JOURNAL OF APPLIED PHYSICS **97**, 093710 (2005)

Nonparabolic macroscopic transport models for device simulation based on bulk Monte Carlo data

T. Grasser[a] and R. Kosik
Christian Doppler Laboratory for TCAD in Microelectronics at the Institute for Microelectronics, TU Vienna, Gußhausstraße 27-29, A-1040 Wien, Austria

C. Jungemann
NST, TU Braunschweig, 38023 Braunschweig, Germany

H. Kosina and S. Selberherr
Institute for Microelectronics, TU Vienna, Gußhausstraße 27-29, A-1040 Wien, Austria

(Received 28 September 2004; accepted 2 February 2005; published online 20 April 2005)

We derive higher-order macroscopic transport models for semiconductor device simulation from Boltzmann's transport equation using the method of moments. To obtain a tractable equation set suitable for numerical implementation the validity of the diffusion limit will be assumed which removes the convective terms from the equation system. The infinite hierarchy of equations is then truncated at the orders two (drift-diffusion model), four (energy-transport model), and six. Nonparabolicity correction factors are included in the streaming terms. Closure relations for the highest-order moments are obtained from a cold Maxwell distribution (drift-diffusion) and a heated Maxwell distribution (energy-transport). For the six moments model this issue is more complicated. In particular, this closure relation is identified to be crucial both in terms of accuracy and in terms of numerical stability. Various possible closure relations are discussed and compared. In addition to the closure of the highest-order moment, various transport parameters such as mobilities and relaxation times appear in the models and need to be accurately modeled. Particularly for higher-order transport models this is a complicated issue and since the analytical models used in our previous attempts did not deliver satisfactory results we extract all these parameters using homogeneous Monte Carlo simulations. Since all macroscopic transport models are based on rather stringent assumptions a practical evaluation is mandatory. Therefore, the proposed six moments model, a corresponding energy-transport model, and the drift-diffusion model are carefully compared to self-consistent Monte Carlo simulations. © *2005 American Institute of Physics*. [DOI: 10.1063/1.1883311]

I. INTRODUCTION

Carrier transport in modern semiconductor devices is commonly investigated using Boltzmann's transport equation, which is solved either directly by means of the Monte Carlo method[1] or by methods based on an expansion of the distribution function in momentum space into a series of spherical harmonics.[2,3] To obtain a physically accurate picture of carrier transport which includes as many effects with as few approximations as possible, the Monte Carlo method is often the method of choice. However, on an engineering level the Monte Carlo method is inconvenient, because it requires large simulation times. In particular, the extraction of certain device parameters is very costly, for instance transit frequencies, small-signal parameters, threshold voltages, and in general situations where only small current levels are involved. Therefore simpler methods have been derived which, for instance, focus only on the first few moments of Boltzmann's transport equation.[1,4,5]

Since the advent of the first simulators that solved the semiconductor equations on one- or two-dimensional geometries, the drift-diffusion model[4] has been the workhorse of Technology CAD (Computer Aided Design) engineers. The drift-diffusion model comprises the first two moments of Boltzmann's equation and is as such the lowest-order truncation. More accurate models which alleviate some of the most stringent assumptions underlying the drift-diffusion equations have already been proposed over 40 years ago.[6,7] Generally speaking, the assumption that the carrier gas is in equilibrium with the lattice is removed by introducing a separate temperature for the carrier gas. However, this requires some additional closure relations which are commonly derived by assuming a heated Maxwellian distribution function.

As of yet, these higher-order transport models have not been widely accepted as a viable substitute. The reasons for this are manifold: A big advantage of the drift-diffusion model is that it contains only one transport parameter, the carrier mobility. This mobility can in principle be measured as a function of various parameters such as the applied electric field, doping concentration, and temperature, fitted by suitable analytical expressions,[8,9] and inserted in the transport model. Unfortunately, this is not the case for higher-order transport models. For instance in the case of the energy-transport model[5,10] we have in addition to the carrier mobility the energy-flux mobility and the energy-relaxation

[a]Electronic mail: Grasser@iue.tuwien.ac.at

 © 2005 American Institute of Physics

time. Since these parameters cannot be directly measured they have to be either modeled[11,12] or extracted from Monte Carlo simulations.[10,13–15] Neither of these options is without problems. Analytical models for the parameters, although accurate on a qualitative level, often fail to reproduce simulation results obtained from Monte Carlo methods. The analytical models use a couple of parameters which are often slightly adjusted to obtain the desired result for a particular application. Unfortunately there is no unique parameter set that fits all requirements, and there is always the danger that the validity of a transport model is extended heuristically by adjusting improper parameters. It is thus difficult to assess whether the model is predictive or not. As an alternative, the extraction of the transport parameters from Monte Carlo simulations has been proposed where we have two options: The first one relies purely on data extracted from homogeneous Monte Carlo simulations,[14,15] whereas other researchers have tried to identify characteristic features of transport parameters by investigating inhomogeneous cases as well.[10,13]

Another subject of intense research is the numerical behavior of higher-order transport models. While the drift-diffusion model is numerically robust and well investigated[4,16] a stable implementation of higher-order transport models is more difficult to obtain.[17–21]

Finally, and most importantly, many doubts regarding the accuracy of fourth-order models like the energy-transport model have been raised. It has often been observed that the additional overhead imposed by these methods is not justifiable when for instance the terminal currents of advanced metal-oxide-semiconductor (MOS) transistors are calculated where even the predictions of the simple drift-diffusion model can be closer to the reference Monte Carlo result.

One of the fundamental motivations to use higher-order models such as energy-transport models is that they provide the average carrier energy which can then be used for modeling hot-carrier effects. Unfortunately, the average carrier energy alone has proven to be insufficient to describe effects like impact-ionization[22,23] and hot-carrier gate currents,[24,25] because the average carrier energy bears no information about the high-energy tail of the distribution function which is critical in these circumstances. Therefore, empirical nonlocal models have to be used.[26] Additional information about the distribution function and the high-energy tail in particular can be obtained by using additional moments for its characterization. Models utilizing six moments have been proposed, which add the kurtosis of the distribution function as an additional solution variable.[27,28] These models have proven to give significantly better descriptions of the distribution function.[29] Here we show that this extended moment hierarchy also improves the predicted velocity profiles and terminal currents.

There are many critical issues concerning the validity of available macroscopic transport models, mainly related to the assumptions made in their derivation. One fundamental assumption underlying basically all macroscopic transport models in practical use for semiconductor device simulation is that some sort of diffusion approximation holds. As a consequence, all convective terms in the transport equations are neglected. There are basically three reasons why this assumptions is so widespread: first, carriers in semiconductors are considered to form a collision dominated system, especially at room temperature. Second, the modeling of these convective terms is rather difficult and third, they introduce hyperbolic modes into the equation system, which makes it considerably more difficult to solve.[30] Monte Carlo simulations show, however, that the convective energy can be of the same order than the thermal energy, particularly in modern submicron devices.[31]

Another critical approximation concerns the modeling of the scattering integral in Boltzmann's equation.[5,32] Rather simple expressions are obtained by applying the macroscopic relaxation time approximation where mobilities are introduced in analogy to the drift-diffusion model. These mobilities are then usually modeled as a function of the average energy only. A more rigorous treatment reveals, however, that the odd moments of the scattering integral depend on the odd moments of the distribution function and thus on all fluxes of the system.[12,33] Unfortunately, closure relations of this type cause an additional coupling between the flux equations and would require a detailed description of the energy-like tensors to obtain an overall improvement of the transport model.[34]

In the following we propose a refined version of our six moments model.[28] The refined version includes the influence of a nonparabolic band structure on the streaming terms. Since the closure of the highest-order moments was found to be crucial for the accuracy of the model and its numerical robustness we discuss various options and identify the most suited one. Considerable effort has been put into studying the numerical behavior and stability and we summarize our findings. Finally, the rather stringent assumptions in the derivation make it difficult to assess the accuracy of the final macroscopic transport model based on theoretical considerations. Therefore, a practical evaluation is required which is attempted in the last sections of this article.

