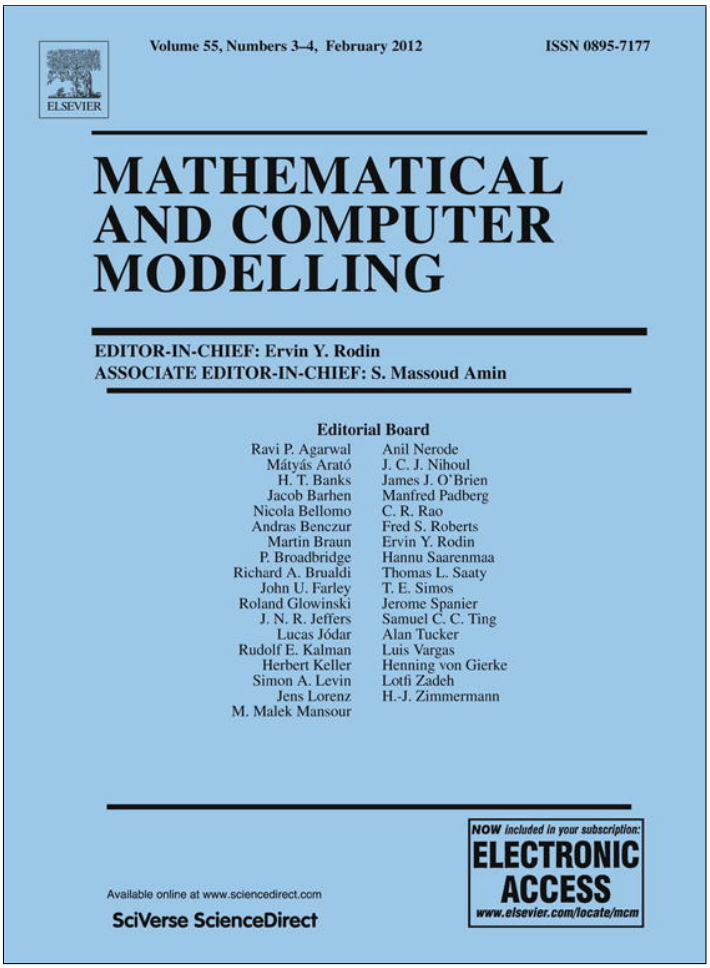

Mathematical and Computer Modelling 55 (2012) 1003–1020

Contents lists available at SciVerse ScienceDirect

Mathematical and Computer Modelling

journal homepage: www.elsevier.com/locate/mcm

A non parabolic hydrodynamical subband model for semiconductors based on the maximum entropy principle

Giovanni Mascali [a], Vittorio Romano [b,*]

[a] Dipartimento di Matematica, Università della Calabria and INFN-Gruppo c. Cosenza, 87036 Cosenza, Italy
[b] Dipartimento di Matematica e Informatica, Università di Catania, viale A. Doria 6, 95125 Catania, Italy

ARTICLE INFO

Article history:
Received 6 April 2011
Received in revised form 16 September 2011
Accepted 19 September 2011

Keywords:
Quantum transport
Semiconductors
Hydrodynamical models

ABSTRACT

A hydrodynamic subband model for semiconductors is formulated by closing the moment system derived from the Schrödinger–Poisson–Boltzmann equations on the basis of the maximum entropy principle, by taking into account non parabolic energy bands of Kane's type. Explicit closure relations for fluxes and production terms are obtained, including scattering of electrons with acoustic and non polar optical phonons and surface scattering. Numerical simulations of a quantum diode show the feasibility of the model. The importance of the non parabolicity is assessed.

1. Introduction

By shrinking the dimension of electronic devices, effects of quantum confinement are observed, e.g. in metal-oxide-semiconductor field-effect transistors (MOSFETs) the gate voltage and the energy barrier at the Si–SiO_2 interface confine the carriers near the oxide–silicon interface. Similarly in double gate MOSFETs (DG-MOSFETs) the potential between the two gates makes the electrons confined. The same effect is present in quantum wells in hetero-structures like AlGa–Ga. For a comprehensive review the reader is referred to [1,2].

In a purely quantum approach electric properties of nanoscale devices, like current–voltage curves, can be computed via the non-equilibrium Green function [3–6]. Another way to tackle the problem is in the framework of quantum kinetic theory via the Wigner function, which gives the macroscopic physical quantities of interest as expectation values [7–9]. Other approaches are based on the master equation [10]. However, in structures like DG-MOSFETs, one has a confining effect in the direction transversal to the oxide, while the electrons flow in the longitudinal direction. This flow can be treated semiclassically when the typical longitudinal length is of the order of a few tens of nanometers. We assume that the electrons as waves achieve equilibrium along the confining direction in a time which is much shorter than the typical transport time, so that one can adopt a quasi-static description along the confining direction by a coupled Schrödinger–Poisson system which leads to a subband decomposition. The transport along the longitudinal direction is described by a semiclassical Boltzmann equation for each subband.

Numerical integration of the Boltzmann–Schrödinger–Poisson system has been performed by Monte-Carlo methods or deterministic schemes for solving the transport part [11–16], but they are very expensive from a computational point of view for computer aided design (CAD) purposes. This has prompted the formulation of macroscopic models like drift-diffusion-Schrödinger–Poisson ones or of energy-transport type [17,18].

Here we use a moment method to describe the electron transport and close the resulting system of equations by resorting to the maximum entropy principle (MEP) and, at variance with [19,20], we include the effects of non parabolicity in the

* Corresponding author. Tel.: +39 095 7383035; fax: +39 095 330094.
E-mail addresses: g.mascali@unical.it (G. Mascali), romano@dmi.unict.it (V. Romano).

0895-7177/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.mcm.2011.09.026

subbands by the Kane dispersion relation. Scattering of electrons with acoustic and non polar optical phonons and surface scattering are taken into account. The resulting moment system forms a quasilinear hyperbolic system and from this latter an energy transport model is deduced under a suitable diffusion scaling. The simulation of a quantum silicon diode is presented in order to validate the model and the robustness of the numerical scheme. It is found that the nonparabolicity has less influence than in the semiclassical case but it produces valuable differences in the current which is found to be about up to 10% lower when the Kane dispersion relation is used instead of the parabolic one. We expect the proposed model to be adequate for simulating devices like MOSFETs or DG-MOSFETs with channel lengths in the range of tens of nanometers where a semiclassical description of transport is still reasonable. Shorter devices, e.g. with channel length of few nanometers, could require a quantum treatment of the transport as well.

The plan of the paper is the following. In Sections 2 and 3 a brief review of quantum confinement and electron transport in subbands is given. Then in Sections 4 and 5 the moment equations are deduced and MEP is used to determine the needed closure relations. In Section 6 a limiting energy-transport model is deduced while in the last section the numerical scheme is presented along with the simulation of a quantum diode. Some details about the constitutive relations are reported in Appendix.

2. Quantum confinement

If electrons are quantized in a direction which we call z and free to move in the x–y plane, it is natural to assume the following ansatz for their wave function

$$\psi(\mathbf{k}, \mathbf{r}) = \psi(k_x, k_y, k_z, x, y, z) = \frac{1}{\sqrt{\mathcal{A}}}\varphi(z)e^{i\mathbf{k}_\parallel \cdot \mathbf{r}_\parallel},$$

with $\mathbf{k}_\parallel = (k_x, k_y)$ and $\mathbf{r}_\parallel = (x, y)$ denoting the longitudinal components of the wave-vector $\mathbf{k}$ and the position vector $\mathbf{r}$, respectively, and $\mathcal{A}$ symbolizing the area of the xy cross-section.

Inserting the previous expression of ψ into the Schrödinger equation in the effective mass approximation

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

gives the following equation for the envelope function $\varphi(z)$

$$\left[-\frac{\hbar^2}{2m^*}\frac{d^2}{dz^2} + E_C(\mathbf{r}_\parallel, z)\right]\varphi(z) = \varepsilon(\mathbf{r}_\parallel)\varphi(z), \quad -L/2 \le z \le L/2, \tag{1}$$

where $\hbar$ is the reduced Planck constant, m^* is the effective electron mass, E_C is the conduction band minimum and L is the extension in the z-direction. ε is the energy associated with the confinement in the z-direction and depends on $\mathbf{r}_\parallel$. ε is constant in each slice $\mathbf{r}_\parallel =$ constant but varies between a slice and another one. One finds a countable set of eigen-pairs (subbands) $(\varphi_\nu, \varepsilon_\nu)$, $\nu = 1, 2, \ldots$. In (1) $E_C = -q(V_C + V)$ where V_C is the confining potential and V is the self-consistent electrostatic potential obtained from the Poisson equation

$$\nabla \cdot (\epsilon \nabla V) = -q(C(\mathbf{r}) - n), \tag{2}$$

where q is the absolute value of the electron charge, ϵ is the dielectric constant, $C(\mathbf{r})$ is the doping concentration, and n is the electron density given by

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

with the areal density of electrons ρ_ν of the ν-th subband.

3. The transport equations

The above Schrödinger–Poisson model is able to describe only the ballistic case, because the scattering is not included. In order to take into account this latter, several approaches are available. One can add in the Hamiltonian a term describing the electron–phonon interaction and solve the corresponding Schrödinger equation for the wave function in the electron–phonon space or use Green's density function methods. However for devices with characteristic length of a few tens of nanometers, which is the case we are going to simulate the transport of electrons in the longitudinal direction can be considered semiclassical within a good approximation. Electrons in each subband are considered as different populations whose dynamics is described by a semiclassical distribution function. Therefore the description of the longitudinal electron transport is included by adding to the Schrödinger–Poisson model the system of coupled Boltzmann equations for the distributions $f_\nu(x, y, k_x, k_y, t)$ of electrons in each subband

$$\frac{\partial f_\nu}{\partial t} + \frac{1}{\hbar}\nabla_{\mathbf{k}_\parallel} E_\nu \cdot \nabla_{\mathbf{r}_\parallel} f_\nu - \frac{q}{\hbar}\mathbf{E}_\nu^{\text{eff}} \cdot \nabla_{\mathbf{k}_\parallel} f_\nu = \sum_{\mu=1}^{\infty} C_{\nu,\mu}[f_\nu, f_\mu], \quad \nu = 1, 2, \ldots \tag{3}$$

where $\mathbf{E}_\nu^{\text{eff}} = \frac{1}{q}\nabla_{\mathbf{r}_\parallel}\varepsilon_\nu(\mathbf{r}_\parallel)$. ρ_ν is expressed in terms of f_ν by

$$\rho_\nu = \int_{B_2} f_\nu(\mathbf{r}_\parallel, \mathbf{k}_\parallel, t)d^2\mathbf{k}_\parallel,$$

with B_2 indicating the 2-D Brillouin zone, which will be approximated with $\mathbb{R}^2$ consistently with the effective mass approximation.

In each subband the energy is the sum of a transversal contribution ε_ν and a longitudinal contribution $\varepsilon_\parallel$. According to previous approaches, in order to include non parabolicity effects in the transport and in the meantime avoid the complexity of the full band model, the longitudinal energy is assumed of Kane's type [2,14] and, therefore, the subband energy is given by

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

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

$$\mathbf{v}_\parallel = \frac{1}{\hbar}\nabla_{\mathbf{k}_\parallel}\varepsilon_\parallel = \frac{\hbar\mathbf{k}_\parallel}{m^*\left(1 + 2\alpha\varepsilon_\parallel\right)}. \tag{4}$$

For the limit case of $\alpha = 0$ the interested reader is referred to [19].

