, the Schrödinger and Poisson equations are coupled and must be solved simultaneously.

For devices with longitudinal characteristic length of a few tens of nanometers, the transport of electrons in the longitudinal direction is semi-classical within a good approximation. Therefore, electrons in each subband can be considered as different populations, the state of each of them being described by a semiclassical distribution function $f_\nu(x, y, k_x, k_y, t)$, and electron transport along the longitudinal direction being determined by adding to the Schrödinger–Poisson system the following system of coupled Boltzmann equations

$$\frac{\partial f_\nu}{\partial t} + \frac{1}{\hbar}\nabla_{\mathbf{k}_{||}}E_\nu \cdot \nabla_{\mathbf{r}_{||}}f_\nu - \frac{q}{\hbar}\mathbf{E}_\nu^{eff} \cdot \nabla_{\mathbf{k}_{||}}f_\nu = \sum_{\mu=1}^{\infty} C_{\nu,\mu}[f_\nu, f_\mu], \quad \nu = 1, 2, \ldots \tag{67}$$

where $\mathbf{E}_\nu^{eff} = \frac{1}{q}\nabla_{\mathbf{r}_{||}}\varepsilon_\nu(\mathbf{r}_{||})$, and the electron energy in each subband E_ν is the sum of a transversal contribution ε_ν and a longitudinal one

$$E_\nu(\mathbf{r}_{||}, \mathbf{k}_{||}) := \varepsilon_\nu(\mathbf{r}_{||}) + \varepsilon_{||}(\mathbf{k}_{||}) = \varepsilon_\nu(\mathbf{r}_{||}) + \frac{1}{2\alpha}\left(\sqrt{1 + 4\alpha\frac{\hbar^2}{2m^*}\left(k_x^2 + k_y^2\right)} - 1 \right),$$

where α is the non-parabolicity parameter. Consequently, the longitudinal electron velocity is

$$\mathbf{v}_{||} = \frac{1}{\hbar}\nabla_{\mathbf{k}_{||}}\varepsilon_{||} = \frac{\hbar\mathbf{k}_{||}}{m^*\left(1 + 2\alpha\varepsilon_{||}\right)}. \tag{68}$$

A formal justification of Equation (67) can be found in [45]. Regarding the collision operator, in the non-degenerate approximation, each contribution has the general form [18–20]

$$C_{\nu,\mu}[f_\nu, f_\mu] = \frac{1}{(2\pi)^2} \int_{B_2} \left[w_{\mu\nu}(\mathbf{k}'_{||}, \mathbf{k}_{||})\, f'_\mu - w_{\nu\mu}(\mathbf{k}_{||}, \mathbf{k}'_{||})\, f_\nu \right] d^2\mathbf{k}'_{||},$$

where B_2 is the two-dimensional Brillouin zone. When $\mu = \nu$, the scattering is intra-subband; when $\mu \neq \nu$, it is the inter-subband. We recall that inter-subband transitions may also happen in the absence of phonons, by quantum mechanical tunneling induced by the presence of a static field as has been shown for example in [46]. The term $w_{\mu\nu}(\mathbf{k}'_{||}, \mathbf{k}_{||})$ is the transition rate from the longitudinal state with wave-vector $\mathbf{k}'_{||}$, belonging to the μth subband, to the longitudinal state with wave-vector $\mathbf{k}_{||}$, belonging to the νth subband, and $f'_\mu \equiv f_\mu(\mathbf{r}_{||}, \mathbf{k}'_{||}, t)$. In Si, the relevant 2D scattering mechanisms are the acoustic phonon scattering, and the non-polar optical phonon scattering. Their expressions are the 2D version of those shown in Section 2 and can be found in [18–20,47]. Attempts to directly solve the Schrödinger–Poisson–Boltzmann system can be found, for example, in [16,17,48] where numerical schemes based on finite differences have been employed. However, the direct numerical approach is a daunting computational task and requires very long computing times. This has again prompted the development of simpler macroscopic models for CAD purposes. These models can be obtained as moment equations of the Boltzmann transport equations under suitable closure relations [18–20]. The moment of the ν-th subband distribution with respect to a weight function $a(\mathbf{k}_{||})$ reads

$$M_{a,\nu} = \int_{B_2} a(\mathbf{k}_{||}) f_\nu(\mathbf{r}_{||}, \mathbf{k}_{||}, t)\, d^2\mathbf{k}_{||}.$$

In particular, analogously to the previous cases, we take as basic moments

the areal density $\quad \rho_\nu = \displaystyle\int_{B_2} f_\nu(\mathbf{r}_{||}, \mathbf{k}_{||}, t)\, d^2\mathbf{k}_{||},$　　　　　　　　(69)

the longitudinal mean velocity $\quad \mathbf{V}_\nu = \dfrac{1}{\rho_\nu} \displaystyle\int_{B_2} \mathbf{v}_{||} f_\nu(\mathbf{r}_{||}, \mathbf{k}_{||}, t)\, d^2\mathbf{k}_{||},$　　　　(70)

the longitudinal mean energy $\quad W_\nu = \dfrac{1}{\rho_\nu} \displaystyle\int_{B_2} \varepsilon_{||} f_\nu(\mathbf{r}_{||}, \mathbf{k}_{||}, t)\, d^2\mathbf{k}_{||},$　　　　(71)

the longitudinal mean energy-flux $\quad \mathbf{S}_\nu = \dfrac{1}{\rho_\nu} \displaystyle\int_{B_2} \varepsilon_{||} \mathbf{v}_{||} f_\nu(\mathbf{r}_{||}, \mathbf{k}_{||}, t)\, d^2\mathbf{k}_{||}.$　　　(72)

The corresponding moment system reads

$$\frac{\partial \rho_\nu}{\partial t} + \nabla_{\mathbf{r}_{||}} \cdot (\rho_\nu \mathbf{V}_\nu) \;=\; \rho_\nu \sum_\mu C_{\rho,\nu\mu}, \tag{73}$$

$$\frac{\partial (\rho_\nu \mathbf{V}_\nu)}{\partial t} + \nabla_{\mathbf{r}_{||}} \cdot (\rho_\nu \mathbf{F}_\nu^{(0)}) + \rho_\nu \mathbf{G}_\nu^{(0)} \nabla_{\mathbf{r}_{||}} \varepsilon_\nu \;=\; \rho_\nu \sum_\mu C_{\mathbf{V},\nu\mu}, \tag{74}$$

$$\frac{\partial (\rho_\nu W_\nu)}{\partial t} + \nabla_{\mathbf{r}_{||}} \cdot (\rho_\nu \mathbf{S}_\nu) + \rho_\nu \nabla_{\mathbf{r}_{||}} \varepsilon_\nu \cdot \mathbf{V}_\nu \;=\; \rho_\nu \sum_\mu C_{W,\nu\mu}, \tag{75}$$

$$\frac{\partial (\rho_\nu \mathbf{S}_\nu)}{\partial t} + \nabla_{\mathbf{r}_{||}} \cdot (\rho_\nu \mathbf{F}_\nu^{(1)}) + \rho_\nu \mathbf{G}_\nu^{(1)} \nabla_{\mathbf{r}_{||}} \varepsilon_\nu \;=\; \rho_\nu \sum_\mu C_{\mathbf{S},\nu\mu}, \tag{76}$$

where

$$
\begin{pmatrix} \mathbf{F}_\nu^{(0)} \\ \mathbf{F}_\nu^{(1)} \end{pmatrix} = \frac{1}{\rho_\nu} \int_{B_2} \begin{pmatrix} 1 \\ \varepsilon_{||} \end{pmatrix} \mathbf{v}_{||} \otimes \mathbf{v}_{||} f_\nu(\mathbf{r}_{||}, \mathbf{k}_{||}, t) \, d^2\mathbf{k}_{||},
\tag{77}
$$

$$
\begin{pmatrix} \mathbf{G}_\nu^{(0)} \\ \mathbf{G}_\nu^{(1)} \end{pmatrix} = \frac{1}{\rho_\nu} \int_{B_2} \begin{pmatrix} \frac{1}{\hbar}\nabla_{\mathbf{k}_{||}} \mathbf{v}_{||} \\ \frac{1}{\hbar}\nabla_{\mathbf{k}_{||}}\left(\varepsilon_{||}\mathbf{v}_{||}\right) \end{pmatrix} f_\nu(\mathbf{r}_{||}, \mathbf{k}_{||}, t) \, d^2\mathbf{k}_{||},
\tag{78}
$$

$$
\begin{pmatrix} C_{\rho,\nu\mu} \\ C_{W,\nu\mu} \end{pmatrix} = \frac{1}{\rho_\nu} \int_{B_2} \begin{pmatrix} 1 \\ \varepsilon_{||} \end{pmatrix} \left[w_{\mu\nu}(\mathbf{k}'_{||}, \mathbf{k}_{||}) f'_\mu - w_{\nu\mu}(\mathbf{k}_{||}, \mathbf{k}'_{||}) f_\nu \right] d^2\mathbf{k}'_{||} \, d^2\mathbf{k}_{||},
\tag{79}
$$

$$
\begin{pmatrix} C_{\mathbf{V},\nu\mu} \\ C_{\mathbf{S},\nu,\mu} \end{pmatrix} = \frac{1}{\rho_\nu} \int_{B_2} \begin{pmatrix} \mathbf{v}_{||} \\ \varepsilon_{||}\mathbf{v}_{||} \end{pmatrix} \left[w_{\mu\nu}(\mathbf{k}'_{||}, \mathbf{k}_{||}) f'_\mu - w_{\nu\mu}(\mathbf{k}_{||}, \mathbf{k}'_{||}) f_\nu \right] d^2\mathbf{k}'_{||} \, d^2\mathbf{k}_{||}.
\tag{80}
$$