II. THE SCALED BOLTZMANN EQUATION

The macroscopic transport models are derived from the single particle Boltzmann transport equation[35]

$$\frac{\partial f}{\partial t} + \mathbf{u} \cdot \nabla_{\mathbf{r}} f + \mathbf{F} \cdot \nabla_{\mathbf{p}} f = Q[f] \quad (1)$$

using the method of moments. Boltzmann's equation is a semiclassical kinetic equation, which assumes that the motion of carriers is governed by Newton's laws. The scattering operator Q represents the rate of change of f due to collisions and is modeled via Fermi's Golden Rule. We limit the discussion to the nondegenerate case and neglect carrier–carrier scattering. The solution of Boltzmann's equation is the time dependent carrier distribution function $f(\mathbf{k},\mathbf{r},t)$ in the six-dimensional phase space. The group velocity $\mathbf{u}$ appearing in (1) is given by $\mathbf{u}(\mathbf{k},\mathbf{r})=\nabla_{\mathbf{p}}\mathcal{E}(\mathbf{k},\mathbf{r})$, where $\mathcal{E}=\mathcal{E}(\mathbf{k},\mathbf{r})$ represents the carrier energy given by the band structure. The force $\mathbf{F}$ exerted on the carriers depends in general on the electric and magnetic fields and the material properties. Here, we only consider materials with a position-independent band

structure $\mathcal{E}=\mathcal{E}(\mathbf{k})$ and omit the influence of magnetic fields which reduces the external force $\mathbf{F}$ to the electrostatic force $\mathbf{F}(\mathbf{r},t)=sq\mathbf{E}(\mathbf{r},t)$, with s being the sign of the carrier's charge. For the band structure and the scattering rates we use a model similar to the one proposed in Ref. 36 with phonon and impurity scattering. The Monte Carlo code which is used as a reference is described in Ref. 37.

As Boltzmann's equation is time consuming to solve directly, simplifications are sought. A common simplification is to investigate only a few moments of the distribution function, such as the carrier concentration and the average carrier energy. A moment is obtained by multiplying the distribution function with a suitable weight function $\phi=\phi(\mathbf{k})$ and integrating over $\mathbf{k}$-space as

$$\langle\phi\rangle(\mathbf{r},t)=\frac{1}{n(\mathbf{r},t)}\int\phi(\mathbf{k})f(\mathbf{k},\mathbf{r},t)\mathrm{d}^3\mathbf{k}, \tag{2}$$

which is normalized by the carrier concentration $n(\mathbf{r},t)$,

$$n(\mathbf{r},t)=\int f(\mathbf{k},\mathbf{r},t)\mathrm{d}^3\mathbf{k}. \tag{3}$$

In the simplest case these weight functions are chosen to be ascending powers of $\mathbf{k}$. However, to obtain macroscopic quantities such as the average velocity and the average energy, commonly used weight functions are powers of the carrier energy $\mathcal{E}^i$ and the fluxes $\mathbf{p}\mathcal{E}^i$ and $\mathbf{u}\mathcal{E}^i$. Only for parabolic bands the moments resulting from the two flux definitions are related via the effective mass as $\langle\mathbf{p}\mathcal{E}^i\rangle=m^*\langle\mathbf{u}\mathcal{E}^i\rangle$. In general, a more complicated relationship exists between the two, which depends on higher-order moments of the distribution function.[5]

In the following we will consider Boltzmann's equation in diffusion scaling[16,38,39]

$$\kappa\frac{\partial f}{\partial t_s}+\mathbf{u}\cdot\nabla_{\mathbf{r}_s}f+\mathbf{F}\cdot\nabla_{\mathbf{k}_s}f=\frac{1}{\kappa}Q_s[f], \tag{4}$$

where the Knudsen number

$$\kappa=\frac{\tau_0 u_0}{x_0} \tag{5}$$

appears as a scaling parameter representing the mean free path $\tau_0 u_0$ relative to the device dimension. The distribution function will be decomposed into its symmetric and antisymmetric parts as

$$f(\mathbf{k},\mathbf{r},t)=f_S(\mathbf{k},\mathbf{r},t)+\kappa f_A(\mathbf{k},\mathbf{r},t). \tag{6}$$

By inserting (6) into (4) and equating symmetric and antisymmetric terms Boltzmann's equation splits into two equations[40]

$$\frac{\partial f_S}{\partial t}+\mathbf{u}\cdot\nabla_{\mathbf{r}}f_A+\mathbf{F}\cdot\nabla_{\mathbf{p}}f_A=\frac{1}{\kappa^2}Q_S[f_S], \tag{7}$$

$$\kappa^2\frac{\partial f_A}{\partial t}+\mathbf{u}\cdot\nabla_{\mathbf{r}}f_S+\mathbf{F}\cdot\nabla_{\mathbf{p}}f_S=Q_A[f_A]. \tag{8}$$

So far no simplifications have been introduced and (7) and (8) are equivalent to (4). In the diffusion limit, however, which will be assumed to hold in the following, we neglect terms of second order in κ, resulting in a simpler equation system.[28,32,40] Recall that by this assumption the hydrodynamic equation system[7] can be transfered into the energy-transport system.[5]

When (7) and (8) are multiplied by the weight function $\mathcal{E}^i$ and $\mathbf{p}\mathcal{E}^i$ and integrated over $\mathbf{k}$-space we obtain the governing equations for the moments $\langle\mathcal{E}^i\rangle$ and $\langle\mathbf{p}\mathcal{E}^i\rangle$. Due to the structure of Boltzmann's equation additional moments appear in these governing equations. It is a nontrivial task to approximate these additional moments by the moments $\langle\mathcal{E}^i\rangle$ and $\langle\mathbf{u}\mathcal{E}^i\rangle$ which are taken as the unknowns of the equation system. This is referred to as closure of the equation system.

III. GENERAL MACROSCOPIC MOMENT MODELS

In the following the equations determining the first N moments will be derived. The macroscopic transport equations are obtained by multiplying Boltzmann's equation with the appropriate weight functions and integrating the product over $\mathbf{k}$ space. We assume that the Brillouin zone extends towards infinity, which is justified because the distribution function declines exponentially.[41] We apply the weight functions $\mathcal{E}^i$ and $\mathbf{p}\mathcal{E}^i$ with $i\in[0,2]$ to the Boltzmann equation given by (7) and (8) but continue with the unscaled forms. The general transport model will be first formulated in terms of the unknowns $w_i=\langle\mathcal{E}^i\rangle$ and $\mathbf{V}_i=\langle\mathbf{u}\mathcal{E}^i\rangle$. Then, for the six moments model, the variable transformation[42] $w_2=(5/3)w_1^2\beta$ will be introduced. The quantity β is the kurtosis of the distribution function and indicates the deviation from the Maxwellian shape. For nonparabolic bands β is approximately in the range [0.75–1.1] under homogeneous conditions. Inside real devices, however, values in the range [0.7–4.0] have been observed.[42] For instance, a value large than one is observed in regions where a mixture of a hot and cold distribution function exists. This occurs when a hot carrier gas coming from a sort of channel meets a cold gas at the other contact and has been observed in both n^+-n-n^+ structures[42] and MOS transistors.[43] One the other hand, a value smaller than one indicates that the high-energy tail of the distribution function is less populated than the one of a Maxwellian distribution with the same temperature.

The balance equations are obtained as the moments of (7) with the even weight functions $\mathcal{E}^i$ as[40]

$$\frac{\partial n w_i}{\partial t}+\nabla\cdot(n\mathbf{V}_i)-i\mathbf{F}\cdot n\mathbf{V}_{i-1}=-n\frac{w_i-w_{i,\mathrm{eq}}}{\tau_i}, \tag{9}$$

where the relaxation times τ_i are defined as

$$\tau_i=-n\frac{w_i-w_{i,\mathrm{eq}}}{\int\mathcal{E}^i Q_S[f_S]\mathrm{d}^3\mathbf{k}}. \tag{10}$$

Note that due to the choice of solution variables only the equilibrium values $w_{i,\mathrm{eq}}$ depend on the band structure.

To formulate the flux equations we apply the odd weight functions $\mathbf{p}\mathcal{E}^i$ to (8) and obtain[40]

$$\nabla \cdot (n\hat{U}_{i+1}) - n\mathbf{F} \cdot (w_i\hat{I} + i\hat{U}_i) = n\mathbf{Q}_i, \quad (11)$$

with the tensors $\hat{U}_i = \langle \mathbf{u} \otimes \mathbf{p}\mathcal{E}^{i-1} \rangle$ and the odd moments of the scattering integral $n\mathbf{Q}_i = \int \mathbf{p}\mathcal{E}^i Q_A[f_A]\, \mathrm{d}^3\mathbf{k}$.

Regarding the modeling of the scattering integral, inverse mobility tensors $\hat{\mu}_i^{-1}$ are often introduced[5] to establish a relation between $\mathbf{Q}_i$ and $\mathbf{V}_i$ analogously to the drift-diffusion model as

$$n\mathbf{Q}_i \doteq -qn\hat{\mu}_i^{-1}\mathbf{V}_i. \quad (12)$$

When $\mathbf{Q}_i$ and $\mathbf{V}_i$ are assumed to be collinear, we can define scalar mobilities by

$$\mu_i = -q\frac{|\mathbf{V}_i|^2}{\mathbf{V}_i \cdot \mathbf{Q}_i}. \quad (13)$$

For more rigorous models which take the flux dependence of the mobilities into account we refer to Refs. 12, 33, and 40 and the discussion in Ref. 34.

