Hydrodynamic and Energy-Transport Models for Semiconductor Device Simulation

T. Grasser, H. Kosina, and S. Selberherr

Abstract— **Hydrodynamic and energy-transport models have emerged as powerful means for gaining additional insight into the complex non-local behavior encountered in state-of-the-art semiconductor devices. However, several different formulations have been proposed which vary considerably in complexity. Furthermore, the handling of these equations is far more complicated than that of the robust and well studied drift-diffusion equations. We give a detailed review of the most important papers published on this subject and try to shed some additional light on the critical issues associated with these transport models.**

I. Introduction

AS THE SIZE of state-of-the-art devices is continually reduced, non-local behavior becomes a critical issue in the simulation of these structures. The well established drift-diffusion (DD) model [1] which is still predominantly used by engineers around the world cannot cover these effects as the electron gas is assumed to be in thermal equilibrium with the lattice temperature. In the DD approach the local energy can be estimated via the homogeneous energy flux equation (e.g., (53) with $\nabla \cdot (n\mathbf{S})$ and ∂_t set to zero). However, for rapidly increasing electric fields the energy lags behind the electric field because it takes the carriers some time to pick up energy from the field. A consequence of the lag is that the maximum energy can be much smaller than the one predicted by the homogeneous energy flux equation. Furthermore, this lag gives rise to an overshoot in the carrier velocity because the mobility depends to first order on the energy and not on the electric field. As the mobility μ has not yet been reduced by the increased energy but the electric field is already large, an overshoot in the velocity $\mathbf{v} = \mu\mathbf{E}$ is observed until the carrier energy comes into equilibrium with the electric field again. Thus, DD simulations predict the same velocity profile as for slowly varying fields which can dramatically underestimate the carrier velocities. Similar to the mobility, many other physical processes are more accurately described by a local energy model rather than a local electric field model. Therefore the assumption of a fixed energy-field relation can cause unphysical results when used to predict, for example impact ionization. To overcome these limitations of the DD model, extensions have been proposed which basically add an additional balance equation for the average carrier energy [2], [3]. Furthermore, an additional driving term is added to the current relation which is proportional to the gradient of the carrier temperature. Several different formulations have been proposed which vary considerably in complexity. Furthermore, these equations have been extended to handle non-homogeneous materials and non-parabolicity effects. In the following, we review some of the basic assumptions underlying these models.

The authors are with the Institute for Microelectronics, Technical University Vienna, A-1040 Vienna, Austria

II. Boltzmann's Transport Equation

Transport equations used in semiconductor device simulation are normally derived from Boltzmann's transport equation (BTE) which reads [4]

$$\partial_t f + \mathbf{u} \cdot \nabla_{\mathbf{r}} f + \frac{\mathbf{F}}{\hbar} \cdot \nabla_{\mathbf{k}} f = C[f] \tag{1}$$

for a general inhomogeneous material with arbitrary band structure [5]. For inclusion of quantum effects equations based on the Wigner-Boltzmann equation have been considered [6]. The group velocity $\mathbf{u}$ is

$$\mathbf{u}(\mathbf{k},\mathbf{r}) = \frac{1}{\hbar}\nabla_{\mathbf{k}}\mathcal{E}(\mathbf{k},\mathbf{r}) \tag{2}$$

which defines the inverse effective mass tensor

$$\hat{\mathrm{m}}^{-1}(\mathbf{k},\mathbf{r}) = \frac{1}{\hbar}\nabla_{\mathbf{k}} \otimes \mathbf{u}(\mathbf{k},\mathbf{r}) = \frac{1}{\hbar^2}\nabla_{\mathbf{k}} \otimes \nabla_{\mathbf{k}}\mathcal{E}(\mathbf{k},\mathbf{r}) \tag{3}$$

where $\otimes$ denotes the tensor product [5]. In the following we will only consider position-independent masses but permit energy-dependent masses. Generalizations to position-dependent band structures will be given in the appropriate context. The force $\mathbf{F}$ exerted on the particles is generally given as

$$\mathbf{F}(\mathbf{k},\mathbf{r}) = -\nabla_{\mathbf{r}}E_{c,0}(\mathbf{r}) - \mathrm{q}(\mathbf{E}(\mathbf{r}) + \mathbf{u}\times\mathbf{B}) - \nabla_{\mathbf{r}}\mathcal{E}(\mathbf{k},\mathbf{r}) \tag{4}$$

and depends both on $\mathbf{k}$ and $\mathbf{r}$. Omitting the influence of $\mathbf{u}\times\mathbf{B}$ (see [7] for a treatment of this term) and assuming homogeneous materials, $\mathbf{F}$ simplifies to

$$\mathbf{F}(\mathbf{r}) = -\mathrm{q}\mathbf{E}(\mathbf{r}) \tag{5}$$

The BTE is an equation in the seven-dimensional phase space which is prohibitive to solve for engineering applications. Monte-Carlo (MC) simulations have been proven to give accurate results but are restrictive time consuming. Furthermore, if the distribution of high-energetic carriers is relevant, or if the carrier concentration is very low in specific regions of the device, MC simulations tend to produce high variance in the results. Therefore, a common simplification is to investigate only some moments of the distribution function, such as the carrier concentration and the carrier temperature. We define the moments of the distribution function as

$$\langle\Phi\rangle = \frac{1}{4\pi^3}\int \Phi f \ \mathrm{d}^3\mathbf{k} \tag{6}$$

with a suitable weight function $\Phi = \Phi(\mathbf{k})$. In the following we will separate the group velocity $\mathbf{u}$ into a random part $\mathbf{c}$ and the mean value $\mathbf{v} = \langle\mathbf{u}\rangle/\langle 1\rangle$ as $\mathbf{u} = \mathbf{c}+\mathbf{v}$. We will write all moment equations introducing the following symbols [8]

$$n = \langle 1\rangle \tag{7}$$

$$\mathbf{p} = \frac{1}{n}\langle\hbar\mathbf{k}\rangle \tag{8}$$

$$\mathbf{v} = \frac{1}{n}\langle\mathbf{u}\rangle = -\frac{\mathbf{J}}{\mathrm{q}n} \tag{9}$$

$$w = \frac{1}{n}\langle\mathcal{E}\rangle \tag{10}$$

$$\mathbf{S} = \frac{1}{n}\langle\mathbf{u}\mathcal{E}\rangle \tag{11}$$

$$\mathbf{Q} = \frac{1}{2n}\langle m(\mathbf{k})\mathbf{c}^2\mathbf{v}\rangle \tag{12}$$

$$\hat{\mathrm{T}} = \frac{1}{\mathrm{k_B}n}\langle m(\mathbf{k})\mathbf{c}\otimes\mathbf{c}\rangle \tag{13}$$

$$\hat{\mathrm{U}} = \frac{1}{n}\langle\hbar\mathbf{u}\otimes\mathbf{k}\rangle \tag{14}$$

$$\hat{\mathrm{R}} = \frac{1}{n}\langle\hbar\mathbf{u}\otimes\mathbf{k}\mathcal{E}\rangle \tag{15}$$

Furthermore, we will employ an isotropic effective mass approximation via the trace of the mass tensor as [9]

$$m^{*-1} = \frac{1}{3}\mathrm{tr}\left(\langle\hat{\mathrm{m}}^{-1}\rangle\right) \tag{16}$$

III. Band Structure

The simplest approximation for the complex band structure is a parabolic relationship between the energy and the crystal momentum

$$\mathcal{E} = \frac{\hbar^2k^2}{2m^*} \tag{17}$$

which is valid for energies close to the band minimum. A first-order non-parabolic relationship was given by Kane [10] as

$$\mathcal{E}(1+\alpha\mathcal{E}) = \frac{\hbar^2k^2}{2m^*} \tag{18}$$

with α being the non-parabolicity correction factor. This gives the following relationship between momentum and velocity [11]

$$\hbar\mathbf{k} = m^*(1+2\alpha\mathcal{E})\mathbf{u} \tag{19}$$

and between energy and velocity

$$\mathcal{E} = \frac{1}{2\alpha}\Big((1-2\alpha m^*\mathbf{u}^2)^{-1/2}-1\Big) \tag{20}$$

which reduce to their parabolic counterparts for $\alpha = 0$. Expansion of the square root in (20) yields terms in ascending powers of velocity which are not negligible when averaged. This is problematic because these quantities are additional unknowns representing higher-order moments of the velocity distribution. Although (18) is an improvement over (17) it is nevertheless a crude approximation for real band structures at higher energies.

IV. Stratton's Approach

One of the first derivations of moment equations was performed by Stratton [2]. First, the distribution function is split into its even and odd parts as

$$f(\mathbf{k},\mathbf{r}) = f_0(\mathbf{k},\mathbf{r}) + f_1(\mathbf{k},\mathbf{r}) \tag{21}$$

From $f_1(-\mathbf{k},\mathbf{r}) = -f_1(\mathbf{k},\mathbf{r})$ it follows that $\langle f_1\rangle = 0$. Assuming that the collision operator C is linear and invoking a microscopic relaxation time approximation for the collision operator

$$C[f] = -\frac{f-f_0}{\tau(\mathcal{E},\mathbf{r})} \tag{22}$$

the BTE can be split into two coupled equations. In particular, f_1 is related to f_0 via

$$f_1 = -\tau(\mathcal{E},\mathbf{r})\Big(\mathbf{u}\nabla_{\mathbf{r}}f_0 - \frac{\mathrm{q}}{\hbar}\mathbf{E}\cdot\nabla_{\mathbf{k}}f_0\Big) \tag{23}$$

The microscopic relaxation time is then expressed using a power law as

$$\tau(\mathcal{E}) = \tau_0\Big(\frac{\mathcal{E}}{\mathrm{k_B}T_L}\Big)^{-p} \tag{24}$$

which allows for an explicit integration over constant energy surfaces. When f_0 is assumed to be a heated Maxwellian distribution, the following equation system is obtained

$$\nabla\cdot\mathbf{J} = \mathrm{q}(\partial_t n + R) \tag{25}$$

$$\mathbf{J} = \mathrm{q}\mu n\mathbf{E} + \mathrm{k_B}\nabla(n\mu T_n) \tag{26}$$

$$\nabla\cdot\mathbf{S} = -\frac{3}{2}\mathrm{k_B}\partial_t(nT_n) + \mathbf{E}\cdot\mathbf{J} - \frac{3}{2}\mathrm{k_B}n\frac{T_n-T_L}{\tau_{\mathcal{E}}} \tag{27}$$

$$\mathbf{S} = -\Big(\frac{5}{2}-p\Big)\Big(\mu n\mathrm{k_B}T_n\mathbf{E} + \frac{\mathrm{k_B^2}}{\mathrm{q}}\nabla(n\mu T_n)\Big) \tag{28}$$

(26) is frequently written as

$$\mathbf{J} = \mathrm{q}\mu\Big(n\mathbf{E} + \frac{\mathrm{k_B}}{\mathrm{q}}T_n\nabla n + \frac{\mathrm{k_B}}{\mathrm{q}}n(1+\nu_n)\nabla T_n\Big) \tag{29}$$

with

$$\nu_n = \frac{T_n}{\mu}\frac{\partial\mu}{\partial T_n} = \frac{\partial\ln\mu}{\partial\ln T_n} \tag{30}$$

which is commonly used as a fit parameter with values $-0.5\ldots-1.0$. For $\nu_n = -1.0$, the thermal diffusion term disappears. Under certain assumptions [2], [12] $p = -\nu_n$. The problem with expression (24) for τ is that p must be approximated by an average value to cover the relevant scattering processes. However, this average depends on the doping profile and the applied field and thus no unique value for p can be given.