In the non degenerate approximation, each contribution to the collision term has the general form

$$C_{\nu,\mu}[f_\nu, f_\mu] = \int_{B_2}\left[S_{\mu\nu}(\mathbf{k}'_\parallel, \mathbf{k}_\parallel)f'_\mu - S_{\nu\mu}(\mathbf{k}_\parallel, \mathbf{k}'_\parallel)f_\nu\right]d^2\mathbf{k}'_\parallel.$$

When $\mu = \nu$ we have intra-subband scatterings; when $\mu \neq \nu$ we have inter-subband scatterings. $S_{\mu\nu}(\mathbf{k}'_\parallel, \mathbf{k}_\parallel)$ is the transition rate per unit time from the longitudinal state with wave-vector $\mathbf{k}'_\parallel$, belonging to the μ-th subband, to the longitudinal state with wave-vector $\mathbf{k}_\parallel$ in the ν-th subband, and $f'_\mu \equiv f_\mu(\mathbf{r}_\parallel, \mathbf{k}'_\parallel, t)$. The relevant 2-D scattering mechanisms in Si are acoustic phonon scattering, nonpolar optical phonon scattering, and surface scattering. Scattering with impurities will be not considered in this paper, but it is relevant only at low temperature or low field.

For the acoustic phonon scattering in the elastic approximation, the transition rate is given by

$$S_{\nu\mu}^{(\text{ac})}(\mathbf{k}_\parallel, \mathbf{k}'_\parallel) = A^{(\text{ac})}G_{\nu\mu}\delta(E_\mu(\mathbf{k}'_\parallel) - E_\nu(\mathbf{k}_\parallel)),$$

with

$$A^{(\text{ac})} = \frac{k_B T_L \Xi_d^2}{4\pi^2\hbar\rho v_s^2},$$

where k_B is the Boltzmann constant, T_L the lattice temperature, which will be considered constant in this paper, ρ the silicon density, Ξ_d the acoustic phonon deformation potential, and v_s the longitudinal sound speed. The $G_{\nu\mu}$'s are the interaction integrals

$$G_{\nu\mu} = \int_{-\infty}^{\infty}|I_{\nu\mu}(q_z)|^2 dq_z, \qquad I_{\nu\mu}(q_z) = \int_{-L/2}^{L/2}\overline{\varphi_\nu(z)}\varphi_\mu(z)e^{iq_z z}\,dz,$$

with $\mathbf{q}$ denoting the 3D-phonon wave vector, and the bar indicating complex conjugation. We note that $G_{\nu\mu} = G_{\mu\nu}$ holds.

Similarly for the non-polar optical phonon scattering one has

$$S_{\nu\mu}^{(\text{no})}(\mathbf{k}_\parallel, \mathbf{k}'_\parallel) = A^{(\text{no})}G_{\nu\mu}\left(N_q + \frac{1}{2} \mp \frac{1}{2}\right)\delta(E_\mu(\mathbf{k}'_\parallel) - E_\nu(\mathbf{k}_\parallel) \mp \hbar\omega),$$

where

$$A^{(\text{no})} = \frac{(D_t K)^2}{8\pi^2\rho\omega},$$

N_q is the Bose–Einstein phonon occupation number, $D_t K$ is the non-polar optical phonon deformation potential, and $\hbar\omega$ is the phonon energy.

At last, we recall that the scattering at a surface is due to the roughness of the interface between oxide and silicon in the MOSFET, which produces fluctuations in the transverse component of the electric potential. It is very relevant at high gate voltage. If the fluctuation of the thickness is assumed to have an exponential autocorrelation function with a root mean square amplitude Δ_{rms} and a correlation length L_C, the transition probability has the form

$$S_{\nu\mu}^{(\text{sur})}(\mathbf{k}_\parallel, \mathbf{k}'_\parallel) = \frac{2\pi}{\hbar}(q\tilde{E}_{\nu\mu})^2\pi\Delta_{\text{rms}}^2 L_C^2\frac{1}{1 + L_C^2(\mathbf{k}_\parallel - \mathbf{k}'_\parallel)^2/2}\delta\left(E_\mu - E_\nu\right),$$

$\tilde{E}_{\nu\mu}$ being the effective electric field given by

$$q\tilde{E}_{\nu\mu} = \int_{-L/2}^{L/2} \overline{\varphi}_\mu q E_z \varphi_\nu dz,$$

with E_z transverse component of the electric field.

The above quantum picture can present further specific features. For example in the Si–SiO_2 interface of a MOSFET, a quantized inversion layer having a (100)-oriented surface has two sets of subbands, called *ladders*: one coming from the projection of the two valleys with longitudinal mass in the direction perpendicular to the interface, the other one originating from the four valleys having a transverse mass in the direction perpendicular to the interface. Here for the sake of simplicity only one ladder will be considered. However our results can be easily generalized in order to include both ladders by taking into account also intervalley scatterings between subbands belonging to different ladders and solving a Schrödinger equation for each effective mass (longitudinal and transversal).

4. The moment system

A complete description of quantum confinement is obtained by solving the system (1)–(3) (see for example [13–15]), but this is a daunting computational task. Therefore simpler macroscopic models are looked at for CAD purposes. These can be obtained as moment equations of the transport Boltzmann equations under suitable closure relations. The moment of the ν-th subband distribution with respect to a weight function $a(\mathbf{k}_\parallel)$ reads

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

In particular we take as basic moments

$$\text{the areal density } \rho^\nu = \int_{B_2} f_\nu(\mathbf{r}_\parallel, \mathbf{k}_\parallel, t) d^2\mathbf{k}_\parallel,$$

$$\text{the longitudinal mean velocity } \mathbf{V}^\nu = \frac{1}{\rho^\nu} \int_{B_2} \mathbf{v}_\parallel f_\nu(\mathbf{r}_\parallel, \mathbf{k}_\parallel, t) d^2\mathbf{k}_\parallel,$$

$$\text{the longitudinal mean energy } W^\nu = \frac{1}{\rho^\nu} \int_{B_2} \varepsilon_\parallel f_\nu(\mathbf{r}_\parallel, \mathbf{k}_\parallel, t) d^2\mathbf{k}_\parallel,$$

$$\text{the longitudinal mean energy-flux } \mathbf{S}^\nu = \frac{1}{\rho^\nu} \int_{B_2} \mathbf{v}_\parallel \varepsilon_\parallel f_\nu(\mathbf{r}_\parallel, \mathbf{k}_\parallel, t) d^2\mathbf{k}_\parallel.$$

The corresponding moment system reads

$$\frac{\partial \rho^\nu}{\partial t} + \nabla_{\mathbf{r}_\parallel} \cdot (\rho^\nu \mathbf{V}^\nu) = \rho^\nu \sum_\mu C_\rho^{\nu,\mu},$$

$$\frac{\partial (\rho^\nu \mathbf{V}^\nu)}{\partial t} + \nabla_{\mathbf{r}_\parallel} \cdot (\rho^\nu \mathbf{F}^{(0)\nu}) + \rho^\nu \mathbf{G}^{(0)\nu} \cdot \nabla_{\mathbf{r}_\parallel} \varepsilon_\nu = \rho^\nu \sum_\mu C_{\mathbf{V}}^{\nu,\mu},$$

$$\frac{\partial \rho^\nu W^\nu}{\partial t} + \nabla_{\mathbf{r}_\parallel} \cdot (\rho^\nu \mathbf{S}^\nu) + \rho^\nu \nabla_{\mathbf{r}_\parallel} \varepsilon_\nu \cdot \mathbf{V}^\nu = \rho^\nu \sum_\mu C_W^{\nu,\mu},$$

$$\frac{\partial (\rho^\nu \mathbf{S}^\nu)}{\partial t} + \nabla_{\mathbf{r}_\parallel} \cdot (\rho^\nu \mathbf{F}^{(1)\nu}) + \rho^\nu \mathbf{G}^{(1)\nu} \cdot \nabla_{\mathbf{r}_\parallel} \varepsilon_\nu = \rho^\nu \sum_\mu C_{\mathbf{S}}^{\nu,\mu},$$

where

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

$$\begin{pmatrix} \mathbf{G}^{(0)\nu} \\ \mathbf{G}^{(1)\nu} \end{pmatrix} = \frac{1}{\rho^\nu} \int_{B_2} \begin{pmatrix} \frac{1}{\hbar} \nabla_{\mathbf{k}_\parallel} \mathbf{v}_\parallel \\ \frac{1}{\hbar} \nabla_{\mathbf{k}_\parallel} (\varepsilon_\parallel \mathbf{v}_\parallel) \end{pmatrix} f_\nu(\mathbf{r}_\parallel, \mathbf{k}_\parallel, t)\, d^2\mathbf{k}_\parallel,$$

$$\begin{pmatrix} C_\rho^{\nu,\mu} \\ C_W^{\nu,\mu} \end{pmatrix} = \frac{1}{\rho^\nu} \int_{B_2} \begin{pmatrix} 1 \\ \varepsilon_\parallel \end{pmatrix} \left[S_{\mu\nu}(\mathbf{k}'_\parallel, \mathbf{k}_\parallel) f'_\mu - S_{\nu\mu}(\mathbf{k}_\parallel, \mathbf{k}'_\parallel) f_\nu \right] d^2\mathbf{k}'_\parallel\, d^2\mathbf{k}_\parallel,$$

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

The moment system is not closed because there are more unknowns than equations. Therefore constitutive relations in terms of the fundamental variables are needed for fluxes and productions terms.

5. The maximum entropy principle and the closure relations

MEP leads to a systematic way for obtaining constitutive relations on the basis of information theory [21–24]. According to MEP, if a given number of moments $M^{\nu}_{a_A}$, $A = 1, \ldots, N$, are known, the distribution functions f_ν can be estimated by the extremal $f^{\mathrm{MEP}} = \left(f_1^{\mathrm{MEP}}, f_2^{\mathrm{MEP}}, \ldots\right)$ of the entropy functional, under the constraints

$$\int a_A f_\nu^{\mathrm{MEP}} d\mathbf{k}_\parallel = M^{\nu}_{a_A}, \quad A = 1, \ldots, N. \tag{5}$$

Actually, in a semiconductor electrons interact with phonons describing the thermal vibrations of the ions placed at the points of the crystal lattice. However, if one considers phonon gas as a thermal bath, one has to extremize only the electron component of the entropy. Moreover, since we are considering the electron gas as sufficiently dilute, one can take in each subband the expression of the entropy obtained as a semiclassical limit of that arising from the Fermi statistics. We define the entropy of the system as

$$\mathcal{S} = -k_B \sum_{\nu=1}^{+\infty} |\varphi_\nu(z,t)|^2 \int_{B_2} \left(f_\nu \log \frac{f_\nu}{y} - f_\nu \right) d^2\mathbf{k}_\parallel, \quad y = \frac{2}{(2\pi)^2},$$

and therefore, according to MEP, the f_ν's are estimated with the distributions f_ν^{MEP}'s that solve the problem

$$\text{extremize } \mathcal{S} \text{ under the constraints } M^{\nu}_{a_A} = \int_{B_2} a_A(\mathbf{k}_\parallel) f_\nu^{\mathrm{MEP}} d^2\mathbf{k}_\parallel,$$

where $M^{\nu}_{a_A}$ are the basic moments that we have previously considered.