The fluxes can finally be written as

$$n\mathbf{V}_i = -\frac{\mu_i}{q}(\nabla \cdot (n\hat{U}_{i+1}) - n\mathbf{F} \cdot (w_i\hat{I} + i\hat{U}_i)). \quad (14)$$

A. Additional closure relations

By applying the diffusion approximation and the macroscopic relaxation time approximation a considerably simpler equation system is obtained in the form of (9) and (14). However, Eq. (14) still contains the tensors $\hat{U}_i$ which have to be approximated using the unknowns of the equation system w_i and $\mathbf{V}_i$. In the diffusion limit, the energylike tensors for nonparabolic bands can be expressed by scalars[40] as $\hat{U}_i = U_i\hat{I}$, where U_i depends only on the even moments w_i. We now introduce nonparabolicity factors H_i as a generalization of the definition introduced in Ref. 44 and obtain U_i via the trace of $\hat{U}_i$ as

$$U_i = \tfrac{1}{3}\,\mathrm{tr}\,\hat{U}_i = \tfrac{2}{3}w_iH_i. \quad (15)$$

These nonparabolicity correction factors equal unity in the case of parabolic bands and have been modeled as either energy-dependent using a simple analytical expression,[44] by the incorporation of bulk Monte Carlo data,[14] or via analytic models for the distribution function.[32]

The equation system is truncated after N equations, where we consider only even numbers N. The highest-order solution variable is $w_{N/2-1}$. However, in the highest-order equation the moment $\hat{U}_{N/2}$ appears which has to be expressed as a function of the available moments. For fourth-order moments, a heated Maxwellian distribution is often used to derive such a relation[6,7,14] which gives for parabolic bands

$$w_2 = \tfrac{5}{3}w_1^2. \quad (16)$$

In the case of the six moments model the closure of the highest-order model is very critical. Even for closure relations that appeared to reasonably describe the sixth moment in inhomogeneous cases nonphysical oscillations in the results were obtained. In the following the closure relations that have been considered are described.

1. Maximum entropy method

The maximum entropy principle yields, for a given set of prior information, a density which contains least additional information in the sense of Shannon. A maximum entropy approach to the closure problem was applied by Levermore.[45] A physical approach based on the maximum entropy principle was initiated by Anile[30] within the framework of extended thermodynamics. Note, however, that the maximum entropy method itself has been criticized by several authors.[46–48]

According to the maximum entropy principle[49–51] a distribution function model can be obtained by requiring that it maximizes the entropy

$$s = -k_B \int (f \ln f - f)\mathrm{d}^3\mathbf{k}. \quad (17)$$

In the case of a six moments model in the diffusion approximation this results in a distribution function with a symmetric part equal to

$$f_S(\mathcal{E}) = A\,\exp(a_1\mathcal{E} + a_2\mathcal{E}^2). \quad (18)$$

The parameters of this analytical distribution function are determined by the unknowns of the transport model n, w_1, and w_2. The moment w_3 can then be calculated as a function of the lower-order moments.[42] Closed form solutions exist only for fourth-order models ($a_2=0$) and parabolic bands and (18) has therefore, to the best of our knowledge, not been used to close higher-order transport models.[51] Despite this, another more stringent restriction becomes evident at closer inspection: for the six moments model the distribution function (18) can only be given for a restricted set of moments w_1 and w_2. For parabolic bands the following inequality has to be fulfilled

$$\tfrac{3}{5} \leqslant \beta \leqslant 1. \quad (19)$$

This means that (18) cannot be used to represent the distribution function inside the drain region of MOS transistors where $\beta > 1$ is observed in Monte Carlo simulations. One way to resolve this issue is to account for a superposition of a hot and cold distribution function explicitly[42] which, however, requires heuristic assumptions[42,52] and makes the closure relations quite complicated.

2. Grad's method

Expansions of the distribution function around a Maxwellian distribution are frequently used in theoretical physics.

$$f_S(\mathcal{E}) = A\,\exp\left(-\frac{\mathcal{E}}{a_1}\right)(1 + a_2\mathcal{E} + a_3\mathcal{E}^2 + \cdots). \quad (20)$$

The parameters a_l can be related to orthogonal Legendre or Hermite polynomials[53,54] or to a Grad-type expansion.[51,55,56] Expression (20) can also be considered as a linearization of the maximum entropy distribution function.[57] Best results are obtained by setting $a_1 = k_BT$. The parameters A, a_2, and a_3 were determined by matching the even moments.[42] For parabolic bands expression (20) can be easily integrated analyti-

cally and w_3 can be approximated by the lower-order moments as

$$w_3 = \tfrac{35}{9} w_1^3 (3\beta - 2). \tag{21}$$

Unfortunately, these expansions around a Maxwellian distribution converge very poorly to realistic distribution functions and require a high order in the polynomial to reproduce features like the thermal tail,[42] which explains the large number of required moments reported for such transport models.[58,59]

Other problems associated with (20) are that due to the polynomial $f_S(\mathcal{E})$ has several roots and thus becomes negative. However, as long as the resulting moments remain positive, this has often been considered a cosmetic problem. In the particular case of (21) we see, that w_3 becomes negative for $\beta < 2/3$. Although such a small value of β has not been observed in Monte Carlo simulations, they do occur in the six moments model closed with (21), in particular during the Newton procedure.

3. Generalized Maxwellian closure

In Ref. 28 we proposed to use a generalized Maxwellian closure

$$w_3 = \tfrac{35}{9} w_1^3 \beta^c, \tag{22}$$

with c an integer in the range [0–3]. Stable implementations were only obtained for $c=3$, whereas values $c \leqslant 2$ show pronounced oscillations in the numerical solution. However, the results obtained from $c=2$, though often unstable, appeared to better reproduce the Monte Carlo results. We now take a somewhat different approach: by requiring consistency with bulk Monte Carlo simulations we obtain c from a best match of w_3 to w_3^{MC}, which gives $c=2.7$. Note that Sonoda *et al.*[27] used $c=1$.

4. Cumulant closure

From a theoretical point of view probability distributions are better described in terms of cumulants than in terms of moments.[60] In particular it can be shown that an exact description of the displaced and heated Maxwellian distribution function requires only the first three cumulants, with all other cumulants being exactly equal to zero. The higher-order moments, one the other hand, do not vanish.[61] This motivated us to apply the method of cumulants to the closure problem. Using a cumulant expansion method for device simulation was also suggested in Ref. 61. In Ref. 56 the cumulant method was used for the solution of the Boltzmann equation in the context of gas dynamics.

In probability theory and statistics, the characteristic function g is defined via the Fourier transform of the distribution function[56,61]

$$g(\boldsymbol{\xi},\mathbf{r},t) = \int f(\mathbf{p},\mathbf{r},t)\exp(i\boldsymbol{\xi}\cdot\mathbf{p})\mathrm{d}^3\mathbf{p}. \tag{23}$$

Considering the Taylor expansion of $\ln(g)$ we get[61]

$$\ln(g) = \sum_{i=0}^{\infty} \frac{i^n}{n!} C^{(n)}_{i_1\cdots i_n} \xi_{i_1}\cdots\xi_{i_n}, \tag{24}$$

where $C^{(n)}_{i_1\cdots i_n}$ in the nth cumulant.

The first cumulant is equal to the logarithm of the total mass of the distribution function (hence zero for probability distribution functions). The second cumulant is the variance. Cumulants of order greater than two are measures of the non-Maxwellian shape. The third and fourth cumulant are related to the skewness and the kurtosis, respectively. For a distribution function with unit variance the kurtosis is equal to the fourth cumulant. Calculations simplify if central moments are used.

In the case of the six moments model we thus obtain for the highest-order moment

$$w_3 = \tfrac{35}{9} w_1^3 (3\beta - 2). \tag{25}$$

The cumulant closure is a type of Maxwellian closure which is distinguished from a theoretical point of view. Interestingly, in the diffusion approximation the sixth moment calculated with the cumulant method (25) is equivalent to the closure derived via Grad's method (21), which is not the case for higher-order moments where additional nonlinear terms appear.[56] Note that these closures are a linear combination of the generalized Maxwellian closures (22) with the values of $c=0$ and $c=1$.