V. Bløtekjær's Approach

Bløtekjær [3] derived conservation equations by taking the moments of the BTE using the weight functions 1, $\hbar\mathbf{k}$,

and $\mathcal{E}$ without imposing any assumptions on the form of the distribution function. These weight functions Φ define the moments of zeroth, first, and second order. The resulting moment equations can be written as follows [8]

$$\partial_t n + \nabla \cdot (n\mathbf{v}) = nC_n \tag{31}$$

$$\partial_t (n\mathbf{p}) + \nabla \cdot (n\hat{\mathrm{U}}) - n\mathbf{F} = n\mathbf{C}_p \tag{32}$$

$$\partial_t (nw) + \nabla \cdot (n\mathbf{S}) - n\mathbf{v} \cdot \mathbf{F} = nC_{\mathcal{E}} \tag{33}$$

Note that these expressions are valid for arbitrary band structures, provided that the carrier mass is position-independent. When $\mathbf{F}$ is allowed to be position-dependent, additional force terms appear in (31)-(33) [13]. The collision terms are usually modeled with a macroscopic relaxation time approximation as

$$C_n = -\frac{1}{n}(R - G) = -\frac{1}{n}U \tag{34}$$

$$\mathbf{C}_p = -\frac{\mathbf{p}}{\tau_p} \tag{35}$$

$$C_{\mathcal{E}} = -\frac{w - w_0}{\tau_{\mathcal{E}}} \tag{36}$$

which introduces the relaxation times τ_p and $\tau_{\mathcal{E}}$. A discussion on this approximation is given in [14]. This equation set is not closed as it contains more unknowns than equations. Closure relations have to be found to express the equations in terms of the unknowns n, $\mathbf{v}$, and w. Due to the strong scattering the temperature tensor is normally assumed to be isotropic and is approximated by a scalar T_n as

$$\hat{\mathrm{T}} \approx T_n \hat{\mathrm{I}} = \frac{T_{xx} + T_{yy} + T_{zz}}{3}\hat{\mathrm{I}} \tag{37}$$

Traditionally, parabolic bands were assumed which gives the following closure relations for $\mathbf{p}$, $\hat{\mathrm{U}}$, and w

$$\mathbf{p} = m^*\mathbf{v} \tag{38}$$

$$\hat{\mathrm{U}} = \frac{m^*}{n}\langle \mathbf{u} \otimes \mathbf{u} \rangle = \mathrm{k_B} T_n \hat{\mathrm{I}} + m^* \mathbf{v} \otimes \mathbf{v} \tag{39}$$

$$w = \frac{3}{2}\mathrm{k_B} T_n + \frac{m^* v^2}{2} \tag{40}$$

Note that the random component of the velocity has zero average ($\langle \mathbf{c} \rangle = 0$). With (38) one obtains the following formulation for $\mathbf{C}_p$

$$\mathbf{C}_p = -\frac{\mathbf{p}}{\tau_p} = -\frac{m^*\mathbf{v}}{\tau_p} = -\frac{\mathrm{q}\mathbf{v}}{\mu} \tag{41}$$

For modeling purposes it is advantageous to lump m^* and τ_p into one new parameter, the mobility μ. As signal frequencies are well below $1/(2\,\pi\,\tau_p) \approx 10^{12}$ Hz the time derivative in (32) can safely be neglected.

Furthermore, a suitable approximation for the energy flux density $n\mathbf{S}$ has to be found and different approaches have been published. Bløtekjær used

$$n\mathbf{S} = (w + \mathrm{k_B} T_n) n\mathbf{v} + n\mathbf{Q} \tag{42}$$

and approximated the heat flux $n\mathbf{Q}$ by Fourier's law as

$$n\mathbf{Q} = -\kappa(T_n)\nabla T_n \tag{43}$$

in which the thermal conductivity is given by the Wiedemann-Franz law as

$$\kappa(T_n) = \Big(\frac{5}{2} - p\Big)\Big(\frac{\mathrm{k_B}}{\mathrm{q}}\Big)^2 \mathrm{q}\mu n T_n \tag{44}$$

where p is a correction factor. As has been pointed out [8], this expression is problematic as (43) only approximates the diffusive component of $n\mathbf{Q}$. For a uniform temperature $\nabla T_n = 0$ and thus $\mathbf{Q} = 0$ which is not plausible. The convective component $\mathbf{Q}_{\mathrm{conv}}$ must be included to obtain physical results when the current flow is not negligible.

With these approximations (31)-(33) can be written in the usual variables as [15]

$$\nabla \cdot \mathbf{J} = \mathrm{q}(\partial_t n + U) \tag{45}$$

$$\mathbf{J} - \frac{\tau_p}{\mathrm{q}}\nabla \cdot \Big(\mathbf{J} \otimes \frac{\mathbf{J}}{n}\Big) = \mu \mathrm{k_B} \nabla (n T_n) + \mathrm{q} n \mu \mathbf{E} \tag{46}$$

$$\nabla \cdot (n\mathbf{S}) = -\partial_t (nw) + \mathbf{E} \cdot \mathbf{J} - n\frac{w - w_0}{\tau_{\mathcal{E}}} \tag{47}$$

$$n\mathbf{S} = -\frac{1}{\mathrm{q}}(w + \mathrm{k_B} T_n)\mathbf{J} - \kappa(T_n)\nabla T_n \tag{48}$$

to give the full hydrodynamic model (FHD) for parabolic band structures. This equation system is similar to the Euler equations of gas dynamics with the addition of a heat conduction term and the collision terms. It describes the propagation of electrons in a semiconductor device as the flow of a compressible, charged fluid. This electron gas has a sound speed $v_c = \sqrt{\mathrm{k_B} T_n / m^*}$, and the electron flow may be either subsonic or supersonic. With $T_n = \xi T_L$ and $T_L = 300$ K, $v_c = \sqrt{\xi}\, 1.3 \cdot 10^7$ cm/s while for $T_L = 77$ K, $v_c = \sqrt{\xi}\, 6.6 \cdot 10^6$ cm/s [16].

In the case of supersonic flow, electron shock waves will in general develop inside the device [16]. These shock waves occur at either short length scales or at low temperatures. As the equation system is hyperbolic in the supersonic regions, special hyperbolic methods have to be used [16], [17], [18], [19]. Furthermore, the traditionally applied Scharfetter-Gummel [20] discretization scheme and its extension to the energy-balance and energy-transport models [21], [22], [23], [24] cannot be used for this type of equation. One approximation is to treat the convective term as a perturbation by freezing its dependence on the state variables at each linearization step and using the values from the last iteration [25]. However, this approach will degrade the convergence in cases where the variation in space or time is important [26]. Thus, to derive a spatial discretization, fluid dynamics methods known as upwinding are used [26]. Furthermore, the handling of the boundary conditions becomes more difficult [19], [27].

When the convective term

$$\frac{\tau_p}{\mathrm{q}}\nabla \cdot \Big(\mathbf{J} \otimes \frac{\mathbf{J}}{n}\Big) \tag{49}$$

is neglected, a parabolic equation system is obtained which only covers the subsonic flow regions. This is a very common approximation in todays device simulators. Furthermore, the contribution of the velocity to the carrier energy is frequently neglected

$$w \approx \frac{3}{2}\mathrm{k_B}T_n \tag{50}$$

which then results in the following equation system

$$\nabla \cdot \mathbf{J} = \mathrm{q}(\partial_t n + U) \tag{51}$$

$$\mathbf{J} = \mu \mathrm{k_B} \nabla(nT_n) + \mathrm{q}n\mu\mathbf{E} \tag{52}$$

$$\nabla \cdot (n\mathbf{S}) = -\frac{3\mathrm{k_B}}{2}\partial_t(nT_n) + \mathbf{E}\cdot\mathbf{J} - n\frac{3\mathrm{k_B}}{2}\frac{T_n - T_L}{\tau_{\mathcal{E}}} \tag{53}$$

$$n\mathbf{S} = -\frac{5\mathrm{k_B}T_n}{2\mathrm{q}}\mathbf{J} - \kappa(T_n)\nabla T_n \tag{54}$$

(51)-(54) form a typical three moment energy-transport (ET) model which has been closed using Fourier's law.

