Finite-element discretizations of semiconductor energy-transport equations

Stefan Holst, Ansgar Jüngel and Paola Pietra

Abstract. Energy-transport models describe the flow of electrons in a semiconductor crystal. Several formulations of these models, in the primal or dual entropy variables or in the drift-diffusion-type variables, are reviewed. A numerical discretization of the steady-state drift-diffusion-type formulation using mixed-hybrid finite elements introduced by Marini and Pietra is presented. The scheme is first applied to the simulation of a one-dimensional ballistic diode with non-parabolic band diagrams. Then a two-dimensional deep submicron MOSFET device with parabolic bands is simulated, using an adaptively refined mesh.

1. Introduction

Energy-transport models are used in the modeling of submicron semiconductor devices to describe the charge transport of electrons in the semiconductor crystal. The models consist of the continuity equations for the mass and energy of the electrons and are coupled to the Poisson equation for the electrostatic potential. The energy-transport equations are more complex than the drift-diffusion equations since they include an energy equation and are thus capable of modeling hot electron effects in submicron devices. On the other hand, they are less complex than the hydrodynamic models which contain hyperbolic modes. This intermediate complexity makes the energy-transport system appropriate for fast and accurate semiconductor applications.

In this paper, we review several formulations of energy-transport models. Furthermore, we describe an adaptively refined mixed finite-element scheme of the drift-diffusion formulation of the energy-transport model and simulate two submicron semiconductor devices.

1991 *Mathematics Subject Classification.* Primary 65N30; Secondary 82D37, 78A35.
Key words and phrases. Mixed finite elements, exponential fitting, adaptive scheme, two-dimensional, non-parabolic bands, ballistic diode, submicron MOSFET, semiconductors.
The authors acknowledge partial support from the IHP Project HPRN-CT-2002-00282 of the European Union and from the German-Italian DAAD-Vigoni Project. The first and second author are partly supported by the Gerhard-Hess Program of the Deutsche Forschungsgemeinschaft, grant JU395/3.

In the physical literature, energy-transport equations, first presented by Stratton in 1962 [37], are been derived from hydrodynamic models usually by neglecting certain convection terms (see, e.g., [33] and references therein). Another approach is to derive energy-transport models from the Boltzmann equation by means of the Hilbert expansion method [2]. Denoting the chemical potential, the electron temperature and the electrostatic potential by μ, T and V, respectively, the scaled stationary equations in the so-called *primal entropy variables* μ/T and $-1/T$ read as follows:

$$-\mathrm{div} J_1 = 0, \tag{1.1}$$

$$-\mathrm{div} J_2 = -J_1 \cdot \nabla V + W(\mu, T), \tag{1.2}$$

$$J_1 = L_{11}\Big(\nabla \frac{\mu}{T} - \frac{\nabla V}{T}\Big) + L_{12} \nabla \Big(\frac{-1}{T}\Big), \tag{1.3}$$

$$J_2 = L_{21}\Big(\nabla \frac{\mu}{T} - \frac{\nabla V}{T}\Big) + L_{22} \nabla \Big(\frac{-1}{T}\Big), \tag{1.4}$$

$$\lambda^2 \Delta V = n - C(x). \tag{1.5}$$

Here, J_1 and J_2 are the particle and energy current densities, respectively, $n = n(\mu, T)$ is the electron density, and $W(\mu, T)$ is the energy relaxation term. The diffusion coefficients $L_{ij} = L_{ij}(\mu, T)$ form a symmetric, positive definite (2×2) matrix. The precise dependence of these dimensionless functions on μ and T will be given in section 2.1. The space-dependent function $C = C(x)$ is the doping profile modeling fixed charged background ions in the crystal, and the positive constant λ is the scaled Debye length.

The above equations have to be solved in a bounded domain $\Omega \subset \mathbb{R}^d$ $(d \leq 3)$. We complement the equations with physically motivated mixed Dirichlet-Neumann boundary conditions:

$$n = n_D, \quad T = T_D, \quad V = V_D \qquad \text{on } \Gamma_D,$$
$$J_1 \cdot \nu = J_2 \cdot \nu = \nabla V \cdot \nu = 0 \qquad \text{on } \Gamma_N,$$

modeling the contacts Γ_D and the insulating boundary parts Γ_N. We have assumed that $\partial\Omega = \Gamma_D \cup \Gamma_N$ and $\Gamma_D \cap \Gamma_N = \emptyset$. The exterior normal unit vector on $\partial\Omega$ is denoted by ν. In the mathematical literature, there are only a few results concerning the existence and uniqueness of solutions to (1.1)-(1.5) (see, e.g., [11, 12, 14, 17]), proving existence of solutions for uniformly bounded diffusion coefficients L_{ij} or in near-equilibrium situations.

The numerical discretization of energy-transport models has been investigated in the physical literature for quite some years [1, 9, 15, 36, 38]. Mathematicians started to pay attention to these models in the 1990s, using different numerical methods. Jerome and Shu employed ENO (essentially non-oscillatory) numerical schemes [22]. Finite-difference discretizations have been used by Ringhofer [32], based on an entropy-decaying property, and by Fournié [16], with compact schemes of fourth order. Bosisio et al. used mixed finite-volume techniques [3]. Mixed finite-element methods applied to different formulations of the energy-transport models have been employed in [13, 18, 19, 25, 29, 30] (see [5] for an overview).

In the first part of the paper we make precise the definitions of L_{ij}, n and W as functions of μ and T for general and parabolic band structures (section 2.1) and give three different formulations of the energy-transport model (i) in the primal entropy variables (see above), (ii) in the dual entropy variables, and (iii) in the drift-diffusion-type variables. We explain the advantages and disadvantages regarding the numerical approximation of the equations for each of these formulations (section 2.2).