Also in this case, the above written fluxes and production terms are extra-variables for which closure relations are needed. In order to employ MEP, a suitable expression of the entropy for the system under consideration has to be found. Neglecting the mutual interaction of electrons in different subbands, considering the phonon gas as a thermal bath and assuming the electron gas is sufficiently dilute, in each subband the expression of the entropy is the semiclassical limit of that arising from the Fermi statistics. Eventually, for obtaining the total entropy, one has to consider that electrons have a quantum behaviour along the *z*-direction, therefore *we define the entropy density of the system as*

$$
\mathcal{S}[f_1, f_2, \cdots] = -k_B \sum_{\nu=1}^{+\infty} \underbrace{|\varphi_\nu(\mathbf{r}_{||}, z, t)|^2}_{\text{quantum effect}} \underbrace{\int_{B_2} \left(f_\nu \log\left(2\pi^2 f_\nu\right) - f_\nu \right) d^2\mathbf{k}_{||}}_{\text{semiclassical contribution}}.
\tag{81}
$$

Remark 7. *The proposed expression of the entropy has been introduced for the first time in [20,47]. It combines quantum effects and semiclassical transport along the longitudinal direction, weighting the contribution of each subband with the square modulus of the $\varphi_\nu(z, t)$'s arising from the Schrödinger–Poisson block.*

According to MEP, the f_ν's are estimated by means of the distributions f_ν^{MEP}'s that $\forall \mathbf{r}_{||} \in [0, L_x] \times [0, L_y]$, $t > 0$, solve the problem:

$$
\max_{\substack{f_1(\mathbf{r}_{||}, \cdot, t) \in \mathcal{F}_{2D} \\ f_2(\mathbf{r}_{||}, \cdot, t) \in \mathcal{F}_{2D} \\ \cdots}} \mathcal{S}[f_1, f_2, \cdots]
$$

$$
\text{under the constraints} \quad M_{a_A,\nu} = \int_{B_2} a_A(\mathbf{k}_{||}) f_\nu(\mathbf{r}_{||}, \mathbf{k}_{||}, t) d^2\mathbf{k}_{||},
$$

where the $M_{a_A,\nu}$'s are the basic moments (69)–(72) and $\mathcal{F}_{2D}$ is the space of the summable function with respect to $\mathbf{k}_{||}$ such that the moments appearing in the system (73)–(76) exist.

With the above choice of the functions $a_A(\mathbf{k}_{||}) = (1, \mathbf{v}_{||}, \varepsilon_{||}, \varepsilon_{||}\mathbf{v}_{||})$, the resulting maximum entropy distribution functions read (the factor k_B has been included into the multipliers)

$$
f_\nu^{MEP} = \exp\left[-\left(\lambda_\nu + \lambda_{\mathbf{V},\nu} \cdot \mathbf{v}_{||} + \left(\lambda_{W,\nu} + \lambda_{\mathbf{S},\nu} \cdot \mathbf{v}_{||} \right) \varepsilon_{||} \right) \right].
$$

In order to complete the procedure, similarly to the previous case, one has to insert the f_ν^{MEP}'s into the constraints (69)–(72) and express the Lagrange multipliers as functions of the basic moments ρ_ν, $\mathbf{V}_\nu$, W_ν, $\mathbf{S}_\nu$. In general, in the case considered in this section as well as in that considered in Section 2,

it is not possible to assure that such a procedure can be accomplished. The solvability of the MEP problem depends on the band structure and on the choice of the moments, as well known already in gas-dynamics [14,49]. For example, if the parabolic case (where the energy is a quadratic function of the wave vector) is considered, the same drawbacks of classical gas-dynamics arise regarding the lack of integrability of the MEP distribution function. In the case of the Kane dispersion relation, the solvability of the MEP problem is guaranteed by the following property, proven in [49].

Proposition 3. *When the Kane model for the energy band is employed, the corresponding maximum entropy models are symmetric hyperbolic systems with convex domains of definition and the equilibria are interior points, guaranteeing the validity of expansions around equilibrium states.*

The proof has been given for a 3D electron gas, but it can be extended to a 2DEG in a straightforward way.

For the 2DEG case, the final expressions of the extra-fluxes and production terms, obtained by using linearized MEP distribution functions, can be found in [18,47] for the parabolic approximation and in [19,20] for the non-parabolic one. The resulting moment system has been used to simulate double-gate MOSFETs in [18,19].

5. Some Numerical Simulations

In this last section, we present some of the numerical results obtained by implementing the models sketched in the previous sections. The information concerning the numerical schemes will be skipped. The interested readers are referred to the indicated papers. For optimization of electron devices using the models presented above, see [50] and references therein.

5.1. p-n Junction

The full model for holes and electrons, presented in Section 2, has been used for the numerical simulation of a Si p-n junction diode with a doping profile given by [22]

$$N_d(x) - N_a(x) = \begin{cases} -c_0 & \text{if } 0 < x < L_0 \\ c_0 & \text{if } L_0 < x < L, \end{cases}$$

with an abrupt junction at L_0. The physical parameters of the device are reported in Table 1. The numerical solution has been obtained by resorting to an extension [51] of the traditional central differencing scheme for one-dimensional balance laws with (possibly *stiff*) source terms, which has been developed on the basis of the Nessyhau and Tadmor scheme [52] for homogeneous hyperbolic systems.

Table 1. Physical parameters of the p-n junction.

Parameter	Physical Meaning	Numerical Value
L	device length	10^{-3} cm
L_0	length of p-region	0.7×10^{-3} cm
c_0	doping concentration	10^{15} cm^{-3}
n_i	intrinsic concentration	1.075×10^{10} cm^{-3}

The bias voltage at the contacts is the superposition of the thermal equilibrium boundary potential (the built in potential V_{bi}) and the applied potential V_a. In Figures 2 and 3, the stationary carrier densities and the electrostatic potential are shown for the following applied voltages: $V_a = -0.2, 0, 0.75, 1$ V, the sign $+$ referring to direct polarization.
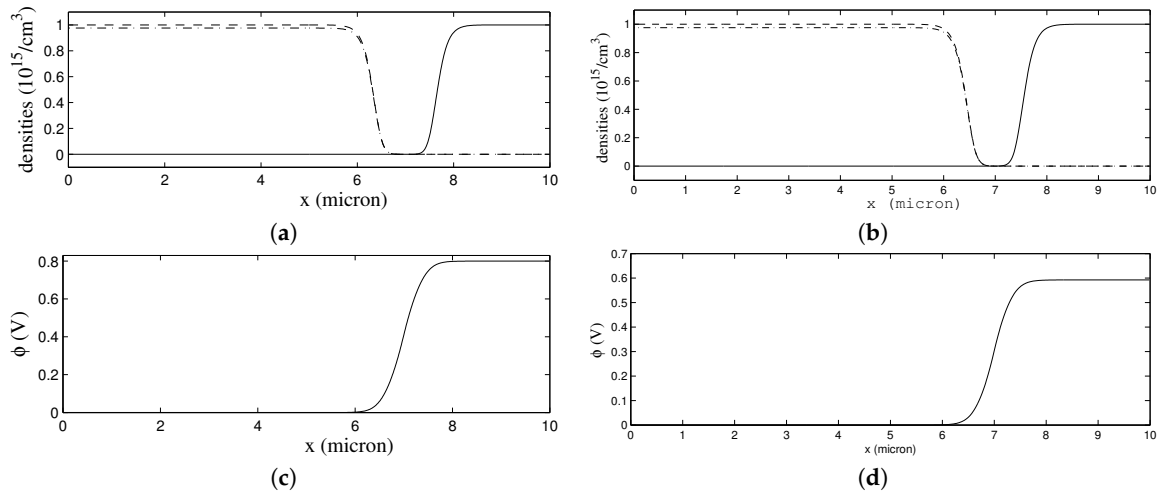
Figure 2. (**a**,**c**) Densities and electric potential for $V_a = -0.2$ V, electrons (continuous line), total holes (dashed line) and heavy holes (dashed–dotted line); (**b**,**d**) Densities and electric potential for $V_a = 0$, electrons (continuous line), total holes (dashed line) and heavy holes (dashed–dotted line).
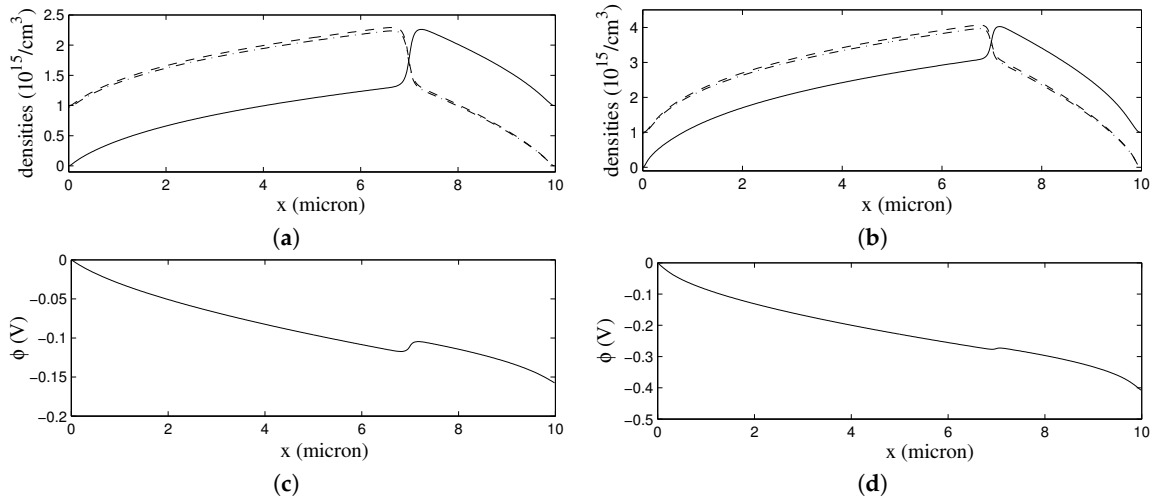
Figure 3. (**a**,**c**) Densities and electric potential for $V_a = 0.75$ V, electrons (continuous line), total holes (dashed line) and heavy holes (dashed–dotted line); (**b**,**d**) Densities and electric potential for $V_a = 1$ V, electrons (continuous line), total holes (dashed line) and heavy holes (dashed–dotted line).

5.2. Simulation of a MESFET

In this subsection, we show the simulation of a bi-dimensional *Metal Semiconductor Field Effect Transistor* (MESFET) made with the MEP model. The numerical method is based on the discretization proposed in [53].