To overcome the difficulties associated with the Fourier law closure (43), the fourth moment of the BTE has been taken into account [28] which gives

$$\nabla \cdot (n\hat{\mathrm{R}}) - n(w\hat{\mathrm{I}} + \hat{\mathrm{U}}) \cdot \mathbf{F} = n\mathbf{C}_{p\mathcal{E}} \tag{55}$$

where the time derivative has been ignored using a similar argument to (32). The collision term in (55) can be modeled in analogy to (41) as

$$\mathbf{C}_{p\mathcal{E}} = -\frac{\mathrm{q}\mathbf{S}}{\mu_S} \tag{56}$$

which gives

$$\mathbf{S} = \frac{\mu_S}{\mu}(w\hat{\mathrm{I}} + \hat{\mathrm{U}}) \cdot \mathbf{v} + \frac{\mu_S}{\mathrm{q}n}\Big((w\hat{\mathrm{I}} + \hat{\mathrm{U}}) \cdot \nabla(n\hat{\mathrm{U}}) - \nabla(n\hat{\mathrm{R}})\Big) \tag{57}$$

Now a closure relation for $\hat{\mathrm{R}}$ has to be introduced, which can be, for example, obtained by assuming a heated Maxwellian distribution which gives

$$\hat{\mathrm{R}} = \frac{5}{2}\mathrm{k_B^2}T_n^2\hat{\mathrm{I}} \tag{58}$$

Using closure (58) and the same approximations that led to the three moments ET model (51)-(54), a more accurate expression for $n\mathbf{S}$ is obtained from the fourth moment of the BTE

$$n\mathbf{S} = -\frac{\mu_S}{\mu}\frac{5}{2}\frac{\mathrm{k_B}T_n}{\mathrm{q}}\mathbf{J} - \frac{\mu_S}{\mu}\frac{5}{2}\Big(\frac{\mathrm{k_B}}{\mathrm{q}}\Big)^2 \mathrm{q}\mu n T_n \nabla T_n \tag{59}$$

which should be used to replace (54) to give a four moments ET model. Comparing (59) with (54) reveals that a consistent three moment ET model can be obtained with $\mu_S/\mu = 1$ and $p = 0$. However, μ_S/μ strongly depends on the carrier temperature and shows a pronounced hysteresis as shown in Fig. 1 where the points B and D are from the rising and decreasing temperature regions, respectively (see Fig. 3 for details). The energy relaxation time and the momentum relaxation time are shown in Fig. 2 and both are not single valued functions of the temperature.

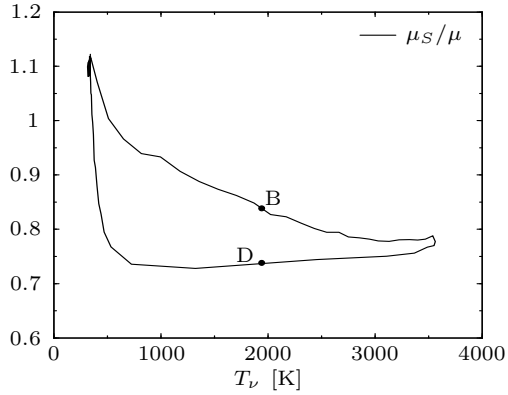

Fig. 1. Ratio of μ_S and μ as a function of the carrier temperature inside the n^+-n-n^+ test-structure obtained from MC simulations.

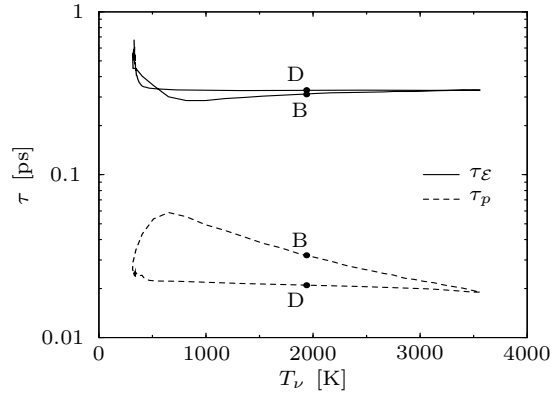

Fig. 2. Relaxation times as a function of the carrier temperature inside the n^+-n-n^+ test-structure obtained from MC simulations.

VI. Comparison

One of the extensively discussed differences between Bløtekjær's (A1) and Stratton's (A2) approach is that in A2 the mobility stands inside the gradient whereas in A1 it stands infront of the gradient in the current relation

Bløtekjær (A1): $\mu_1 \nabla(nT_n)$
Stratton (A2): $\nabla(n\mu_2 T_n)$

This issue was addressed by Stratton himself [29] and by Landsberg [30], [31]. It is important to note, that although this parameter is called mobility in both approaches, their definition differs significantly. Tang *et al.* [32] compared both approaches and found that both formulations are justified, provided that the respective mobilities are modeled accordingly. For bulk simulations the mobilities are equal and can be properly modeled using conventional energy-dependent expressions [33], [34]. However, in inhomogeneous samples where the electric field varies rapidly, the mobilities are no longer single-valued functions of the average carrier energy. The advantage of the μ_1 formulation lies in the fact, that for increasing values of the electric field, it can be roughly approximated by its bulk value whereas μ_2 is always different. Thus μ_1 can be expected to be more suitable because in most commercial simulators the mobility is modeled as a function of the carrier energy only. By expressing $\mathbf{C}_p$ empirically as

$$\mathbf{C}_p = \mathbf{C}_p^* + \lambda_p \nabla \cdot \hat{\mathrm{U}} \tag{60}$$

where $\mathbf{C}_p^*$ is the homogeneous component and λ_p a dimensionless transport coefficient, Tang *et al.* [35] showed that A2 can be obtained from A1 with $\lambda_p = -\nu_n$. Other comparisons of the two approaches can be found in [12], [32], [36], [37], [38].

VII. Non-Parabolicity Extensions

The general hydrodynamic equations (31)-(33) are valid for any band structure as $\mathbf{F}$ depends only on the spatial gradient of the dispersion relation. However, parabolicity assumptions are invoked to derive the closure relations (38)-(40). On the other hand, non-parabolicity effects enter the HD equations through the models used for the collision terms. A good example is the mobility whose homogeneous values are frequently obtained through measured $\mathbf{v}(\mathbf{E})$ characteristics. This mobility contains the full information of a real band structure, something which is much more difficult to obtain with MC simulations where the mobility has to be modeled using microscopic scattering rates [39]. As pointed out in the discussion of (20), there is no analytic relationship between the $\langle\mathcal{E}\rangle$ and $\mathbf{v}$ in the general case. For parabolic bands the carrier temperature is normally defined via the average carrier energy as

$$T_n = \frac{2}{3\mathrm{k_B}}\Big(\langle\mathcal{E}\rangle - \frac{m^*\mathbf{v}^2}{2}\Big) \tag{61}$$

Unfortunately, there is no similar equation for non-parabolic bands. Another possibility is to define the temperature via the variance of the velocity as [9]

$$T^* = \frac{m^*}{3\mathrm{k_B}}\langle\mathbf{u}^2\rangle \tag{62}$$

Definitions (61) and (62) are consistent with the thermodynamic definition of the carrier temperature in thermodynamic equilibrium and both are identical for non-equilibrium cases when a constant carrier mass is assumed which in turn corresponds to the assumption of parabolic energy bands. However, large differences are observed when a more realistic band structure is considered [9], [40].

A. The Generalized Hydrodynamic Model

Thoma *et al.* [9], [40] proposed a model which they termed 'generalized hydrodynamic model'. Instead of using the average energy *and* the temperature as variables in their formulation, they opted for a temperature-only description. To obtain a form similar to standard models, they defined the temperature according to (62) which differs significantly from (61) for non-parabolic bands. Instead of the momentum weight functions $\hbar\mathbf{k}$ and $\hbar\mathbf{k}\mathcal{E}$, they used $\mathbf{u}$ and $\mathbf{u}\mathcal{E}$ to derive the moment equations of order one and three. Without assuming a Kane dispersion they derived the following equations for the current and energy flux density

$$\mathbf{J} = \frac{\tau_i}{\tau_i^*}\mu^*\mathrm{k_B}\nabla(nT^*) + \mathrm{q}\mu^* n\mathbf{E} \tag{63}$$

$$\nabla\cdot(n\mathbf{S}) = -\frac{3\mathrm{k_B}}{2}\partial_t(nT^*) + \mathbf{E}\cdot\mathbf{J} - n\frac{3\mathrm{k_B}}{2}\frac{T^*-T_L}{\tau_\mathcal{E}^*} \tag{64}$$

$$n\mathbf{S} = -\frac{\mu_S^*}{\mu^*}\frac{5}{2}\frac{\mathrm{k_B}T^*}{\mathrm{q}}\Big(\mathbf{J} + \mu^*\frac{\tau_i}{\tau_i^*}n\mathrm{k_B}\nabla T^*\Big) \tag{65}$$

All relaxation times and mobilities are modeled as a function of T^* and explicit formulas were given in [41]. The advantage of this formulation is that it can be applied to arbitrary band structures. Thoma *et al.*, however, used parameters extracted from MC simulations employing the Kane dispersion relation.

B. Model of Bordelon

Another non-parabolic formulation was derived by Bordelon *et al.* [42], [11] which explicitly assumed a Kane dispersion. They used the weight functions 1, $\hbar\mathbf{k}$ and $\mathcal{E}$ and closed the system by ignoring the heat flux. To avoid the problem with the missing energy-temperature relation, they formulate their equation system solely in w. By introducing the function $H(w) = (1+\alpha w)/(1+2\alpha w)$ in the approximation for $\hat{\mathrm{U}}$ [37] they obtain

$$\mathbf{J} = \mu\frac{2}{3}\nabla(nwH(w)) + \mathrm{q}\mu n\mathbf{E} \tag{66}$$

$$n\mathbf{S} = -\Omega(w)w\frac{\mathbf{J}}{\mathrm{q}} \tag{67}$$

with $\Omega(w) \approx 1.3$. In the comparison made by Ramaswami *et al.* [37] the predicted $\mathbf{v}$ and $\mathbf{S}$ curves agree quite well with the MC data, even with this simplified model for $\mathbf{S}$.