5. Comparison of the closure relations

A comparison of the various closure relations is given in Fig. 1 using w_1 and w_2 from Monte Carlo simulations. The top figure shows the silicon bulk result with a doping concentration of $N_D = 10^{18}$ cm^{-3}. The cumulant closure and the Grad expansion give reasonable results only for low bias conditions. Good results are obtained for the empirical closure relation with $c=2.7$, which was found by a least square fit over the energy region 0–1.5 eV. In the bottom figure the results for a 100 nm n^+-n-n^+ structure are shown. A least square fit again delivered a value close to $c=2.7$. Best results where obtained in the "channel" region, while clearly all models become inaccurate inside the "drain" region. This is due to the complicated form of the distribution function which differs significantly from the bulk case.[42] Further research on this topic might be in order.

B. The nonparabolic six moments model

Ignoring the terms ∇H_i as in Ref. 14 the fluxes can finally be written as

$$n\mathbf{V}_i = -\frac{2\mu_i H_{i+1}}{3q}\left(\nabla(nw_{i+1}) - n\mathbf{F}w_i\frac{3+2iH_i}{2H_{i+1}}\right). \tag{26}$$

We now introduce the variable transformation[42]

$$w_1 = \tfrac{3}{2}k_B T, \quad w_2 = \tfrac{15}{4}k_B^2 T^2 \beta, \quad w_3 = \tfrac{105}{8}k_B^3 T^3 \gamma. \tag{27}$$

The quantity γ depends on the applied closure relation and equals β^c for the generalized Maxwellian closure and $(3\beta - 2)$ for the cumulant closure. We then obtain

$$n\mathbf{V}_0 = -A_0(\nabla(nk_BT) - n\mathbf{F}h_0), \tag{28}$$

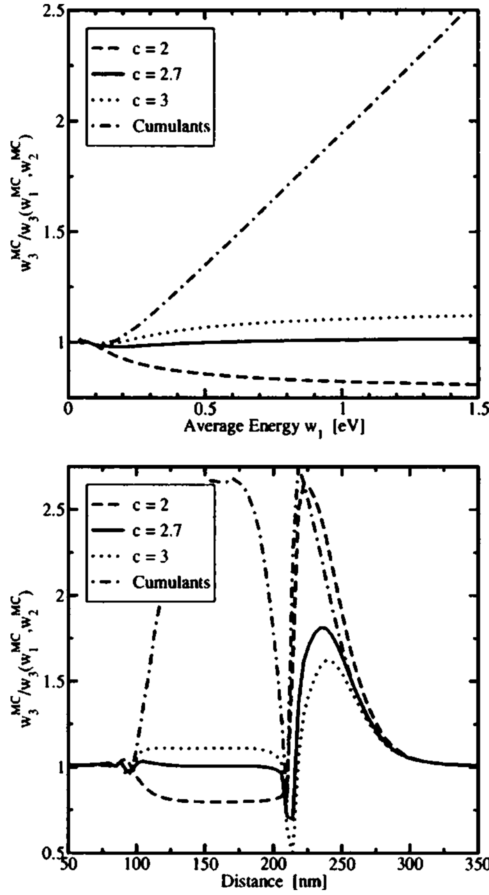

FIG. 1. Comparison of the various closure relations calculated by a Monte Carlo postprocessing step. The top figure shows the silicon bulk result for a doping concentration of $N_D=10^{18}$ cm^{-3} while the bottom figure shows the result for a 100 nm n^+-n-n^+ structure. The generalized Maxwellian closure gives better results than the cumulant closure.

$$n\mathbf{V}_1 = -A_1(\nabla(nk_B^2T^2\beta) - n\mathbf{F}h_1k_BT), \tag{29}$$

$$n\mathbf{V}_2 = -A_2(\nabla(nk_B^3T^3\gamma) - n\mathbf{F}h_2k_B^2T^2\beta), \tag{30}$$

with the auxiliary quantities

$$A_i = \frac{\mu_i H_{i+1}}{2^i q}\frac{1}{3}\prod_{j=0}^{i+1}(1+2j), \tag{31}$$

$$h_i = \frac{3+2iH_i}{(3+2i)H_{i+1}}. \tag{32}$$

Note that for parabolic bands $h_i=1$ holds.

The balance equations are obtained as

$$\frac{\partial n}{\partial t} + \nabla \cdot (n\mathbf{V}_0) = 0, \tag{33}$$

$$C_1\frac{\partial nT}{\partial t} + \nabla \cdot (n\mathbf{V}_1) - \mathbf{F}\cdot n\mathbf{V}_0 = -nC_1\frac{T-T_{\mathrm{eq}}}{\tau_1}, \tag{34}$$

$$C_2\frac{\partial nT^2\beta}{\partial t} + \nabla \cdot (n\mathbf{V}_2) - 2\mathbf{F}\cdot n\mathbf{V}_1 = -nC_2\frac{T^2\beta - T_{\mathrm{eq}}^2\beta_{\mathrm{eq}}}{\tau_2}, \tag{35}$$

with $C_1=3k_B/2$ and $C_2=15k_B^2/4$. The equilibrium values T_{eq} and β_{eq} are calculated from the equilibrium solution of Boltzmann's equation for the nondegenerate case which is the Maxwell distribution. Note that only for parabolic bands the familiar relation $w_1=3k_BT_L/2$ is obtained. Due to the nonparabolic band structure the equilibrium carrier temperature defined via (27) is different from the lattice temperature and we have $T_{\mathrm{eq}}\approx 309.452$ K and $\beta_{\mathrm{eq}}\approx 1$ at $T_L=300$ K. These are also the values used for the Dirichlet boundary conditions for T and β.

IV. MODELING OF THE PHYSICAL PARAMETERS

In addition to the mobilities μ_0, μ_1, and μ_2, the flux equations (28)–(30) contain the nonparabolicity factors H_1, H_2, and H_3, while the balance equations (33)–(35) contain the relaxation times τ_1 and τ_2. These parameters are difficult to model since they all depend on the actual shape of the distribution function and on the band structure. They therefore contain information on hot-carrier and nonparabolicity effects. In addition, the mobilities and relaxation times depend on the scattering rates. Simple empirical models as the ones used in Ref. 28 did not deliver satisfactory results. In particular, a consistent comparison with Monte Carlo simulations is difficult, because the resulting transport model does not reproduce the Monte Carlo results in the homogeneous case. It is therefore problematic to judge the validity of these models in the inhomogeneous case. To avoid uncertainties arising from this issue we extract all physical parameters as a function of the doping concentration and the average energy from homogeneous Monte Carlo simulations. The ratios of the higher-order mobilities μ_1 and μ_2 to the carrier mobility μ_0, the relaxation times τ_1 and τ_2, and the nonparabolicity factors H_i are shown in Fig. 2 for silicon with a doping concentration of 10^{18} cm^{-3}.

All these transport parameters are then plugged into the macroscopic transport models as a function of the average energy. Since all model parameters are obtained from bulk Monte Carlo simulations the transport models are free of fit-parameters which leaves us with "no knobs to turn."[62] Having too many adjustable parameters is a particular inconvenience inherent in many energy-transport models based on analytical models for the mobilities and relaxation times.[5] To avoid fit-parameters is essential for higher-order models, since the interplay between the various parameters is highly complex and the numerical stability of the whole transport model depends significantly on the choice of these parameters. In particular, the model based on the Monte Carlo data outperforms its counterparts based on analytical mobility models[28] significantly, both in terms of numerical stability and in the quantitative agreement of the simulation results with Monte Carlo device simulations.

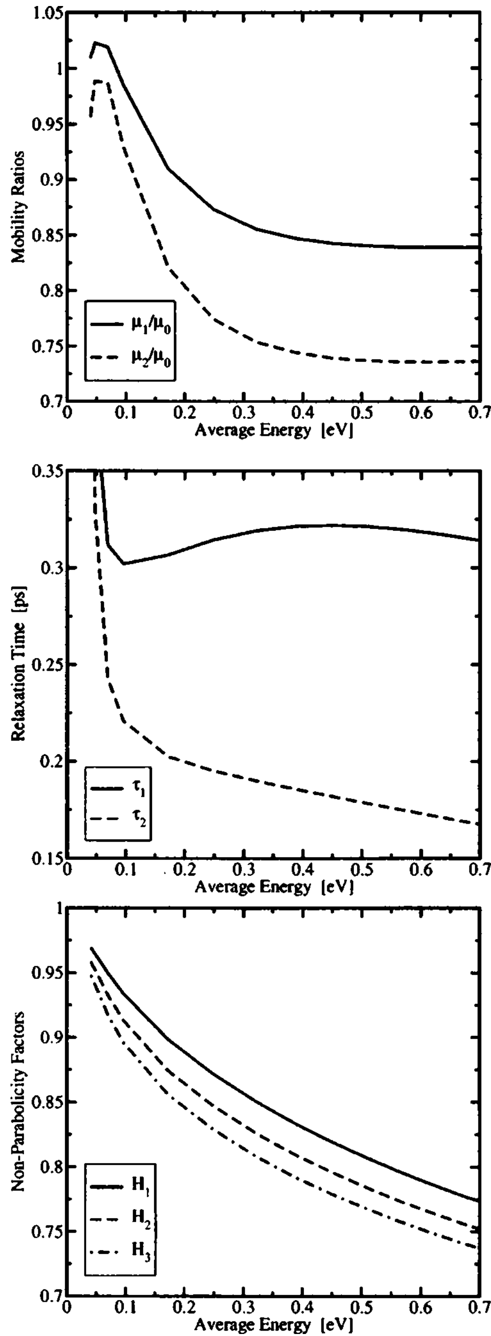

FIG. 2. Ratio of the higher-order mobilities μ_1 and μ_2 to the carrier mobility μ_0 (top), the relaxation times τ_1 and τ_2 (middle); and the nonparabolicity factors H_i (bottom) for a doping concentration of 10^{18} cm^{-3}.