The proposed expression of the entropy combines quantum effects and semiclassical transport along the longitudinal direction, weighting the contribution of each f_ν with the squared modulus of the $\varphi_\nu(z,t)$'s arising from the Schrödinger–Poisson block. In order to determine the f_ν^{MEP}'s, we introduce the Lagrange transform

$$\mathcal{S}' = \mathcal{S} + \sum_{\nu=1}^{\infty} \sum_A \lambda^{\nu}_{a_A} |\varphi_\nu(z,t)|^2 \left[M^{\nu}_{a_A} - \int a_A f_\nu^{\mathrm{MEP}} d^2\mathbf{k}_\parallel \right].$$

The condition that the first variation must be zero, $\delta\mathcal{S}' = 0$, gives

$$0 = \sum_{\nu=1}^{+\infty} |\varphi_\nu(z,t)|^2 \int_{B_2} \left[k_B \log \frac{f_\nu}{y} + \sum_A \lambda^{\nu}_{a_A} a_A(\mathbf{k}_\parallel) \right] \delta f_\nu \, d^2\mathbf{k}_\parallel, \quad \forall \delta f_\nu.$$

With the above choice of the functions $a_A(\mathbf{k}_\parallel) = (1, \mathbf{v}_\parallel, \varepsilon_\parallel, \varepsilon_\parallel \mathbf{v}_\parallel)$, one has (including the factors k_B and y into the multipliers)

$$f_\nu^{\mathrm{MEP}} = \exp\left[-\left(\lambda^\nu + \lambda^\nu_{\mathbf{V}} \cdot \mathbf{v}_\parallel + \left(\lambda^\nu_W + \lambda^\nu_{\mathbf{S}} \cdot \mathbf{v}_\parallel \right) \varepsilon_\parallel \right) \right]. \tag{6}$$

In order to complete the procedure one has to insert the f_ν^{MEP}'s into the constraint relations (5) and express the Lagrangian multipliers as functions of the basic moments ρ^ν, $\mathbf{V}^\nu$, W^ν, $\mathbf{S}^\nu$. However such a procedure requires a numerical inversion, which is not practical for numerical simulations of electron devices, since it must be performed at each time or iteration step (see [25] for the semiclassical case). Following the same approach as in [26–28], we assume a small anisotropy of the distribution functions and expand them up to first order

$$f_\nu^{\mathrm{MEP}} \approx \exp\left(-\lambda^\nu - \lambda^\nu_W \varepsilon_\parallel \right) \left[1 - \left(\lambda^\nu_{\mathbf{V}} \cdot \mathbf{v}_\parallel + \lambda^\nu_{\mathbf{S}} \cdot \mathbf{v}_\parallel \varepsilon_\parallel \right) \right].$$

Inserting the above-written expansions into the constraints for the densities and the energies, each measured from the bottom of the corresponding subband, one finds in polar coordinates[1]

$$\rho = \frac{m^*}{\hbar^2} \int_0^{2\pi} \int_0^{+\infty} \exp\left(-\lambda - \lambda_W \varepsilon_\parallel \right) \left(1 + 2\alpha\varepsilon_\parallel \right) d\varepsilon_\parallel \, d\theta,$$

$$\rho W = \frac{2\pi m^*}{\hbar^2} \exp(-\lambda) \int_0^{+\infty} \varepsilon_\parallel \exp\left(-\lambda_W \varepsilon_\parallel \right) \left(1 + 2\alpha\varepsilon_\parallel \right) d\varepsilon_\parallel,$$

from where

$$\lambda_W = \frac{1 - 2\alpha W + \sqrt{(1 - 2\alpha W)^2 + 16\alpha W}}{2W}, \qquad \lambda = -\ln\left(\frac{\hbar^2}{2\pi m^*} \rho g(W) \right), \tag{7}$$

with

$$g(W) = \left(\frac{1}{\lambda_W} + \frac{2\alpha}{(\lambda_W)^2} \right)^{-1}.$$

[1] Here and whenever possible we omit the index subband, for simplicity.

Similarly, substituting (6) into the remaining constraints for the velocities and the energy-fluxes and inverting, one has

$$\lambda_{\mathbf{V}} = b_{11}(W)\mathbf{V} + b_{12}(W)\mathbf{S}, \qquad \lambda_{\mathbf{S}} = b_{21}(W)\mathbf{V} + b_{22}(W)\mathbf{S}, \tag{8}$$

where

$$b_{ij}(W) = \frac{(-)^{i+j-1}m^*}{b(W)g(W)}\left[\gamma_{(5-i-j)}(W,0) + \alpha\gamma_{(6-i-j)}(W,0)\right],$$
$$b(W) = [\gamma_1(W,0) + \alpha\gamma_2(W,0)][\gamma_3(W,0) + \alpha\gamma_4(W,0)] - [\gamma_2(W,0) + \alpha\gamma_3(W,0)]^2.$$

Note the symmetry of the coefficients, $b_{12} = b_{21}$, which reminds us of the Onsager reciprocity conditions [29]. The γ's are reported in Appendix.

5.1. Fluxes

The obtained distribution functions are used to get the needed closure relations for the fluxes and the production terms. For the fluxes we find

$$\mathbf{F}^{(i)} = \frac{g(W)}{m^*}\left[\gamma_{1+i}(W,0) + \alpha\gamma_{2+i}(W,0)\right]\mathbf{I}, \quad i = 0, 1, \tag{9}$$

$$\mathbf{G}^{(i)} = i\mathbf{F}^{(0)} + \frac{g(W)}{m^*}\left[\left(\frac{1}{\lambda_W}\right)^{i+1} - \zeta_{i+1}(W) - \alpha\zeta_{i+2}(W)\right]\mathbf{I}, \quad i = 0, 1 \tag{10}$$

where $\mathbf{I}$ is the identity. The ζ's are reported in Appendix.

5.2. Production terms

Concerning the production terms, for the acoustic phonon scattering one has

$$C^{\nu}_{W_{(i)}} = \frac{2\pi m^*}{\hbar^2}\sum_{\mu=1}^{+\infty} C_{\nu\mu}\left[\frac{\rho_\mu g(W^\mu)}{\rho_\nu}\exp\left(-\Delta_{\nu\mu}\lambda^\mu_W\right)B^{(i)}(W^\mu, a_{\nu\mu}, \Delta_{\nu\mu}) - g(W^\nu)B^{(i)}(W^\nu, a_{\nu\mu}, \Delta_{\nu\mu})\right], \tag{11}$$

$$C^{\nu}_{\mathbf{V}_{(i)}} = \frac{2\pi g\,(W^\nu)}{\hbar^2}\sum_{\mu=1}^{+\infty} C_{\nu\mu}\left\{\lambda^\nu_{\mathbf{V}}\left[\eta^{(1+i)}(W^\nu, a_{\nu\mu}, \Delta_{\nu\mu})\right] + \lambda^\nu_{\mathbf{S}}\left[\eta^{(2+i)}(W^\nu, a_{\nu\mu}, \Delta_{\nu\mu})\right]\right\}, \quad i = 0, 1, \tag{12}$$

where $W_{(0)} = \rho$, $W_{(1)} = W$, $\mathbf{V}_{(0)} = \mathbf{V}$, $\mathbf{V}_{(1)} = \mathbf{S}$, $\Delta_{\nu\mu} = \varepsilon_\nu - \varepsilon_\mu$, $a_{\nu\mu} = \max(0, \varepsilon_\mu - \varepsilon_\nu)$, and $C_{\nu\mu} = A^{(\mathrm{ac})}G_{\nu\mu}$. The $B^{(i)}$'s as well as the $\eta^{(i)}$'s are reported in Appendix.

For the non-polar optical phonon scattering one has

$$\begin{aligned} C^{\nu}_{W_{(i)}} = {} & \frac{2\pi m^*}{\hbar^2}N_{op}\sum_{\mu=1}^{+\infty} D_{\nu\mu}\left\{\frac{\rho_\mu}{\rho_\nu}g\,(W^\mu)\left[\exp\left(\frac{\hbar\omega}{k_BT_L} - \Delta^+_{\nu\mu}\lambda^\mu_W\right)B^{(i)}\left(W^\mu, a^-_{\nu\mu}, \Delta^+_{\nu\mu}\right)\right.\right. \\ & \left. + \exp\left(-\Delta^-_{\nu\mu}\lambda^\mu_W\right)B^{(i)}\left(W^\mu, a^+_{\nu\mu}, \Delta^-_{\nu\mu}\right)\right] \\ & \left. - g(W^\nu)\left[\exp\left(\frac{\hbar\omega}{k_BT_L}\right)B^{(i)}\left(W^\nu, a^+_{\nu\mu}, \Delta^-_{\nu\mu}\right) + B^{(i)}\left(W^\nu, a^-_{\nu\mu}, \Delta^+_{\nu\mu}\right)\right]\right\}, \end{aligned} \tag{13}$$

$$\begin{aligned} C^{\nu}_{\mathbf{V}_{(i)}} = {} & \frac{2\pi}{\hbar^2}N_{op}g\,(W^\nu)\sum_{\mu=1}^{+\infty} D_{\nu\mu}\left\{\lambda^\nu_{\mathbf{V}}\left[\eta^{(1+i)}\left(W^\nu, a^-_{\nu\mu}, \Delta^+_{\nu\mu}\right) + \exp\left(\frac{\hbar\omega}{k_BT_L}\right)\eta^{(1+i)}\left(W^\nu, a^+_{\nu\mu}, \Delta^-_{\nu\mu}\right)\right]\right. \\ & \left. + \lambda^\nu_{\mathbf{S}}\left[\eta^{(2+i)}\left(W^\nu, a^-_{\nu\mu}, \Delta^+_{\nu\mu}\right) + \exp\left(\frac{\hbar\omega}{k_BT_L}\right)\eta^{(2+i)}\left(W^\nu, a^+_{\nu\mu}, \Delta^-_{\nu\mu}\right)\right]\right\}, \quad i = 0, 1, \end{aligned} \tag{14}$$

where $a^{\mp}_{\nu\mu} = \max(0, \varepsilon_\mu - \varepsilon_\nu \mp \hbar\omega)$, $\Delta^{\pm}_{\nu\mu} = \varepsilon_\nu - \varepsilon_\mu \pm \hbar\omega$, $D_{\nu\mu} = A^{(\mathrm{no})}G_{\nu\mu}$.

For the scattering at a surface one has

$$C^{\nu}_{W_{(i)}} = \frac{2\pi m^*}{\hbar^2}D_S\sum_{\mu=1}^{+\infty}(q\tilde{E}_{\nu\mu})^2\left(\frac{\rho_\mu}{\rho_\nu}g\,(W^\mu)\exp\left(-\Delta_{\nu\mu}\lambda^\mu_W\right)B^{(i)}_1\left(W^\mu, a_{\nu\mu}, \Delta_{\nu\mu}\right) - g\,(W^\nu)\,B^{(i)}_1\left(W^\nu, a_{\nu\mu}, \Delta_{\nu\mu}\right)\right), \tag{15}$$

$$C^{\nu}_{\mathbf{V}_{(i)}} = \frac{2\pi}{\hbar^2} D_S \sum_{\mu=1}^{+\infty} (q\tilde{E}_{\nu\mu})^2 \Bigg\{ g\left(W^{\nu}\right) \left[\lambda^{\nu}_{\mathbf{V}} B_2^{(i+1)} \left(W^{\nu}, a_{\nu\mu}, \Delta_{\nu\mu}\right) + \lambda^{\nu}_{\mathbf{S}} B_2^{(i+2)} \left(W^{\nu}, a_{\nu\mu}, \Delta_{\nu\mu}\right) \right] - \frac{\rho_{\mu}}{\rho_{\nu}} g\left(W^{\mu}\right) \exp\left(-\Delta_{\nu\mu}\lambda^{\mu}_{W}\right) \Bigg[\lambda^{\mu}_{\mathbf{V}} B_3^{(i)} \left(W^{\mu}, a_{\nu\mu}, \Delta_{\nu\mu}\right) + \lambda^{\mu}_{\mathbf{S}} \Bigg(\Delta_{\nu\mu} B_3^{(i)} \left(W^{\mu}, a_{\nu\mu}, \Delta_{\nu\mu}\right) + B_3^{(i+1)} \left(W^{\mu}, a_{\nu\mu}, \Delta_{\nu\mu}\right) \Bigg) \Bigg] \Bigg\}, \tag{16}$$

where $i = 0, 1$, $D_S = \frac{2\pi^2}{\hbar} \Delta^2 L_C^2$. The $B_j^{(i)}$'s are reported in Appendix.