C. Model of Chen

In [43], Chen *et al.* published a model which they termed 'energy transport model'. They tried to include non-parabolic and non-Maxwellian effects to a first order. Their approach is based on Stratton's model and the use of Kane's dispersion relation. Their Ansatz contains a non-Maxwellian factor γ which, however, does not show up in the final equations which read

$$\mathbf{J} = \mathrm{k_B}\nabla(n\mu T_m) + \mathrm{q}\mu n\mathbf{E} \tag{68}$$

$$n\mathbf{S} = -C_e\Big(\mu n\mathrm{k_B}T_m\mathbf{E} + \frac{\mathrm{k_B^2}}{\mathrm{q}}\nabla(n\mu T_m^2)\Big) \tag{69}$$

with

$$C_e = \Big(\frac{5}{2} - p\Big)\Big(1 - \frac{\mathrm{k_B}\alpha T_m}{2}\Big) \tag{70}$$

$$\langle\mathcal{E}\rangle = \Big(1 + \frac{5}{2}\alpha\mathrm{k_B}\mathrm{T}_n\Big)\frac{3}{2}\mathrm{k_B}T_m \tag{71}$$

Sadovnikov *et al.* [41] showed that Chen's model fails to predict proper velocity profiles and is not consistent with homogeneous simulation results.

D. Model of Tang

Tang *et al.* [35] gave very elaborate expressions for $\hat{\mathrm{U}}$, $\mathbf{p}$ and $\hat{\mathrm{R}}$ to close the equation system. Their discussion aims at a correct handling of the inhomogeneity effects normally

ignored. By observing that $(\hat{\mathrm{U}}-\mathbf{v}\otimes\mathbf{p})$ and $(\hat{\mathrm{R}}+2.5w\mathbf{v}\otimes\mathbf{p})$ show nearly no hysteresis for an n^+-n-n^+ test-structure they proposed the following closure relations

$$\hat{\mathrm{U}} = \frac{2}{3}w\hat{\mathrm{I}} + \mathbf{v}\otimes\mathbf{p} + u(w)\hat{\mathrm{I}} \tag{72}$$

$$\mathbf{p} = m^*\mathbf{v} + 2\alpha m^*\mathbf{S} \tag{73}$$

$$\hat{\mathrm{R}} = \frac{10}{9}w^2\hat{\mathrm{I}} - 2.5w\mathbf{v}\otimes\mathbf{p} + r(w)\hat{\mathrm{I}} \tag{74}$$

with $u(w)$ and $r(w)$ being single-valued fit-functions. They used a Kane dispersion for the MC simulation which might somehow limit the validity of the expressions above. Unfortunately, the additional convective terms are likely to cause numerical problems in an actual multi-dimensional implementation.

E. Model of Smith

In [44] Smith *et al.* derived two non-parabolic equation sets for inhomogeneous and degenerate semiconductors (see also [45], [46]). They used both the Kane dispersion and a simpler power-law approximation after Cassi and Riccò [47] because the Kane dispersion relation cannot be integrated analytically. They showed that the typically employed binomial expansion of the Kane-integrands looses its validity and physically not consistent results are obtained. The power-law approximation, on the other hand, approaches the parabolic limit and has a larger range of validity.

F. Model of Anile

Anile *et al.* [48] and Muscato [49] derived expressions for the closure $\hat{\mathrm{U}}$ and $\hat{\mathrm{R}}$ using the maximum entropy principle. In addition, they were able to derive expressions for the collision terms. They found that their model fulfills Onsagers reciprocity principle and gave a comparison with other hydrodynamic models.

G. Comparison

A comparison of the simple ET model with the expressions given by Thoma *et al.*, Lee *et al.*, Chen *et al.*, and Tang *et al.* for Si bipolar transistors is given by Sadovnikov *et al.* [41] who observed no significant differences in the simulated output characteristics.

VIII. Extensions for Semiconductor Alloys

The derivations given above are restricted to homogeneous materials where the effective carrier masses and the band edge energies do not depend on position. Over the last years extensive research has been made concerning III-V materials and SiGe heterostructure devices. Especially for III-V materials inclusion of the carrier temperature in the transport equations is considered a must. State-of-the-art III-V heterostructure transistors employ many different combinations of materials. In addition differently graded profiles have been used. SiGe bipolar transistors with graded Ge profiles in the base have also been widely investigated. To properly account for the additional driving forces due to changes in the effective masses and the band edge energies the energy-transport models have been extended accordingly. The foundation for these extensions was laid in the pioneering work by Marshak for the drift-diffusion equations [50], [51]. These concepts have been applied to the energy-transport models by Azoff in [5], [52], [13]. In the case of a position-dependent parabolic band structure, the force exerted on an electron is given as

$$\mathbf{F} = -\nabla E_c + \mathcal{E}\frac{\nabla m^*}{m^*} \tag{75}$$

These additional forces give rise to an additional component in the current relation and the electric field is replaced by an effective electric field which also contains the gradient of the band edges.

$$\mathbf{J} = \mu n\nabla E_c + \frac{2}{3}\mu\nabla(nw) - \mu nw\nabla\ln(m^*) \tag{76}$$

$$\mathbf{S} = -\frac{5}{3}\frac{\mathbf{J}w}{\mathrm{q}} - \frac{10}{9}\mu\mathrm{k_B}nw\nabla w \tag{77}$$

An extension to non-parabolic bandstructures has been presented by Smith *et al.* [44], [45].

IX. Multiple Band Models

Bløtekjær's [3] equations were originally devised for semiconductors with multiple bands. Woolard *et al.* [53], [54] extended these expressions for multiple non-parabolic bands in GaAs. Other GaAs models can be found in [55], [56]. Wilson [57] gave an alternate form of the hydrodynamic model which he claims to be more accurate than [3]. Another multivalley non-parabolic energy-transport model was proposed in [58].

X. Band Splitting Models

As device geometries are further reduced without according reduction of the supply voltages, the electric fields occurring inside the devices increase rapidly. Furthermore, strong gradients in the electric field are observed. These highly non-homogeneous field distributions give rise to distribution functions which deviate significantly from the frequently assumed Maxwellian distribution. Furthermore, as has been pointed out in [59], the distribution function is not uniquely described using just the average carrier energy. This is depicted in Fig. 3 which shows some electron distribution functions inside an n^+-n-n^+ test-structure obtained by MC simulation. Points ABC are in the channel while the points D and E are taken from the drain region. In the drain region, the overpopulation of the high-energy tail is obvious, whereas in the channel it is underpopulated, showing a significant thermal tail [60].

Several moment based models have been proposed so far which aim at obtaining some additional information about the DF to the average energy. One approach is to split the energy range at some characteristic energy and handle both energy ranges with a two-population and two-temperature model [61], [62]. As these models were aimed at modeling impact ionization the band gap energy was taken as the characteristic energy. This approach leads to various

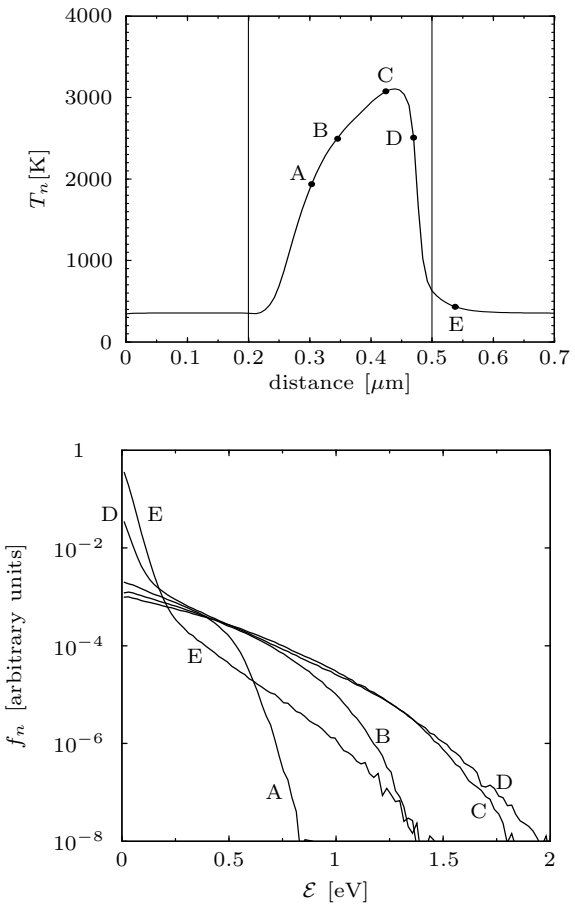

Fig. 3. Electron temperature and distribution function five characteristics points inside the n^+-n-n^+ test-structure. Note that the average energies at the points B and D are the same.

additional macroscopic parameters which model the transitions between the two energy regions. Determination of these parameters relies on carefully set up MC simulations. Due to this specialization to impact ionization, this model would have to be reformulated if another energy range is of interest as is the case for the calculation of gate currents. Thus this approach is difficult to generalize if both effects need to be captured at the same time which is demanded for state-of-the-art devices. A special formulation using two electron populations has been proposed in [63] for those regions where the high-energy tail is heavily populated. In [64] Tang gave a simplified version of the two energy model [61] which used assumptions similar to those made by Cook and Frey [65].

XI. Electro-Thermal Extensions

One of the problems resulting from the reduction in device geometries is that the generated heat has to be kept to a minimum. To capture these self-heating effects, the moment equations have to be extended to account for non-constant lattice temperature. A detailed treatment of this subject was given by Wachutka [66] for the classical DD equations. Chen *et al.* gave an extension for the energy-transport models in [67]. Benvenuti *et al.* introduced a thermal-fully hydrodynamic model in [26]. A detailed discussion can be found in [68].

XII. Critical Issues

The models given above employ various approximations of different severity. As these approximations have been discussed extensively in literature, they will be summarized in the following sections.

A. Closure

The method of moments transforms the BTE into an equivalent, infinite set of equations. One of the severest approximation is the truncation to a finite number of equations (normally three or four). The equation of highest order contains the moment of the next order which has to be suitably approximated using available information, typically the lower order moments. Even though no form of the distribution function needs to be assumed in the derivation, an implicit coupling of the highest order moment and the lower order moments is enforced by this closure. For their generalized HD model, Thoma *et al.* [9], [40] give a maximum error of 30% which can be quite significant. One approach to derive a suitable closure relation is to assume a distribution function and calculate the fourth order moment. Geurts [69] expanded the distribution function around drifted and heated Maxwellian distribution using Hermite polynomials. This gives a closure relation which generalizes the standard Maxwellian closure. However, these closures proved to be numerically unstable for higher electric fields. Liotta *et al.* [70] investigated a closure using an equilibrium Maxwellian which proved to be numerically very efficient but with unacceptable errors for higher electric fields. For a discussion on heated Fermi-Dirac distributions see [44], [46]. Ramaswami and Tang [37] gave a comparison of different closure relations available in literature.