The shape of the device is pictured in Figure 4. The device has a 0.4 μm channel. The source and drain lengths are 0.1 μm and the contact at the gate is 0.2 μm long. The distance between the gate and the other two contacts is 0.1 μm. The lateral subdiffusion of the source and the drain region is about 0.05 μm. The same doping concentration $C(x)$ as in [53] is considered

$$C(x) = \begin{cases} n_+ = 3 \times 10^{17} \, \text{cm}^{-3} & \text{in the } n^+ \text{ regions} \\ n_- = 10^{17} \, \text{cm}^{-3} & \text{in the } n \text{ region} \end{cases}$$

with abrupt junctions. We take a reference frame with axes parallel to the edges of the device. The numerical domain representing the MESFET is

$$\Omega = [0, 0.6] \times [0, 0.2]$$

where the unit length is the micron.
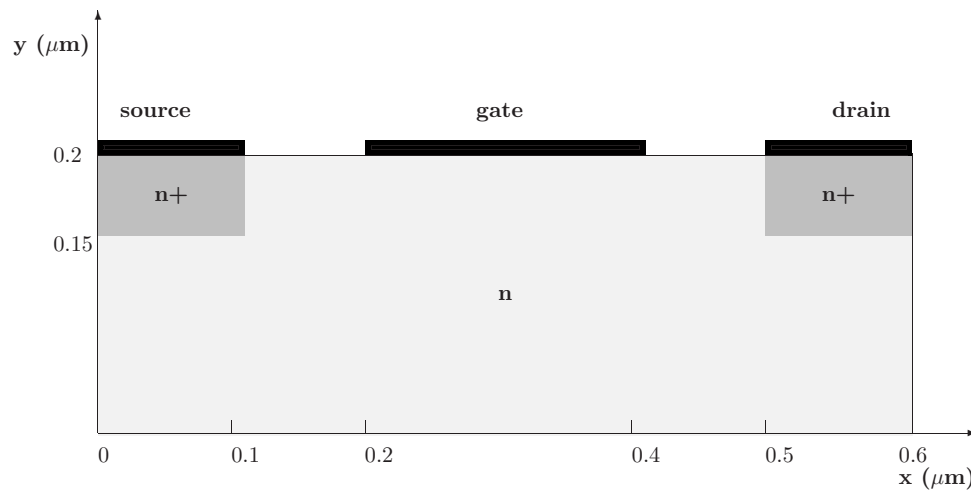
Figure 4. Schematic representation of a bidimensional MESFET.

The regions of high doping n^+ are the subset

$$[0, 0.1] \times [0.15, 0.2] \cup [0.5, 0.6] \times [0.15, 0.2].$$

Γ_D represents the source, gate and drain

$$\Gamma_D = \{(x, y) : y = 0.2, \ 0 \le x \le 0.1, \ 0.2 \le x \le 0.4, \ 0.5 \le x \le 0.6\}.$$

The remaining part of $\partial\Omega$ is Γ_N. The boundary conditions are assigned as follows. We have ohmic contacts at the source and drain:

$$n = n_+, \quad W = \frac{3}{2}k_B T_L, \quad V = \begin{cases} \Phi_b & \text{at drain} \\ 0 & \text{at source.} \end{cases} \tag{82}$$

At the gate we have a Schottky contact

$$n = n_g, \quad W = \frac{3}{2}k_B T_L, \quad V = \Phi_g. \tag{83}$$

Indeed, the potential at the contacts should include the built-in potential and the density at the gate should be related to the potential. Here, we do not enter into the details of the modeling of the Schottky contacts and, by using the invariance of the electric field with respect to changes of the potential for additive constants, we choose as in [54]

$$n_g = 3.9 \times 10^5 \text{ cm}^{-3}, \quad \Phi_g = -0.8 \text{ V}$$

In the remaining part Γ_N of the boundary, we have

$$\boldsymbol{\nu} \cdot \nabla n = 0, \quad \boldsymbol{\nu} \cdot \nabla W = 0, \quad \boldsymbol{\nu} \cdot \nabla V = 0. \tag{84}$$

Here, ∇ is the bi-dimensional gradient operator while ν is the unit outward normal vector to $\partial\Omega$ in the considered points.

A uniform mesh has been used with 33×97 grid points. The stationary solution is reached after about 5 picoseconds. The numerical results are shown in Figures 5–8. The typical depletion region close to the gate with the presence of steep gradients is numerically well described.
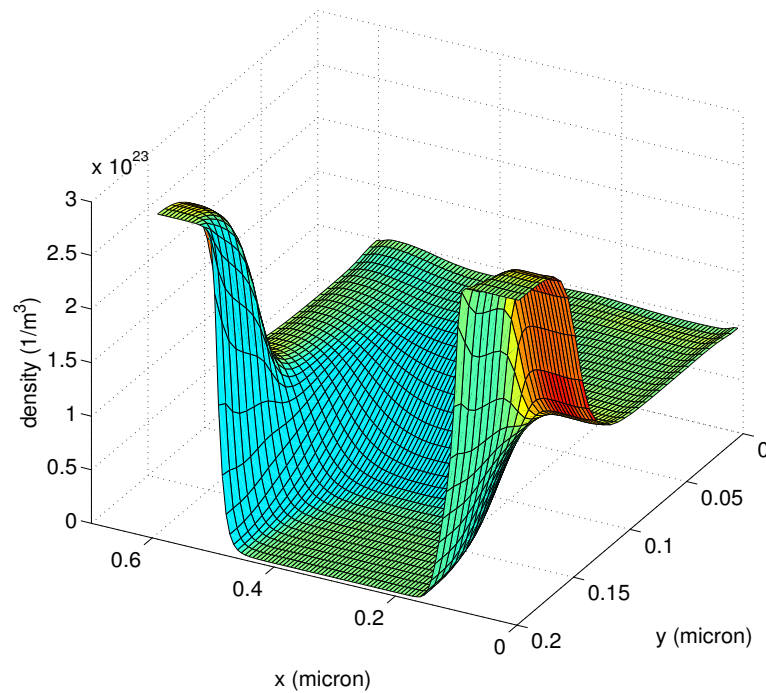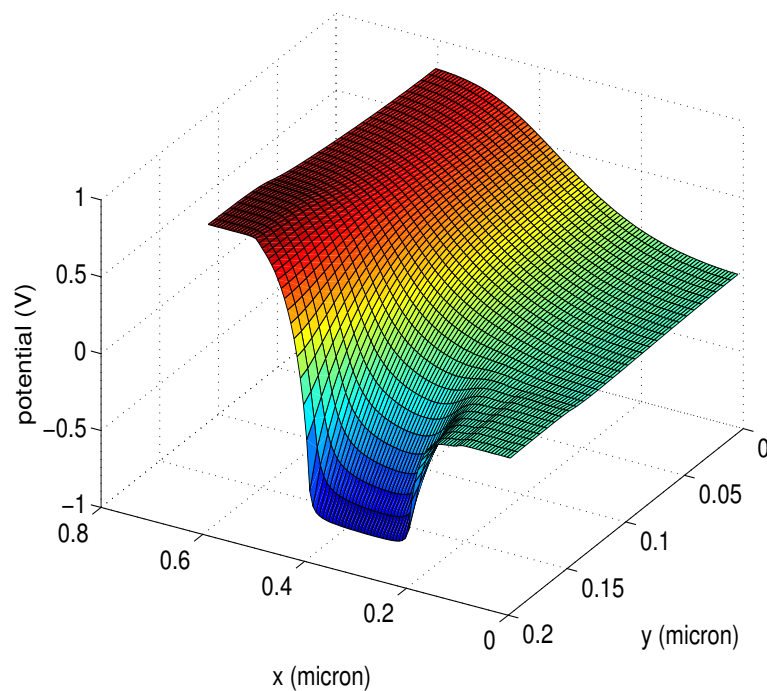

Figure 5. Stationary charge density in the MESFET in the case $\Phi_b = 1$ V.
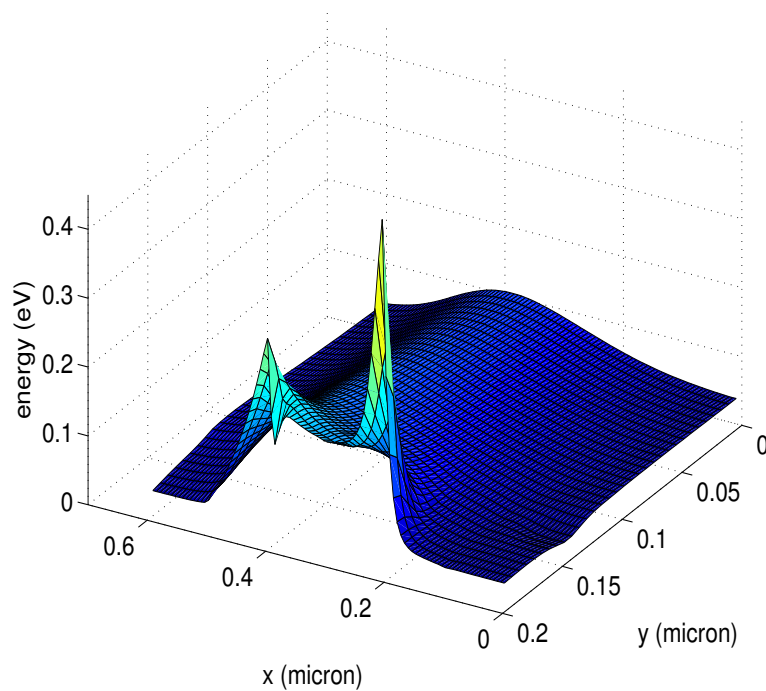
Figure 6. Stationary energy per charge carrier in the MESFET in the case $\Phi_b = 1$ V.
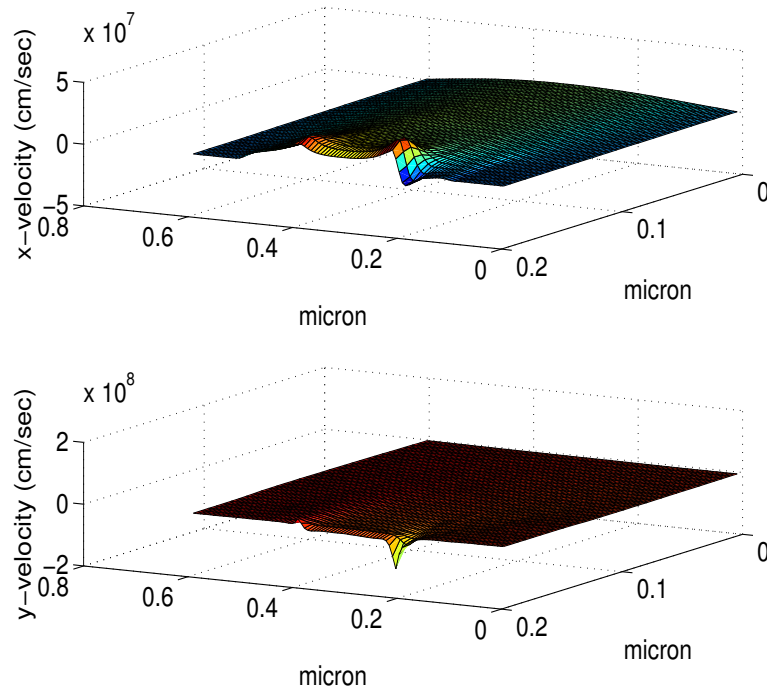
Figure 7. Stationary velocity of the carriers in the MESFET in the case $\Phi_b = 1$ V.