We remark that only the subbands with the lower energies are significantly populated and therefore only a finite number of moment equations will be relevant.

6. Energy-transport limit

If we neglect the surface scattering, the production terms of the velocities and the energy fluxes can be written for each subband in the compact form

$$\begin{pmatrix} C^{\nu}_{\mathbf{V}} \\ C^{\nu}_{\mathbf{S}} \end{pmatrix} = \begin{pmatrix} c^{\nu}_{11}(W^{\nu}) & c^{\nu}_{12}(W^{\nu}) \\ c^{\nu}_{21}(W^{\nu}) & c^{\nu}_{22}(W^{\nu}) \end{pmatrix} \begin{pmatrix} \mathbf{V}^{\nu} \\ \mathbf{S}^{\nu} \end{pmatrix}.$$

The coupling among the subbands in the production of velocities and energy-fluxes is only through the differences of the subband bottom energies.

Under the diffusion scaling [30–32] $t = \mathcal{O}\left(\frac{1}{\delta^2}\right)$, $\mathbf{r}_{\parallel} = \mathcal{O}\left(\frac{1}{\delta}\right)$, $\mathbf{V}_{\parallel} = \mathcal{O}(\delta)$, $\mathbf{S}_{\parallel} = \mathcal{O}(\delta)$, the following *energy transport model* is obtained

$$\frac{\partial \rho^{\nu}}{\partial t} + \nabla_{\mathbf{r}_{\parallel}} \cdot (\rho^{\nu} \mathbf{V}^{\nu}) = \rho^{\nu} C^{\nu}_{\rho}(\mathbf{W}), \tag{17}$$

$$\frac{\partial \rho^{\nu} W^{\nu}}{\partial t} + \nabla_{\mathbf{r}_{\parallel}} \cdot (\rho^{\nu} \mathbf{S}^{\nu}) + \rho^{\nu} \nabla_{\mathbf{r}_{\parallel}} \varepsilon_{\nu} \cdot \mathbf{V}^{\nu} = \rho^{\nu} C^{\nu}_{W}(\mathbf{W}), \tag{18}$$

where the index ν runs over the considered subbands, $\mathbf{W} = (W^1, W^2, \ldots)$ and

$$\mathbf{V}^{\nu} = D^{\nu}_{11}(\mathbf{W}) \nabla_{\mathbf{r}_{\parallel}} \log \rho^{\nu} + D^{\nu}_{12}(\mathbf{W}) \nabla_{\mathbf{r}_{\parallel}} W^{\nu} - D^{\nu}_{13}(\mathbf{W}) \nabla_{\mathbf{r}_{\parallel}} \varepsilon_{\nu},$$

$$\mathbf{S}^{\nu} = D^{\nu}_{21}(\mathbf{W}) \nabla_{\mathbf{r}_{\parallel}} \log \rho^{\nu} + D^{\nu}_{22}(\mathbf{W}) \nabla_{\mathbf{r}_{\parallel}} W^{\nu} - D^{\nu}_{23}(\mathbf{W}) \nabla_{\mathbf{r}_{\parallel}} \varepsilon_{\nu}.$$

The coefficients D^{ν}_{ij} are given by

$$D^{\nu}_{11} = \frac{c^{\nu}_{22} F^{(0)\nu} - c^{\nu}_{12} F^{(1)\nu}}{c^{\nu}}, \qquad D^{\nu}_{12} = \frac{c^{\nu}_{22} (F^{(0)\nu})' - c^{\nu}_{12} (F^{(1)\nu})'}{c^{\nu}}, \qquad D^{\nu}_{13} = \frac{c^{\nu}_{12} G^{(1)\nu} - c^{\nu}_{22} G^{(0)\nu}}{c^{\nu}},$$

$$D^{\nu}_{21} = \frac{c^{\nu}_{11} F^{(1)\nu} - c_{21} F^{(0)\nu}}{c^{\nu}}, \qquad D^{\nu}_{22} = \frac{c_{11} (F^{(1)\nu})' - c_{21} (F^{(0)\nu})'}{c^{\nu}}, \qquad D^{\nu}_{23} = \frac{c^{\nu}_{21} G^{(0)\nu} - c^{\nu}_{11} G^{(1)\nu}}{c^{\nu}},$$

with $c^{\nu} = c^{\nu}_{11} c^{\nu}_{22} - c^{\nu}_{12} c^{\nu}_{21}$ and $F^{(r)\nu}$, $G^{(r)\nu}$ longitudinal components of $\mathbf{F}^{(r)\nu}$, $\mathbf{G}^{(r)\nu}$, $r = 0, 1$.

The system (17)–(18) must be coupled with the Schrödinger–Poisson block. Indeed the energy-transport formulation can be given also in the presence of the roughness scattering, but in the applications we will show, such a scattering is not present and therefore for the sake of conciseness we skip it. The limits of validity of approximating the Boltzmann equation with the energy-transport model have been assessed in [33] in the fully semiclassical case.

7. Numerical simulations

In order to asses the validity of the proposed model, a quantum silicon diode is simulated (see Fig. 1). The choice of the device is motivated by the fact that in the present article the main emphasis is on the model rather than on the numerical issues. In particular we are interested to see the differences in the description of the transport between the two considered kinds of energy bands. The diode of Fig. 1 allows us to introduce some analytical approximations regarding the confinement and to reduce the computational effort. Of course the quantization along the transversal direction is likely to be the dominant effect, but this can be handled in the general with approaches already known in the literature (see for example [34,35]). More complex 2-D cases, like MOSFETs and DG-MOSFETs, are under current investigation and will be presented elsewhere. We assume that the oxide gives rise to an infinitely deep potential barrier. This justifies the use of the following analytical expressions as bottom energies and envelope functions

$$\varepsilon_{\nu} = \frac{\nu^2 \pi^2 \hbar^2}{2L^2 m^*}, \qquad \varphi_{\nu}(z) = \sqrt{\frac{2}{L}} \sin \frac{\nu\pi}{L} z, \quad z \in \left[-\frac{L}{2}, \frac{L}{2}\right], \ \nu = 1, 2, \ldots$$

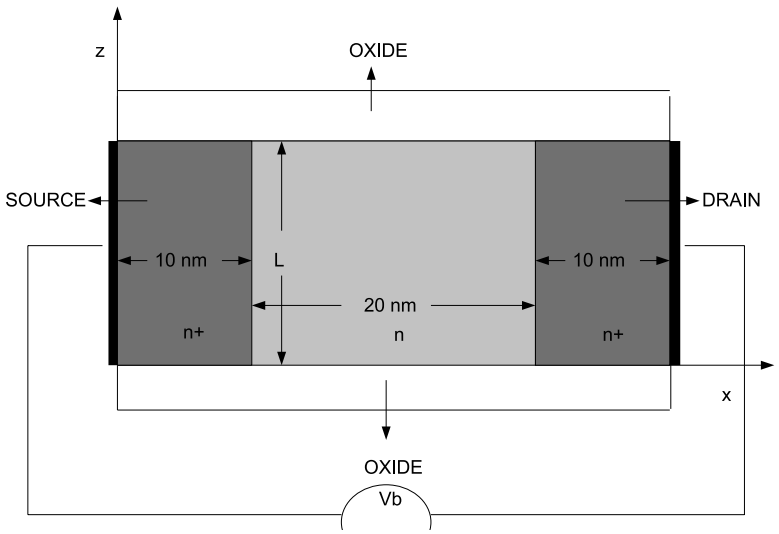

Fig. 1. Simulated diode.

and the approximation of the driving potential with the mean electrostatic potential

$$\overline{V}(x) = \frac{1}{L}\int_{-\frac{L}{2}}^{\frac{L}{2}} V(x,z)\,dz.$$

By taking homogeneous Neumann boundary conditions at the Si/Si–O_2 interfaces, $\overline{V}(x)$ satisfies the 1-D Poisson equation

$$L\epsilon\overline{V}(x)_{xx} = -q\left(\overline{N}_D(x) - \sum_{\nu}\rho^{\nu}(x)\right), \tag{19}$$

where

$$\overline{N}_D(x) = \int_{-\frac{L}{2}}^{\frac{L}{2}} N_D(x,z)\,dz,$$

with N_D donor doping profile. Therefore we have to solve Eqs. (17), (18) with $-q\overline{V}$ as potential energy instead of ε_ν

$$\frac{\partial\rho^{\nu}}{\partial t} + \nabla_{\mathbf{r}_\parallel}\cdot(\rho^{\nu}\mathbf{V}^{\nu}) = \rho^{\nu}C_{\rho}^{\nu}(\mathbf{W}), \tag{20}$$

$$\frac{\partial\rho^{\nu}W^{\nu}}{\partial t} + \nabla_{\mathbf{r}_\parallel}\cdot(\rho^{\nu}\mathbf{S}^{\nu}) - q\rho^{\nu}\nabla_{\mathbf{r}_\parallel}\overline{V}\cdot\mathbf{V}^{\nu} = \rho^{\nu}C_{W}^{\nu}(\mathbf{W}), \quad \nu = 1,2,\ldots, \tag{21}$$

coupled to Eq. (19).

Six equivalent valleys are considered with a single effective mass $m^* = 0.32m_e$, m_e being the free electron mass. A possible generalization could include both longitudinal and transverse masses. Due to the symmetries of the problem and the boundary conditions, the transverse component of the electric field is zero. As a consequence the surface scattering vanishes and only scattering of electrons with acoustic phonons and non-polar optical phonons will be retained.

The length of the diode is 40 nm and the width is $L = 10$ nm. The doping in the n^+ regions is $N_D(x) = N_D^+ = 10^{20}$ cm^{-3} and in the n region is $N_D(x) = N_D^- = 10^{16}$ 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 $s = 2$ nm, $d = \frac{N_D^+}{2}\left(1 - \frac{N_D^-}{N_D^+}\right)$, $x_1 = 10$ nm and $x_2 = 30$ nm. Several values of the bias voltage, V_b, between source and drain are considered. Homogeneous Neumann boundary conditions at source and drain are assumed:

$$\frac{\partial}{\partial x}\rho^{\nu} = \frac{\partial}{\partial x}W^{\nu} = 0.$$

We note that imposing Dirichlet boundary conditions for the energy at source and drain, although common in the literature, leads to an inconsistency with MC simulations in the semiclassical case [33]. The following initial data

$$\rho^{\nu}(x,0) = \frac{e^{-\varepsilon_\nu/k_BT_L}}{\sum_{\mu} e^{-\varepsilon_\mu/k_BT_L}}\overline{N}_D, \qquad W^{\nu}(x,0) = k_BT_L$$

are taken. The numerical experiments indicate that it is sufficient to take into account only the first three subbands, since the other ones are very scarcely populated. The steady state is reached after about 5 ps.