V. NUMERICAL PROPERTIES

A Scharfetter–Gummel style discretization[63] was initially used for the discretization of the parabolic six moments model proposed in Ref. 28. Several suggestions for the discretization of energy-transport equations which generalize the Scharfetter–Gummel method can be found in literature. Methods of this kind which we studied recently are Meinerzhagen's model,[64] Forghieri's method,[18] and variants thereof.[19] All these methods can be unified in the framework of an optimal artificial diffusion method.[21] In addition, a double grid variant was implemented.[65] All these variants were evaluated,[65] but the observed differences were negligible (see the Appendix for details on some of the discretization schemes). In particular for the n^+-n-n^+ structure studied in this article, the convergence behavior was found to be excellent. For instance, the number of Newton iterations is comparable to those required for the solution of the energy-transport models while the condition of the equation system was found to be only slightly degraded. For the limited number of devices considered so far we found that the six moments model converged at maximum bias with roughly the same number of iterations as the energy-transport models starting from the equilibrium initial guess. Differences in the overall simulation time are related to the increased system matrix size and the slightly poorer condition of the equation system. A comparison of the required computation time for the six moments model gave approximately a factor of 4 over the drift-diffusion model and a factor of 2 for the energy-transport model. We consider this a reasonable price for the improved accuracy to expect from the six moments model. For the sake of completeness it is worth mentioning that the simulations based on the macroscopic models took a couple of seconds while the self-consistent Monte Carlo simulations required several hours.

In a numerical implementation some caution is required regarding the nonparabolicity factors. In equilibrium no current is allowed to flow and all other fluxes must be zero as well. As can be easily derived from (28)–(30) this requires the conditions

$$H_2 = \frac{H_1}{5\beta_{\mathrm{eq}}}(3 + 2H_1), \quad H_3 = \frac{H_1\beta_{\mathrm{eq}}}{7\gamma_{\mathrm{eq}}}(3 + 4H_2) \tag{36}$$

to be fulfilled in equilibrium. In this work these parameters are extracted from Monte Carlo simulations. At low fields they show large stochastic errors and H_2 and H_3 calculated from H_1^{MC} using (36) differ from the Monte Carlo results by roughly 1%. Unfortunately, the transport equations are very sensitive to small changes in the nonparabolicity factors. For instance, for the 50 nm device this discrepancy results in a current flow of 20 nA/μm and a temperature difference of 20 K inside the junction regions instead of the expected equilibrium values. Therefore, H_2 and H_3 are calculated from H_1 using (36) for carrier temperatures close to equilibrium, thereby enforcing the proper equilibrium response.

VI. EVALUATION AND DISCUSSION

To investigate the accuracy of the six moments model we consider a series of one-dimensional n^+-n-n^+ structures. These structures display similar features as contemporary MOS and bipolar transistors like a pronounced velocity overshoot and a mixture of a hot and a cold distribution function in the "drain" region. With these structures it is possible to study the basic behavior of macroscopic transport models for very small devices without the additional levels of complexity introduced by two-dimensional MOS devices. The doping concentrations were set to 5×10^{19} cm^{-3} and 10^{17} cm^{-3} and the channel-length was varied from 1000 nm down to 20 nm. Since the accuracy of the transport models was found to be bias-dependent, the bias-conditions were chosen in such a way that a maximum electric field of 300 kV/cm or

600 kV/cm was maintained. In addition to the six moments model we consider the corresponding energy-transport model where the energy-flux relation is closed with a heated Maxwellian distribution. That corresponds with setting β to its equilibrium value in (29). By this measure the equation for the kurtosis becomes decoupled from the other equations and can be calculated in a postprocessing step.[27] One might argue that this closure relation is not that important and that for the modeling of hot-carrier effects this decoupled equation system gives a sufficiently accurate estimate for β. We will see in the sequel that this is not the case. For the sake of completeness, we will also show the results of the corresponding drift-diffusion model where the tabulated mobility is used as a function of the local electric field.

A comparison of the average velocity $\mathbf{V}_0$ and the kurtosis β obtained from the macroscopic models with the Monte Carlo results is shown in Figs. 3 and 4 for three devices. The spurious velocity overshoot is significantly reduced in the six moments model, consistent with previous results,[66] while the kurtosis produced by the decoupled six moments (energy-transport) model is only a poor approximation to the Monte Carlo results for shorter channel-lengths. An accurate kurtosis, however, is a prerequisite for the modeling of hot carrier effects.

The strong influence of the closure relation on the resulting velocity profile and the kurtosis is depicted in Fig. 5. Since convergence could not be obtained for the cumulant and Grad closures for bias voltages larger than a few thermal voltages only the generalized Maxwellian closure (22) is shown. It was found that when the value of c approached 2, the convergence of the equation system was heavily affected until convergence could not be obtained at all. The result shown in Fig. 5 for $c=2$ was obtained by gradually reducing c starting from $c=2.7$. Although convergence was reached (quadratic behavior in the Newton method) and the residua and updates where reduced to very small values, wiggles in the solution can be observed. These wiggles disappear when c is increased. The optimum bulk value $c=2.7$, however, delivered satisfactory results also in the inhomogeneous case.

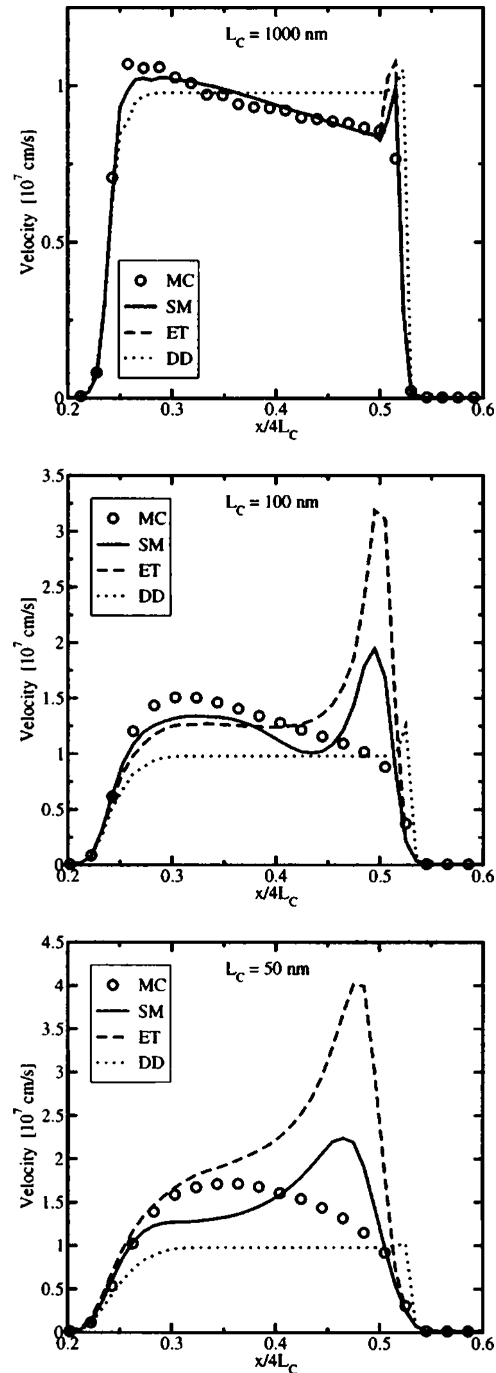

FIG. 3. Comparison of the velocity profile delivered by the macroscopic models with the Monte Carlo results. The six moments model (SM), the energy-transport model (ET), and the drift-diffusion model (DD) are shown.