B. Tensor Quantities

An issue which has only been vaguely dealt with is the approximation of the tensors by scalar quantities, such as the carrier mass and the carrier temperature. One-dimensional simulations have been carried out in [8] which indicate that the longitudinal temperature component T_l is larger than the transverse temperature component T_t, indicating that the distribution function is elongated along the field direction and thus that the normally assumed equipartition of the energy is invalid. A rigorous approach has been taken by Pejčinović *et al.* [71] who model four components of the temperature tensor. They observed no significant difference between the scalar temperature and $\mathrm{tr}\,(\hat{\mathrm{T}}_\mathrm{n})/3$ for ballistic diodes and bipolar transistors but a 15 % difference for aggressively scaled MOSFETs in the linear region of the transfer characteristics.

Tang *et al.* [35] observed that the energy tensor is not a single valued function of the average energy and give models using available moments (see (72) and (74)).

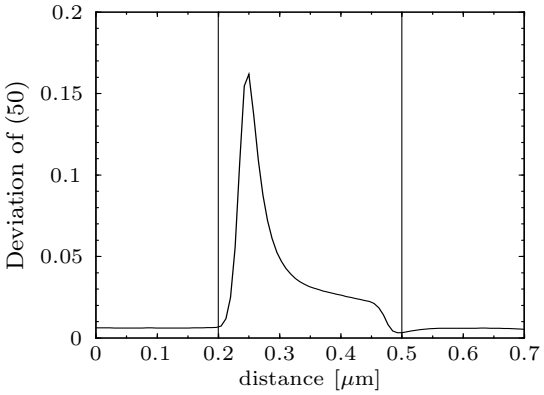

Fig. 4. Effect of approximation (50) on the carrier temperature and the deviation introduced by it obtained by a MC simulation of the n^+-n-n^+ test-structure.

C. Drift Energy vs. Thermal Energy

Another common approximation is the neglection of the drift energy in the average carrier energy [65]

$$w = \frac{m\mathbf{v}^2}{2} + \frac{3}{2}\mathrm{k_B}T_n \approx \frac{3}{2}\mathrm{k_B}T_n \tag{78}$$

As has been pointed out by Baccarani and Wordeman [72], the convective energy can reach values comparable to thermal energy. A plot of the difference of these temperature definitions inside an n^+-n-n^+ test-structure is given in Fig. 4. As can be seen, the error introduced by this approximation is larger than 15 % in the beginning of the channel where the carrier temperature is still low and a velocity overshoot is observed. This effect has been studied in [73].

D. Relaxation Times

The relaxation times have traditionally been derived from homogeneous field measurements or MC simulations. For homogeneous fields, there is a unique relationship between the electric field and the carrier temperature via (47) which can be used as a definition for $\tau_\mathcal{E}$. However, due to the modeling of the collision terms, the relaxation times depend on the distribution function. Since the distribution function is not uniquely described by the average energy, models based on the average energy are bound to fail. Furthermore, the band structure plays a dominant role. Nevertheless, all models should be able to correctly reproduce the homogeneous limit. In the following, some models for Si are reviewed.

D.1 Mobility

Two models for the energy dependence of the mobility are frequently used, the model after Baccarani *et al.* [33], [72]

$$\frac{\mu(T_n)}{\mu_0} = \frac{T_L}{T_n} \tag{79}$$

and the model after Hänsch [34], [74]

$$\frac{\mu(T_n)}{\mu_0} = \Big(1 - \frac{3}{2}\frac{\mu_0}{\tau_\mathcal{E} v_s^2}\Big(\frac{\mathrm{k_B}T_L}{\mathrm{q}} + \frac{2}{5}\frac{nS}{J}\Big)\Big)^{-1} \tag{80}$$

For homogeneous materials $S/J = 5\mathrm{k_B}T_n/(2\mathrm{q})$ which can be used to simplify (80) as

$$\frac{\mu(T_n)}{\mu_0} = \Big(1 + \frac{3}{2}\frac{\mu_0 \mathrm{k_B}}{\mathrm{q}\tau_\mathcal{E} v_s^2}\Big(T_n - T_L\Big)\Big)^{-1} \tag{81}$$

As has been shown in [8], [75] expression (81) reproduces the mobility quite well in the regions with increasing $\mathbf{E}$. However, for decreasing $\mathbf{E}$, (80) should be used [8], [32].

Tang *et al.* [35] proposed another expression by separating the homogeneous from the inhomogeneous part of the mobility. They suggest to model the collision term $\mathbf{C}_p$ as

$$n\mathbf{C}_p = \frac{\mathbf{J}}{\mu} = \frac{\mathbf{J}}{\mu^*} + \lambda_p n \nabla \cdot \hat{\mathrm{U}} \tag{82}$$

with μ^* being the homogeneous mobility. The second term of (82) can then be moved to the left-hand-side of (32) to give a Stratton-like energy gradient expression $(1-\lambda_p)n\nabla\hat{\mathrm{U}}$. The quantity λ_p has been extracted from MC simulations as

$$\lambda_p = \begin{cases} 0.15 & \text{for} \quad \nabla\hat{\mathrm{U}} > 0 \\ 0.50 & \text{for} \quad \nabla\hat{\mathrm{U}} < 0 \end{cases} \tag{83}$$

In [76] Tang *et al.* give an improved expression.

D.2 Energy Relaxation Time

The simplest approach at modeling $\tau_\mathcal{E}$ a constant approximation with values in the range $0.3 - 0.4$ ps. Baccarani *et al.* [33], [72] proposed the expression

$$\tau_\mathcal{E}(T_n) = \frac{3}{2}\frac{\mathrm{k_B}\mu_0}{\mathrm{q}v_s^2}\frac{T_n T_L}{T_n + T_L} + \frac{m^*\mu_0}{2\mathrm{q}}\frac{T_n}{T_L} \tag{84}$$

Note that when

$$\tau_\mathcal{E} = \frac{3\mathrm{k_B}\mu_0 T_L}{2\mathrm{q}v_s^2} \tag{85}$$

is used in the Hänsch mobility model [34], the Baccarani and Hänsch models are equivalent in the homogeneous case [77]. It is also to note that (79) should be used together with (84) whereas in the Hänsch approach $\tau_\mathcal{E}$ is only required to be independent of the temperature for (80) to correctly predict the homogeneous limit. A comparison of these two models is given in Fig. 5 where the differences for the non-homogeneous case are visible. A discussion of the inconsistencies resulting from mixing arbitrary energy-dependent mobility and energy relaxation time models can be found in [78].

Agostinelli *et al.* [79] proposed a model which is fit to the data of Fischetti [80]

$$\frac{\tau_\mathcal{E}(W)}{1\ \mathrm{ps}} = \begin{cases} 0.172 + 2.656W - 3.448W^2 & \text{for}\, W \le 0.4 \\ 0.68 & \text{for}\, W > 0.4 \end{cases} \tag{86}$$

with $W = w/(1\mathrm{eV})$. Another more elaborate fit to the data of Fischetti is given in [81]. A maximum value of 0.68 ps

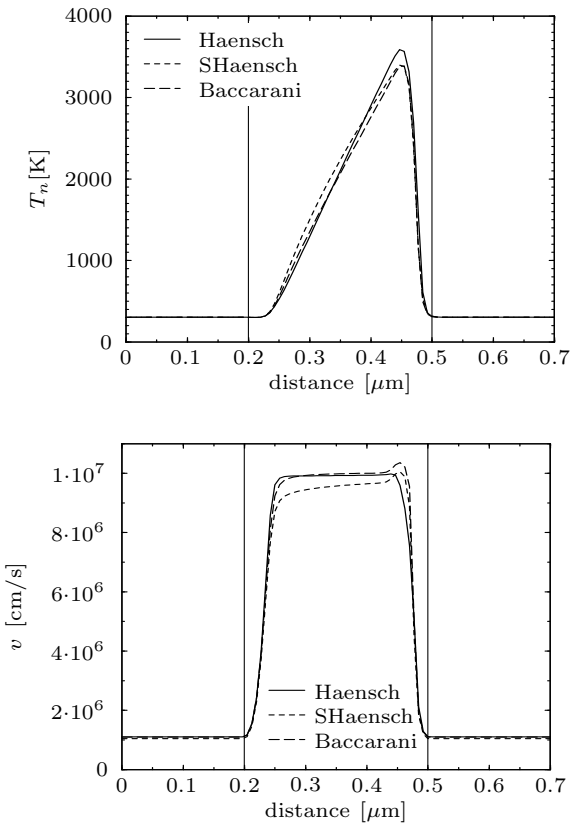

Fig. 5. Electron temperature and velocity distribution inside the n^+-n-n^+ test-structure. μ and $\tau_\mathcal{E}$ were modeled according to (80)+(85) and (79)+(84). SHaensch stands for the simplified Hänsch model (81).

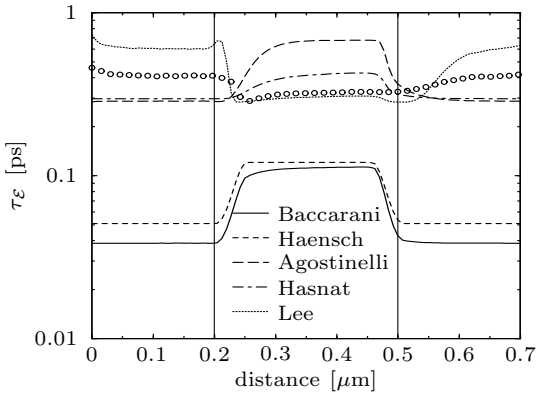

Fig. 6. Comparison of different energy relaxation time models for a n^+-n-n^+ test-structure. The open circles are from a MC simulation.

seems to be too high, and yet another fit to newer data from Fischetti has been published by Hasnat *et al.* [82] as

$$\frac{\tau_\mathcal{E}(W)}{1\text{ ps}} = 0.27 + 0.62W - 0.63W^2 + 0.13W^3 + 0.01W^4 \tag{87}$$

with a maximum value of approximately 0.42 ps. Another expression was given by Lee *et al.* [8] as

$$\tau_\mathcal{E}(w) = t_0 + t_1\eta(w) + t_2\exp(-\beta\eta(w)) \tag{88}$$

with $\eta = w/w_0 - 1$ and the parameters $t_0 = 0.28$ ps, $t_1 = 3$ fs, $t_2 = 2.2$ ps, and $\beta = 10$.