Figure 8. Stationary electric potential in the MESFET in the case $\Phi_b = 1$ V.

5.3. Simulation of a MOSFET

The *Metal Oxide Semiconductor Field Effect Transistor* (MOSFET) is a largely used device in modern electronics. The same mathematical model and numerical scheme as for the MESFET is adopted; the shape of the device is pictured in Figure 9. The device has a 0.2 µm channel. The source and drain lengths are 0.1 µm and the contact at the gate is 0.15 µm long. The distance between the gate and the other two contacts is 0.025 µm. The lateral subdiffusion of the source and the drain region is about

0.05 μm. The gate oxide is 0.15 μm long and 20 nm thick. Of course, smaller devices can be considered without any substantial modification to the used approach.
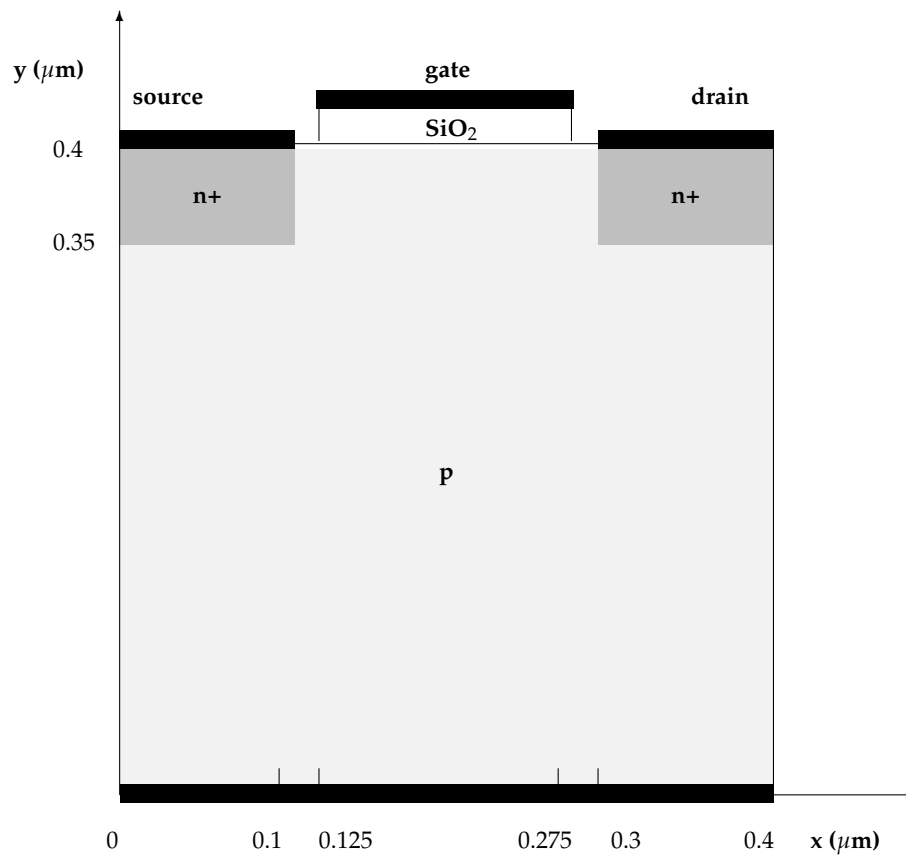
Figure 9. Schematic representation of a bidimensional MOSFET.

The doping concentration is

$$C(x) = \left\{ \begin{array}{l} n_+ = 10^{17}\,\mathrm{cm}^{-3}\ \text{in the } n^+ \text{ regions} \\ p_- = 10^{14}\,\mathrm{cm}^{-3}\ \text{in the } p \text{ region} \end{array} \right.$$

with abrupt junctions.

At variance with MESFET, there are different built-in potentials which we explicitly take into account by using the simple model

$$\Phi_b^D = \Phi_b^S = \frac{k_B T_L}{q} \ln \frac{n_+}{n_i}$$

at drain and source,

$$\Phi_b^B = -\frac{k_B T_L}{q} \ln \frac{p_-}{n_i}$$

at bulk contact. Here, n_i is the intrinsic electron concentration (10^{10} cm^{-3}). The reference axes are chosen parallel to the edges of the device. The silicon part of the MOSFET is represented by the numerical domain

$$[0, 0.4] \times [0, 0.4]$$

and at the top of the silicon part the silicon oxide domain is

$$[0.125, 0.275] \times [0.4, 0.42]$$

where the unit length is the micron.

The regions of high-doping n^+ are the subset

$$[0, 0.1] \times [0.35, 0.4] \cup [0.3, 0.4] \times [0.35, 0.4].$$

A uniform mesh of 64×64 grid points has been used in the silicon part. The Poisson equation is solved on the entire (silicon and oxide) domain. Of course, in the oxide, the Poisson equation becomes the Laplace one.

We have assumed ohmic contacts at the source, drain, gate and bulk contacts, to be homogeneous Neumann conditions on the remaining part of the boundary. The surface charge at the oxide interface is neglected and the continuity of the electric potential and electric field is imposed. The values of density and energy at the interface are obtained by the interior grid points with a linear interpolation in the direction orthogonal to the boundary.

In order to reach the desired bias, we have needed to resort to a continuation method on applied potential. First, we iterate with respect to the difference of the built-in potential between drain and bulk contacts, keeping at zero V_{DS}. Then, we iterate with respect to the drain–gate potential and finally we increase the drain–source potential.

All the main features of the electron dynamics are well described, in particular the charge accumulation beside the oxide, and the pronounced depletion at the drain contact due to the strong electric field.

Again, the density current presents a singularity at the first edge of the drain and therefore we evaluated the current by considering, as for MESFET, the regularization from the interior.