The errors in the simulated terminal currents of the macroscopic transport models relative to the Monte Carlo simulation are shown in Fig. 6 as a function of the channel-length. With the occurrence of nonlocal effects the accuracy of the drift-diffusion model begins to gradually degrade for channel-lengths smaller than 250 nm. Interestingly, for a maximum bias of 300 kV/cm, a maximum error of 27% is reached at 65 nm and for a further reduction of the channel-length the error begins to gradually decrease. For a maximum bias of 600 kV/cm, however, the maximum error is 31% and occurs at 36 nm and for a further reduction of the channel-length the error decreases only moderately. The reason for this peculiar behavior will be explained shortly.

The higher-order transport models, on the other hand, show a different behavior: They remain accurate down to a certain channel-length but continuously lose their accuracy for further decreased channel-lengths. In general, however, contrary to the drift-diffusion model, the accuracy of the higher-order moments increases for increasing bias. The energy-transport model shows an error of 10% at 65 nm and 55 nm for 300 kV/cm and 600 kV/cm, respectively. For the six moments model we found the 10% error boundary at 42 nm and 30 nm, respectively, while the drift-diffusion model reaches the 10% mark at 200 nm. However, while the error in the drift-diffusion model shows a conservative behavior by staying within reasonable bounds, the error in higher-order transport models deteriorates below a certain channel-length. In particular, for a maximum electric field of 600 kV/cm, the energy-transport model delivers the same

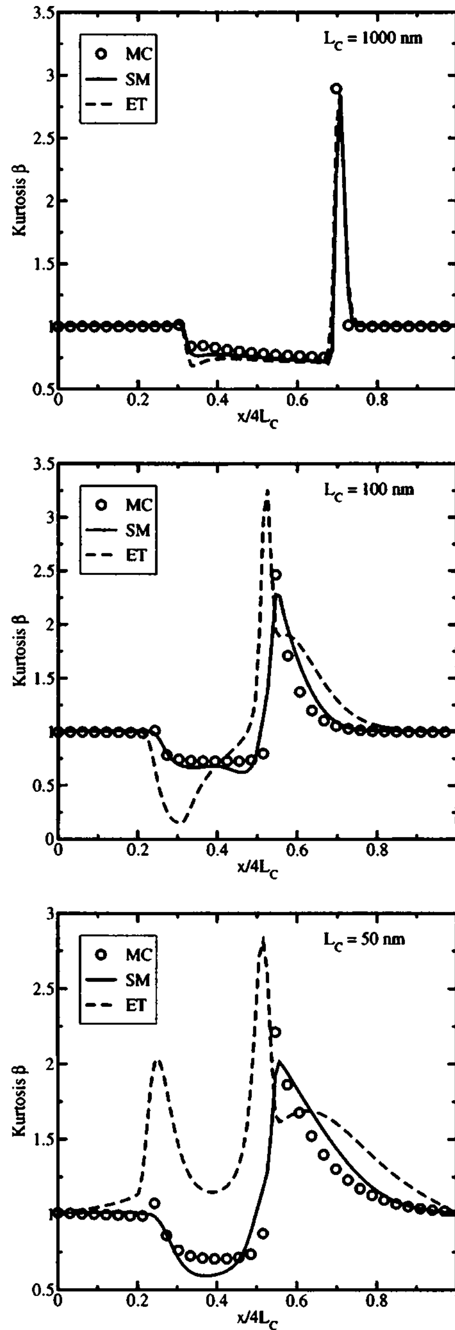

FIG. 4. Comparison of the kurtosis delivered by the macroscopic models with the Monte Carlo results. The six moments model (SM), the energy-transport model (ET), and the drift-diffusion model (DD) are shown.

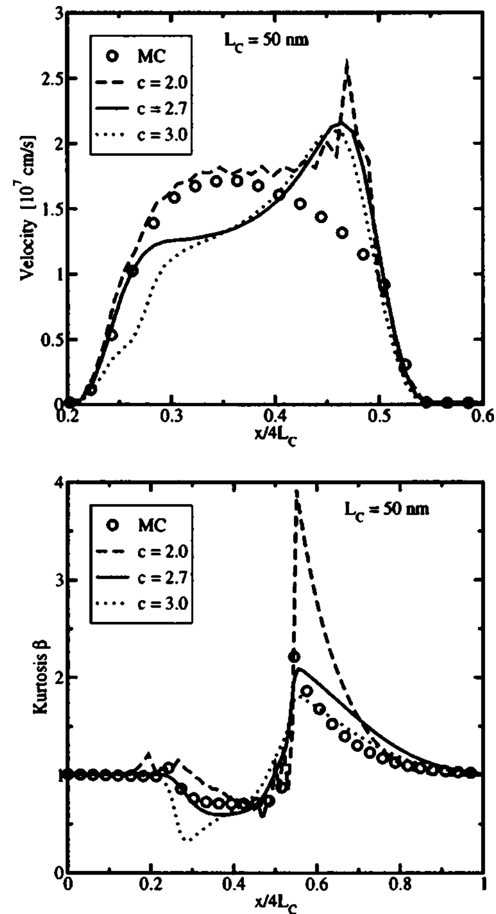

FIG. 5. Influence of the closure relation on the velocity and the kurtosis. Clearly visible are the numerical instabilities arising from $c=2$.

error as the drift-diffusion model at 41 nm (31%) while the error curve of the six moments model intersects the curve of the drift-diffusion model at 24 nm (29%).

A common perception is that the drift-diffusion model delivers smaller currents than the Monte Carlo model. This is only partly correct. To understand this peculiar behavior we have to look at the terminal currents as a function of the terminal voltage as shown in Fig. 7. The bias-dependence of the accuracy can be clearly seen. In particular, all macroscopic transport models give too large currents close to equilibrium. This overestimation of the near-equilibrium conductance is largest for the drift-diffusion model, as can be seen in Fig. 8, and becomes more important for smaller devices in accordance with the findings of Ref. 58. When the bias voltage is increased, however, the drift-diffusion model gradually begins to underestimate the terminal current since nonlocal effects are not accounted for. In total, we have two different counteractive effects and only for larger applied bias we get the expected underestimation. As a consequence, there is an intersection point where terminal currents calculated by the drift-diffusion model equal those of the Monte Carlo model. This fact becomes more apparent when the channel-length is reduced. In the intermediate bias range the accuracy of the currents is thus astonishingly good.

Such an intersection point seems to exist for higher-order models as well, as indicated in Fig. 7, where the terminal current of the 30 nm device predicted by the six moments model intersects the Monte Carlo curve for a maximum electric field of approximately 900 kV/cm. Another indicator is the increased accuracy of the currents for higher bias as apparent in Fig. 6.

VII. CONCLUSIONS

We have derived nonparabolic higher-order transport models for semiconductor device simulation. All physical parameters are taken from bulk Monte Carlo simulations. The particular focus of this work is on the nonparabolic six moments model. In that context the closure of the highest-order moment was identified as being critical for the numeri-

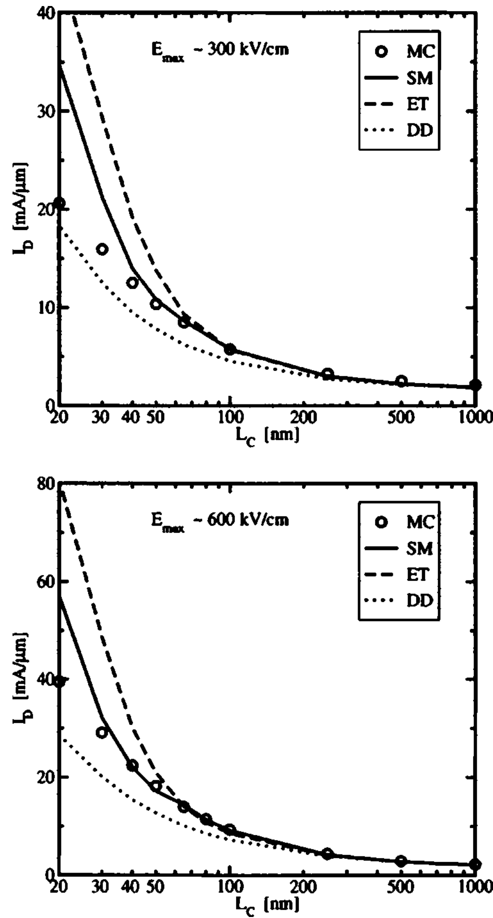

FIG. 6. Comparison of the terminal currents delivered by the macroscopic models as a function of the channel-length. The bias voltage was adapted for each device in such a way that the maximum electric field is approximately kept constant.

