Characterization of the hot electron distribution function using six moments

T. Grasser,[a] H. Kosina, C. Heitzinger, and S. Selberherr
Institute for Microelectronics, TU Vienna Gusshausstrasse 27–29, A-1040 Vienna, Austria

(Received 4 June 2001; accepted for publication 21 December 2001)

The shape of the hot electron distribution function in semiconductor devices is insufficiently described using only the first four moments. We propose using six moments of the distribution function to obtain a more accurate description of hot carrier phenomena. An analytic expression for the symmetric part of the distribution function as a function of the even moments is given which shows good agreement with Monte Carlo data for both the bulk case and inside n^+-n-n^+ test structures. The influence of the band structure on the parameters of the distribution function is studied and proven to be of importance for an accurate description. © *2002 American Institute of Physics.* [DOI: 10.1063/1.1450257]

I. INTRODUCTION

A very common assumption in device simulation is that the distribution function (DF) can be modeled with some variant of a Maxwellian distribution. This may either be a displaced, heated, or displaced and heated Maxwellian shape.[1] As the drift velocity, which gives the displacement, is normally small,[2] a heated Maxwellian distribution

$$f(\mathcal{E}) \approx \exp\left(-\frac{\mathcal{E}}{k_B T_n}\right) \tag{1}$$

is commonly used to describe the symmetric part of the DF where $\mathcal{E}$ is the energy, T_n is the carrier temperature, and k_B Boltzmann's constant. As has been frequently pointed out, this is at best a very poor approximation in state-of-the-art devices where the gradients of the electric field are large. Even for the bulk case, as is less frequently noted, this assumption is poor.

Two main deviations from the Maxwellian shape have been described by many authors. First, it has been observed that beyond a certain energy, the slope of the DF decreases rapidly. This has been called the thermal tail of the DF because its effective temperature equals the lattice temperature. Abramo and Fiegna[3] discussed this thermal behavior of the high-energy tail and showed that it is not a band structure effect by reproducing it using a single isotropic and parabolic band, including only acoustic and optical phonon scattering. Furthermore, they showed that the effective temperature of the thermal tail is increased when electron–electron scattering (EES) is taken into account. The influence of EES has been investigated in detail by Chang *et al.*[4] who also evaluated the influence of the band structure (parabolic versus fullband). In addition impact ionization has been shown to affect the high-energy tail.[5,6]

Another important deviation from the Maxwellian shape occurs when hot and cold carrier populations mix, as in the drain region of a metal–oxide–semiconductor transistor. These populations coexist for some time and have been approximated by a superposition of a hot and a cold Maxwellian distribution.[7–9] In those regions, the relaxation times are largely determined by the average energy of the hot population. Since the number of hot carriers is normally much smaller than the number of cold carriers, the energy of the hot carriers has only negligible influence on the average energy of the total electron gas and, therefore, models using the total average energy are bound to fail.

Due to these strong deviations from the Maxwellian shape, the shape of the distribution function can not be uniquely characterized by the average carrier energy alone. For the same average energies, the shape of the DF is completely different depending on whether the absolute value of the electric field increases or decreases.[7,10] As the relaxation times which are commonly used in macroscopic models depend on the shape of the distribution function via the scattering operator, problems are to be expected when they are modeled as a function of the average carrier energy only.

To characterize the DF, we include in addition to the carrier temperature T_n the next higher moment of the distribution function, the kurtosis β_n, which represents the normalized moment of fourth order in $\mathbf{k}$.[11]

$$T_n = \frac{2}{3}\frac{\langle \mathcal{E} \rangle}{k_B} \tag{2}$$

$$\beta_n = \frac{3}{5}\frac{\langle \mathcal{E}^2 \rangle}{\langle \mathcal{E} \rangle^2} \tag{3}$$

For a heated Maxwell–Boltzmann distribution and parabolic bands we get $\beta_n = \beta_{\mathrm{MB}} = 1$. Thus, a $\beta_n \neq 1$ quantifies the deviation from the Maxwellian shape in the parabolic case. When nonparabolicity is taken into account, a different value for β_{MB} is obtained, as is shown in the following.

II. QUALITATIVE DESCRIPTION OF THE DISTRIBUTION FUNCTION

Monte Carlo (MC) simulations indicate that the shape of the DF inside an n^+-n-n^+ structure behaves qualitatively

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

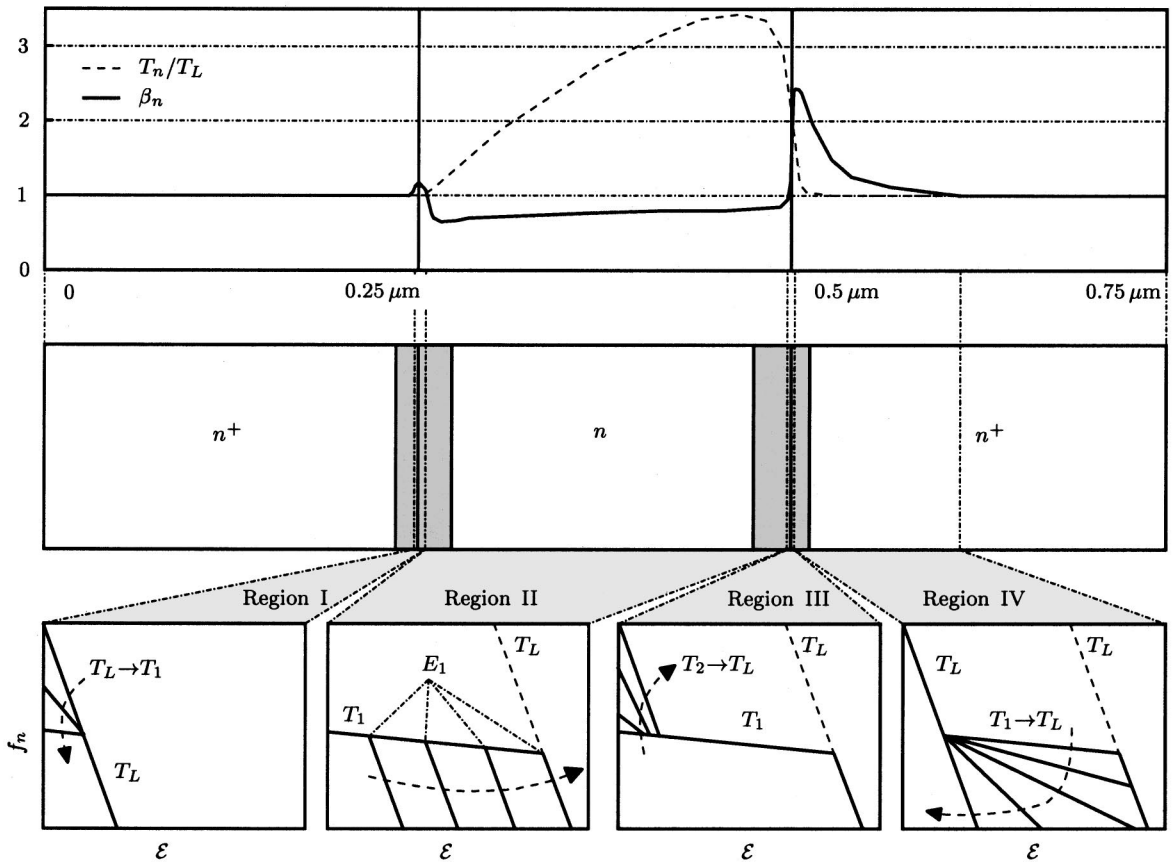


FIG. 1. Schematic evolution of the distribution function inside an n^+-n-n^+ structure.