The second part of the paper is devoted to the numerical approximation of the equations in the drift-diffusion formulation using mixed-hybrid finite elements introduced by Marini and Pietra [26, 27] (section 3). The resulting algebraic system is solved employing a variant of the Gummel method. In the two-dimensional situation we adaptively refine the mesh using an error estimator motivated by the results of [20] and [39]. Finally, we apply the resulting scheme to the numerical simulation of a 1-D ballistic diode with non-parabolic bands (section 4.1) and to a 2-D MOSFET device with parabolic bands (section 4.2).

2. Formulation of the model

2.1. Physical assumptions and non-parabolic band approximations

In order to give precise definitions for the coefficients L_{ij} in terms of μ and T, we impose the following physical assumptions:

- The energy band diagram ε of the semiconductor crystal is spherically symmetric and a strictly monotone function of the modulus $k = |\vec{k}|$ of the wave vector $\vec{k}$.
- The momentum relaxation time is defined by
$$\tau(\varepsilon) = \left(\phi_0 \varepsilon^\beta N(\varepsilon)\right)^{-1}, \quad \beta > -2,$$
where $\phi_0 > 0$ is some constant and $N(\varepsilon) = 4\pi k^2/|\varepsilon'(k)|$ is the density of states of energy $\varepsilon = \varepsilon(k)$ [2].
- The electron density n is given by non-degenerate Boltzmann statistics.

For the physical background, we refer to the textbooks [4, 35], for instance.

The energy-transport model (1.1)-(1.5) is completed by giving the definitions of the functions $L_{ij}(\mu, T)$, $n(\mu, T)$, and $W(\mu, T)$. In [2] the following expressions have been derived:

$$L_{ij}(\mu, T) = e^{\mu/T} \int_0^\infty d(\varepsilon)\varepsilon^{i+j-2} e^{-\varepsilon/T} d\varepsilon, \quad i, j = 1, 2, \tag{2.1}$$

$$n(\mu, T) = e^{\mu/T} \int_0^\infty e^{-\varepsilon/T} N(\varepsilon) d\varepsilon, \tag{2.2}$$

where

$$d(\varepsilon) = \frac{4\pi}{3}\tau(\varepsilon)|\varepsilon'(k)|k^2 \quad \text{and} \quad \varepsilon = \varepsilon(k).$$

The energy relaxation term in the Fokker-Planck approximation of the phonon collision operator is given by [13]

$$W(\mu, T) = -\phi_1 e^{\mu/T} \int_0^\infty \varepsilon^\beta N(\varepsilon)^2 e^{-\varepsilon/T} d\varepsilon \left(1 - \frac{T_0}{T}\right),$$

where $T_0 > 0$ is the (constant) lattice temperature and $\phi_1 > 0$ is some constant.

The advantages of the model are that the assumptions on the band diagram are rather general (the above hypotheses can be even weakened, see [2]) and that *analytic* expressions can be obtained under the assumption of non-parabolic bands in the sense of Kane [24], where $\varepsilon(k)$ is defined by

$$\varepsilon(1 + \alpha\varepsilon) = k^2/2,$$

and $\alpha > 0$ is the non-parabolicity parameter. We call the band structure *parabolic* if $\alpha = 0$; in this situation $\varepsilon(k) = k^2/2$.

For non-parabolic bands in the sense of Kane, we can rewrite the expressions for L_{ij}, n and W in terms of the following integrals:

$$L_{ij}(\mu, T) = \frac{2}{3\phi_0} T^{i+j-\beta} e^{\mu/T} \int_0^\infty \frac{1 + \alpha T u}{(1 + 2\alpha T u)^2} u^{i+j-\beta-1} e^{-u} du, \tag{2.3}$$

$$n(\mu, T) = 2\pi(2T)^{3/2} e^{\mu/T} \int_0^\infty (1 + \alpha T u)^{1/2}(1 + 2\alpha T u) u^{1/2} e^{-u} du, \tag{2.4}$$

$$\begin{aligned} W(\mu, T) &= 8(2\pi)^2 \phi_1 e^{\mu/T} T^{1+\beta}(T_0 - T) \\ &\quad \times \int_0^\infty (1 + \alpha T u)(1 + 2\alpha T u)^2 u^{1+\beta} e^{-u} du. \end{aligned} \tag{2.5}$$

We obtain *explicit* expressions for parabolic band diagrams. In fact, we have just to set $\alpha = 0$ in the above equations which yields

$$\begin{aligned} L_{ij}(\mu, T) &= \frac{2}{3\phi_0} \Gamma(i + 1 - \beta) T^{i+j-\beta} e^{\mu/T}, \\ n(\mu, T) &= (2\pi)^{3/2} T^{3/2} e^{\mu/T}, \\ W(\mu, T) &= 8(2\pi)^2 \phi_1 \Gamma(2 + \beta) T^{1+\beta} e^{\mu/T}(T_0 - T), \end{aligned} \tag{2.6}$$

where Γ denotes the Gamma function. Chen et al. [9] used $\beta = 1/2$ in the simulation of MOS devices. Then it follows

$$(L_{ij}) = \mu_0 n \begin{pmatrix} 1 & 3T/2 \\ 3T/2 & 15T^2/4 \end{pmatrix} \tag{2.7}$$

and

$$W(\mu, T) = \frac{3}{2} \frac{n(T_0 - T)}{\tau_0}, \tag{2.8}$$

where $\mu_0 = 1/(6\sqrt{2\pi}\phi_0)$ and $\tau = 1/(4\sqrt{2}\pi\phi_1)$. Generally, we call an energy-transport model, in which the expressions (2.3)-(2.5) and $\beta = 1/2$ are used, the *Chen model*.