As for the MESFET device, current gain and miniaturization are crucial tasks to take into account when designing this kind of device. However, miniaturization plays a crucial role for MOSFETs for various reasons; first, a shorter MOSFET helps the current flow, since a small device decreases the resistance. Second, a small MOSFET helps in reducing gate capacity and, in general, to have a faster switch on/off operation. The numerical results are presented in Figures 10–12.
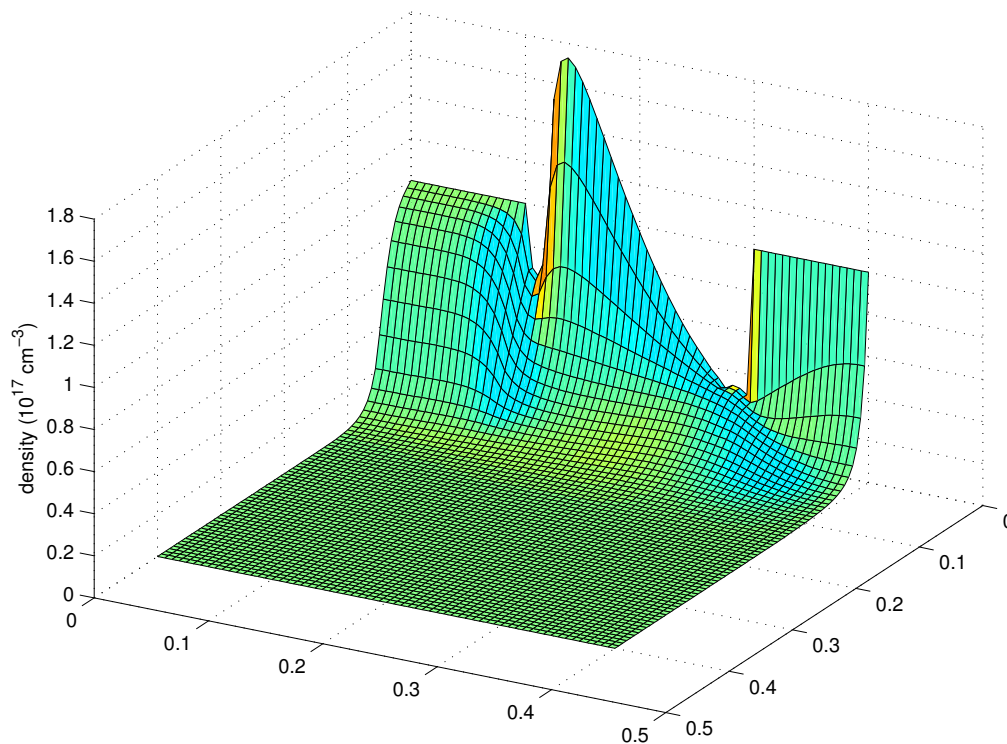
Figure 10. Stationary electron density in the MOSFET in the case $V_b = 1$ V and $V_g = 0.8$ V.
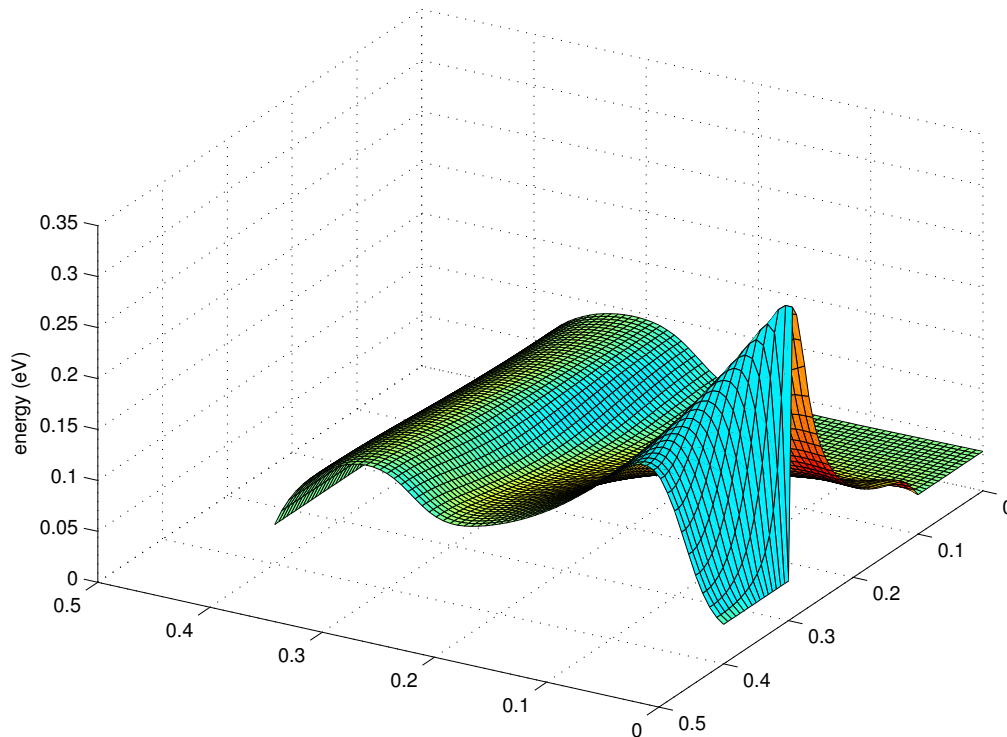
Figure 11. Stationary energy per charge carrier in the MOSFET in the case $V_b = 1$ V and $V_g = 0.8$ V.
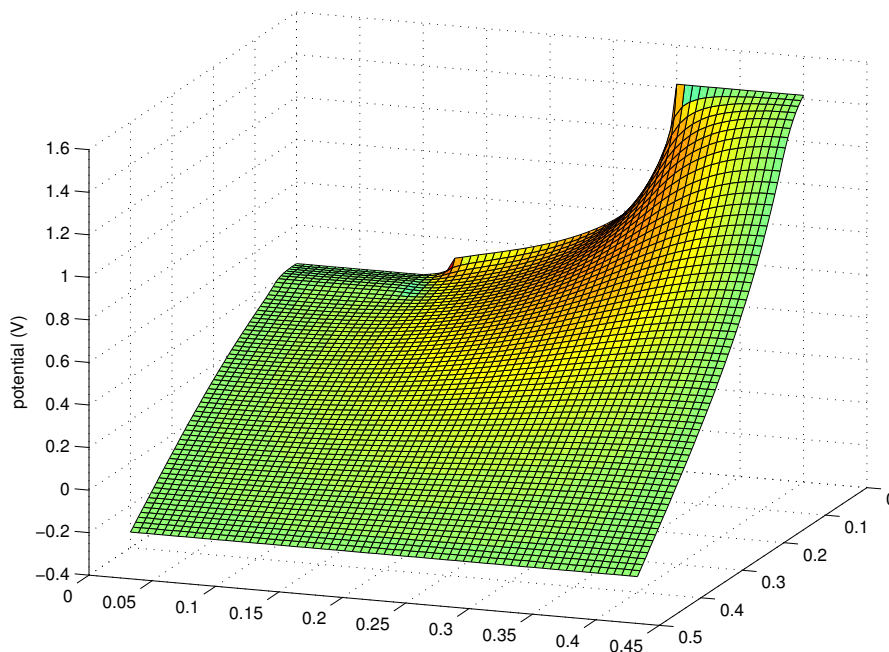
Figure 12. Stationary electric potential in the MOSFET in the case $V_b = 1$ V and $V_g = 0.8$ V.

5.4. Simulation of a DG-MOSFET

The 2DEG model, illustrated in Section 4, has been employed to simulate the nanoscale silicon DG-MOSFET represented in Figure 13, including the non-parabolicity effects [18,19]. The length of the diode is $L_x = 40$ nm, the width of the silicon layer is $L_z = 8$ nm and the thickness of each oxide layer is $t_{ox} = 1$ nm. The highly doped n^+ regions are 10 nm long. The gate contacts have the same length as the n region and are above it. The device is supposed to be infinite in the y direction.
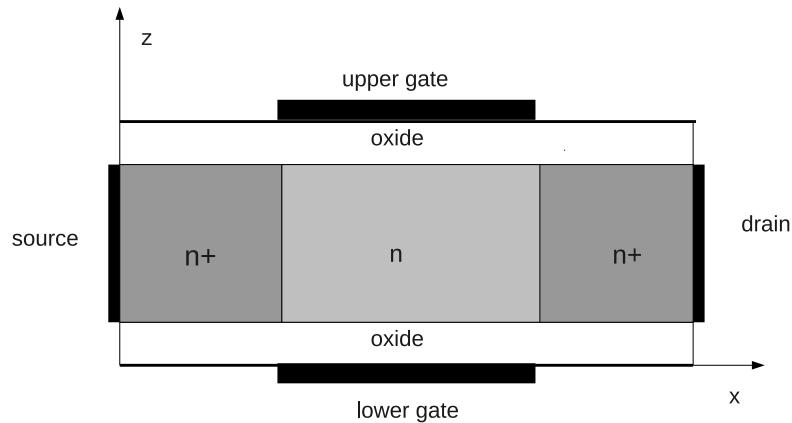
Figure 13. Simulated double gate MOSFET. Along the y axis, the device is considered as infinite.

The doping in the n^+ regions is $N_d(x) = N_d^+ = 10^{20}$ cm^{-3} and in the n region is $N_d(x) = N_d^- = 10^{15}$ cm^{-3}, with a regularization at the two junctions by a hyperbolic tangent profile

$$N_d(x) = N_d^+ - d\left(\tanh\frac{x-x_1}{s} - \tanh\frac{x-x_2}{s}\right),$$

where $s = 0.2$ nm, $d = \dfrac{N_d^+}{2}\left(1 - \dfrac{N_d^-}{N_d^+}\right)$, $x_1 = 10$ nm and $x_2 = 30$ nm.

Due to the symmetries and dimensions of the device, the transport is, within a good approximation, one dimensional and along the longitudinal direction with respect to the two oxide layers, while the electrons are quantized in the transversal direction. The oxide gives rise, with a good approximation, to an infinitely deep potential barrier; in fact, realistic values of the potential barrier are more than 3 eV high and it is very unlikely to find electrons with such an energy in the device under consideration. Six equivalent valleys are considered with a single effective mass $m^* = 0.32 m_e$. For details about the appropriate boundary and initial conditions, as well as the numerical method, the interested reader is referred to [18,19]. The numerical experiments indicate that it is sufficient to take into account only the first three subbands.

As the first case, a symmetric situation with $V_D = 0.5$ V and $V_{gl} = V_{gu} = -3$ V is considered, where V_D is the voltage applied at the drain with respect to that at the source, and V_{gl} and V_{gu} are the voltages applied at the lower and the upper gate, respectively. In Figures 14 and 15, the steady state density and the potential are plotted. The solution does not present any spurious oscillation or boundary layer and reflects the symmetry of the problem.

As the second case, $V_D = 0.5$ V, $V_{gl} = -3$ V and $V_{gu} = 3$ V are taken. In Figures 16 and 17, the density and the potential are respectively plotted, while in Figure 18b the first three subband bottoms are shown. One can note the depletion region beneath the upper gate.

In Figure 18, the first three subband bottoms are shown. There is a good qualitative agreement with the other numerical simulations known in the literature [17].

From Figure 19, a very accurate current conservation is evident, proving the robustness of the numerical method. In the second case, the current is reduced by one half, due to the gate voltage in agreement with the behaviour of the density. The areal density, the average velocity and the energy measured from the subband bottom, and the current in the first three subbands are shown in Figures 20–22. The areal density is not symmetrical between the source and drain within each subband but the total areal density is so.
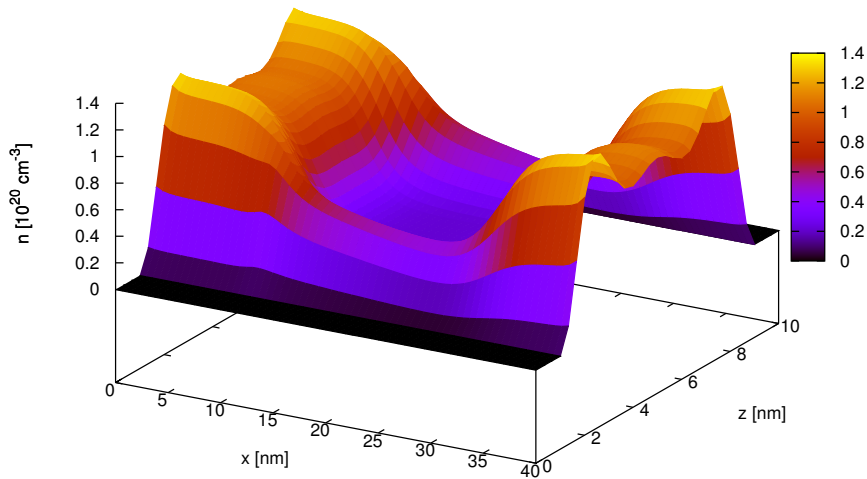
Figure 14. Stationary density in the case $V_D = 0.5$ V and $V_{gl} = V_{gu} = -3$ V.
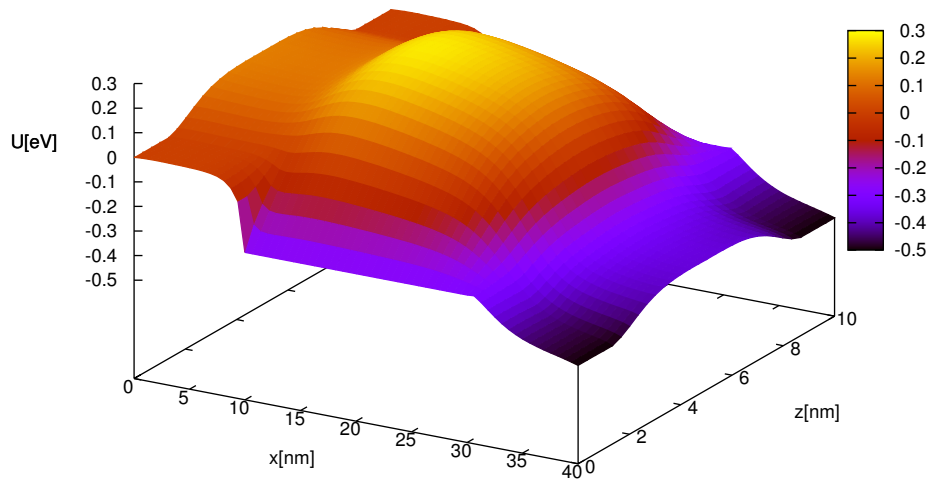