as shown in Fig. 1. For cold carriers which are injected at the contacts, the Maxwellian shape provides a good description. In Region I, carriers diffuse against the built-in energy barrier. While moving along Region I, the amount of low energetic carriers in the DF decreases due to reflection at the energy barrier. Therefore, unlike the prediction by the Maxwellian approximation, only the low-energy range of the energy distribution is affected whereas nearly no changes in the high-energy tail are observed. It is interesting to note that the slope in the low-energy range is already nearly the same as at the end of the channel. However, the knee-energy E_1 changes, shifting towards higher energies as the carriers travel through the channel in Region II. In Region III, the small number of hot carriers from the channel meet the large pool of cold carriers in the drain which is visible in the DF by a rapid increase of the low-energy part. Region III was found to be very small in our simulations and left the high-energy tail nearly unchanged. As the hot carriers travel through Region IV, the temperature of the high-energy tail relaxes to the equilibrium temperature.

After reviewing previously published models for the DF, we will clarify the importance of the kurtosis for the shape of the DF by deriving an analytical expression for the DF as a function of both the carrier temperature and the kurtosis. Furthermore, we will show that the kurtosis provides the required information to distinguish between Regions I+II and Regions III+IV.

III. PREVIOUS ANALYTICAL MODELS

A lot of effort has been put into the development of analytical expressions for the symmetric part of the distribution function which underlines the importance of the matter. Many different expressions have been published so far, see e.g. Refs. 12–16. All analytical expressions for the DF contain parameters which have to be calculated in one of the following ways.

(1) By considering some simplified form of Boltzmann's equation, analytic expressions for the distribution function have been given, see for instance Refs. 13, 15, and 17. The parameters are calculated in an *a priori* manner, for instance as a function of the electric field, the band structure, and some dominant scattering mechanisms. Although these approaches are highly interesting from a theoretical point of view and also provide insight into transport phenomena, they give poor results in practical applications, because the assumptions made during their derivation are normally heavily violated in real devices. In particular, the expression for the analytic DFs obtained this way does not correctly reproduce the moments as obtained by MC simulations or macroscopic transport models. This a source of inconsistencies.
(2) A macroscopic transport model can be derived by putting an Ansatz for a DF into Boltzmann's transport equation. However, errors made in the Ansatz of the DF di-

rectly enter the transport model and models based on this method have shown to perform poorly.[18] Note that these models do *not* consider the moments of the DF but the parameters of the DF, a fact that is frequently confused.

(3) By generalizing Blǿtekjær's method of moments a transport model can be derived without assuming any particular DF.[11,19] This procedure results in balance equations for the even moments in **k**, typically $\langle \mathcal{E}^l \rangle$, and in flux equations for the odd order moments, typically $\langle \mathbf{u}\mathcal{E}^l \rangle$ or $\langle \mathbf{k}\mathcal{E}^l \rangle$. The method of moments delivers an infinite set of equations which has to be truncated at a certain order to give a tractable equation set. In particular, the highest order equation contains the moment of next higher order which has to be suitably approximated using available moments. This *can* be done by assuming a particular DF, for instance a Maxwellian distribution. However, different closures have been suggested.[11,20] Note that the four-moments based model which has been closed with a Maxwellian DF is structurally equivalent to the transport model obtained when a Maxwellian DF is put into Boltzmann's transport equation and parabolic bands are assumed. This is not the case with a general band structure and other Ansatz functions for the DF. In particular, the moments based model is independent of the DF shape and the temperature appearing in the transport equations is defined via the second-order moment. In contrast, the temperature appearing in models based on an Ansatz for the DF is the parameter of the DF itself rather than the mean energy, which becomes obvious when for instance a Fermi–Dirac distribution is assumed.[21]

In the following, some typical Ansatz functions for the DF and their properties are reviewed. One common feature is that all Ansatz functions must be able to reproduce a cold Maxwellian DF which is the solution of Boltzmann's transport equation in equilibrium (assuming nondegenerate semiconductors). Furthermore, we will only consider the even moments of the DF $\langle \mathcal{E}^l \rangle$ because the contribution of odd order moments is considered to be of minor importance for modeling the symmetric part of the DF. Odd order moments like the average carrier velocity **v** have been used to displace the symmetric part of the DF.[22] However, the odd order moments are vector quantities in the multidimensional case and they are not state variables in macroscopic transport models which complicates the handling of the model considerably. In the following, we will show that excellent agreement can be obtained without considering the odd order moments.

A. Expansion around a heated Maxwellian DF

Expansions of the DF around a heated Maxwellian DF are commonly used in theoretical physics:

$$f(\mathcal{E}) \approx a_0 \exp\left(-\frac{\mathcal{E}}{a_1}\right)(1 + a_2\mathcal{E} + a_3\mathcal{E}^2 + \ldots + a_{N-1}\mathcal{E}^{N-2}). \tag{4}$$

The parameters a_l are frequently related to orthogonal Legendre or Hermite polynomials[23,24] or to a Grad-type expansion.[25] They can be determined in one of the aforementioned ways where N is the number of parameters. Expression (4) can be easily integrated analytically and has nice mathematical properties due to the orthogonality of the polynomials. However, convergence of Eq. (4) is known to be poor and a high order is required to reproduce features like the thermal tail. This could be the reason for the poor properties of the associated transport model studied in Refs. 18 and 26.

Other problems associated with Eq. (4) are that $f(\mathcal{E})$ may become negative and that for higher orders oscillations may occur. This can be shown for $N=3$ and parabolic bands where a_2 can be analytically expressed as

$$a_2 = -\frac{3}{k_B T_n \beta_n}\left(1 \pm \sqrt{1 - \frac{2}{3}\beta_n}\right). \tag{5}$$

The term in parenthesis is always positive and real valued for β_n in the range [0,1.5], which covers the whole of Region II and the beginning of Region III. Thus, a_2 will always be negative which implies that $f(\mathcal{E}) < 0$ for $\mathcal{E} > 1/a_2$ which is clearly unphysical.

B. Polynomial in exponential function

Even though expressions of type (4) are still frequently used, better approximations have been considered. These are obtained by putting the polynomial into the exponent of the exponential function

$$f(\mathcal{E}) \approx a_0 \exp(a_1\mathcal{E} + a_2\mathcal{E}^2 + \ldots a_{N-1}\mathcal{E}^{N-1}). \tag{6}$$

A theoretical derivation which justifies this Ansatz can be found in Ref. 12. Theoretical considerations based on the maximum entropy principle deliver the same result.[25,27] For $N>2$, a deviation from the Maxwellian DF is obtained. Other authors have used this Ansatz with $N=3$ but determined the parameters a_l via an energy-transport model,[14] which can only supply two parameters and has to be supplemented with heuristic considerations, or via a six-moments transport model[27] which provides all required parameters and thus gives the best results. Unfortunately, Eq. (6) is difficult to handle analytically, except for $N=2$, which corresponds to a Maxwellian DF.