2.2. Formulations of the energy-transport models

We describe three different formulations of the energy-transport equations:

- formulation in the *primal entropy variables* μ/T, $-1/T$ (see (1.1)-(1.5)),
- formulation in the *dual entropy variables* $(\mu - V)/T$, $-1/T$, and
- *drift-diffusion* formulation in the variables L_{11}, L_{12}.

The first formulation has been already presented in section 1. The second formulation of the energy-transport equations equations is obtained by introducing the *dual entropy variables*

$$w_1 = \frac{\mu - V}{T}, \quad w_2 = -\frac{1}{T},$$

yielding the system

$$-\mathrm{div} I_1 = 0, \qquad I_1 = D_{11}\nabla w_1 + D_{12}\nabla w_2, \tag{2.9}$$

$$-\mathrm{div} I_2 = W(\mu, T), \qquad I_2 = D_{21}\nabla w_1 + D_{22}\nabla w_2, \tag{2.10}$$

where the new diffusion coefficients are given by

$$D_{11} = L_{11}, \quad D_{12} = D_{21} = L_{12} - VL_{11}, \quad D_{22} = L_{22} - 2VL_{12} + V^2 L_{11}.$$

The diffusion matrix consisting of the coefficients D_{ij}, $i, j = 1, 2$, is again symmetric and positive definite. In this formulation, the Joule heating term $J_1 \cdot \nabla V$ is eliminated and the problem is symmetrized. The advantage of this formulation is twofold. First, the theory of systems of quasilinear elliptic equations can be directly applied. This has been used in [11, 12] in order to prove the existence of solutions to the transient and stationary equations (however, under the assumption of *uniformly bounded* coefficients $D_{ij}(w)$). Second, the system can be discretized by a mixed finite-element method. However, as the diffusion matrix is not diagonal, the resulting nonlinear equations have to be solved generally using the Newton method which makes the final code unflexible to changes of the parameters or the semiconductor band structure.

This disadvantage can be avoided in the third formulation which is given in the *drift-diffusion variables* g_1 and g_2:

$$-\mathrm{div} J_1 = 0, \tag{2.11}$$

$$-\mathrm{div} J_2 = -J_1 \cdot \nabla V + W(\mu, T), \tag{2.12}$$

$$J_i = \nabla g_1 - g_i \frac{\nabla V}{T}, \quad i = 1, 2, \tag{2.13}$$

where $g_1 = L_{11}$ and $g_2 = L_{12}$. Indeed, from (1.3)-(1.4) and (2.1) follows

$$J_i = \int_0^\infty d(\varepsilon) \left(\nabla e^{(\mu-\varepsilon)/T} + \nabla V \frac{\partial}{\partial \varepsilon} e^{(\mu-\varepsilon)/T} \right) \varepsilon^{i-1} d\varepsilon$$

$$= \nabla \int_0^\infty d(\varepsilon) e^{(\mu-\varepsilon)/T} \varepsilon^{i-1} d\varepsilon - \frac{\nabla V}{T} \int_0^\infty d(\varepsilon) e^{(\mu-\varepsilon)/T} \varepsilon^{i-1} d\varepsilon, \quad i = 1, 2,$$

which equals (2.13), setting

$$g_1 = \int_0^\infty d(\varepsilon) e^{(\mu-\varepsilon)/T} d\varepsilon, \quad g_2 = \int_0^\infty d(\varepsilon) e^{(\mu-\varepsilon)/T} \varepsilon d\varepsilon,$$

and from (2.1) follows the assertion.

We have to show that n and T can be uniquely defined as functions of g_1 and g_2. We define the function $f(T) = L_{12}(\mu, T)/L_{11}(\mu, T) = g_2/g_1$. In [13] it is shown that the function f can be always inverted since $f'(T) = \det(L_{ij})_{ij}/(Tg_1)^2 > 0$. Thus, the function $T = T(g_1, g_2) = f^{-1}(g_2/g_1)$ can be uniquely defined. Now, the density n is defined through (2.2), i.e.

$$n(g_1, g_2) = g_1 \int_0^\infty N(\varepsilon) e^{-\varepsilon/T(g_1,g_2)} d\varepsilon \Big/ \int_0^\infty d(\varepsilon) e^{-\varepsilon/T(g_1,g_2)} d\varepsilon.$$

In the parabolic Chen model, the temperature can be explicitely computed from g_1 and g_2. Indeed, since $g_1 = L_{11} = \mu_0 n$ and $g_2 = L_{12} = (3/2)\mu_0 nT$, it follows $T = 2g_2/3g_1$.

The relaxation term can be written as follows [13]:

$$W = \gamma_1(T) g_1 - \gamma_2(T) g_2,$$

where

$$\frac{1}{\gamma_1(T)} = \frac{T}{\phi_1 T_0} \int_0^\infty d(\varepsilon) N(\varepsilon) d\varepsilon, \quad \frac{1}{\gamma_2(T)} = \frac{1}{\phi_1} \int_0^\infty d(\varepsilon) N(\varepsilon) \varepsilon d\varepsilon.$$

Notice that the quotients $e^{\mu/T}/L_{1i}(\mu, T)$ do not depend on μ (see (2.1)).

The formulation (2.11)-(2.13) has the advantages that first, the equations can be easily discretized using mixed finite elements as already employed for the drift-diffusion model (see, e.g., [6, 7]) and second, a Gummel iteration technique can be applied to the final nonlinear system (see section 3).