We use a numerical method based on the scheme proposed in [33]. The Poisson equation (19) is discretized with standard central finite differences, while for each subband the balance equations of areal density and energy are discretized with a Scharfetter–Gummel type scheme as follows. Along the x direction we introduce the grid points x_i, $i = i_a, \ldots, i_b$, with $x_{i+1} - x_i = h =$ constant, and the middle points $x_{i\pm1/2} = x_i \pm h/2$. Moreover we take a uniform time step Δt and set $u_i^n = u(x_i, n\Delta t)$, for a generic function u. We discretize the balance equations (17) and (18) in the one dimensional case as

$$\frac{\rho_i^{n+1} - \rho_i^n}{\Delta t} + \frac{(J)_{i+1/2}^n - (J)_{i-1/2}^n}{h} - \rho_i^n C_{\rho_i}^n + O(h^2, \Delta t) = 0, \tag{22}$$

$$\frac{(\rho W)_i^{n+1} - (\rho W)_i^n}{\Delta t} + \frac{Z_{i+1/2}^n - Z_{i-1/2}^n}{h} - \frac{J_{i+1/2}^n + J_{i-1/2}^n}{2}\frac{\overline{V}_{i+1}^n - \overline{V}_{i-1}^n}{2h} - \rho_i^n C_{W_i}^n + O(h^2, \Delta t) = 0. \tag{23}$$

What we need to complete the scheme is an approximation of the currents $\mathbf{J} \equiv \rho\mathbf{V}$ and $\mathbf{Z} \equiv \rho\mathbf{S}$. The key point is the observation that in each subband the current density $\mathbf{J}$ and the energy-flux density $\mathbf{Z}$ can be rewritten as (the subband index is omitted)

$$\mathbf{J} = \mathbf{J}^{(1)} - \mathbf{J}^{(2)}, \qquad \mathbf{Z} = \mathbf{Z}^{(1)} - \mathbf{Z}^{(2)}, \tag{24}$$

where each term has a *drift-diffusion form*

$$\mathbf{J}^{(1)} = \frac{c_{22}}{c}\left[\nabla_{\mathbf{r}_\parallel}(\rho F^{(0)}) - \rho\lambda_W F^{(0)}\nabla_{\mathbf{r}_\parallel}\overline{V}\right], \qquad \mathbf{J}^{(2)} = \frac{c_{12}}{c}\left[\nabla_{\mathbf{r}_\parallel}(\rho F^{(1)}) - \rho\lambda_W F^{(1)}\nabla_{\mathbf{r}_\parallel}\overline{V}\right], \tag{25}$$

$$\mathbf{Z}^{(1)} = \frac{c_{11}}{c}\left[\nabla_{\mathbf{r}_\parallel}(\rho F^{(1)}) - \rho\lambda_W F^{(1)}\nabla_{\mathbf{r}_\parallel}\overline{V}\right], \qquad \mathbf{Z}^{(2)} = \frac{c_{21}}{c}\left[\nabla_{\mathbf{r}_\parallel}(\rho F^{(0)}) - \rho\lambda_W F^{(0)}\nabla_{\mathbf{r}_\parallel}\overline{V}\right], \tag{26}$$

such a form being evident if one identifies $\frac{c_{ij}}{c}F^{(r)}$, $r = 0, 1$, as generalized mobilities and $q\lambda_W$ as the inverse of a sort of thermal potential. The previous considerations allow us to discretize $\mathbf{J}^{(i)}$ and $\mathbf{Z}^{(i)}$, $i = 1, 2$, in the 1D case as (see [33] for more details)

$$(J^{(r)})_{i+1/2} = -\sigma_{i+1/2}\coth\sigma_{i+1/2}\frac{(g_{1r})_{i+1} - (g_{1r})_i}{h} + \sigma_{i+1/2}\frac{(g_{1r})_{i+1} + (g_{1r})_i}{h}, \quad r = 1, 2 \tag{27}$$

$$(Z^{(r)})_{i+1/2} = -\sigma_{i+1/2}\coth\sigma_{i+1/2}\frac{(g_{2r})_{i+1} - (g_{2r})_i}{h} + \sigma_{i+1/2}\frac{(g_{2r})_{i+1} + (g_{2r})_i}{h}, \quad r = 1, 2 \tag{28}$$

where $\sigma_{i+1/2} = \frac{\overline{V}_{i+1} - \overline{V}_i}{2q\overline{U}_T}$. $U_T = \frac{1}{q\lambda_W}$ plays the role of a thermal potential in the drift-diffusion model and we indicate by $\overline{U}_T$ its piecewise constant approximation, $\overline{U}_T = \frac{1}{2}\left[\frac{1}{\lambda_W(W_{i+1})} + \frac{1}{\lambda_W(W_i)}\right]$ in the cell $I_{i+1/2} = [x_i, x_{i+1}]$. g_{ij} are the *local* mobilities

$$g_{11} = \frac{\overline{c}_{22}}{\overline{c}}\rho F^{(0)}, \qquad g_{12} = \frac{\overline{c}_{12}}{\overline{c}}\rho F^{(1)}, \tag{29}$$

$$g_{21} = \frac{\overline{c}_{11}}{\overline{c}}\rho F^{(1)}, \qquad g_{22} = \frac{\overline{c}_{21}}{\overline{c}}\rho F^{(0)}, \tag{30}$$

where $\overline{c}_{pq}$ is a piecewise constant approximation of c_{pq}, $p, q = 1, 2$, given by $\overline{c}_{pq} = c_{pq}(W_i)$ in the cell $I_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$, and $\overline{c} = \overline{c}_{11}\overline{c}_{22} - \overline{c}_{12}\overline{c}_{21}$. The local spatial error in formulas (27), (28) is $O(h^2)$ for sufficiently smooth solutions.

In order to study also the global error and to choose an appropriate number of grid points, we compare the results obtained with 16, 32 and 64 cells (17, 33, 65 grid points respectively) in the case $V_b = 0.3$ V. In Figs. 2 and 3 the results for the subband areal densities and energies, measured from the subband energy bottoms, in the first three subbands are plotted. One can see that the difference between the solutions with 32 and 64 cells is much smaller than that between 16 and 32 cells, indicating a good performance of the numerical scheme. An estimate of the convergence rate c_R, whose theoretical value is 2, is given by the relation

$$c_R = \log_2\frac{\|u_{(16)} - u_{(32)}\|_p}{\|u_{(32)} - u_{(64)}\|_p}$$

where $\|\cdot\|_p$ indicates a suitable norm and $u_{(k)}$ represents the solution with k cells. In Table 1 the estimated convergence c_R for the areal densities and energies is reported for several types of norms: norm 1, norm 2, and infinity norm. c_R is always better then the theoretical value. On the basis of the previous considerations, a spatial grid of 65 points has been used in the other simulations.

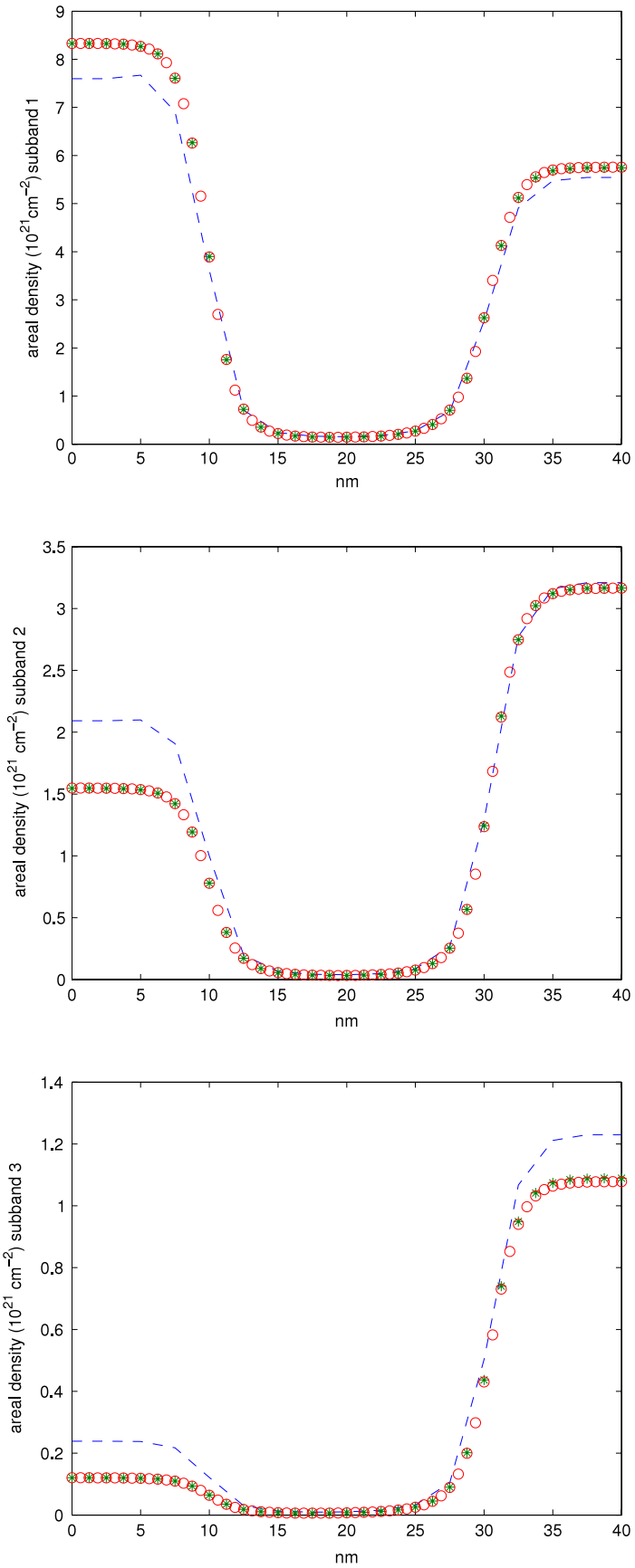

Fig. 2. Areal subband densities for $V_b = 0.3$ V by using different grids: "–" 16 cells, "∗" 32 cells, "∘" 64 cells.