C. Generalization of Cassi's expression

Cassi and Riccò[13] derived an analytical expression for the distribution function assuming that the diffusion term in Boltzmann's equation is negligible compared to the drift term and by fitting the resulting model to MC simulations. Generalizing their result gives the Ansatz

$$f(\mathcal{E}) = a_0 \exp(a_1\mathcal{E}^{a_2} + a_3\mathcal{E}^{a_4} + \ldots + a_{N-2}\mathcal{E}^{a_{N-1}}). \tag{7}$$

Cassi and Riccò used $a_1 = -\chi/|\mathbf{E}|^{3/2}$ and $a_2 = 3$ with χ being a fit factor. A comparison with MC data shows that this is a reasonable expression for the bulk case under high electric fields which corresponds to their assumptions. However, in the devices of interest, the electric field is never homogeneous and strong gradients exist. Therefore, to fit their experimental data, other authors have used this expression in modified forms[28–30] to calculate gate and substrate

currents of submicron devices. For example, Hasnat *et al.*[29] replaced the electric field in a_1 by a function of the carrier temperature to better account for the nonlocal behavior and used $a_1=\eta^{-1}(k_B T_n)^{-\zeta}$ and $a_2=1.3$ with $\eta=0.265$ and $\zeta=0.75$.

D. Comparison of analytical expressions

To test how well the given approximation can reproduce realistic DFs, we calculate the parameters a_l in such a way that the analytic DFs exactly reproduce a given set of moments. This is considered the optimal parameter set, favorable to any *ad hoc* calculated parameters. The number of given moments equals the number of unknowns N. As a test device we used an n^+-n-n^+ structure with a channel length of 200 μm.

A comparison of expression (4) with MC data is shown in Fig. 2(a). We favor to plot $f(\mathcal{E})$ directly instead of the frequently used product $f(\mathcal{E})g(\mathcal{E})$ which only obscures details of the DF as the density of states $g(\mathcal{E})$ is fixed. Note that the DF becomes negative at around $\mathcal{E}=1.5$ eV and the oscillations for $N=7$. Altogether, the agreement is poor. In Fig. 2(b), the expressions (6) and (7) are compared to MC data. Even for $N=3$, both expressions give accurate results where Eq. (7) performs even better in the high-energy tail. In addition, expression (7) has the advantage that it can be analytically integrated for $N=3$ which is one order higher than Eq. (6).

IV. MODELS FOR REGION III AND IV

Expressions (6) and (7) deliver accurate results for Region II. When the hot carriers coming from Region II meet the cold carriers in Region III+IV, two populations coexist and the contribution of the cold carriers is clearly visible in the DF.

A. Superposition of two Maxwellian DFs

A straightforward approach would be to assume a superposition of two Maxwellian distribution functions. In Refs. 7 and 8, the parameters of the two Maxwellian DFs have been determined using two coupled transport models for the cold and the hot populations. Another attempt[9] was to calculate the parameters from a six-moments transport model. Even though these approaches give an approximate description, they overestimate the number of carriers in the high-energy tail which does not show the thermal behavior.

B. Sonoda's model

Sonoda *et al.*[27] used the following expression to model the contribution of the cold carriers

$$f(\mathcal{E})=\underbrace{a_0\exp(a_1\mathcal{E}+a_2\mathcal{E}^2)}_{f_1(\mathcal{E})}+\underbrace{a_3\exp\left(-\frac{\mathcal{E}}{k_B T_2}\right)}_{f_2(\mathcal{E})}. \tag{8}$$

The parameters a_l were calculated via given moments of the DF. Equation (8) contains five unknowns and to make it suitable for a six-moments transport model, two assumptions

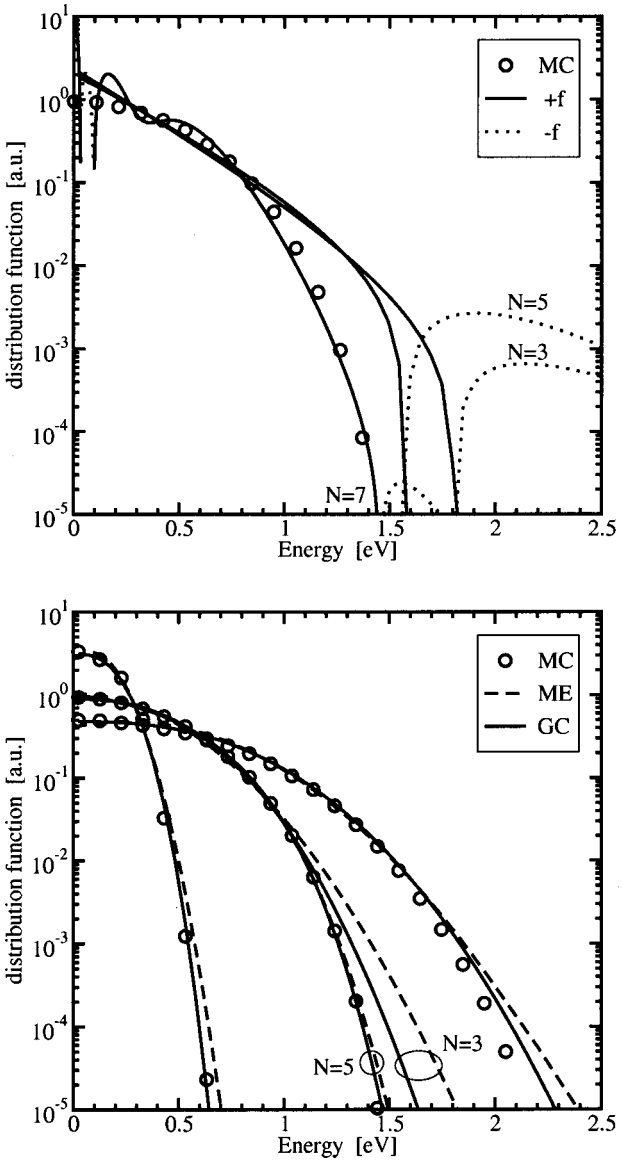


FIG. 2. Comparison of different analytic distribution function models for various orders N inside Region II (channel): (a) The polynomial expansion around a heated Maxwellian. Note the sign change in $f(\mathcal{E})$. (b) Maximum entropy (ME) and generalized Cassi's expression (GC) for order $N=3$. Also shown is $N=5$ in the middle of Region II.

were made: First, the temperature of the Maxwellian DF T_2 was assumed to equal the lattice temperature and second, it was assumed that the kurtosis of the hot DF $f_1(\mathcal{E})$ does not change within Region III+IV and that thus a constant value of $\beta_1=\beta_h$ can be used. The kurtosis of the hot DF is defined as

$$\beta_1=\frac{3}{5}\frac{\int f_1 g d\mathcal{E}\int\mathcal{E}^2 f_1 g d\mathcal{E}}{(\int\mathcal{E} f_1 g d\mathcal{E})^2}. \tag{9}$$

As a value for β_h the saturated high-field value for bulk was used. In addition, β_h was used as a threshold value for the kurtosis to distinguish between Region II and Region III+IV. In particular, when β_n was smaller than β_h the prefactor of the Maxwellian DF a_3 was assumed to be zero (Region II), and only for $\beta_n>\beta_h$ the more complex expression (8) was used. Although this approach looks reasonable,