A comparison for these energy relaxation time models is given in Fig. 6 where the differences between the models is obvious. Also shown is the result of our MC simulation. Note that for the Baccarani and Hänsch models the low-field mobility has been calculated with the expression of Caughey and Thomas [83] using the given doping profile and $v_s = 10^7$ cm/s. The other models offer no adjustment of the low-field mobility which can be considered a serious drawback and a cause for inconsistencies. A detailed comparison of the effects of both relaxation times and transport models on the performance of Si bipolar transistors is given in [41].

As the temperature profile occurring inside the device is very sensitive to $\tau_\mathcal{E}$ this disagreement is rather astonishing and further research on this topic is in order.

D.3 Energy Flux Relaxation Time

The ratio of the energy flux mobility and the mobility τ_S/τ_p is usually modeled as a constant with values in the range 0.79-1 [8], [35]. Tang *et al.* [35] proposed an expression in analogy to (82) by separating the homogeneous and inhomogeneous parts. They suggest to model the collision term $\mathbf{C}_{p\mathcal{E}}$ as

$$\mathbf{C}_{p\mathcal{E}} = -\frac{\mathrm{q}\mathbf{S}}{\mu_S^*} + \lambda_{p\mathcal{E}}\nabla\cdot\hat{\mathrm{R}} \tag{89}$$

with μ_S^* being the homogeneous energy flux mobility. Expressions for μ_S^* and $\lambda_{p\mathcal{E}}$ can be found in [35].

E. Spurious Velocity Overshoot

Models based on Bløtekjær's approach have been frequently associated with spurious velocity overshoot (SVO), that is, non plausible spikes in the velocity characteristics which do not occur in MC simulations. This effect can be seen in Fig. 5 where the SVO is clearly visible. Several theories have been put forward to explain this effect. Some authors argue that it is related to the hysteresis in the mobility [59], whereas others relate it to non-parabolicity effects [42]. Still others argue that it is related to the closure of the energy-transport equation system [73]. The improvement obtained by the non-parabolic model [42] is probably due to the improved closure relation for $\hat{\mathrm{R}}$. As already argued by [73], SVO is not likely to be caused only by the mobility because the mobility is not properly modeled in the whole $\nabla\cdot\hat{\mathrm{U}} < 0$ region and SVO is restricted to a very small area. Lee *et al.* [8] investigated SVO using different mobility models and found that improvement is possible when proper mobility models are used. For example, with the Hänsch mobility model (80) these spikes are strongly diminished but not completely removed. Unfortunately, our own MC simulations show that (80) also overestimates the real velocity overshoot at the beginning of the channel. Chen *et al.* [84] proposed a model based on Stratton's approach. In their simplified analysis they used Baccarani's mobility model which gives $\nu_n = -1$ in (30)

and thus removes any thermal diffusion current inside the whole device which is definitely unphysical [35].

XIII. Validity of Moment Based Models

When the critical dimensions of devices shrink below a certain value (around 100 nm for Si at room temperature) MC simulations reveal strong off-equilibrium transport effects such as velocity overshoot and quasi-ballistic transport. So the range of validity for moment based models has been extensively examined. Furthermore, with shrinking device geometries quantum effects gain more importance and limit the validity for the BTE itself [85]. Banoo *et al.* [86] compared the results obtained by (51)-(54) with a DD model and a solution of the BTE obtained by using the scattering matrix approach. They found that this energy-transport model dramatically overestimates both the drain current and the velocity inside the device. Tomizawa *et al.* [87] found through a comparison with MC simulations that relaxation time based models tend to overestimate non-stationary carrier dynamics, especially the energy distribution. Nekovee *et al.* [88] compared moment hierarchy based models with a solution of the BTE and found that moment based models fail in the prediction of ballistic diodes because the moment hierarchy converges to slowly. A similar conclusion was drawn by Liotta *et al.* [70] who found that a hierarchy containing 12 moment equations was needed to reproduce results similar to those obtained by spherical harmonics expansions.

XIV. Simplified Models

Despite of the limitations and approximations contained in the moment equations given above, the solution procedure can be quite involved. Thus several authors tried to find suitable approximations to simplify the analysis. These approximations were frequently used in post-processors to account for an average energy distribution different from the local approximation. Slotboom *et al.* [89] used this technique to calculate energy-dependent impact ionization rates via a post-processing model. Cook *et al.* [65], [90] proposed a simplified model by using the approximations $v_x \gg v_y$ and $E_x \gg E_y$ in a two-dimensional Si MESFET to yield

$$\frac{\partial w}{\partial x} = \frac{21}{20}\mathrm{q}E_x - \frac{9}{20}\Big(\frac{40}{9}\frac{m^*(w-w_0)}{v_x\tau_{\mathcal{E}}\tau_p} + \mathrm{q}^2E_x^2\Big)^{1/2} \qquad (90)$$

Thus, the energy balance equation and the continuity equation become decoupled and the complexity of the problem is considerably reduced. Approximations for GaAs were also given. Although these approximations might have delivered promising results, progress in the size-reduction of state-of-the-art devices makes the assumptions $v_x \gg v_y$ and $E_x \gg E_y$ questionable. In particular, for deep-submicron MOSFETs, velocity overshoot influences the electric field distribution for a given bias condition, and effectively defines a higher drain saturation voltage which in turn defines a higher current [91].

To bring the energy-transport equations into a self-adjoint form Lin *et al.* [92] approximated the carrier temperature in the diffusion coefficient by the lattice temperature as

$$\mathbf{J} = \mu\mathrm{k_B}\underbrace{T_n}_{\approx T_L}\nabla n + \mu\mathrm{k_B}n\nabla T_n + \mathrm{q}n\mu\mathbf{E} \qquad (91)$$

which will underestimate the diffusion current by a factor of $\approx 10-20$ for state-of-the-art devices and can therefore not be recommended.

XV. Conclusions

Many different hydrodynamic and energy-transport models have been published so far. They rely on either Stratton's or Bløtekjær's approach to find a suitable set of balance and flux equations. In Stratton's approach there is no need to invoke Fourier's law to close the equation system due to the relationship (23). Bløtekjær used only three moments and closed the equation system by approximating the heat flux with Fourier's law. This closure has frequently been replaced by equations obtained from the fourth moment of the BTE.

Uncertainties are introduced by the approximation of the collision terms which are modeled via relaxation times and by the derivation of closure relations. Expression for these are normally extracted from homogeneous MC simulations. As has been clearly shown, homogeneous MC simulation data are not sufficient for the simulation of state-of-the-art devices as neither the relation times nor the closure relations are single-valued functions of the average energy. This used to be one of the advantages of the macroscopic transport models over the MC method because measured $\mu(E)$ characteristics could be directly incorporated into the simulation which is not possible for the microscopic approach taken in the MC method. Unfortunately, data for inhomogeneous situations are difficult to extract from measurements due to the complex interaction between the various parameters. Therefore, MC simulations of n^+-n-n^+ test-structure were performed to extract the desired data.

Another problem is directly related to the MC simulations itself. As has been frequently reported, the results obtained by available MC codes differ significantly [93]. Especially impurity scattering is difficult to model [94] and any error in the mobility influences the simulated energy relaxation times were large differences were found in the published data.

References

[1] S. Selberherr, *Analysis and Simulation of Semiconductor Devices*, Springer, Wien–New York, 1984.
[2] R. Stratton, "Diffusion of Hot and Cold Electrons in Semiconductor Barriers," *Physical Review*, vol. 126, no. 6, pp. 2002–2014, 1962.
[3] K. Bløtekjær, "Transport equations for electrons in two-valley semiconductors," *IEEE Trans.Electron Devices*, vol. ED-17, no. 1, pp. 38–47, 1970.
[4] D.K. Ferry, *Semiconductors*, Macmillan, New York, 1991.
[5] E.M. Azoff, "Generalized Energy-Momentum Conservation Equation in the Relaxation Time Approximation," *Solid-State Electron.*, vol. 30, no. 9, pp. 913–917, 1987.

[6] C.L. Gardner, "The Classical and Quantum Hydrodynamic Models," in *Proceedings of the International Workshop on Computational Electronics*, University of Leeds, Aug. 1993, pp. 25–36.

[7] K. Hess, *Theory of Semiconductor Devices*, IEEE Press, 2000.

[8] S.-C. Lee and T.-W. Tang, "Transport Coefficients for a Silicon Hydrodynamic Model Extracted from Inhomegenous Monte-Carlo Calculations," *Solid-State Electron.*, vol. 35, no. 4, pp. 561–569, 1992.

[9] R. Thoma, A. Emunds, B. Meinerzhagen, H.J. Peifer, and W.L. Engl, "Hydrodynamic Equations for Semiconductors with Nonparabolic Band Structure," *IEEE Trans.Electron Devices*, vol. 38, no. 6, pp. 1343–1353, June 1991.

[10] E.O. Kane, "Band Structure of Indium Antimonide," *J.Phys.Chem.Solids*, vol. 1, pp. 249–261, 1957.

[11] T.J. Bordelon, X.-L. Wang, C.M. Maziar, and A.F. Tasch, "Accounting for Bandstructure Effects in the Hydrodynamic Model: A First-Order Approach for Silicon Device Simulation," *Solid-State Electron.*, vol. 35, no. 2, pp. 131–139, 1992.

[12] Y. Apanovich, E. Lyumkis, B. Polsky, A. Shur, and P. Blakey, "Steady-State and Transient Analysis of Submicron Devices Using Energy Balance and Simplified Hydrodynamic Models," *IEEE Trans.Computer-Aided Design*, vol. 13, no. 6, pp. 702–711, 1994.

[13] E.M. Azoff, "Energy Transport Numerical Simulation of Graded AlGaAs/GaAs Heterojunction Bipolar Transistors," *IEEE Trans.Electron Devices*, vol. ED-36, no. 4, pp. 609–616, 1989.