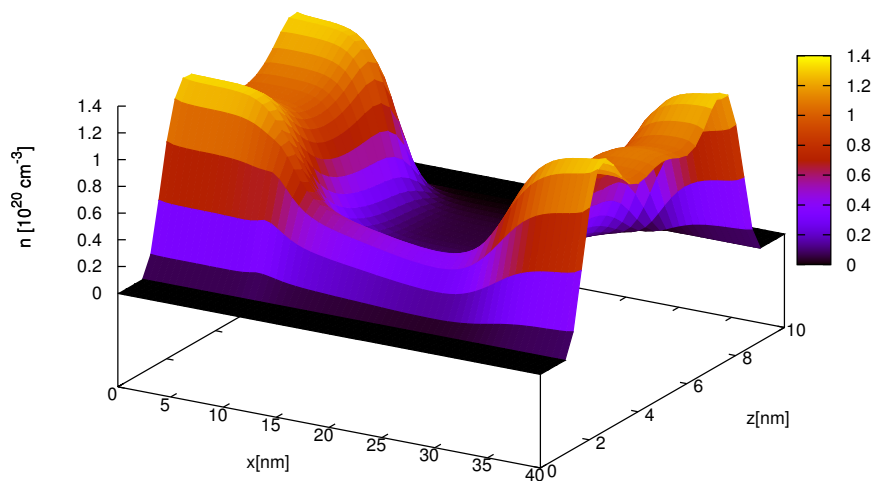**Figure 15.** Stationary electrostatic potential energy in the case $V_D = 0.5$ V and $V_{gl} = V_{gu} = -3$ V.

Figure 16. Stationary density in the case $V_D = 0.5$ V and $V_{gl} = -3$ V, $V_{gu} = 3$ V.
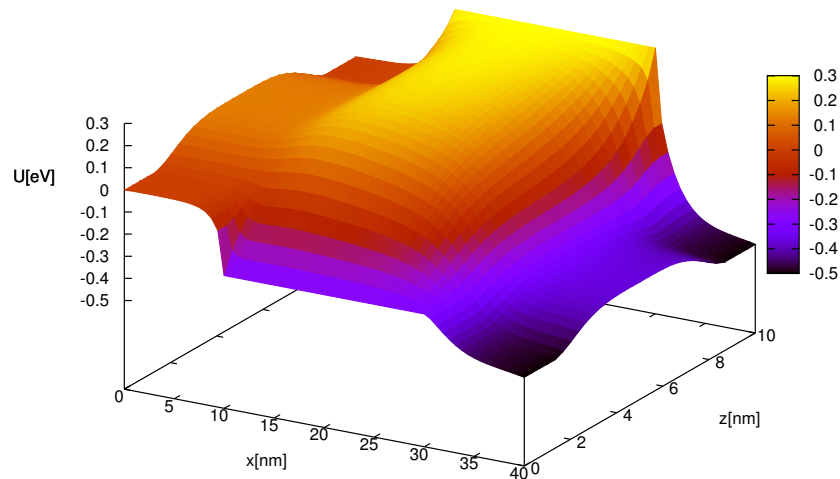
Figure 17. Stationary electrostatic potential energy in the case $V_D = 0.5$ V and $V_{gl} = -3$ V, $V_{gu} = 3$ V.
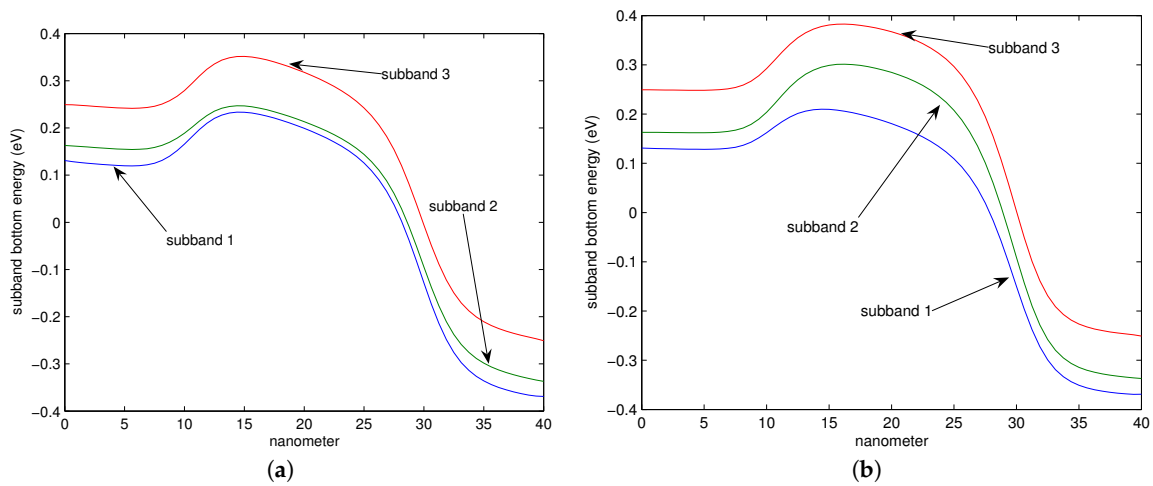

Figure 18. First three subbands at the steady state in the case $V_D = 0.5$ V and $V_{gl} = V_{gu} = -3$ V (**a**); $V_D = 0.5$ V and $V_{gl} = -3$ V, $V_{gu} = 3$ V (**b**).
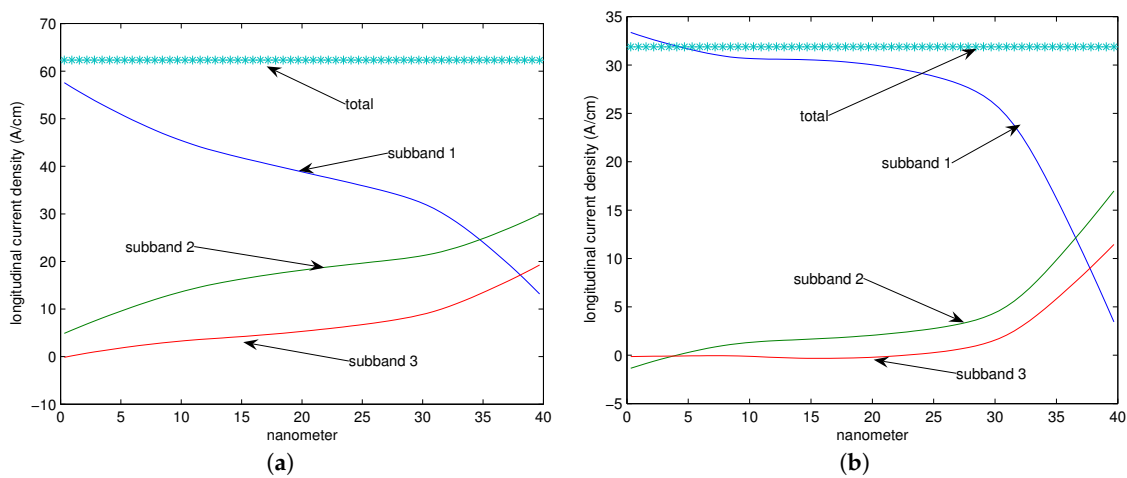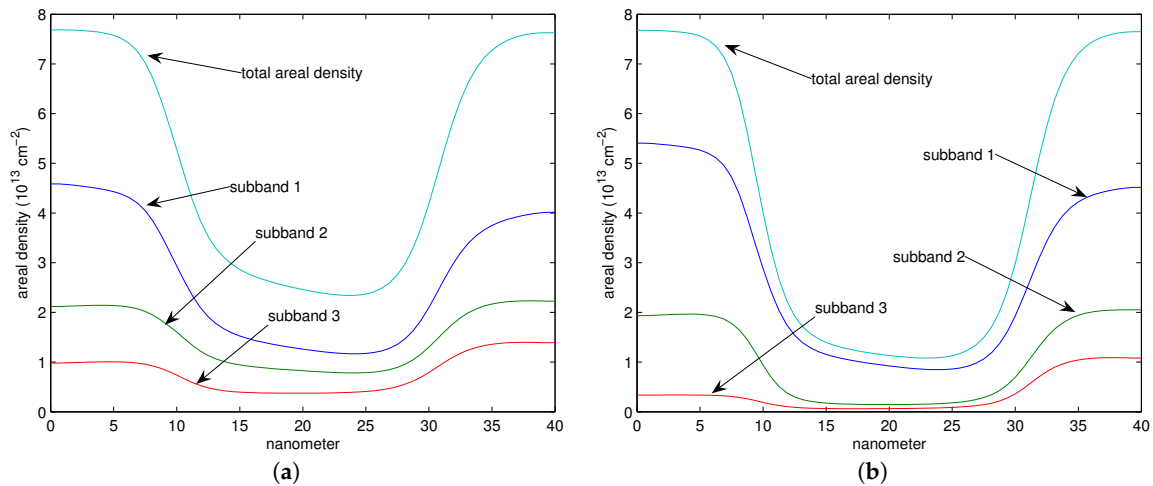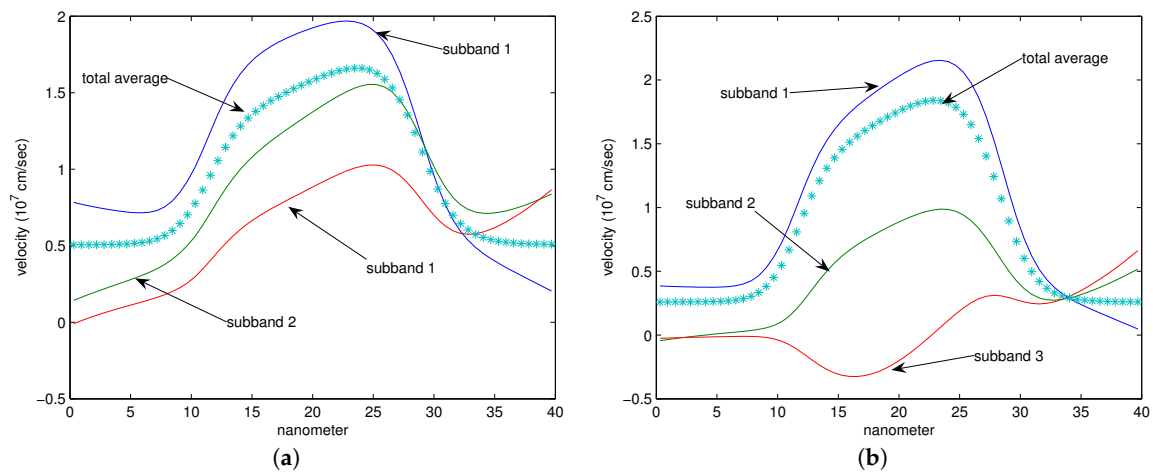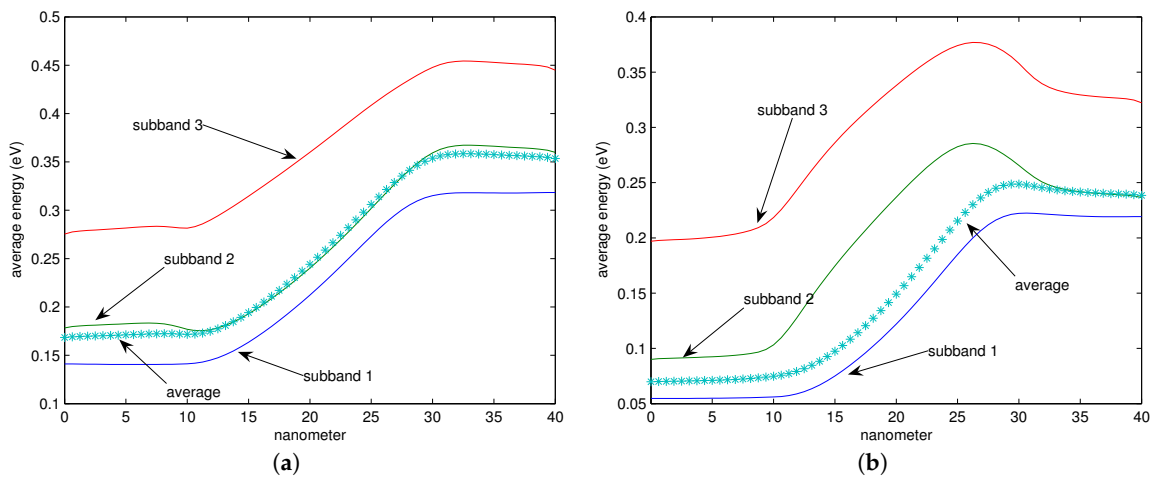