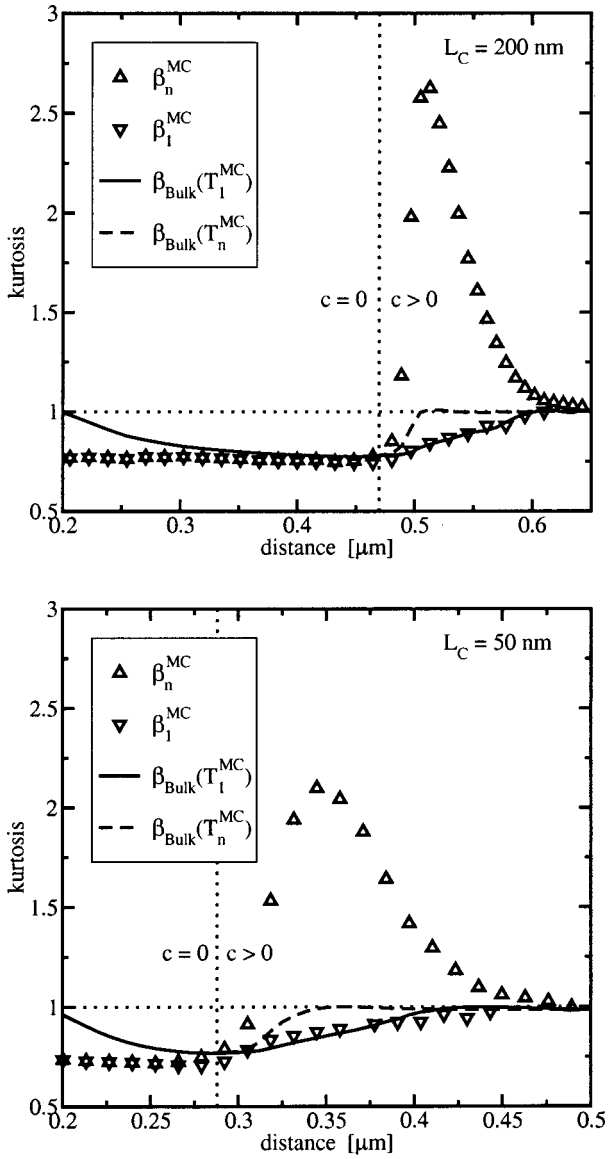


FIG. 3. Comparison of β_1^{MC} with two analytical models. When the temperature of the hot DF T_1 is used in the bulk characteristic, accurate results are obtained for n^+-n-n^+ structures with $L_c=200$ nm and $L_c=50$ nm. Note that $\beta_{Bulk}(T_n)$ does not properly describe the behavior of β_1. Also shown is the kurtosis of the total distribution function β_n.

there are two problems with assuming a constant value for β_h: First, β_1 *does* change inside Region III and approaches unity in Region IV. With β_1, the shape of the hot DF changes and approaches a Maxwellian shape somewhere in Region IV. Second, for lower applied fields, the kurtosis might not reach β_h at all and the model erroneously predicts a spurious cold population throughout the whole device. For larger applied biases, a spurious cold population is always predicted in the larger part of Region II where $\beta_n>\beta_h$.

C. Generalization of Cassi's expression

When the parameters of Eq. (7) are calculated from the moments of the DF, a_2 becomes smaller than unity in Region III+IV which is in contrast to Region II where $a_2>1$ holds.[31] This approach gives reasonable approximations in Region III+IV. There are two problems, however: First, the overestimation of the high-energy tail with $a_2<1$ is even stronger than the overestimation observed within the models using two Maxwellian DFs. Second, during the transition between Region II and Region III, a value of 1 is obtained for the exponent a_2. Unfortunately, this results in a Maxwellian DF which is not confirmed by MC simulations. This error is amplified when, for instance, impact ionization rates are calculated where only the high-energy tail of the DF is required.[31]

V. IMPROVED MODEL

Weighing the pros and cons of the previously published models, we can now construct an improved model with the following properties.

(1) First, at least six moments ($N=3$) are needed to improve the Maxwellian approximation. Even though expressions have been used which go beyond the Maxwellian shape approximation using lower order transport models, there would be no way to predict Region III+IV without resorting to heuristic criteria. The kurtosis provides both a change in the shape and a differentiation between Region II and Region III+IV. Therefore, if details of the DF are required, at least a six-moments transport model is required which is one order higher than conventional hydrodynamic and energy-transport models.

(2) Our model should give closed form solutions so that, for instance, closed form expressions of impact ionization rates can be given.[31] Therefore, we use Eq. (7) with $N=3$ as a basis.

(3) To avoid the problems which arise from using Eq. (7) in Region III+IV, we allow for a cold Maxwellian DF in the way Sonoda *et al.* did.[27] However, the kurtosis of the hot DF, β_1, must be modeled dynamically, that is, depending on bias and position to capture the high-energy tail correctly.

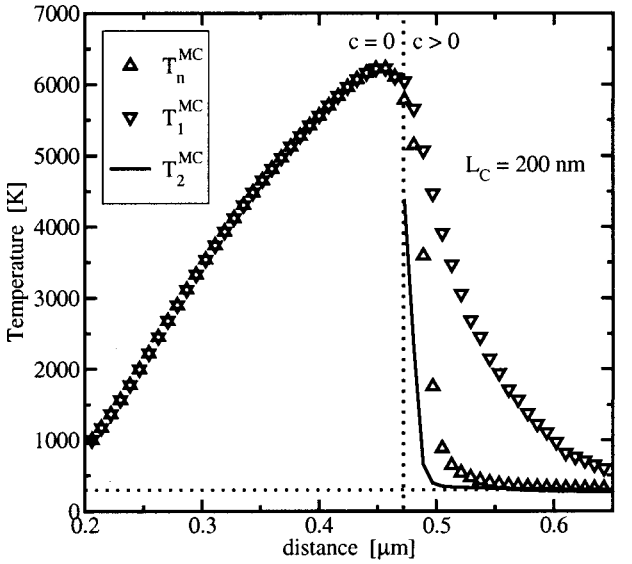
Therefore, we use the Ansatz

$$f(\mathcal{E})=A\left\{\underbrace{\exp\left[-\left(\frac{\mathcal{E}}{k_BT_{\text{ref}}}\right)^b\right]}_{f_1(\mathcal{E})}+c\underbrace{\exp\left[-\frac{\mathcal{E}}{k_BT_2}\right]}_{f_2(\mathcal{E})}\right\} \tag{10}$$

for the symmetric part of the DF. We now have to determine the five parameters A, T_{ref}, b, c, and T_2 which describe the DF, that is, we need two additional equations to the three parameters n, T_n, and β_n provided by the six-moments model.

To get an idea about the behavior of the DF in Region III+IV, we look at the moments β_1 and T_1 (Figs. 3 and 4). T_1 is the temperature of the hot DF $f_1(\mathcal{E})$ which is defined analogously to β_1. Note that T_1 is different from T_{ref}. For the extraction of these parameters, see Appendix A. From Fig. 4, we see that the temperature of the cold Maxwellian DF in Region III rapidly relaxes to the lattice temperature T_L and will be modeled as $T_2=T_L$ in this work.