The results are plotted in Figs. 4–8. In Fig. 4 we plot the reconstruction of the electron density from the surface density and the envelope function for $V_b = 0.5$ V. In Figs. 5–8 the densities, the velocities, the energies, measured from the bottom of the first subband, and the currents in the first three subbands are shown. The surface density in the third subband is

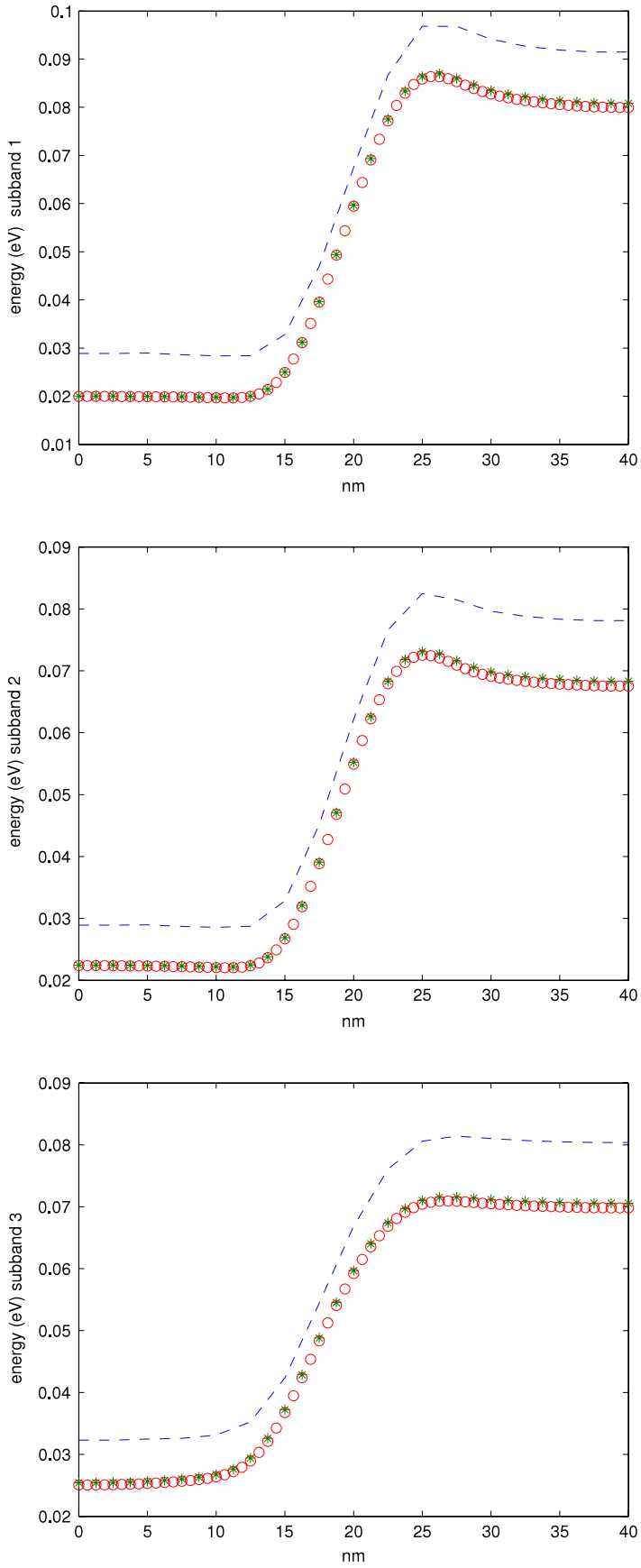

Fig. 3. Subband energies measured from the subband bottoms for $V_b = 0.3$ V by using different grids: "–" 16 cells, "∗" 32 cells, "∘" 64 cells.

about 2% of the total surface density as a confirmation that the inclusion of further subbands has a negligible effect. It is worth mentioning that the energy has an evidently different value between source and drain as happens in the semiclassical

Table 1
Convergence rate.

Norm	ρ^1	ρ^2	ρ^3	W^1	W^2	W^3
$p = 1$	6.7229	6.5867	6.4355	6.5241	6.3030	6.0440
$p = 2$	4.8566	4.6097	4.4837	4.5309	4.2849	4.0033
$p = \infty$	3.7625	3.0708	3.0903	2.8710	2.5853	2.0973

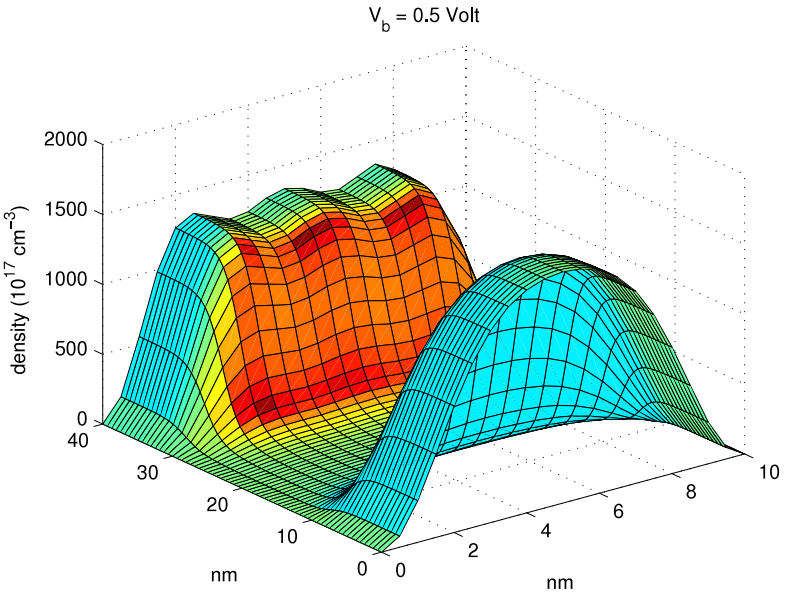

Fig. 4. 2-D reconstruction of the density for $V_b = 0.5$ V.

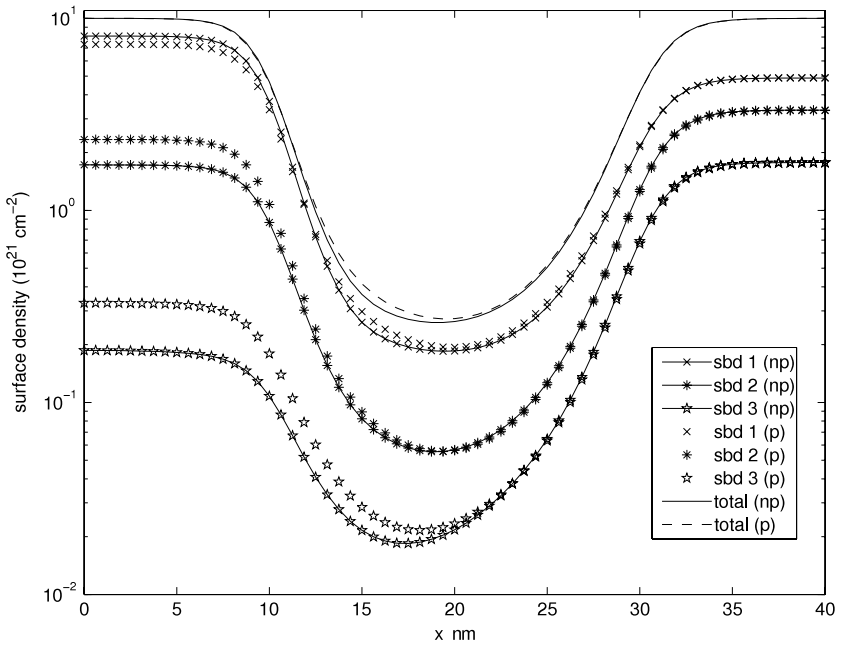

Fig. 5. Total and subband densities for $V_b = 0.5$ V. (p) stands for parabolic model, (np) stands for non parabolic model, (sbd) stands for subband.

case. The use of Dirichlet conditions for the energy at the contacts misses such an effect. We also observe that the current conservation is very accurate, proving the robustness of the numerical method.

The mean velocity, evaluated according to the formula

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

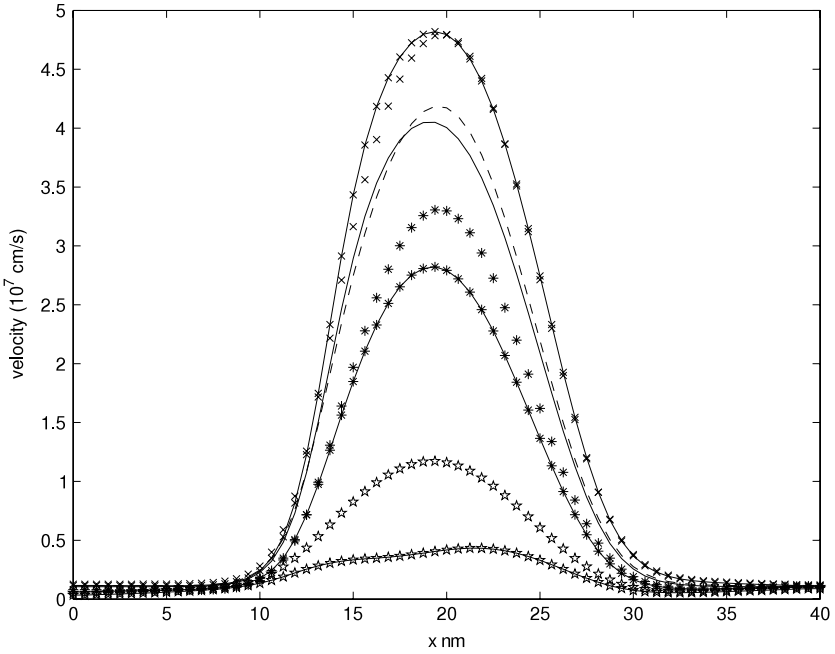

Fig. 6. Velocity for $V_b = 0.5$ V. The legend is the same as the previous figure.

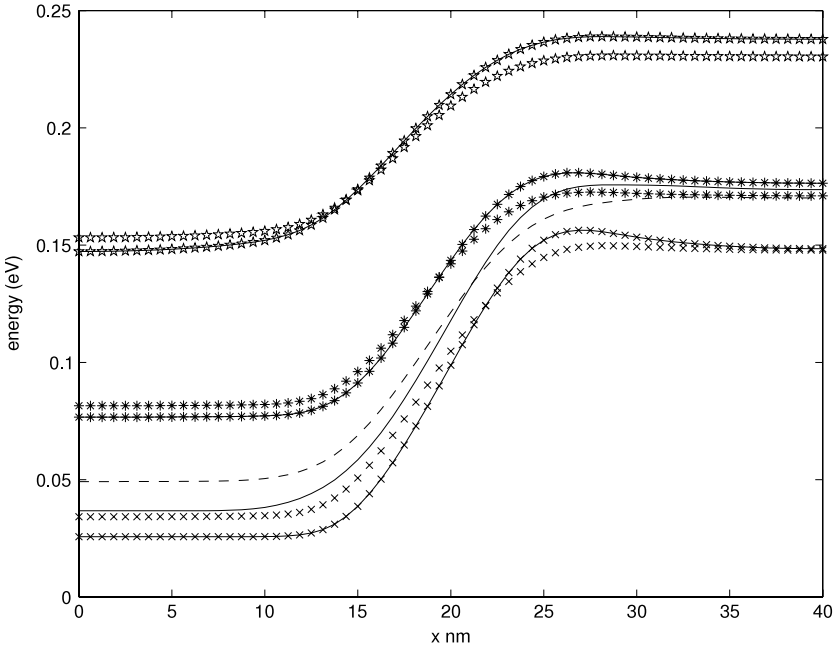

Fig. 7. Energy (eV) for $V_b = 0.5$ V. The legend is as in Fig. 5.

is shown along with the mean energy

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

One observes that the maximum velocity in the channel is up to more than four times the saturation velocity, but the differences between the parabolic and non parabolic bands are less than in the semiclassical case [33]. The more evident difference in the velocity is in the second and third subband.