Figure 19. Average areal current in the first three subbands and global areal current in the case $V_D = 0.5$ V and $V_{gl} = V_{gu} = -3$ V (**a**); $V_D = 0.5$ V and $V_{gl} = -3$ V, $V_{gu} = 3$ V (**b**).

Figure 20. Areal density in the first three subbands in the case $V_D = 0.5$ V and $V_{gl} = V_{gu} = -3$ V (**a**); $V_D = 0.5$ V and $V_{gl} = -3$ V, $V_{gu} = 3$ V (**b**).

Figure 21. Average velocity in the first three subbands and global mean velocity in the case $V_D = 0.5$ V and $V_{gl} = V_{gu} = -3$ V (**a**); $V_D = 0.5$ V and $V_{gl} = -3$ V, $V_{gu} = 3$ V (**b**).

Figure 22. Average total energy measured from the bottom of the first subband $W^\nu + \varepsilon_\nu - \varepsilon_1$ in the first three subbands and global mean energy in the case $V_D = 0.5$ V and $V_{gl} = V_{gu} = -3$ V (**a**); $V_D = 0.5$ V and $V_{gl} = -3$ V, $V_{gu} = 3$ V (**b**).

The drift (mean) velocity has been evaluated according to the formula

$$\mathbf{V} = \frac{\sum_\nu \rho^\nu \mathbf{V}^\nu}{\sum_\mu \rho^\mu}.$$

Similarly, the global longitudinal mean energy has been evaluated taking as reference value the bottom of the first subband according to the formula

$$W = \frac{\sum_\nu \rho^\nu \left(W^\nu + \varepsilon_\nu - \varepsilon_1 \right)}{\sum_\mu \rho^\mu}.$$

The maximum drift velocity in the channel is one and half times the saturation velocity when $V_{gl} = V_{gu} = -3$ V, while it is about two times the saturation velocity when $V_{gl} = -3$ V and $V_{gu} = 3$ V. Moreover, in the first case, the velocity in the first subband is lower than that in the first subband in the second case. Instead, the velocity in the second and the third subbands is higher in the first case, but with a resulting total longitudinal current which is lower. The energy has an evidently different value between the source and drain as happens in the semiclassical case.

Eventually, the characteristic curves are shown in Figure 23 by fixing $V_{gl} = -3$ V and varying V_{gu} from -3 V to 3 V. With increasing V_{gu}, the average longitudinal current increases as a consequence of the controlling effect of the gate voltage on the electric characteristics of the device.
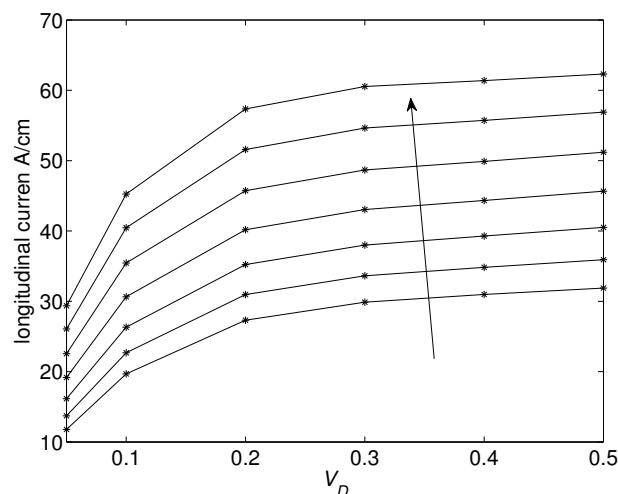
Figure 23. Longitudinal mean current (A/cm) versus the source–drain voltage V_D witht $V_{gl} = -3$ V and V_{gu} ranging from -3 V to $+3$ V according to the arrow.

6. Conclusions

A review on the exploitation of the Maximum Entropy Principle in the formulation of macroscopic models able to describe the charge and heat transport in semiconductor devices has been presented. The models are obtained starting from the Boltzmann transport equations for the charge and the phonon distribution functions, by taking—as macroscopic variables—suitable moments of the distributions and resorting to MEP in order to close the evolution equations for the chosen moments. Firstly, semiclassical models have been reviewed and eventually an approach to also take into account quantum effects has been shown. The latter consists of using the moment system arising from the Wigner equation.

In the cases where quantum effects are relevant only along one direction, another strategy has been introduced. A quasi-static description has been adopted, based on a coupled Schrödinger–Poisson system, along the confining direction, leading to a subband decomposition of the electron energy

levels, while the transport along the longitudinal directions is described by a semiclassical Boltzmann equation for each subband.

Several physical situations, which have been investigated in the last two decades, have been reported in the paper. For the sake of conciseness, some further applications have been omitted. A MEP based model for 2D–3D electron gases, of particular interest in the simulation of nanoscale MOSFET, can be found in [55]. A hydrodynamical model for charge transport in graphene has been formulated in [56,57].

Apparently, MEP is revealed to be a sound method for devising, in a systematic way, macroscopic models, e.g., energy–transport and hydrodynamical ones, which are less complex to tackle numerically but still retain a good accuracy with respect to the results based on the direct solutions of the transport equations.

Author Contributions: Both authors have jointly worked to the manuscript with an equal contribution. Both authors have read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

MEP	Maximum Entropy Principle
BTE	Boltzmann Transport Equation
CAD	Computer Aided Design
DG-MOSFET	Double Gate Metal Oxide Semiconductor Field Effect Transistors
BBP	Boltzmann–Bloch–Peierls

References

1. Jacoboni, C. *Theory of Electron Transport in Semiconductors: A Pathway from Elementary Physics to Nonequilibrium Green Functions*; Springer: Berlin/Heidelberg, Germany, 2010.

2. Carrillo, J.A.; Majorana, A.; Vecil, F. A semi-Lagrangian deterministic solver for the semiconductor Boltzmann-Poisson system. *Commun. Comput. Phys.* **2007**, *2*, 1027–1054.

3. Baccarani, G.; Wordeman, M.R. An investigation on steady-state velocity overshoot in silicon. *Solid State Electron.* **1982**, *29*, 970–977.

4. Blotekjaer, K. Transport equations for electron in two-valley semiconductors. *IEEE Trans. Electron. Devices* **1970**, *17*, 38–47.

5. Hänsch, W. *The Drift-Diffusion Equation and Its Applications in MOSFET Modeling*; Springer: Berlin/Heidelberg, Germany, 1991.

6. Dreyer, W.J. Maximisation of the entropy in non-equilibrium. *Phys. A Math. Gen.* **1987**, *20*, 6505.

7. Dreyer, W.J.; Struchtrup, H. Heat pulse experiment revisited. *Contin. Mech. Thermodyn.* **1993**, *5*, 3–50.

8. Alì, G.; Mascali, G.; Romano, V.; Torcasio, C.R. A Hydrodynamical Model for Covalent Semiconductors with Applications to GaN and SiC. *Acta Appl. Math.* **2012**, *122*, 335–348.