In contrast to T_2, modeling of the kurtosis β_1 is crucial for an accurate description of the high-energy tail. One approach would be to describe β_1 via the bulk relation



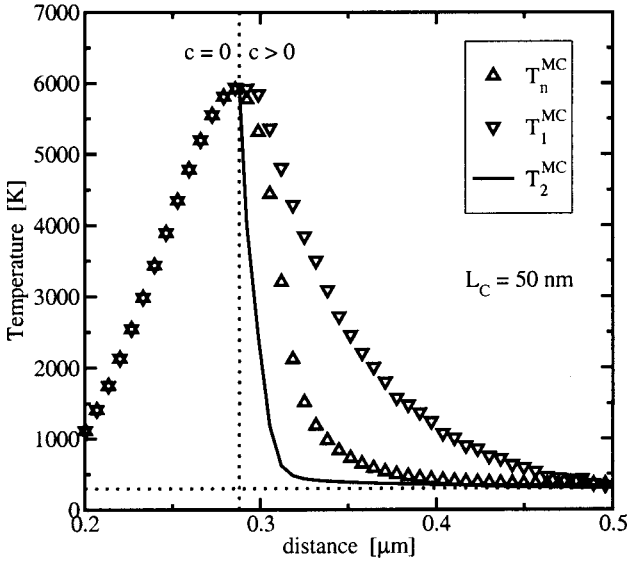


FIG. 4. The three different temperatures T_n, T_1, and T_2 extracted from MC simulations for two n^+-n-n^+ structures. Note that in Region II, $T_n=T_1$ holds.

$\beta_{\text{Bulk}}(T_n)$ which can be derived from the homogeneous six-moments model[11] as

$$\beta_{\text{Bulk}}(T_n)=\frac{T_L^2}{T_n^2}+2\frac{\tau_\beta}{\tau_\mathcal{E}}\frac{\mu_S}{\mu_n}\left(1-\frac{T_L}{T_n}\right), \tag{11}$$

where $\tau_\mathcal{E}$, τ_β, μ_n, and μ_S are the energy relaxation time, the kurtosis relaxation time, the electron mobility, and the energy flux mobility, respectively. Unfortunately, still no sufficiently accurate models for these parameters exist which work in the present context. Therefore, we use the fit to MC data

$$x(T_n)=2\frac{\tau_\beta}{\tau_\mathcal{E}}\frac{\mu_S}{\mu_n}=x_0+x_1\left[1-\exp\left(-x_2\frac{T_L}{T_n}\right)\right], \tag{12}$$

with $x_0=0.69$, $x_1=1.34$, and $x_2=1.89$. This expression is

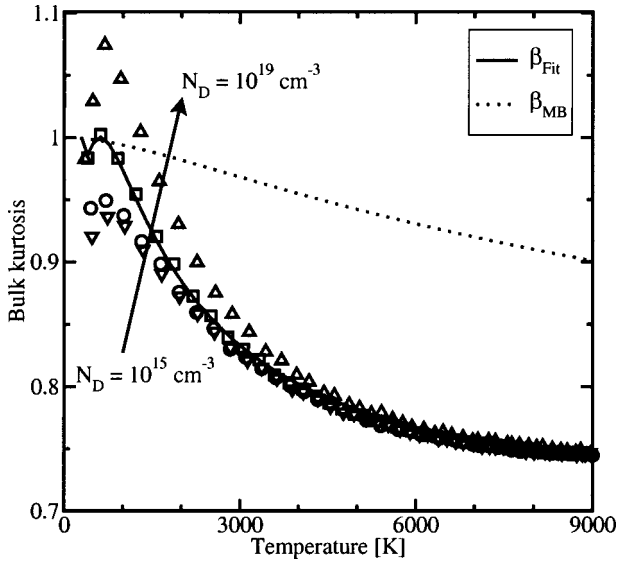


FIG. 5. The kurtosis β_n plotted over the temperature T_n for bulk silicon with the doping concentration as a parameter ($N_D=10^{15},10^{17},10^{18}$, and 10^{19} cm^{-3}). Also shown is the fit expressions (12) and the kurtosis β_{MB} which would be obtained from a fictitious Maxwellian DF.

accurate for doping concentrations around 10^{18} cm^{-3} but the doping dependence of β_{Bulk} is only relevant at lower doping concentrations (Fig. 5).

A comparison of the model $\beta_1=\beta_{\text{Bulk}}(T_1)$ with MC data is shown in Fig. 3. Note that $\beta_{\text{Bulk}}(T_n)$ approaches unity too quickly as also shown in Fig. 3 which underlines the idea of modeling the hot and cold electrons as separate populations. Furthermore, we see that $\beta_{\text{Bulk}}(T_n)$ provides a separation between Region II and Region III+IV. This is reasonable because inside Region II the high-energy tail at any point can be expected to be less populated than in a comparable bulk case, resulting in a kurtosis smaller than β_{Bulk}. In Region III+IV, on the other hand, the cold population dominates the kurtosis, resulting in values larger than β_{Bulk}.

Inside Region II, there is no cold population and thus $c=0$. This implies that T_1 equals the total temperature T_n and β_1 the total kurtosis β_n. At the beginning of Region III, a cold population appears which causes $c\neq 0$. Note that at the transition point, $T_1=T_n$ holds and thus $\beta_{\text{Bulk}}(T_1)=\beta_{\text{Bulk}}(T_n)$, which guarantees a continuous transition.

Thus, for each grid point, the following nonlinear equation system is solved using Newton's method

$$\begin{pmatrix} T_n(T_{\text{ref}},b,c) \\ \beta_n(T_{\text{ref}},b,c) \\ \beta_1(T_{\text{ref}},b,c) \end{pmatrix}=\begin{pmatrix} T_n^{\text{MC}} \\ \beta_n^{\text{MC}} \\ \beta_{\text{Bulk}}[T_1(T_{\text{ref}},b,c)] \end{pmatrix}. \tag{13}$$

Note that T_n, β_n, and β_1 are analytic expressions derived from the moments of Eq. (10) and T_n^{MC} and β_n^{MC} are taken from MC simulations. As just stated, in Region II, $c=0$ is assumed and the last row of Eq. (13) is dropped.

A. Expressions for the density of satates

For the calculation of the moments

$$m_l=C\int \mathcal{E}^l g(\mathcal{E})f(\mathcal{E})d\mathcal{E} \tag{14}$$

TABLE I. Coefficients for the Taylor expansion of Kane's DOS.[a]

i	0	1	2	3	4	5
d_i	1	5/2	7/8	−3/16	11/128	−13/256

[a]See Ref. 17.

of the analytic distribution function (10), an expression for the density of states (DOS) $g(\mathcal{E})$ is required. Besides the simple parabolic band approximation, Kane's dispersion relation,[32]

$$\mathcal{E}(1+\alpha\mathcal{E})=\frac{\hbar^2 k^2}{2\,m_d}, \tag{15}$$

is a frequently used expression to incorporate nonparabolicity effects to first order, with α being the nonparabolicity factor. This factor is generally considered a fitting parameter, with $\alpha\approx 0.5\ \mathrm{eV}^{-1}$ in Si. The dispersion relation (15) is generally accepted as valid for energies up to 0.5 eV.