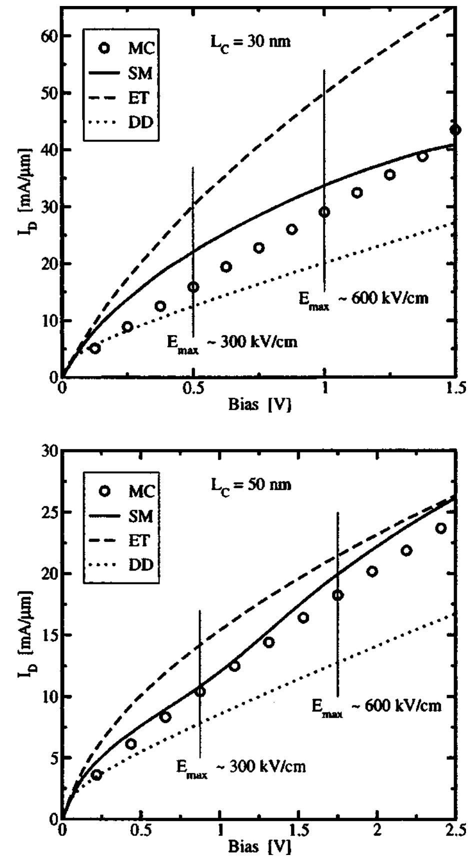

FIG. 7. Comparison of the terminal currents delivered by the macroscopic models for two short devices. As can be seen, the accuracy depends on the bias condition.

cal stability and a model consistent with the homogenous Monte Carlo data has been proposed. A comparison of the terminal currents predicted by these models with the results from self-consistent Monte Carlo simulations reveals some interesting properties. First of all, the heated Maxwellian closure commonly used in energy-transport models has an important impact on the accuracy of the resulting transport model. In particular, the six moments model, which avoids this closure relation, predicts more accurate terminal currents. For example, if a 10% error compared to the Monte Carlo result is defined as acceptable, the drift-diffusion model reaches this error around a channel-length of approximately 200 nm, the energy-transport model at 55 nm, and the six-moments model at 30 nm. Two findings regarding the drift-diffusion model are of particular interest: first, the overestimated near-equilibrium conductance which is counteracted by the missing velocity overshoot gives a smaller underestimation of the terminal currents as expected. Secondly, this underestimation behaves quite conservatively, that is, it remains below 30% for typical bias conditions. The higher-order moment models, on the other hand, behave differently because below a minimum channel-length the errors increase. Above this critical channel-length which is somewhere around 25 nm for the six moments model, the six moments model delivers considerably better results than the drift-diffusion model. In particular, the calculated concentration, average energy, and average of the square of the energy, show good agreement with the Monte Carlo result, which allows accurate modeling of hot-carrier effects like impact ionization and hot-carrier injection into the gate. This is not possible to such a degree with the energy-transport model, not to mention the drift-diffusion model.

APPENDIX: DISCRETIZATION SCHEMES

Here some of the discretization schemes that have been considered are briefly described. For example, we used a generalization of the discretization proposed in Ref. 28 which is based on the general form of the flux relations

$$n\mathbf{V}_i = -A_i\left(\nabla(\xi_i W_i) - Fh_i(\xi_i W_i)\frac{1}{W_i}\right), \tag{A1}$$

with

$$\xi_0 = n, \quad W_0 = k_B T, \tag{A2}$$

$$\xi_1 = nk_B T, \quad W_1 = k_B T\beta, \tag{A3}$$

$$\xi_2 = nk_B^2 T^2\beta, \quad W_2 = k_B T\gamma/\beta. \tag{A4}$$

Applying the standard assumption of Scharfetter–Gummel-type discretization schemes[63] that the projected flux $n\mathbf{V}_i/A_i$ between two gridpoints m and l is constant and that the en-

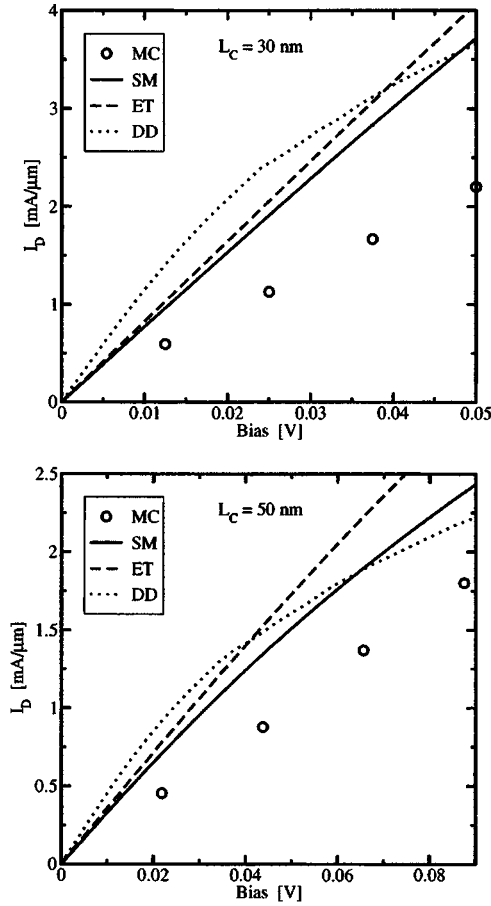

FIG. 8. Near-equilibrium small signal response of the macroscopic models compared to the Monte Carlo results. All macroscopic models overestimate the near-equilibrium conductivity but the inclusion of additional moments improves the result. Note how this effect becomes stronger for decreasing channel-length.

ergy associated to each flux W_i varies linearly on the edge gives the following discretization:

$$nV_i^{m,l} = -\frac{\overline{A}_i}{\Delta x}\widetilde{W}_i(\xi_i^l\mathcal{B}(Y_i) - \xi_i^m\mathcal{B}(-Y_i)), \qquad \text{(A5)}$$

$$Y_i = -\frac{sq\overline{h}_i\Delta\psi + \Delta W_i}{\widetilde{W}_i}, \qquad \text{(A6)}$$

$$\widetilde{W}_i = \frac{\Delta W_i}{\ln(W_i^l/W_i^m)}, \qquad \text{(A7)}$$

where $\mathcal{B}$ is the Bernoulli function. The mobility μ_i and the nonparabolicity factors H_{i+1} appearing in $\overline{A}_i$ and $\overline{h}_i$ are evaluated as a function of $\overline{T}=(T^m+T^l)/2$. For the balance equations which have to be only marginally modified for the nonparabolic case the discretization from Ref. 28 is used. A slightly different variant is obtained by assuming that $n\mathbf{V}_i/(A_iW_i)$ is constant between two gridpoints. This assumption might be justified by considering that the carrier mobilities roughly behave like $1/W_i$ which makes A_iW_i roughly constant. We thus obtain

$$nV_i^{m,l} = -\frac{\overline{A}_i}{\Delta x}\widetilde{W}_i^2\left(\frac{\xi_i^l}{W_i^l}\mathcal{B}(Y_i) - \frac{\xi_i^m}{W_i^m}\mathcal{B}(-Y_i)\right), \qquad \text{(A8)}$$

$$Y_i = -\frac{sq\overline{h}_i\Delta\psi + 2\Delta W_i}{\widetilde{W}_i}. \qquad \text{(A9)}$$

However, the differences in the results when using (A5) and (A8) were found to be marginal for the devices considered.

A somewhat different variant which included the ∇H_i term was also considered. The flux equations are then of the form

$$n\mathbf{V}_i = B_i\left(\nabla(\xi_iW_i) - F(\xi_iW_i)\frac{1}{W_i}\right), \qquad \text{(A10)}$$

with $B_i=A_i/H_{i+1}$ and

$$\xi_0 = n, \quad W_0 = k_BTH_1, \qquad \text{(A11)}$$

$$\xi_1 = nk_BTh_1H_2, \quad W_1 = k_BT\beta/h_1, \qquad \text{(A12)}$$

$$\xi_2 = nk_B^2T^2\beta h_2H_3, \quad W_2 = k_BT\gamma/(\beta h_2). \qquad \text{(A13)}$$

For the discretization we obtain with the assumption $n\mathbf{V}_i/B_i \approx \text{const}$

$$nV_i^{m,l} = -\frac{\overline{B}_i}{\Delta x}\widetilde{W}_i(\xi_i^l\mathcal{B}(Y_i) - \xi_i^m\mathcal{B}(-Y_i)), \qquad \text{(A14)}$$

$$Y_i = -\frac{sq\Delta\psi + \Delta W_i}{\widetilde{W}_i}. \qquad \text{(A15)}$$

The discretization for $n\mathbf{V}_i/(B_iW_i) \approx \text{const}$ can be obtained accordingly. Although more accurate in theory, as the averaging of the nonparabolicity factors is avoided, neither of these variants gave as good results as the variants based on the neglected ∇H_i terms.