3. Numerical discretization

We discretize the energy-transport model in the drift-diffusion formulation (2.11)-(2.13). For this, we recast the equations in the following form

$$-\mathrm{div} J_i + \sigma_i g_i = f_i, \quad J_i = \nabla g_i - \beta g_i \quad \text{in } \Omega, \ i = 1, 2, \tag{3.1}$$

$$g_i = g(n_D, T_D) \quad \text{on } \Gamma_D, \quad J_i \cdot \nu = 0 \quad \text{on } \Gamma_N, \tag{3.2}$$

where

$$\sigma_1 = 0, \quad \sigma_2 = \gamma_2, \quad f_1 = 0, \quad f_2 = -J_1 \cdot \nabla V + \gamma_1 g_1, \quad \beta = \nabla V / T(g_1, g_2).$$

By freezing the coefficients (obtained from the previous iteration in the global iteration procedure), the energy-transport model can be written (in each iteration step) as a system of convection-diffusion equations. It is well known in semiconductor modeling that this system is usually convection dominated due to high electric fields. Therefore, special care has to be taken in the numerical approximation. We discretize the equations using an exponentially fitted mixed finite-element method.

We describe now the discretization of the energy conservation equation (3.1) with $i = 2$ in detail. The discretization of the mass conservation equation (3.1) with $i = 1$ is similar but simpler since the zeroth-order term and the right-hand side vanish.

Let $\mathcal{T}_h$ be a regular family of decompositions of Ω into triangles K such that there is no element across the interface between Γ_D and Γ_N. Furthermore, let $\mathcal{E}_h$ be the set of edges e of $\mathcal{T}_h$. We discretize the Poisson equation using nonconforming Crouzeix-Raviart elements yielding V^h [10]. Furthermore, we assume that the temperature is a piecewise constant function $\overline{T}$ obtained in the iteration procedure. As ∇V^h is piecewise constant, also $\beta^h = \nabla V^h/\overline{T}$ is constant in each triangle. Therefore, we can suppose that β in (3.1) is piecewise constant. Consequently, we can symmetrize the problem in each triangle K by means of the local Slotboom variable $y = e^{-\beta\cdot\mathbf{x}}u$. The symmetric form of the current density equation reads

$$e^{-\beta\cdot\mathbf{x}}J - \nabla y = 0 \quad \text{in } K. \tag{3.3}$$

The vector-valued approximation space for the mixed scheme developed in [27] is built on the following set of polynomial vectors:

$$\Sigma(K) = \operatorname{span}(\tau_1, \tau_2, \tau_3),$$

with $\tau_1 = (1, 0)$, $\tau_2 = (0, 1)$, and $\tau_3 = (\omega_1, \omega_2)$. To define the properties that determine the polynomials ω_i we need the following notations. The edges of each triangle K are numbered counter-clockwise starting with e_1 that connects the vertices with the smallest values of $\beta\cdot\mathbf{x}$ in K. We denote by ν_i the exterior unit normal vector of K associated to the edge e_i ($i = 1, 2, 3$). Then we choose $\tau_3 = (\omega_1, \omega_2)$ with ω_1, $\omega_2 \in P_2(K)$ fulfilling the following conditions:

$$\tau_3\cdot\nu_2|_{e_2} = \tau_3\cdot\nu_3|_{e_3} = 0, \quad \tau_3\cdot\nu_1|_{e_1} = 1, \tag{3.4}$$

$$\int_K \tau_3\cdot\operatorname{curl}(\lambda_1\,\lambda_2\,\lambda_3)dxdy = 0, \tag{3.5}$$

$$\int_K \omega_1 dxdy = \int_K \omega_2 dxdy = 0, \tag{3.6}$$

where λ_i denotes the i-th barycentric coordinate. It can be easily checked that $\dim(\operatorname{div}\Sigma(K)) = 1$ (see (3.4)).

We introduce the following finite dimensional spaces:

$$\begin{aligned}
\Sigma_h &= \{\tau \in [L^2(\Omega)]^2 : \tau|_K \in \Sigma(K)\ \forall K \in \mathcal{T}_h\},\\
W_h &= \{\phi \in L^2(\Omega) : \phi|_K \in P_0(K)\ \forall K \in \mathcal{T}_h\},\\
\Lambda_{h,\xi} &= \{\mu \in L^2(\mathcal{E}_h) : \mu|_K \in P_0(e)\ \forall e \in \mathcal{E}_h;\ \textstyle\int_e(\mu - \xi)ds = 0\ \forall e \in \Gamma_D\},
\end{aligned}$$

where ξ is any function in $L^2(\Gamma_D)$. The spaces $P_i(K)$ and $P_i(e)$ denote the sets of polynomials of degree i in K and on e, respectively. The mixed-hybrid formulation of (3.1)-(3.2) then reads as follows:

(3.7)
$$\begin{cases} \text{Find } J_2^h \in \Sigma_h,\ \bar{g}_2^h \in W_h \text{ and } g_2^h \in \Lambda_{h,u_D} \text{ such that:} \\ \displaystyle\sum_{K\in\mathcal{T}_h} \Big(\int_K A_K J_2^h \cdot \tau dx + \int_K B_K \bar{g}_2^h \mathrm{div}\tau dx - \sum_{e_i(K)} \int_{e_i} B_{e_i} g_2^h \tau \cdot \nu ds \Big) = 0, \\ \displaystyle\sum_{K\in\mathcal{T}_h} \Big(-\int_K \mathrm{div} J_2^h \phi dx + \int_K \sigma_2 \bar{g}_2^h \phi dx \Big) = \int_K f_2 \phi dx, \\ \displaystyle\sum_{K\in\mathcal{T}_h} \int_{\partial K} \mu J_2^h \cdot \nu ds = 0 \end{cases}$$

for all $\tau \in \Sigma_h$, $\phi \in W_h$, and $\mu \in \Lambda_{h,0}$. The functions A and B are piecewise constant on a triangle K or on one of its edges $e_i(K)$, respectively, and are defined by the mean values

$$A_K = \frac{1}{|K|} \int_K e^{-\beta\cdot\mathbf{x}} dx, \tag{3.8}$$

$$B_{e_i(K)} = \frac{1}{|e_i|} \int_{e_i} e^{-\beta\cdot\mathbf{x}} ds, \qquad B_K = \max_i B_{e_i(K)} = B_{e_1(K)}. \tag{3.9}$$

These terms account for the convective part of the flux. One important property is that strong convective effects are reproduced by the scheme with the correct order of magnitude [27].