[14] M. Lundstrom, *Fundamentals of Carrier Transport*, vol. X of *Modular Series on Solid State Device*, Addison-Wesley, 1990.

[15] K. Souissi, F. Odeh, H.H.K. Tang, and A. Gnudi, "Comparative Studies of Hydrodynamic and Energy Transport Models," *COMPEL*, vol. 13, no. 2, pp. 439–453, 1994.

[16] C.L. Gardner, "Numerical Simulation of a Steady-State Electron Shock Wave in a Submicrometer Semiconductor Device," *IEEE Trans.Electron Devices*, vol. 38, no. 2, pp. 392–398, Feb. 1991.

[17] A.M. Anile, C. Maccora, and R.M. Pidatella, "Simulation of n^+-n-n^+ Devices by a Hydrodynamic Model: Subsonic and Supersonic Flows," *COMPEL*, vol. 14, no. 1, pp. 1–18, 1995.

[18] E. Fatemi, J.Jerome, and S. Osher, "Solution of the hydrodynamic device model using high-order nonoscillatory shock capturing algorithms," *IEEE Trans.Computer-Aided Design*, vol. 10, no. 2, pp. 232–244, Feb. 1991.

[19] E. Thomann and F. Odeh, "On the Well-Posedness of the Two-Dimensional Hydrodynamic Model for Semiconductor Devices," *COMPEL*, vol. 9, no. 1, pp. 45–57, 1990.

[20] D.L. Scharfetter and H.K. Gummel, "Large-Signal Analysis of a Silicon Read Diode Oscillator," *IEEE Trans.Electron Devices*, vol. ED-16, no. 1, pp. 64–77, 1969.

[21] C.C. McAndrew, K. Singhal, and E.L. Heasell, "A Consistent Nonisothermal Extension of the Scharfetter-Gummel Stable Difference Approximation," *IEEE Electron Device Lett.*, vol. EDL-6, no. 9, pp. 446–447, 1985.

[22] 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.Computer-Aided Design*, vol. 7, no. 2, pp. 231–242, 1988.

[23] W.-S. Choi, J.-G. Ahn, Y.-J. Park, H.-S. Min, and C.-G. Hwang, "A Time Dependent Hydrodynamic Device Simulator SNU-2D With New Discretization Scheme and Algorithm," *IEEE Trans.Computer-Aided Design*, vol. 13, no. 7, pp. 899–908, 1994.

[24] A. Gnudi and F. Odeh, "An Efficient Discretization Scheme for the Energy-Continuity Equation in Semiconductors," in *Simulation of Semiconductor Devices and Processes*, G. Baccarani and M. Rudan, Eds., Bologna, 1988, vol. 3, pp. 387–390, Tecnoprint.

[25] M. Rudan and F. Odeh, "Multi-Dimensional Discretization Scheme for the Hydrodynamic Model of Semiconductor Devices," *COMPEL*, vol. 5, no. 3, pp. 149–183, 1986.

[26] A. Benvenuti, W.M. Coughran, and Mark R. Pinto, "A Thermal-Fully Hydrodynamic Model for Semiconductor Devices and Applications to III-V HBT Simulation," *IEEE Trans.Electron Devices*, vol. ED-44, no. 9, pp. 1349–1359, 1997.

[27] N.R. Aluru, K.H. Law, P.M. Pinsky, and R.W. Dutton, "An Analysis of the Hydrodynamic Semiconductor Device Model – Boundary Conditions and Simulations," *COMPEL*, vol. 14, no. 2/3, pp. 157–185, 1995.

[28] Y. Zhang and M.E. Nokali, "A Hydrodynamic Transport Model and its Applications in Semiconductor Device Simulation," *Solid-State Electron.*, vol. 36, no. 12, pp. 1689–1696, 1993.

[29] R. Stratton, "Semiconductor Current-Flow Equations (Diffusion and Degeneracy)," *IEEE Trans.Electron Devices*, vol. ED-19, no. 12, pp. 1288–1292, 1972.

[30] P.T. Landsberg and S.A. Hope, "Two Formulations of Semiconductor Transport Equations," *Solid-State Electron.*, vol. 20, pp. 421–429, 1977.

[31] P.T. Landsberg, "Dgradν or grad(Dν) ?," *IEEE Trans.Electron Devices*, 1983.

[32] T.-W. Tang and Haitao Gan, "Two Formulations of Semiconductor Transport Equations Based on Spherical Harmonic Expansion of the Boltzmann Transport Equation," *IEEE Trans.Electron Devices*, vol. ED-47, no. 9, pp. 1726–1732, 2000.

[33] G. Baccarani, M. Rudan, R. Guerrieri, and P. Ciampolini, "Physical Models for Numerical Device Simulation," in *Process and Device Modeling*, W.L. Engl, Ed. 1986, vol. 1 of *Advances in CAD for VLSI*, pp. 107–158, North-Holland.

[34] W. Hänsch, *The Drift Diffusion Equation and its Application in MOSFET Modeling*, Springer, Wien, New York, 1991.

[35] T.W. Tang, S. Ramaswamy, and J. Nam, "An Improved Hydrodynamic Transport Model for Silicon," *IEEE Trans.Electron Devices*, vol. 40, no. 8, pp. 1469–1476, 1993.

[36] M. Ieong and T.-W. Tang, "Influence of Hydrodynamic Models on the Prediction of Submicrometer Device Characteristics," *IEEE Trans.Electron Devices*, vol. 44, no. 12, pp. 2242–2251, 1997.

[37] S. Ramaswamy and T.-W Tang, "Comparison of Semiconductor Transport Models Using a Monte Carlo Consistency Check," *IEEE Trans.Electron Devices*, vol. ED-41, no. 1, pp. 76–83, 1994.

[38] M.C. Vecchi and L.G. Reyna, "Generalized Energy Transport Models for Semiconductor Device Simulation," *Solid-State Electron.*, vol. 37, no. 10, pp. 1705–1716, 1994.

[39] M.V. Fischetti and S.E. Laux, "Monte Carlo Analysis of Electron Transport in Small Semiconductor Devices Including Band-Structure and Space-Charge Effects," *Physical Review B*, vol. 38, no. 14, pp. 9721–9745, 1988.

[40] R. Thoma, A. Emunds, B. Meinerzhagen, H. Peifer, and W.L. Engl, "A Generalized Hydrodynamic Model Capable of Incorporating Monte Carlo Results," in *Intl.Electron Devices Meeting*, 1989, pp. 139–142.

[41] A.D. Sadovnikov and D.J. Roulston, "A Study of the Influence of Hydrodynamic Model Effects on Characteristics of Silicon Bipolar Transistors," *COMPEL*, vol. 12, no. 4, pp. 245–262, 1993.

[42] T.J. Bordelon, X.-L. Wang, Ch.M. Maziar, and A.F. Tasch, "An Efficient Non-Parabolic Formulation of the Hydrodynamic Model for Silicon Device Simulation," in *Intl.Electron Devices Meeting*, 1990, pp. 353–356.

[43] D. Chen, E.C. Kan, U. Ravaioli, C.-W. Shu, and R.W. Dutton, "An Improved Energy Transport Model Including Nonparabolicity and Non-Maxwellian Distribution Effects," *IEEE Electron Device Lett.*, vol. 13, no. 1, pp. 26–28, 1992.

[44] A.W. Smith and K.F. Brennan, "Non-Parabolic Hydrodynamic Formulations for the Simulation of Inhomogeneous Semiconductor Devices," *Solid-State Electron.*, vol. 39, no. 11, pp. 1659–1668, 1996.

[45] A.W. Smith and A. Rohatgi, "Non-Isothermal Extension of the Scharfetter-Gummel Technique for Hot Carrier Transport in Heterostructure Simulations," *IEEE Trans.Electron Devices*, vol. ED-12, pp. 1515–1523, 1993.

[46] A.W. Smith and K.F. Brennan, "Comparison of Non-Parabolic Hydrodynamic Simulations for Semiconductor Devices," *Solid-State Electron.*, vol. 39, no. 7, pp. 1055–1063, 1996.

[47] D. Cassi and B. Riccò, "An Analytical Model of the Energy Distribution of Hot Electrons," *IEEE Trans.Electron Devices*, vol. 37, no. 6, pp. 1514–1521, June 1990.

[48] A. Anile and V. Romano, "Hydrodynamical Modeling of Charge Carrier Transport in Semiconductors," http://www.dipmat.unict.it/~anile/preprint.html, 1999, Summer school on Industrial Mathematics, IST Lisboa, Portugal.

[49] O. Muscato, "The Onsager Reciprocity Principle as a check of consistency for semiconductor carrier transport models," accepted for publication in Physica A, preprint under http://www.dipmat.unict.it/~anile/preprint.html.

[50] A.H. Marshak and K.M. van Vliet, "Electrical Current in Solids with Position-Dependent Band Structure," *Solid-State Electron.*, vol. 21, pp. 417–427, 1978.

[51] A.H. Marshak and C.M. VanVliet, "Electrical Current and Carrier Density in Degenerate Materials with Nonuniform Band Structure," *Proc.IEEE*, vol. 72, no. 2, pp. 148–164, 1984.

[52] E.M. Azoff, "Closed-Form Method for Solving the Steady-State Generalized Energy-Momentum Conservation Equations," *COMPEL*, vol. 6, no. 1, pp. 25–30, 1987.

[53] D.L. Woolard, R.J. Trew, and M.A. Littlejohn, "Hydrodynamic Hot-Electron Transport Model With Monte Carlo-Generated Transport Parameters," *Solid-State Electron.*, vol. 31, no. 3/3, pp. 571–574, 1988.

[54] D. L. Woolard, H. Tian, R. J. Trew, M. A. Littlejohn, and K. W. Kim, "Hydrodynamic Electron-Transport Model: Nonparabolic Corrections to the Streaming Terms," *Physical Review B*, vol. 44, no. 20, pp. 11119–11132, 1991.

[55] R.A. Stewart and J.N. Churchill, "A Fully Nonparabolic Hydrodynamic Modle for Describing Hot Electron Transport in GaAs," *Solid-State Electron.*, vol. 33, no. 7, pp. 819–829, 1990.