The DOS evaluates for Kane's expression to

$$g(\mathcal{E})=g_0\sqrt{\mathcal{E}}\sqrt{1+\alpha\mathcal{E}}(1+2\alpha\mathcal{E}) \tag{16}$$

Unfortunately, using Eq. (16), the moments can not be evaluated analytically for arbitrary T_{ref}, b, and c. A straightforward approach is to consider the Taylor expansion of the nonparabolic correction term in Eq. (16) as

$$g(\mathcal{E})=g_0\sqrt{\mathcal{E}}\sum_{i=0}^{G} d_i(\alpha\mathcal{E})^i \tag{17}$$

with the coefficients given in Table I. This expansion converges rather poorly and is difficult to handle in a numerical implementation for higher order truncations. Accurate results were obtained for $G=3$.

To obtain a more tractable expression, Cassi and Riccò[13] approximated the dispersion relation as

$$x\mathcal{E}^y=\frac{\hbar^2 k^2}{2\,m_d}, \tag{18}$$

fitting the parameters x and y for different energy ranges. From Eq. (18), the DOS follows as[13]

$$g(\mathcal{E})=\frac{8\pi}{h^3}(2\,m_d\,x)^{3/2}y\,\mathcal{E}^{(3/2)y-1}=g_0\,\mathcal{E}^{\lambda}. \tag{19}$$

Note that for $y=1$, the parabolic DOS is obtained. As pointed out in Ref. 21, this expression must be used with care. In particular, physically meaningful results could only be obtained by fitting Eq. (18) to the energy range [0,0.2 eV]. This can be explained by looking more closely at the DOS. A comparison of different fits to the Kane expression is shown in Fig. 6 together with the numerical data used by Fischetti and Laux.[33] The fitted values were taken from Ref. 21 and are $x=1.365$ and $y=1.52$ when fitted to the low-energy range [0,0.2 eV] and $x=1.185$ and $y=1.052$ when fitted to the high-energy range [1.5,3.0 eV], where x has the unusual dimension of eV^{1-y}. For $\lambda=1$, the shape of $g(\mathcal{E})$ changes from convex to concave and thus either the low- or the high-energy range can be fitted properly but not both simultaneously. As $g(\mathcal{E})$ is normally needed in a context similar to

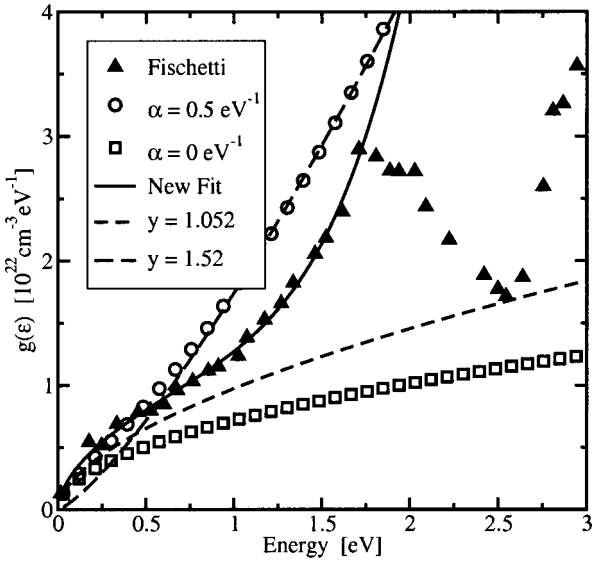


FIG. 6. Comparison of different expressions for the DOS.

Eq. (14), a value smaller than 1 is required for λ to accurately fit the low-energy region, because this is where $f(\mathcal{E})$ has its maximum. However, the resulting DOS shows a ''parabolic-like'' behavior.

To include ''nonparabolic-like'' behavior without losing accuracy in the low-energy region, we propose using

$$g(\mathcal{E})=g_0\sqrt{\mathcal{E}}(1+(\eta\mathcal{E})^{\zeta}). \tag{20}$$

This expression has the advantage of fitting Kane's expression (16) and the numerical data (see Fig. 6), the shape of which can not be reproduced by adjusting α in Eq. (16). The fitted parameters for these two cases are shown in Table II.

Using Eq. (20), we can calculate the moments analytically as

$$\begin{aligned} m_l=m_{1,l}+m_{2,l}=A\,g_0\Bigg\{&\frac{(k_BT_{ref})^{l+3/2}}{b}\Bigg[\Gamma\left(\frac{2l+3}{2b}\right)\\ &+(\eta k_BT_{\mathrm{ref}})^{\zeta}\Gamma\left(\frac{2l+2\zeta+3}{2b}\right)\Bigg]\\ &+c(k_BT_2)^{l+3/2}\Bigg[\Gamma\left(\frac{2l+3}{2}\right)\\ &+(\eta k_BT_2)^{\zeta}\Gamma\left(\frac{2l+2\zeta+3}{2}\right)\Bigg]\Bigg\}, \end{aligned} \tag{21}$$

TABLE II. Coefficients for Ref. 20.

Fit to	η (eV^{-1})	ζ	Energy range (eV)
Kane	1.371 53	1.044 59	[0,0.5]
Kane	1.401 32	1.081 28	[0,1.0]
Kane	1.404	1.112 07	[0,1.5]
Kane	1.394 77	1.138 35	[0,2.0]
Fischetti	0.884 09	0.9096	[0,1.5]
Fischetti	0.884 09	1.3742	[0,1.71]

with $m_{1,l}$ and $m_{2,l}$ being the moments of the hot and cold DF, respectively. The normalized moments are calculated as

$$k_B T_n = \frac{2}{3}\frac{m_0}{m_1}, \quad \beta_n = \frac{3}{5}\frac{m_0 m_2}{m_1^2},$$

and

$$\beta_1 = \frac{3}{5}\frac{m_{1,0} m_{1,2}}{m_{1,1}^2}. \tag{22}$$

VI. COMPARISON WITH MONTE CARLO DATA

In the following, we give a comparison of the analytical model with results obtained by rigorous MC simulations. For the MC model, we employed optical and acoustic phonon scattering in addition to impurity scattering. Furthermore, nonparabolicity was considered using Kane's dispersion relation with $\alpha = 0.5\ \mathrm{eV}^{-1}$ and Keldysh's impact ionization model.[34]

A. Bulk

A comparison of Cassi's model and the improved model with MC data for bulk is shown in Fig. 7 for $N_D = 10^{15}\ \mathrm{cm}^{-3}$. Cassi's model, where we used $\chi = 1.1 \times 10^8$ for the fitting parameter, does not work properly for low electric fields ($E < 200$ kV/cm) where it gives poor agreement in the low-energy region. This is due to the fixed curvature of the distribution function which does not allow for a better resolution. Excellent results have been obtained with the improved model as shown in Fig. 7, where the analytical expression accurately reproduces the whole energy range, although the high-energy tail is slightly overestimated.