The first equation in (3.7) is a weak discrete version of (3.3), the second equation corresponds to the discrete version of the first equation of (3.1), and the third equation imposes the continuity requirement of the normal component of J_2^h at the interelement boundaries.

Due to the piecewise regularity of the approximation spaces, a static condensation procedure applied to the linear system associated with (3.7) allows to eliminate the unknowns for J_2^h and $\bar{g}_2^h$ from the algebraic system leading to a final matrix acting only on the Lagrange multipliers g_2^h. This matrix is an M-matrix if the triangulation is weakly acute. The nonlinear algebraic system then is solved using a variant of the Gummel method. We recall that the Gummel iteration procedure can be regarded as an approximate Newton method, where the information about the strong coupling of the unknowns is incorporated into the Poisson equation, and the Jacobian of the system is replaced by a diagonal matrix [21].

In order to refine the mesh adaptively, we have to define a suitable error estimator. We use an error estimator consisting of two parts: an indicator for the primal variables of the mixed-hybrid formulation in the L^2 norm, motivated by the error estimators used for Raviart-Thomas elements [20], and an indicator for the Poisson equation, which is solved using Crouzeix-Raviart elements, taken from [39]. The error estimator for the Crouzeix-Raviart elements is a residual based error indicator. The estimator derived for the Raviart-Thomas elements can be interpreted as a residual based error indicator. It is motivated by a saturation assumption involving the difference $g_2^h - \bar{g}_2^h$. In [19] the saturation assumption has been verified numerically for the above finite elements introduced by Marini and Pietra, applied to some convection-dominated test problems.

After convergence of the Gummel method, the triangles belonging to the set $\{K \in \mathcal{T}_h : \eta_{K,ET} \leq 0.7 \max_K \eta_{K,ET}\}$ are regularly refined, removing any hanging nodes by using the algorithm of Bank and regularizing the mesh. Then new vectors of nodal values are generated by interpolating the solution from the previous mesh and the Gummel algorithmus is again applied until the error estimator is smaller than a prescribed value or until the maximal number of elements is reached (see [19] for details).

4. Numerical examples

In this section we present numerical simulations of a silicon ballistic diode and a silicon MOSFET device. The ballistic diode is simulated using the energy-transport model (2.11)-(2.13) including the physical parameters (see, e.g., [23]) with non-parabolic band diagrams in the sense of Kane (see section 2.1) and the coefficients (2.3)-(2.5) with $\beta = 1/2$. In this case, the temperature T is a nonlinear function of g_1 and g_2. The MOSFET device is solved using the parabolic Chen model of section 2.1 (i.e. using (2.6)-(2.8)) including electron-hole recombination-generation effects. More precisely, the equations (2.11) and (2.12) now read

$$\begin{aligned} -\mathrm{div} J_1 &= -R(n,p), \\ -\mathrm{div} J_2 &= -J_1 \cdot \nabla V + W(n,T) - \frac{3}{2} T R(n,p), \end{aligned}$$

where p denotes the hole particle density, and $R(n,p)$ is the Shockley-Read-Hall recombination-generation term given by

$$R(n,p) = \frac{np - n_i^2}{\tau_p(n+n_i) + \tau_n(p+n_i)},$$

with the (scaled) intrinsic density n_i and the electron and hole life times τ_n, τ_p, respectively. The hole density is computed using the drift-diffusion model

$$-\mathrm{div} J_p = R(n,p), \quad J_p = -\nabla p - p\nabla V.$$

The relevant numerical values of the physical parameters used for both devices are collected in Table 1.

4.1. A 1-D ballistic diode with non-parabolic bands

First, we present numerical simulations of a one-dimensional n^+nn^+ ballistic diode which is a simple model for the channel of a MOS transistor. The semiconductor domain is given by the interval $(0, \ell)$. The maximal doping concentration in the n^+-regions equals $5 \cdot 10^{17}\mathrm{cm}^{-3}$; in the n-channel the minimal doping profile is $2 \cdot 10^{15}\mathrm{cm}^{-3}$. The length of the n^+-regions is 0.1μm, whereas the length of the channel region equals 0.4μm. On the boundary points we assume that the total space charge $C - n$ vanishes and that the temperature equals the ambient temperature, i.e. (in unscaled variables) $n = C$ and $T = T_0$. Denoting by $U > 0$ the applied potential

Parameter	Physical meaning	Numerical value
$\tau_p \,/\, \tau_n$	carrier life times	$10^{-5} \,/\, 10^{-6}$s
$\mu_p \,/\, \mu_n$	low-field carrier mobilities	$450 \,/\, 1500$ cm^2/Vs
n_i	intrinsic density	$1.4 \cdot 10^{10}$ cm^{-3}
τ_0	energy relaxation time	0.4 ps
$\varepsilon_r(Si)$	relative permittivity of silicon	11.7
$\varepsilon_r(SiO_2)$	relative permittivity of the oxide	3.8
T_0	ambient temperature	300K

TABLE 1. Material and model parameters.

we set $V(0) = 0$ and $V(\ell) = -U$. The non-parabolicity parameter is taken to be $\alpha = 0.5$/eV.