In Figs. 9 and 10 the mean velocity and energy are plotted with bias voltage varying from 0.1 to 0.5 V. We notice that the peak in the velocity lowers if the bias overcomes a certain threshold value, showing a negative differential conductivity which is typical of the presence of more subbands and it is due to the inter-subband scattering. In fact according to (4), the

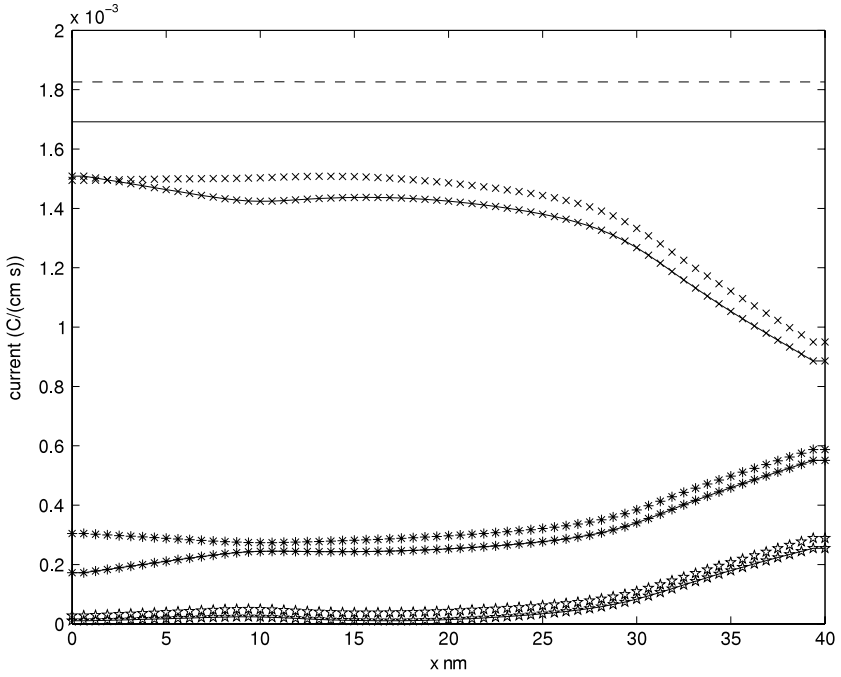

Fig. 8. Current for $V_b = 0.5$ V. The legend is as in Fig. 5.

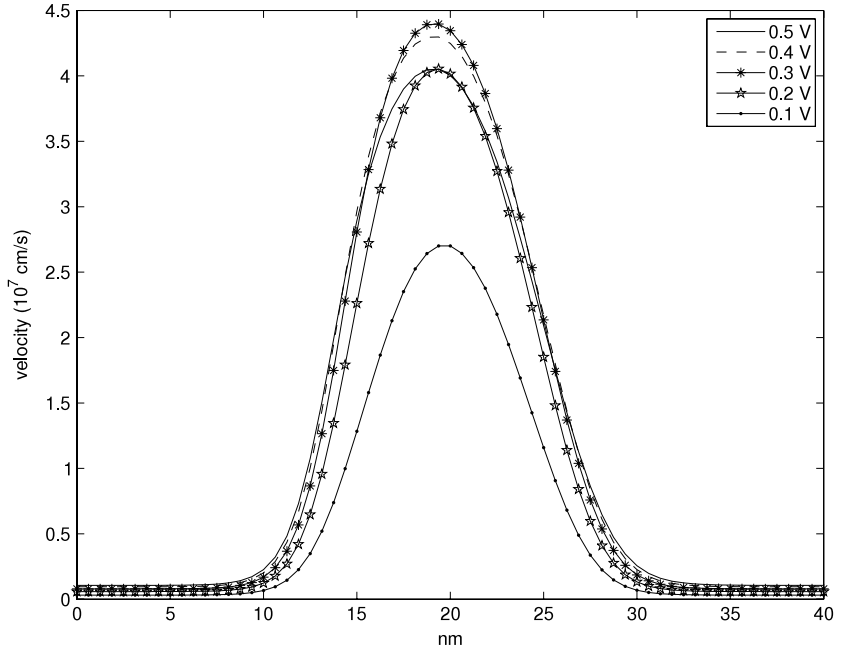

Fig. 9. Mean velocity for $V_b = 0.1, 0.2, 0.3, 0.4, 0.5$ V.

microscopic longitudinal velocity depends only on the longitudinal subband energy

$$|\mathbf{v}_\parallel| = \frac{\sqrt{m^*\left(\varepsilon_\parallel + \alpha\varepsilon_\parallel\right)}}{m^*\left(1 + 2\alpha\varepsilon_\parallel\right)}.$$

At high energy if an electron moves from a subband to a higher one (see Fig. 11), the longitudinal contribution decreases (the absorption or emission energy is less then the gap between subbands) and as a consequence $\mathbf{v}_\parallel$ has a smaller modulus.

Regarding the total current, in the non parabolic band approximation it is about 10% lower than that in the parabolic band one, indicating that the characteristic curves are influenced by the nonparabolicity.

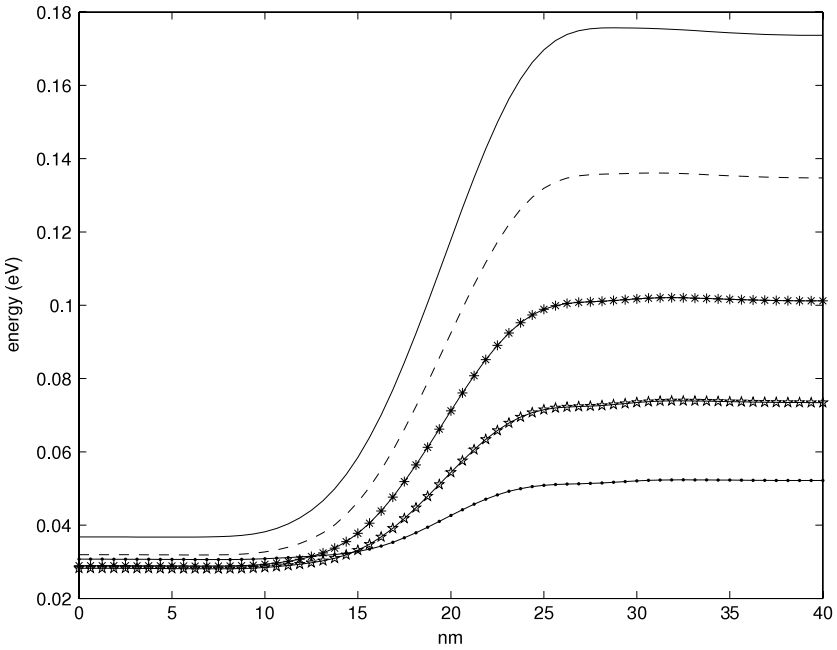

Fig. 10. Mean energy for $V_b = 0.1, 0.2, 0.3, 0.4, 0.5$ V. The legend is the same as in the previous figure.

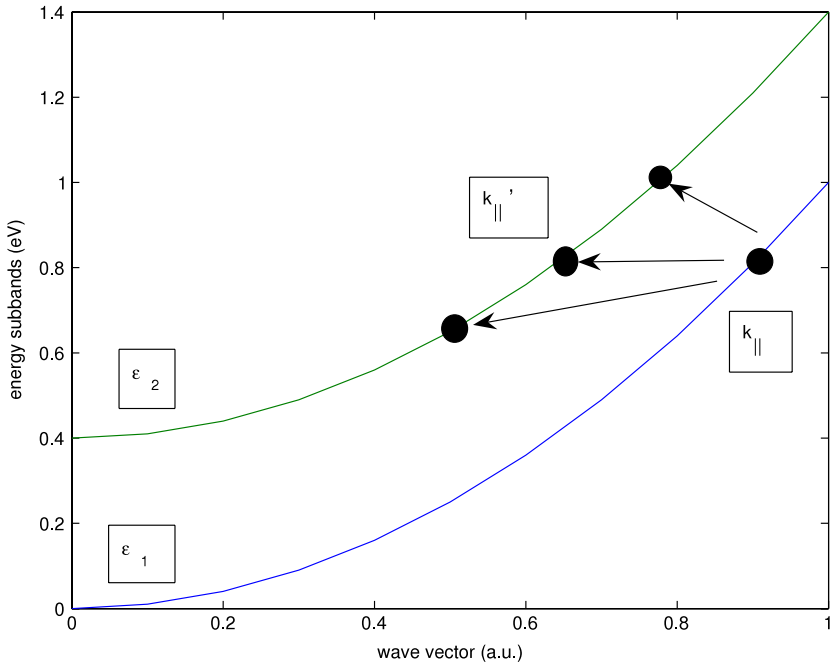

Fig. 11. Schematic representation of an inter subband scattering. Just as example we consider the first two bands. An electron having wave-vector $\mathbf{k}_\parallel$ in subband 1 may move to subband 2 with an elastic scattering, which conserves the energy, or with an absorption scattering, which leads to a final energy $\varepsilon_f = \varepsilon_i + \hbar\omega + \varepsilon_1 - \varepsilon_2$, or with an emission scattering, which leads to a final energy $\varepsilon_f = \varepsilon_i - \hbar\omega + \varepsilon_1 - \varepsilon_2$. In any case $\mathbf{k}'_\parallel$, the wave-vector after the transition, is smaller in modulus than $\mathbf{k}_\parallel$ as well as $\varepsilon_\parallel$ and this implies a smaller $|\mathbf{v}_\parallel|$.

8. Conclusion

An energy-transport subband model has been formulated with the use of the maximum entropy principle by taking into account also the non parabolic band effects. The numerical simulations of a quantum diode, under some simplifying assumptions, give results which capture the main confining effects. A comparison between the parabolic and the non parabolic case is presented and discussed.

More realistic simulations require the set up of a numerical code for the solution of the energy-transport equations coupled to the Schrödinger–Poisson block. This issue is under current investigation by the authors and will be the subject of a forthcoming article.

Acknowledgments

G. M. and V. R. acknowledge the financial support by P.R.A., University of Calabria and University of Catania, respectively. G. M. also acknowledges support by Gruppo Nazionale di Fisica Matematica, Progetto giovani 2011/12.

Appendix

In this section we report the explicit expressions of the functions present in the fluxes and in the productions terms and their limit as $\alpha \to 0$, which lets us retrieve the parabolic case presented in [19]. Let us define

$$\gamma_n(x, y) := \int_0^\infty \frac{t^n}{1 + 2\alpha(t + y)} \exp\left[-\lambda_W(x)(t + y)\right] dt, \quad n = 0, 1, \ldots, \tag{31}$$

where λ_W is the function in (7). Noting that

$$\gamma_n(x, y) = (-)^n \exp\left(-\lambda_W(x)y\right) \frac{\partial^n}{\partial \lambda_W^n} \int_0^\infty \frac{1}{1 + 2\alpha(t + y)} \exp\left[-\lambda_W t\right] dt \Big|_{\lambda_W = \lambda_W(x)},$$

one has

$$\gamma_1(x, y) = -\frac{1}{2\alpha}\gamma_0(x, y) + \frac{\exp\left(-\lambda_W(x)y\right)}{2\alpha\lambda_W(x)},$$

from which the following formula, useful for computation, can be proved

$$\gamma_n(x, y) = \left(-\frac{1}{2\alpha}\right)^n \gamma_0(x, y) + \frac{1}{2\alpha}\left(\sum_{i=0}^{n-1}\left(-\frac{1}{2\alpha}\right)^{n-i-1} \sum_{k=0}^{i} \frac{i!}{k!} \frac{y^k}{(\lambda_W(x))^{i-k}}\right) \frac{e^{-\lambda_W(x)y}}{\lambda_W(x)}. \tag{32}$$

In particular, γ_0 can be expressed in terms of the exponential integral

$$E_n(x') := \int_1^\infty \frac{\exp\left(-x't\right)}{t^n} dt, \quad n = 0, 1, \ldots$$

as

$$\gamma_0(x, y) = \frac{1}{2\alpha} \exp\left(\frac{\lambda_W(x)}{2\alpha}\right) E_1\left(\frac{\lambda_W(x)}{2\bar{y}}\right),$$

with $\bar{y} = \frac{\alpha}{1+2\alpha y}$.