To judge the accuracy of the models qualitatively, we can calculate the moments of the analytic DFs. We require that these moments reproduce the MC moments and we quantify the deviation as

$$\mathfrak{E} = \sqrt{\left(\frac{T_n^A - T_n^{\mathrm{MC}}}{T_n^{\mathrm{MC}}}\right)^2 + \left(\frac{\beta_n^A - \beta_n^{\mathrm{MC}}}{\beta_n^{\mathrm{MC}}}\right)^2}, \tag{23}$$

where the superscript A indicates the moments of the analytic DF. The Cassi model is quite accurate for electric fields larger than 200 kV/cm with an error as small as 3.5%. At lower fields, however, the error increases to over 70% (at 50 kV/cm), because the curvature has been fixed to the high-field case. Making the curvature of the DF field dependent might improve the accuracy for the bulk case. However, Cassi's model is still not suitable for the inhomogeneous case, because it only uses the electric field as a parameter.

The results obtained with the parameters given by Hasnat *et al.*,[29] who replaced the electric field dependence with a carrier temperature dependence, are even worse with an error larger than 70% over nearly the whole electric field range. This suggests that the parameters for their distribution function model were fitted to resemble some average characteristic for the whole device.

B. Inhomogeneous case

To evaluate the accuracy of the improved model for the inhomogeneous case, two $n^+ - n - n^+$ structures have been simulated. The basic structure was taken from Ref. 27 where the channel length L_c has been adjusted to 200 and 50 nm. To obtain comparable results, the devices were biased to give a maximum electric field of 300 kV/cm in both cases.

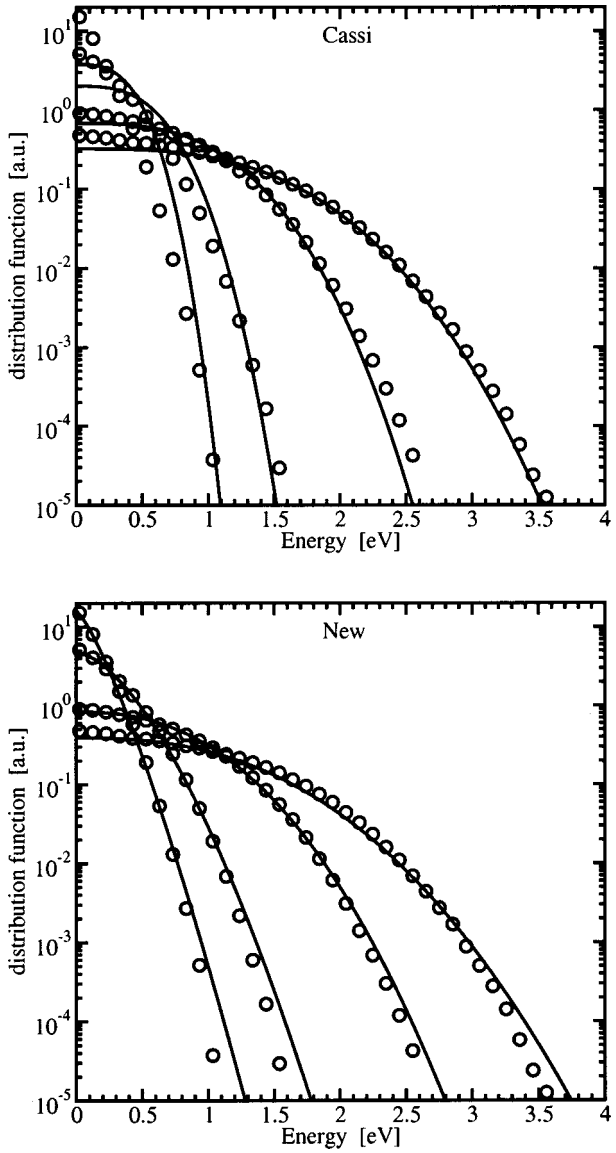
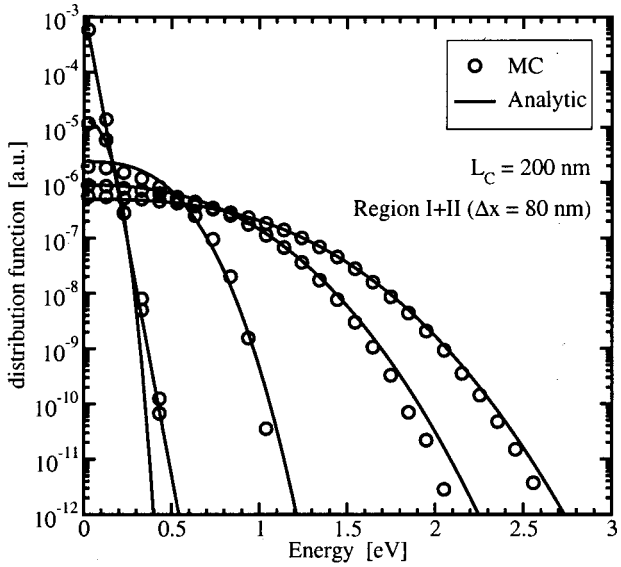
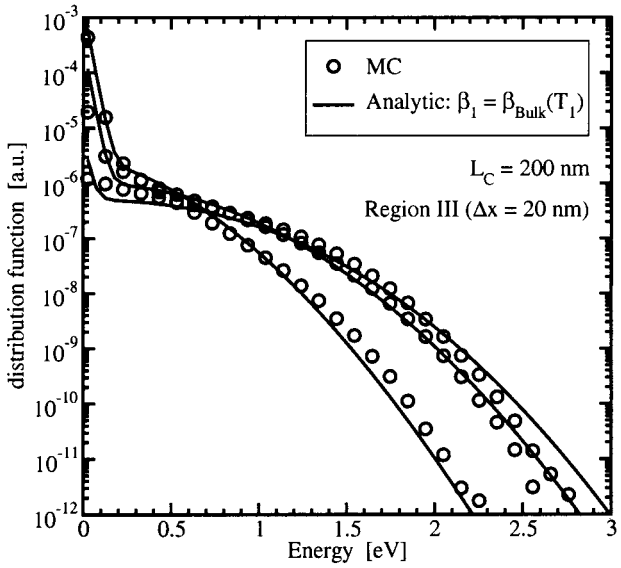


FIG. 7. Comparison of bulk distribution functions obtained by Cassi's model and the improved model with MC data for electric field values of 50, 100, 300, and 600 kV/cm.

As expected, the distribution functions given by the Cassi formula show no visible correlation with the MC data and are thus not shown. Also not shown are the results obtained with Hasnat's correction which gave a minimum error of 65% compared to the 109% of Cassi's formula. Interestingly, assuming a Maxwellian distribution function only gives a maximum error of 57% where the introduced error results from the assumption $\beta_n = 1$. This implies that probably neither Cassi's expression nor Hasnat's corrections give an improvement over the simple Maxwellian-shape assumption which makes their application to submicron devices questionable.