The numerical experiments are performed with a uniform mesh of 100 grid points. In Figure 1 (left) we present the electron temperature for vanishing and non-vanishing non-parabolicity parameter. As expected the temperature in the n-channel is large, corresponding to "hot" electrons. The temperature is reduced due to the non-parabolicity effects. Another effect of the non-parabolicity is the reduction of the spurious velocity overshoot peak in Figure 1 (right).

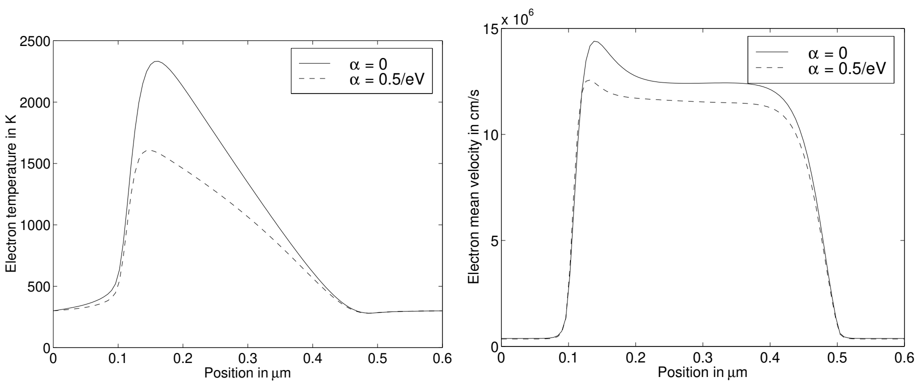

FIGURE 1. Electron temperature T (left) and mean velocity J/n versus position in a ballistic diode.

4.2. A 2-D MOSFET device with parabolic bands

A MOSFET (metal-oxide semiconductor field-effect transistor) device can be used as a voltage-driven switch and is the most used device in computer technology. We simulate a transistor of size 420nm$\times$210nm with an oxide thickness of 1.5nm. The length of the source and drain contacts is 30nm and the length of the channel varies between 70nm and 110nm (see Figure 2). The doping profile is given by a step function with values 10^{19}cm^{-3} in the n^+-region and -10^{17}cm^{-3} in the p bulk region. The voltage between source and drain is $V_{SD} = 1$V. The geometry and data of this device are adapted from the work of Cassan et al. [8] where the use of the energy-transport model for such device geometries has been justified.

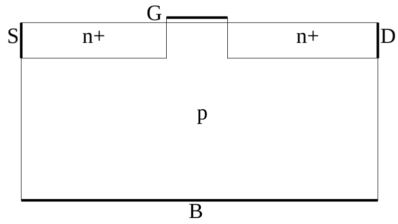

FIGURE 2. Geometry of the MOSFET with source S, drain D, gate G and bulk B contacts.

We model the influence of the oxide by assuming that the particles do not penetrate the oxide region Ω_O. Thus, the particle and energy densities are only computed in the semiconductor domain Ω_S. We choose Dirichlet boundary conditions at the source, drain and bulk contacts, i.e., the densities are set to their equilibrium values (see [28]). On the remaining boundary parts (including the silicon/silicon oxide interface), homogeneous Neumann boundary conditions $J_i \cdot \nu = 0$, $i = 1, 2, p$, are imposed. The Poisson equation is solved in the domain $\Omega_S \cup \Omega_O$ with a space-dependent permittivity which is constant in each Ω_i, $i = S, O$, and we specify homogeneous Neumann boundary conditions on the interface.

Before we present the numerical results we show the meshes generated by our refinement strategy in Figure 3. The initial mesh is constructed without any knowledge of the location of the junctions. After the adaptive procedure, the final grid is refined near the junctions and near the gate oxide. The importance of the adaptive scheme becomes apparent when comparing the current-voltage curves for different meshes (Figure 4 (left)). The characteristics seem to stabilize for meshes with about 4000 triangles.

Figure 4 (right) shows the dependence of the drain current on the effective channel length. For smaller channel lengths, not only the values of the drain current are larger but also the slope in the nearly linear region (for $V > 0.3$V) increases, due to higher electric fields in the channel. The electron density in a MOSFET with 70nm and 110nm channel lengths is presented in Figure 5. The voltage between gate and bulk is chosen to be $V_{GB} = 1$V. Our numerical scheme is able to resolve the large gradients near the junctions. Due to the high temperature region at the drain junction, the effective electron mobility $\mu(T) = \mu_n/T$ decreases and a region with a higher electron concentration is formed near the gate. The temperature T in a 110nm MOSFET is depicted in Figure 6. The values for T in the channel are larger in the closed state than in the open state. However, notice that in the closed state, the channel consists only of a few electrons.

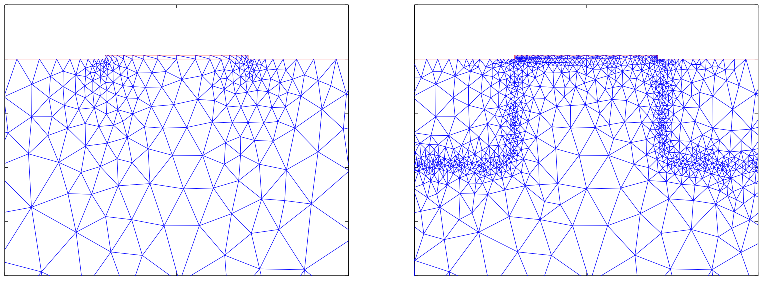

FIGURE 3. Adaptively refined triangulations for the MOSFET (zoom). Left: initial mesh (about 1000 triangles); right: final mesh (about 4000 triangles).