[1] M. Lundstrom, *Fundamentals of Carrier Transport* (Cambridge University Press, Cambridge, 2000).
[2] M. Vecchi and M. Rudan, IEEE Trans. Electron Devices **45**, 230 (1998).
[3] C.-K. Lin *et al.*, in *Proceedings of the Simulation of Semiconductor Processes and Devices* (Kyoto, Japan, 1999), pp. 167–170.
[4] S. Selberherr, *Analysis and Simulation of Semiconductor Devices* (Springer, Wien–New York, 1984).
[5] T. Grasser, T.-W. Tang, H. Kosina, and S. Selberherr, Proc. IEEE **91**, 251 (2003).
[6] R. Stratton, Phys. Rev. **126**, 2002 (1962).
[7] K. Bløtekjær, IEEE Trans. Electron Devices **17**, 38 (1970).
[8] D. Caughey and R. Thomas, Proc. IEEE **52**, 2192 (1967).
[9] S. Selberherr, W. Hänsch, M. Seavey, and J. Slotboom, Solid-State Electron. **33**, 1425 (1990).
[10] S.-C. Lee and T.-W. Tang, Solid-State Electron. **35**, 561 (1992).
[11] G. Baccarani and M. Wordeman, Solid-State Electron. **28**, 407 (1985).
[12] W. Hänsch, *The Drift Diffusion Equation and its Application in MOSFET Modeling* (Springer, Wien–New York, 1991).
[13] T.-W. Tang, S. Ramaswamy, and J. Nam, IEEE Trans. Electron Devices **40**, 1469 (1993).
[14] R. Thoma *et al.*, IEEE Trans. Electron Devices **38**, 1343 (1991).
[15] C. Jungemann and B. Meinerzhagen, *Hierarchical Device Simulation: The Monte Carlo Perspective* (Springer, Wien–New York, 2003).
[16] P. Markowich, C. Ringhofer, and C. Schmeiser, *Semiconductor Equations* (Springer, Wien–New York, 1990).
[17] M. Rudan and F. Odeh, Compel **5**, 149 (1986).
[18] A. Forghieri *et al.*, IEEE Trans. Comput.-Aided Des. **7**, 231 (1988).
[19] W.-S. Choi *et al.*, IEEE Trans. Comput.-Aided Des. **13**, 899 (1994).

[20]Q. Lin, N. Goldsman, and G.-C. Tai, Solid-State Electron. **37**, 359 (1994).
[21]T.-W. Tang and M.-K. Ieong, IEEE Trans. Comput.-Aided Des. **14**, 1309 (1995).
[22]J.-G. Ahn *et al.*, IEEE Electron Device Lett. **15**, 348 (1994).
[23]T. Grasser, H. Kosina, and S. Selberherr, J. Appl. Phys. **90**, 6165 (2001).
[24]K. Hasnat *et al.*, IEEE Trans. Electron Devices **44**, 129 (1997).
[25]A. Gehring, T. Grasser, H. Kosina, and S. Selberherr, J. Appl. Phys. **92**, 6019 (2002).
[26]C. Jungemann *et al.*, Solid-State Electron. **42**, 647 (1998).
[27]K. Sonoda *et al.*, J. Appl. Phys. **80**, 5444 (1996).
[28]T. Grasser, H. Kosina, M. Gritsch, and S. Selberherr, J. Appl. Phys. **90**, 2389 (2001).
[29]T. Grasser, H. Kosina, C. Heitzinger, and S. Selberherr, Appl. Phys. Lett. **80**, 613 (2002).
[30]A. M. Anile, V. Romano, and G. Russo, SIAM J. Appl. Math. **61**, 74 (2000).
[31]M. Stettler, M. Alam, and M. Lundstrom, IEEE Trans. Electron Devices **40**, 733 (1993).
[32]T. Grasser, H. Kosina, and S. Selberherr, in *Proceedings of the Simulation of Semiconductor Processes and Devices* (IEEE Operations Center, Piscataway, NJ, 2003), pp. 63–66.
[33]A. Anile and O. Muscato, Phys. Rev. B **51**, 16728 (1995).
[34]T. Grasser, H. Kosina, and S. Selberherr, in *Proceedings of the Simulation of Semiconductor Processes and Devices* (Springer, Wien–New York, 2004), pp. 109–112.
[35]D. Ferry, *Semiconductors* (Macmillan, New York, 1991).
[36]C. Jacoboni and P. Lugli, *The Monte Carlo Method for Semiconductor Device Simulation* (Springer, Wien–New York, 1989).
[37]*VMC 1.0 User's Guide*, Institut für Mikroelektronik, Technische Universität Wien, Austria, 2003, http://www.iue.tuwien.ac.at/software/vmc
[38]A. Anile, W. Allegretto, and C. Ringhofer, *Mathematical Problems in Semiconductor Physics* (Springer, Wien–New York, 1988).
[39]C. Ringhofer, C. Schmeiser, and A. Zwirchmayer, SIAM (Soc. Ind. Appl. Math.) J. Numer. Anal. **39**, 1078 (2001).
[40]T. Grasser, H. Kosina, and S. Selberherr, in *Advanced Device Modeling and Simulation*, edited by T. Grasser (World Scientific, Singapore, 2003), pp. 173–201.
[41]D. Woolard *et al.*, Phys. Rev. B **44**, 11119 (1991).
[42]T. Grasser, H. Kosina, C. Heitzinger, and S. Selberherr, J. Appl. Phys. **91**, 3869 (2002).
[43]T. Grasser *et al.*, in *Proceedings of the 31th European Solid-State Device Research Conference*, edited by H. Ryssel, G. Wachutka, and H. Grünbacher (Frontier Group, Nuremberg, Germany, 2001), pp. 215–218.
[44]T. Bordelon, X.-L. Wang, C. Maziar, and A. Tasch, Solid-State Electron. **35**, 131 (1992).
[45]C. Levermore, J. Stat. Phys. **83**, 1021 (1996).
[46]M. Junk and A. Unterreiter, Math. Methods Appl. Sci. **10**, 1001 (2000).
[47]K. Bandyopadhyay, K. Bhattacharyya, and A. Bhattacharya, Pramana **54**, 365 (2000).
[48]A. Tagliani, Appl. Math. Lett. **16**, 519 (2003).
[49]K. Sonoda *et al.*, Jpn. J. Appl. Phys., Part 1 **35**, 818 (1996).
[50]A. Anile and V. Romano, Hydrodynamical Modeling of Charge Carrier Transport in Semiconductors, Summer School on Industrial Mathematics, IST Lisboa, Portugal, http://www.dipmat.unict.it/anile/preprint.html, 1999.
[51]H. Struchtrup, Physica A **275**, 229 (2000).
[52]T.-W. Tang, Q. Cao, and J. Nam, Jpn. J. Appl. Phys., Part 1 **42**, 2137 (2003).
[53]D. Matz, J. Phys. Chem. Solids **28**, 373 (1967).
[54]B. Geurts, M. Nekovee, H. Boots, and M. F. H. Schuurmans, J. Appl. Phys. **59**, 1743 (1991).
[55]H. Grad, Commun. Pure Appl. Math. **2**, 311 (1949).
[56]S. Seeger and K. Hoffmann, Continuum Mech. Thermodyn. **12**, 403 (2000).
[57]W. Dreyer, M. Junk, and M. Kunik, Nonlinearity **14**, 881 (2001).
[58]M. Nekovee, B. Geurts, H. Boots, and M. Schuurmans, Phys. Rev. B **45**, 6643 (1992).
[59]D. Chen *et al.*, IEEE Electron Device Lett. **13**, 26 (1992).
[60]G.-C. Rota, in *Algebraic Combinatorics and Computer Science* (Springer Italia, Milan, 2001), pp. 57–93.
[61]E. Wang, M. Stettler, S. Yu, and C. Maziar, in *Proceedings of the International Workshop on Computational Electronics* (Piscataway, NJ, 1998), pp. 234–237.
[62]T.-W. Tang, in *Semiconductor TCAD Workshop & Exhibition* (Hsinchu, Taiwan, 1999), pp. 1–19.
[63]T.-W. Tang, IEEE Trans. Electron Devices **31**, 1912 (1984).
[64]B. Meinerzhagen and W. Engl, IEEE Trans. Electron Devices **35**, 689 (1988).
[65]R. Kosik, dissertation, Technische Universität Wien, 2004, http://www.iue.tuwien.ac.at.
[66]T. Grasser, H. Kosina, and S. Selberherr, Appl. Phys. Lett. **79**, 1900 (2001).