Taking the limit for $\alpha \to 0$, one gets

$$\gamma_n(x, y) \to \lambda_W(x)^{-1-n} \exp\left(-\lambda_W(x)y\right) \Gamma(n + 1),$$

Γ being the gamma function. Analogously, we have

$$\zeta_n(x) := \int_0^\infty \frac{2\alpha t^n}{(1 + 2\alpha t)^2} \exp\left(-\lambda_W(x)t\right) dt = -2\alpha \left(n\gamma_n(x, 0) - \lambda_W \gamma_{n+1}(x, 0)\right). \tag{33}$$

Of course $\lim_{\alpha\to 0} \zeta_n(x) = 0$.

Regarding the $B^{(n)}$'s, one has

$$\begin{aligned} B^{(n)}(x, y, z) &:= \int_0^\infty (t + y)^n \exp\left(-\lambda_W(x)(t + y)\right) \left(1 + 2\alpha(t + y)\left[1 + 2\alpha\left(t + y + z\right)\right]\right) dt \\ &= \sum_{i=0}^{n} (-1)^i \exp\left(-\lambda_W(x)y\right) \binom{n}{n-i} \frac{\partial^i p\left(\lambda_W, y, z\right)}{\partial \lambda_W^i}\Big|_{\lambda_W = \lambda_W(x)} y^{n-i}, \end{aligned} \tag{34}$$

with

$$p\left(x, y, z\right) = \frac{8\alpha^2}{\lambda_W^3(x)} + \frac{4\alpha\left(1 + \alpha\left(2y + z\right)\right)}{\lambda_W^2(x)} + \frac{\left[1 + 2\alpha\left(2y + z\right) + 4\alpha^2\left(y^2 + yz\right)\right]}{\lambda_W(x)}.$$

In the parabolic limit, we find

$$\lim_{\alpha\to 0} p(x) = \frac{1}{\lambda_W(x)}.$$

The functions $\eta^{(n)}(x, y, z)$ are defined by

$$\eta^{(n)}(x, y, z) = (1 + 2\alpha z)\gamma_n(x, y) + (3\alpha + 2\alpha^2 z)\gamma_{n+1}(x, y) + 2\alpha^2\gamma_{n+2}(x, y), \tag{35}$$

and in the parabolic limit

$$\eta^{(n)}(x, y, z) \longrightarrow \lambda_W(x)^{-1-n} \exp\left(-\lambda_W(x)y\right) \Gamma(n + 1). \tag{36}$$

Table 2
Values of the physical parameters.

m_e	Electron rest mass	9.1095×10^{-28} g
m^*	Effective electron mass	$0.32\, m_e$
T_L	Lattice temperature	300 K
ρ	Density	2.33 g/cm^3
v_s	Longitudinal sound speed	9.18×10^5 cm/s
Ξ_d	Acoustic-phonon deformation potential	9 eV
ϵ_r	Si relative dielectric constant	11.7
ϵ_{r_O}	SiO_2 relative dielectric constant	3.9
ϵ_0	Vacuum dielectric constant	8.85×10^{-18} C/V μm

Table 3
Coupling constants and phonon energies for the inelastic scatterings in silicon.

A	Z_f	$\hbar\omega$ (meV)	D_tK (10^8 eV/cm)
1	1	12	0.5
2	1	18.5	0.8
3	4	19.0	0.3
4	4	47.4	2.0
5	1	61.2	11
6	4	59.0	2.0

While the functions $B_j^{(n)}(x, y, z)$ are given by

$$B_1^{(n)}(x,y,z) = \int_0^\infty (t+y)^n e^{-(t+y)\lambda_W(x)} \frac{[1+2\alpha\,(t+y+z)]\,[1+2\alpha\,(t+y)]}{\beta_0(t+y,z)}\,dt, \tag{37}$$

$$B_2^{(n)}(x,y,z) = \int_0^\infty (t+y)^n e^{-(t+y)\lambda_W(x)} \frac{[1+\alpha\,(t+y)]\,[1+2\alpha\,(t+y+z)]}{[1+2\alpha\,(t+y)]\,\beta_0(t+y,z)}\,dt, \tag{38}$$

$$B_3^{(n)}(x,y,z) = \int_0^\infty (t+y)^n e^{-(t+y)\lambda_W(x)} \beta_1(t+y,z)\,dt, \tag{39}$$

with

$$\beta_0(t,z) = \left\{1 + \frac{m^* L_C^2}{\hbar^2}\,[t(1+\alpha t) + (t+z)(1+\alpha(t+z))] + \frac{(m^*)^2 L_C^4}{\hbar^4}\,[t(1+\alpha t) - (t+z)(1+\alpha(t+z))]^2\right\}^{1/2},$$

$$\beta_1(t,z) = \left\{1 + \frac{m^* L_C^2}{\hbar^2}\,[(t+z)(1+\alpha(t+z)) + t(1+\alpha t)] - \beta_0(t,z)\right\}\left(\frac{2m^* L_C^2}{\hbar^2}\beta_0(t,z)\right)^{-1}.$$

We notice that for $\alpha \to 0$

$$\lambda_W \to \frac{1}{W}, \qquad g(W) \to \lambda_W \qquad \lambda \to -\ln\left(\frac{\hbar^2\,\rho}{2\pi m^* W}\right).$$

Eventually, in Tables 2 and 3 we report the values of physical parameters used in the simulations.

References

[1] S. Datta, Quantum Phenomena, in: The Modular Series on Solid State Devices, vol. VIII, Addison-Wesley Publishing, 1989.
[2] M. Lundstrom, Fundamentals of Carrier Transport, Cambridge University Press, 2000.
[3] G.D. Mahan, Many Particles Physics, Kluwer Academic/Plenum Publishers, 2000.
[4] H.J.W. Haug, A.-P. Jauho, Quantum Kinetics in Transport and Optics of Semiconductors, Springer, 2008.
[5] J. Wang, E. Polizzi, M. Lundstrom, A three dimensional quantum simulation of silicon nanowire transistors with the effective-mass approximation, J. Appl. Phys. 96 (2004) 2192–2203.
[6] D. Chen, G.-W. Wei, Modeling and simulation of electronic structure, material interface and random doping in nano-electronic devices, J. Comput. Phys. 229 (2010) 4431–4460.
[7] P. Markowich, C.A. Ringhofer, C. Schmeiser, Semiconductor Equations, Springer, Wien, 1990.
[8] A. Jüngel, Transport Equations for Semiconductors, in: Lect. Notes in Phys., vol. 773, Springer, 2009.
[9] V. Romano, Quantum corrections to the semiclassical hydrodynamical model of semiconductors based on the maximum entropy principle, J. Math. Phys. 48 (2007) 123504.
[10] M.V. Fischetti, Master equation approach to the study of electronic transport in small semiconductor devices, Phys. Rev. B 59 (1999) 4901–4917.
[11] T. Ando, B. Fowler, F. Stern, Electronic properties of two-dimensional systems, Rev. Modern Phys. 54 (1982) 437–672.
[12] E. Polizzi, N. Ben Abdallah, Self-consistent three dimensional models for quantum ballistic transport in open systems, Phys. Rev. B 66 (2002) 245301-1–245301-9.
[13] E. Polizzi, N.B. Abdallah, Subband decomposition approach for the simulation of quantum electron transport in nanostructures, J. Comput. Phys. 202 (1) (2004) 150–180.

[14] M. Galler, F. Schuerrer, A deterministic solver to the Boltzmann–Poisson system including quantization effects for silicon-MOSFETs, in: Progress in Industrial Mathematics at ECMI 2006, in: Mathematics in Industry, Springer, 2008, pp. 531–536.
[15] N.B. Abdallah, M.J. Caceres, J.A. Carrillo, F. Vecil, A deterministic solver for a hybrid quantum-classical transport model in nanoMOSFETs, J. Comput. Phys. 228 (2009) 6553.
[16] A. Majorana, O. Muscato, C. Milazzo, Charge transport in 1D silicon devices via Monte Carlo simulation and Boltzmann–Poisson solver, COMPEL 23 (2) (2004) 410–425.
[17] C. De Falco, E. Gatti, A. Lacaita, R. Sacco, Quantum-corrected drift-diffusion models for transport in semiconductor devices, J. Comput. Phys. 204 (2004) 533.
[18] N. Ben Abdallah, F. Méhats, N. Vauchelet, Diffusive transport of partially quantized particles: existence, uniqueness and long-time behaviour, Proc. Edinb. Math. Soc. 2 (49) (2006) 513–549.
[19] G. Mascali, V. Romano, Hydrodynamic subband model for semiconductors based on the maximum entropy principle, II Nuovo Cimento C 33 (2010) 155.
[20] G. Mascali, V. Romano, Numerical simulation of a subband model based on the maximum entropy principle, in: Proceedings of the 14th International Workshop on Computational Electronics, IWCE, 2010, pp. 1–4.
[21] E.T. Jaynes, Information theory and statistical mechanics, Phys. Rev. B 106 (1957) 620.
[22] N. Wu, The Maximum Entropy Method, Springer-Verlag, 1997.
[23] G. Mascali, V. Romano, Simulation of Gunn oscillations with a non-parabolic hydrodynamical model based on the maximum entropy principle, COMPEL 24 (1) (2005) 35–54.
[24] G. Mascali, M. Trovato, A non-linear determination of the distribution function of degenerate gases with an application to semiconductors, Physica A 310 (1–2) (2002) 121–138.
[25] S. La Rosa, G. Mascali, V. Romano, Exact maximum entropy closure of the hydrodynamical model for Si semiconductors: the 8-moment case, SIAM J. Appl. Math. 70 (2009) 710.
[26] A.M. Anile, V. Romano, Non parabolic band transport in semiconductors: closure of the moment equations, Contin. Mech. Thermodyn. 11 (1999) 307.
[27] V. Romano, Non parabolic band transport in semiconductors: closure of the production terms in the moment equations, Contin. Mech. Thermodyn. 12 (2000) 31.
[28] G. Mascali, V. Romano, Hydrodynamical model of charge transport in GaAs based on the maximum entropy principle, Contin. Mech. Thermodyn. 14 (2002) 405.
[29] O. Muscato, The Onsager reciprocity principle as a check of consistency for semiconductor carrier transport models, Physica A 289 (3–4) (2001) 422–458.
[30] V. Romano, Nonparabolic band hydrodynamical model of silicon semiconductors and simulation of electron devices, Math. Methods Appl. Sci. 24 (2001) 439–471.
[31] G. Mascali, V. Romano, Si and GaAs mobility derived from a hydrodynamical model for semiconductors based on the maximum entropy principle, Physica A 352 (2005) 459–476.
[32] G. Mascali, V. Romano, A hydrodynamical model for holes in silicon semiconductors: the case of non-parabolic warped bands, Math. Comput. Modelling 53 (01–2) (2011) 213–229.
[33] V. Romano, 2D numerical simulation of the MEP energy-transport model with a finite difference scheme, J. Comput. Phys. 221 (2007) 439.
[34] A.S. Spinelli, A. Benvenuti, A. Pacelli, Self-consistent 2-D model for quantum effects in *n*-MOS transistor, IEEE Trans. Electron Devices 45 (6) (1998) 1342–1349.
[35] G. Curatola, G. Doornbos, J. Loo, Y.V. Ponomarev, G. Iannaccone, Detailed modeling of sub-100-nm MOSFETS based on Scrödinger DD per subband and experiments and evaluation of the performance gap to ballistic transport, IEEE Trans. Electron Devices 52 (8) (2005) 1851–1858.