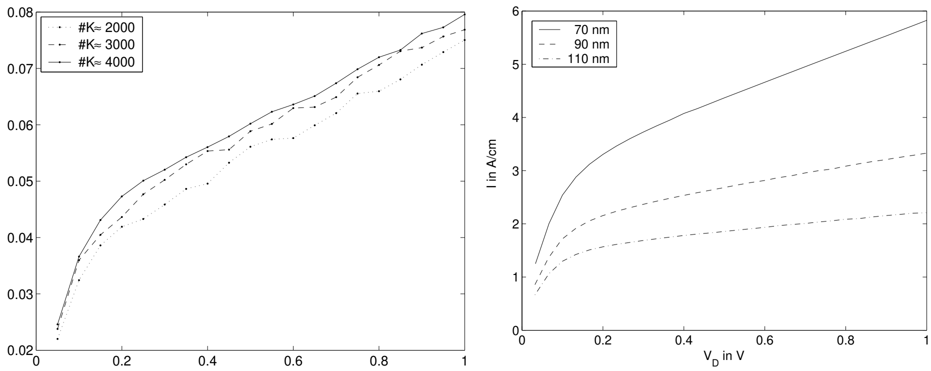

FIGURE 4. Drain current depending on the drain voltage for different meshes with $V_{GB} = 1\text{V}$ (left) and for different channel lengths (right).

References

[1] Y. Apanovich, P. Blakey, R. Cottle, E. Lyumkis, B. Polsky, A. Shur, and A. Tcherniaev. Numerical simulations of submicrometer devices including coupled nonlocal transport and nonisothermal effects. *IEEE Trans. Electr. Dev.*, 42:890–897, 1995.

[2] N. Ben Abdallah and P. Degond. On a hierarchy of macroscopic models for semiconductors. *J. Math. Phys.*, 37:3308–3333, 1996.

[3] F. Bosisio, R. Sacco, F. Saleri, and E. Gatti. Exponentially fitted mixed finite volumes for energy balance models in semiconductor device simulation. In H. Bock et al., editors, *Proceedings of ENUMATH 97*, pages 188–197, Singapore, 1998. World Scientific.

[4] K. Brennan. *The Physics of Semiconductors.* Cambridge University Press, Cambridge, 1999.

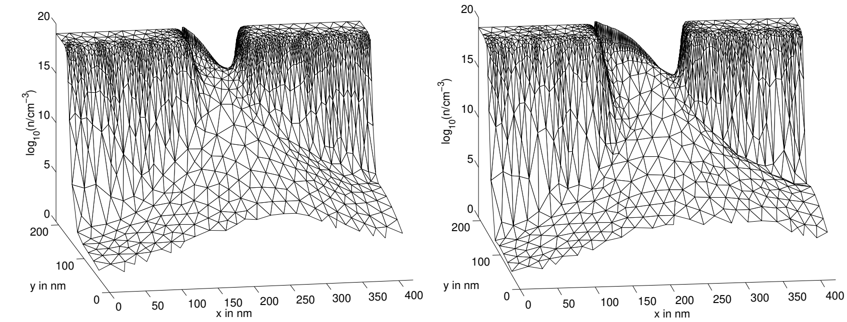

FIGURE 5. Electron density in a MOSFET with 70nm channel length (left) and 110nm channel length (right). Notice that a part of the bulk region is not shown.

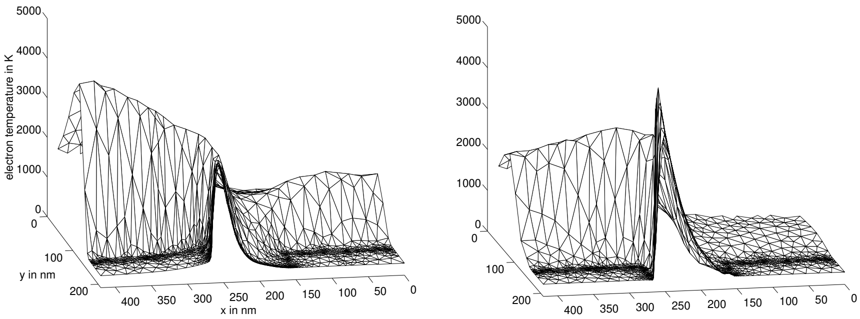

FIGURE 6. Electron temperature in a MOSFET with 110nm channel length in an open state ($V_{GB} = 1$V, left) and in a closed state ($V_{GB} = 0$V, right).

[5] F. Brezzi, L. Marini, S. Micheletti, P. Pietra, R. Sacco, and S. Wang. Discretization of semiconductor device problems. In W. Schilders and E. ter Maten, editors, *Handbook of Numerical Analysis. Numerical Methods for Electrodynamical Problems*, to appear, 2002. North-Holland-Elsevier Science Publishers.

[6] F. Brezzi, L. Marini, and P. Pietra. Numerical simulation of semiconductor devices. *Comp. Meth. Appl. Mech. Engrg.*, 75:493–514, 1989.

[7] F. Brezzi, L. Marini, and P. Pietra. Two-dimensional exponential fitting and applications to drift-diffusion models. *SIAM J. Num. Anal.*, 26:1342–1355, 1989.

[8] E. Cassan, S. Galdin, P. Dollfus, and P. Hesto. Comparison between device simulators for gate current calculation in ultra-thin gate oxide n-MOSFETs. *IEEE Trans. Elect. Dev.*, 83:1194–1202, 2000.

[9] D. Chen, E. Kan, U. Ravaioli, C. Shu, and R. Dutton. An improved energy transport model including nonparabolicity and non-Maxwellian distribution effects. *IEEE Electr. Dev. Letters*, 13:26–28, 1992.

[10] M. Crouzeix and P.A. Raviart. Conforming and nonconforming finite element methods for solving the stationary Stokes equation. *RAIRO*, 7:33–76, 1973.

[11] P. Degond, S. Génieys, and A. Jüngel. A system of parabolic equations in nonequilibrium thermodynamics including thermal and electrical effects. *J. Math. Pures Appl.*, 76:991–1015, 1997.