9. Alì, G.; Mascali, G.; Romano, V.; Torcasio, C.R. A Hydrodynamic Model for covalent semiconductors with a generalized energy dispersion relation. *Eur. J. Appl. Math.* **2014**, *25*, 255–276.

10. Anile, A.M.; Romano, V. Non parabolic band transport in semiconductors: Closure of the moment equations. *Contin. Mech. Thermodyn.* **1999**, *11*, 307–325.

11. Mascali, G.; Trovato, M. A non-linear determination of the distribution function of degenerate gases with an application to semiconductors. *Physica A* **2002**, *310*, 121–138.

12. Muscato, O.; Di Stefano, V. Modeling heat generation in a sub-micrometric $n^+ - n - n^+$ silicon diode. *J. Appl. Phys.* **2008**, *104*, 124501.

13. Muscato, O.; Di Stefano, V. Local equilibrium and off-equilibrium thermoelectric effects in silicon semiconductors. *J. Appl. Phys.* **2011**, *110*, 093706.

14. Levermore, C.D. Moment Closure Hierarchies for kinetic Theories. *J. Stat. Phys.* **1996**, *83*, 1021–1167.

15. Romano, V. Quantum corrections to the semiclassical hydrodynamical model of semiconductors based on the maximum entropy principle. *J. Math. Phys.* **2007**, *48*, 123504.

16. Polizzi, E.; Ben Abdallah, N. Subband decomposition approach for the simulation of quantum electron transport in nanostructures. *J. Comput. Phys.* **2004**, *202*, 150–180.

17. Ben Abdallah, N.; Caceres, M.J.; Carrillo, J.A.; Vecil, F. A deterministic solver for a hybrid quantum-classical transport model in nanoMOSFETs. *J. Comput. Phys.* **2009**, *228*, 6553–6571.

18. Camiola, V.D.; Mascali, G.; Romano, V. Numerical Simulation of a Double-Gate Mosfet with a Subband Model for Semiconductors Based on the Maximum Entropy Principle. *Contin. Mech. Thermodyn.* **2012**, *24*, 417–436.

19. Camiola, V.D.; Mascali, G.; Romano, V. Simulation of a double-gate MOSFET by a non-parabolic energy-transport model for semiconductors based on the maximum entropy principle. *Math. Comput. Model.* **2013**, *58*, 321–343.

20. Mascali, G.; Romano, V. A non parabolic hydrodynamical subband model for semiconductors based on the maximum entropy principle. *Math. Comput. Model.* **2012**, *55*, 1003–1020.

21. Mascali, G.; Romano, V. A hydrodynamical model for hole transport in silicon semiconductors: The case of non-parabolic warped bands. *Math. Comput. Model.* **2011**, *53*, 213–229.

22. Mascali, G.; Romano, V. A hydrodynamical model for hole transport in silicon semiconductors: The case of parabolic warped bands. *Compel* **2012**, *31*, 552–582.

23. Mascali, G. A hydrodynamical model for silicon semiconductors includind crystal heating. *Eur. J. Appl. Math.* **2015**, *26*, 477–496.

24. Auer, C.; Schürrer, F.; Koller, W. A semicontinuous formulation of the Bloch–Boltzmann–Peierls equations. *SIAM J. Appl. Math.* **2004**, *64*, 1457–1475.

25. González, P.; Carrillo, J.A.; Gámiz, F. Deterministic Numerical Simulation of 1d kinetic descriptions of Bipolar Electron Devices'. In *Scientific Computing in Electrical Engineering, Series: Mathematics in Industry, Subseries: The European Consortium for Mathematics in Industry*; Anile, A.M., Alì, G., Mascali, G., Eds.; Springer: Berlin/Heidelberg, Germany, 2006; pp. 339–344.

26. Mascali, G. Maximum entropy principle in relativistic radiation hydrodynamics II: Compton and double Compton scattering. *Contin. Mech. Thermodyn.* **2002**, *14*, 549.

27. Anile, A.M.; Mascali, G. Theoretical foundations for tail electron hydrodynamical models in semiconductors. *Appl. Math. Lett.* **2001**, *14*, 245–252.

28. Morandi, O.; Barletti, L. Particle dynamics in graphene: Collimated beam limit. *J. Comput. Theor. Transp.* **2015**, *43*, 418–432.

29. La Rosa, S.; Mascali, G.; Romano, V. Exact maximum entropy closure of the hydrodynamical model for Si semiconductors: The 8-moment case. *SIAM J. Appl. Math.* **2009**, *70*, 710–734.

30. Romano, V. Non parabolic band transport in semiconductors: Closure of the production terms in the moment equations. *Contin. Mech. Thermodyn.* **2000**, *12*, 31–51.

31. Mascali, G. A new formula for thermal conductivity based on a hierarchy of hydrodynamical models. *J. Stat. Phys.* **2016**, *163*, 1268–1284.

32. Jaynes, E.T. Information Theory and Statistical Mechanics. *Phys. Rev. B* **1957**, *106*, 620.

33. Wu, N. *The Maximum Entropy Method*; Springer: Berlin/Heidelberg, Germany, 1997.

34. Degond, P.; Ringhofer, C. Quantum moment hydrodynamics and the entropy principle. *J. Stat. Phys.* **2003**, *112*, 587–628.

35. Degond, P.; Méhats, F.; Ringhofer, C. Quantum energy-transport and drift-diffusion models. *J. Stat. Phys.* **2005**, *118*, 625–627.

36. Barker, J.; Ferry, D. Self-Scattering Path-Variable Formulation of High-Field, Time-Dependent, Quantum Kinetic Equations for Semiconductor Transport in the Finite-Collision-Duration Regime. *Phys. Rev. Lett.* **1979**, *42*, 1779.

37. Fromlet, F.; Markowich, P.A.; Ringhofer, C. A Wigner function Approach to Phonon Scattering. *VLSI Des.* **1999**, *9*, 339–350.

38. Camiola, V.D.; Romano, V. Quantum BGK model for electron transport in semiconductors. In Proceedings of the 15th Conference on Waves and Stability in Continuous Media, Palermo, Italy, 28 June–1 July 2009.

39. Romano, V. The equilibrium Wigner function in the case of nonparabolic energy band. In *Progress in Industrial Mathematics at ECMI 2008*; Fitt, A.D., Norbury, J., Ockendon, H., Wilson, E., Eds.; Springer: Berlin/Heidelberg, Germany, 2010; Volume 15, pp. 135–140.

40. Majorana, A. Space homogeneous solutions of the Boltzmann equation describing electron-phonon interactions in semiconductors. *Transp. Theory Stat. Phys.* **1991**, *20*, 261–279.

41. Majorana, A. Conservation laws from the Boltzmann equation describing electron-phonon interactions in semiconductors. *Transp. Theory Stat. Phys.* **1993**, *22*, 849–859.

42. Majorana, A. Equilibrium solutions of the non-linear Boltzmann equation for an electron gas in a semiconductor. *Il Nuovo Cimento B* **1993**, *108*, 871–877.

43. Gardner, C. The quantum hydrodynamic model for semiconductor devices. *SIAM J. Appl. Math.* **1994**, *54*, 409–427.

44. Muscato, O.; Castiglione, T. A Hydrodynamic Model for Silicon Nanowires Based on the Maximum Entropy Principle. *Entropy* **2016**, *18*, 368.

45. Camiola, V.D.; Mascali, G.; Romano, V. *The Moment Method for Charge Transport in Low Dimensional Semiconductor Devices*; Springer: Berlin/Heidelberg, Germany, 2017, in press.

46. Morandi, O. Multiband Wigner-function formalism applied to the band to band transition phenomena. *Phys. Rev. B* **2009**, *80*, 024301.

47. Mascali, G.; Romano, V. Hydrodynamic subband model for semiconductors based on the maximum entropy principle. *Nuovo Cimento C* **2010**, *33*, 155–163.

48. Galler, M.; Schuerrer, F. A deterministic Solver to the Boltzmann-Poisson System Including Quantization Effects for Silicon-MOSFETs. In *Progress in Industrial Mathematics at ECMI 2006, Series: Mathematics in Industry*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 531–536.

49. Junk, M.; Romano, V. Maximum entropy moment system of the semiconductor Boltzmann equation using Kane dispersion relation. *Contin. Mech. Thermodyn.* **2005**, *17*, 247–267.

50. Straquadanio, G.; Romano, V.; Nicosia, G. Semiconductor device design using the BIMADS algorithm. *J. Comput. Phys.* **2013**, *242*, 304–320.

51. Liotta, S.F.; Romano, V.; Russo, G. Central schemes for balance laws of relaxation type. *SIAM J. Numer. Anal.* **2000**, *38*, 1337–1356.

52. Nessyahu, H.; Tadmor, E. Non-oscillatory central differencing for hyperbolic conservation law. *J. Comput. Phys.* **1990**, *87*, 408–463.

53. Romano, V. 2D Simulation of a Silicon MESFET with a Nonparabolic Hydrodynamical Model Based on the Maximum Entropy Principle. *J. Comput. Phys.* **2002**, *176*, 70–92.

54. Chen, Z.; Cockburn, B.; Jerome, J.; Shu, C.-W. Finite element computation of the hydrodynamic model of semiconductor devices. *VLSI Des.* **1995**, *3*, 145–158.

55. Camiola, V.D.; Romano, V. 2DEG-3DEG charge transport model for MOSFET based on the maximum entropy principle. *SIAM J. Appl. Math.* **2013**, *73*, 1439–1459.

56. Camiola, V.D.; Romano, V. Hydrodynamical Model for Charge Transport in Graphene. *J. Stat. Phys.* **2014**, *157*, 1114–1137.

57. Coco, M.; Majorana, A.; Mascali, G.; Romano, V. Comparing kinetic and hydrodynamical models for electron transport in monolayer graphene. In Proceedings of the VI International Conference on Computational Methods for Coupled Problems in Science and Engineering, Coupled Problems 2015, Venice, Italy, 18–20 May 2015.