[56] R.A. Stewart, L. Ye, and J.N. Churchill, "Improved Relaxation-Time Formulation of Collision Terms for Two-Band Hydrodynamic Models," *Solid-State Electron.*, vol. 32, no. 6, pp. 497–502, 1989.

[57] C.L. Wilson, "Hydrodynamic Carrier Transport in Semiconductors with Multiple Band Minima," *IEEE Trans.Electron Devices*, vol. ED-35, no. 2, pp. 180–187, 1988.

[58] J.C. Cao and X.L. Lei, "Simulation of and $Al_xGa_{1-x}As$ Ballistic Diode Using Mutlivalley Nonparabolic Hydrodynamic Balance Equations," *Solid-State Electron.*, vol. 41, no. 11, pp. 1781–1785, 1997.

[59] G. Wolokin and J. Frey, "Overshoot Effects in the Relaxation Time Approximation," in *Proc.NASECODE VIII*, Vienna, 1992, pp. 107–108.

[60] A. Abramo and C. Fiegna, "Electron Energy Distributions in Silicon at Low Applied Voltages and High Electric Fields," *J.Appl.Phys.*, vol. 80, no. 2, pp. 889–893, 1996.

[61] P.G. Scrobohaci and T.-W. Tang, "Modeling of the Hot Electron Subpopulation and its Application to Impact Ionization in Submicron Silicon Devices–Part I: Transport Equations," *IEEE Trans.Electron Devices*, vol. 41, no. 7, pp. 1197–1205, 1994.

[62] J.-G. Ahn, C.-S. Yao, Y.-J. Park, H.-S. Min, and R.W. Dutton, "Impact Ionization Modeling Using Simulation of High Energy Tail Distributions," *IEEE Electron Device Lett.*, vol. 15, no. 9, pp. 348–350, 1994.

[63] T.J. Bordelon, V.M. Agostinelli, X.-L. Wang, C.M. Maziar, and A.F. Tasch, "Relaxation Time Approximation and Mixing of Hot and Cold Electron Populations," *Electron.Lett.*, vol. 28, no. 12, pp. 1173–1175, 1992.

[64] T.-W. Tang and J. Nam, "A Simplified Impact Ionization Model Based on the Average Energy of Hot-Electron Subpopulation," *IEEE Electron Device Lett.*, vol. 19, no. 6, pp. 201–203, 1998.

[65] R.K. Cook and J. Frey, "An Efficient Technique for Two-Dimensional Simulation of Velocity Overshoot Effects in Si and GaAs Devices," *COMPEL*, vol. 1, no. 2, pp. 65–87, 1982.

[66] G.K. Wachutka, "Rigorous Thermodynamic Treatment of Heat Generation and Conduction in Semiconductor Device Modeling," *IEEE Trans.Computer-Aided Design*, vol. 9, no. 11, pp. 1141–1149, Nov. 1990.

[67] D. Chen, Z. Yu, K.-C Wu, R. Goossens, and R.W. Dutton, "Dual Energy Transport Model with Coupled Lattice and Carrier Temperatures," in *Simulation of Semiconductor Devices and Processes*, S. Selberherr, H. Stippel, and E. Strasser, Eds., Wien, 1993, vol. 5, pp. 157–160, Springer.

[68] S. Sikorski, "A thermodynamical approach to the laws of current carrier and thermal energy transport in semiconductors," *Electron Technology*, vol. 31, no. 2, pp. 227–241, 1998.

[69] B.J. Geurts, "An Extended Scharfetter-Gummel Scheme for High Order Momentum Equations," *COMPEL*, vol. 10, no. 3, pp. 179–194, 1991.

[70] S.F. Liotta and H. Struchtrup, "Moment Equations for Electrons in Semiconductors: Comparison of Spherical Harmonics and Full Moments," *Solid-State Electron.*, vol. 44, pp. 95–103, 2000.

[71] B. Pejčinović, H.H.K. Tang, J.L. Egley, L.R. Logan, and G.R. Srinivasan, "Two-Dimensional Tensor Temperature Extension of the Hydrodynamic Model and Its Applications," *IEEE Trans.Electron Devices*, vol. 42, no. 12, pp. 2147–2155, 1995.

[72] G. Baccarani and M.R. Wordeman, "An Investigation of Steady-State Velocity Overshoot in Silicon," *Solid-State Electron.*, vol. 28, no. 4, pp. 407–416, 1985.

[73] M.A. Stettler, M.A. Alam, and M.S. Lundstrom, "A Critical Examination of the Assumptions Underlying Macroscopic Transport Equations for Silicon Devices," *IEEE Trans.Electron Devices*, vol. 40, no. 3, pp. 733–740, Apr. 1993.

[74] W. Hänsch and M. Miura-Mattausch, "The Hot-Electron Problem in Small Semiconductor Devices," *J.Appl.Phys.*, vol. 60, no. 2, pp. 650–656, 1986.

[75] S.-C. Lee, T.-W. Tang, and D.H. Navon, "Transport Models for MBTE," in *NASECODE VI - Numerical Analysis of Semiconductor Devices and Integrated Circuits*, J.J.H. Miller, Ed., Dublin, 1989, pp. 261–265, Boole Press.

[76] T.-W. Tang, X. Wang, H. Gan, and M.K. Leong, "An Analytic Expression of Thermal Diffusion Coefficient for the Hydrodynamic Simulation of Semiconductor Devices," in *IWCE*, Glasgow, UK, Oct. 2000, pp. 55–56.

[77] T. Grasser, *Mixed-Mode Device Simulation*, Dissertation, Technische Universität Wien, 1999, http://www.iue.tuwien.ac.at.

[78] K. Rahmat, J. White, and D.A. Antoniadis, "Computation of Drain and Substrate Current in Ultra-Short-Channel nMOSFET's Using the Hydrodynamic Model," *IEEE Trans.Computer-Aided Design*, vol. 12, no. 6, pp. 817–824, 1993.

[79] V.M. Agostinelli, T.J. Bordelon, X.L. Wang, C.F. Yeap, C.M. Maziar, and A.F. Tasch, "An Energy-Dependent Two-Dimensional Substrate Current Model for the Simulation of Submicrometer MOSFETs," *IEEE Electron Device Lett.*, vol. 13, no. 11, pp. 554–556, 1992.

[80] M.V. Fischetti, "Monte Carlo Simulation of Transport in Technologically Significant Semiconductors of the Diamond and Zinc-Blende Structures–Part I: Homogeneous Transport," *IEEE Trans.Electron Devices*, vol. 38, no. 3, pp. 634–649, 1991.

[81] B. Gonzales, V. Palankovski, H. Kosina, A. Hernandez, and S. Selberherr, "An Energy Relaxation Time Model for Device Simulation," *Solid-State Electron.*, vol. 43, pp. 1791–1795, 1999.

[82] K. Hasnat, C.-F. Yeap, S. Jallepalli, S.A. Hareland, W.-K Shih, V.M. Agostinelli, A.F. Tasch, and C.M. Maziar, "Thermionic Emission Model of Electron Gate Current in Submicron NMOSFETs," *IEEE Trans.Electron Devices*, vol. ED-44, no. 1, pp. 129–138, 1997.

[83] D.M. Caughey and R.E. Thomas, "Carrier Mobilities in Silicon Empirically Related to Doping and Field," *Proc.IEEE*, vol. 52, pp. 2192–2193, 1967.

[84] D. Chen, E. Sangiorgi, M.R. Pinto, E.C. Kan, U. Ravaioli, and R.W. Dutton, "Analysis of Spurious Velocity Overshoot in Hydrodynamic Simulations," in *Intl. Workshop on Numerical Modeling of Processes and Devices for Integrated Circuits NUPAD IV*, Seattle, Washington, May 1992, pp. 109–114.

[85] Z. Yu, R.W. Dutton, and R.A. Kiehl, "Circuit/Device Modeling at the Quantum Level," *IEEE Trans.Electron Devices*, vol. ED-47, no. 10, pp. 1819–1825, 2000.

[86] K. Banoo and M.S. Lundstrom, "Electron Transport in a Model Si Transistor," *Solid-State Electron.*, vol. 44, pp. 1689–1695, 2000.

[87] M. Tomizawa, K. Yokoyama, and A. Yoshii, "Nonstationary Carrier Dynamics in Quarter-Micron Si MOSFET's," *IEEE Trans.Computer-Aided Design*, vol. 7, no. 2, pp. 254–258, 1988.

[88] M. Nekovee, B.J. Geurts, H.M.J. Boots, and M.F.H. Schuurmans, "Failure of Extended Moment Equation Approaches to Describe Ballistic Transport in Submicron Structures," *Physical Review B*, vol. 45, no. 10, pp. 6643–6651, 1992.

[89] J.W. Slotboom, G. Streutker, M.J. v. Dort, P.H. Woerlee, A. Pruijmboom, and D.J. Gravesteijn, "Non-Local Impact Ionization in Silicon Devices," in *Intl.Electron Devices Meeting*, 1991, pp. 127–130.

[90] R.K. Cook and J. Frey, "Two-Dimensional Numerical Simulation of Energy Transport Effects in Si and GaAs MESFET's," *IEEE Trans.Electron Devices*, vol. ED-29, no. 6, pp. 970–977, 1982.

[91] D. Chang and J.G. Fossum, "Simplified Energy-Balance Model for Pragmatic Multi-Dimensional Device Simulation," *Solid-State Electron.*, vol. 41, no. 11, pp. 1795–1802, 1997.

[92] Q. Lin, N. Goldsman, and G.-C. Tai, "Highly Stable and Routinely Convergent 2-Dimensional Hydrodynamic Device Simulation," *Solid-State Electron.*, vol. 37, no. 2, pp. 359–371, 1994.

[93] A. Abramo *et al.*, "A Comparison of Numerical Solutions of the Boltzmann Transport Equation for High-Energy Electron Transport Silicon," *IEEE Trans.Electron Devices*, vol. 41, no. 9, pp. 1646–1654, 1994.

[94] H. Kosina and G. Kaiblinger-Grujin, "Ionized-impurity scattering of majority electrons in silicon," *Solid-State Electron.*, vol. 42, no. 3, pp. 331–338, 1998.