[12] P. Degond, S. Génieys, and A. Jüngel. A steady-state system in nonequilibrium thermodynamics including thermal and electrical effects. *Math. Meth. Appl. Sci.*, 21:1399–1413, 1998.

[13] P. Degond, A. Jüngel, and P. Pietra. Numerical discretization of energy-transport model for semiconductors with non-parabolic band structure. *SIAM J. Sci. Comp.*, 22:986–1007, 2000.

[14] W. Fang and K. Ito. Existence of stationary solutions to an energy drift-diffusion model for semiconductor devices. *Math. Models Meth. Appl. Sci.*, 11:827–840, 2001.

[15] A. Forghieri, R. Guerrieri, P. Ciampolini, A. Gnudi, M. Rudan, and G. Baccarani. A new discretization strategy of the semiconductor equations comprising momentum and energy balance. *IEEE Trans. Comp. Aided Design Integr. Circuits Sys.*, 7:231–242, 1988.

[16] M. Fournié. Numerical discretization of energy-transport model for semiconductors using high-order compact schemes. *Appl. Math. Lett.*, 15:727–734, 2002.

[17] J. Griepentrog. An application of the implicit function theorem to an energy model of the semiconductor theory. *Z. Angew. Math. Mech.*, 79:43–51, 1999.

[18] S. Holst, A. Jüngel, and P. Pietra. A mixed finite-element discretization of the energy-transport model for semiconductors. To appear in *SIAM J. Sci. Comp.*, 2002.

[19] S. Holst, A. Jüngel, and P. Pietra. An adaptive mixed scheme for energy-transport simulations of field-effect transistors. Submitted for publication, 2002.

[20] R. Hoppe and B. Wohlmuth. A comparison of a posteriori error estimators for mixed finite element discretization by Raviart-Thomas elements. *Math. Comp.*, 68:1347-1378.

[21] J. Jerome. *Analysis of charge transport. A mathematical study of semiconductor devices.* Springer, Berlin, 1996.

[22] J. Jerome and C.-W. Shu. Energy models for one-carrier transport in semiconductor devices. In W. Coughran et al., editors, *Semiconductors, Part II*, volume 59 of *IMA Volumes in Math. and Its Appl.*, pages 185–207, New York, 1994. Springer.

[23] A. Jüngel. *Quasi-hydrodynamic Semiconductor Equations.* Birkhäuser, Basel, 2001.

[24] E. Kane. Band structure of indium-antimonide. *J. Phys. Chem. Solids*, 1:249–261, 1957.

[25] C. Lab and P. Caussignac. An energy-transport model for semiconductor heterostructure devices: application to AlGaAs/GaAs MODFETs. *COMPEL*, 18:61–76, 1999.

[26] L. D. Marini and P. Pietra. An abstract theory for mixed approximations of second order elliptic equations. *Mat. Aplic. Comp.*, 8:219–239, 1989.

[27] L. D. Marini and P. Pietra. New mixed finite element schemes for current continuity equations. *COMPEL*, 9:257–268, 1990.

[28] P. A. Markowich, C. A. Ringhofer, and C. Schmeiser. *Semiconductor Equations.* Springer, Vienna, 1990.

[29] A. Marrocco and P. Montarnal. Simulation des modèles energy-transport à l'aide des éléments finis mixtes. *C. R. Acad. Sci. Paris*, 323:535–541, 1996.

[30] A. Marrocco, P. Montarnal, and B. Perthame. Simulation of the energy-transport and simplified hydrodynamic models for semiconductor devices using mixed finite elements. In *Proceedings ECCOMAS 96*, London, 1996. John Wiley.

[31] P. Raviart and J. Thomas. A mixed finite element method for second order elliptic equations. In *Mathematical Aspects of the Finite Element Method*, volume 606 of *Lecture Notes in Math.*, pages 292–315. Springer, 1977.

[32] C. Ringhofer. An entropy-based finite difference method for the energy transport system. *Math. Models Meth. Appl. Sci.*, 11:769–796, 2001.

[33] M. Rudan, A. Gnudi, and W. Quade. A generalized approach to the hydrodynamic model of semiconductor equations. In G. Baccarani, editor, *Process and Device Modeling for Microelectronics*, Amsterdam, 1993. Elsevier.

[34] C. Schmeiser and A. Zwirchmayr. Elastic and drift-diffusion limits of electron-phonon interaction in semiconductors. *Math. Models Meth. Appl. Sci.*, 8:37–53, 1998.

[35] K. Seeger. *Semiconductor Physics. An Introduction.* 6th edition, Springer, Berlin, 1997.

[36] K. Souissi, F. Odeh, H. Tang, and A. Gnudi. Comparative studies of hydrodynamic and energy transport models. *COMPEL*, 13:439–453, 1994.

[37] R. Stratton. Diffusion of hot and cold electrons in semiconductor barriers. *Phys. Rev.*, 126:2002–2014, 1962.

[38] P. Visocky. A method for transient semiconductor device simulation using hot-electron transport equations. In J. Miller, editor, *Proc. of the Nasecode X Conf.*, Dublin, 1994. Boole Press.

[39] B. Wohlmuth. A residual based error estimator for mortar finite element discretizations. *Num. Math.*, 84:143–171, 1999.

S.H. and A.J.: Fachbereich Mathematik und Informatik, Universität Mainz, Staudingerweg 9, 55099 Mainz, Germany
E-mail address: {holst,juengel}@mathematik.uni-mainz.de

P.P.: Istituto di Matematica Applicata e Tecnologie Informatiche, C.N.R., Via Ferrata 1, 27100 Pavia, Italy
E-mail address: pietra@dragon.imati.cnr.it