In contrast, the improvecd model gives accurate results for all four regions of the devices. The results for both n^+





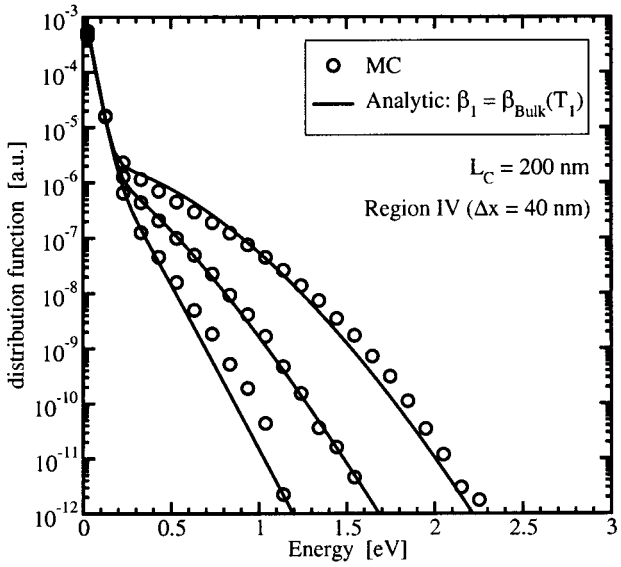
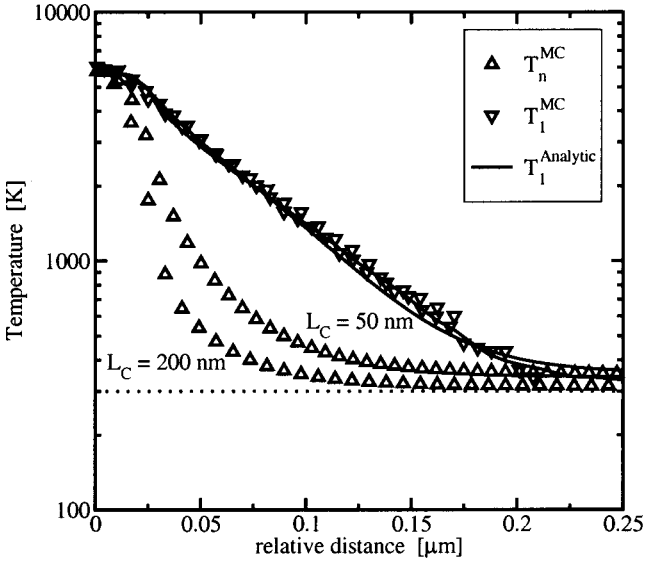


FIG. 8. The analytic distribution function inside the four regions of a n^+-n-n^+ structure with $L_c=200$ nm. The spacing between the DFs is given as Δx. Note the different lengths of each region.



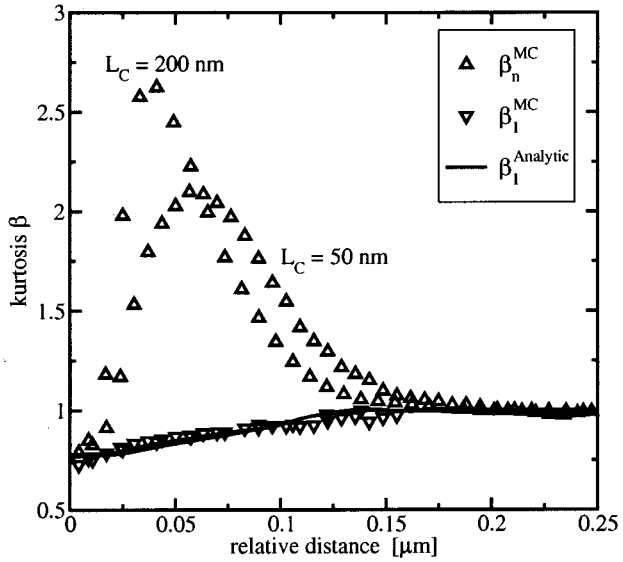


FIG. 9. Comparison of the moments of the hot analytic DF with MC data. The temperature T_1 is obtained implicitly whereas β_1 is modeled via the bulk characteristic. Note that T_1 appears to be independent of the channel length and the total temperature T_n. Furthermore, T_1 relaxes exponentially towards the lattice temperature. The origin of the x axis has been moved to the transition point between Region II and III.

$-n-n^+$ structures are similar and the $L_c=200$ nm case is shown Fig. 8. The accuracy of the improved model is confirmed when we look at the normalized moments of the hot DF only. This is shown in Fig. 9 where the moments β_1 and T_1 are compared to the values extracted from the MC simulation. Note that the temperature T_1 is obtained implicitly whereas β_1 is modeled via the bulk characteristic. The accuracy of both parameters is astonishing. It is interesting to note that neither β_1 nor T_1 seem to depend on the channel length and the value of the total temperature T_n and the total kurtosis β_n which behave differently in both devices. Furthermore, T_1 relaxes exponentially towards the lattice temperature.

The importance of a proper model for β_1 is demonstrated in Fig. 10 where a constant value of $\beta_1=\beta_h$ is as-

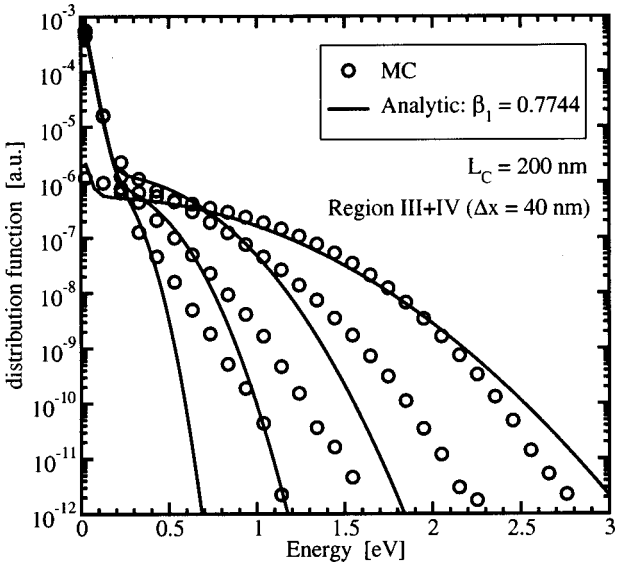


FIG. 10. The analytic distribution function inside Region III+IV when a constant value is used for β_1 as in Ref. 27.

sumed as in Ref. 27. Because of the small value of β_h, the curvature at the end of Region IV is too high, resulting in a suppression of the high-energy tail. For larger electric fields, the influence of the band structure model used in the moment calculation becomes important. This is shown in Fig. 11 where a parabolic DOS has been assumed during the evaluation of Eq. (14). In particular, a maximum error of 50% was observed at the end of Region II.

VII. CONCLUSION

We have developed an analytical description for the distribution function which goes beyond the assumption of a Maxwellian shape. Of fundamental importance to this model is the information provided by the kurtosis of the distribution function. We have shown that the kurtosis provides the information to differentiate between the channel region (Region II) and the transition region from channel to drain (Region III+IV). Furthermore, it can be used to describe the changing shape of the distribution function throughout the whole device. To obtain the kurtosis, a six-moments transport model[11] can be used which contains one additional balance equation as compared to the conventional energy-transport models.

Another important factor is the inclusion of band structure effects. We propose a model for the DOS which can be fit to both Kane's expression and to data obtained from pseudopotential calculations. Although these relatively simple analytical expressions can not capture all of the details of a realistic band structure, they provide a considerable improvement over the parabolic band assumption and are proven to be suitable for efficient simulation of submicron devices. With the improved models of the distribution function and the DOS, models based on microscopic scattering descriptions can be developed for the inclusion in macroscopic transport models. Candidates are impact ionization, as reported in Ref. 31, and gate current modeling.

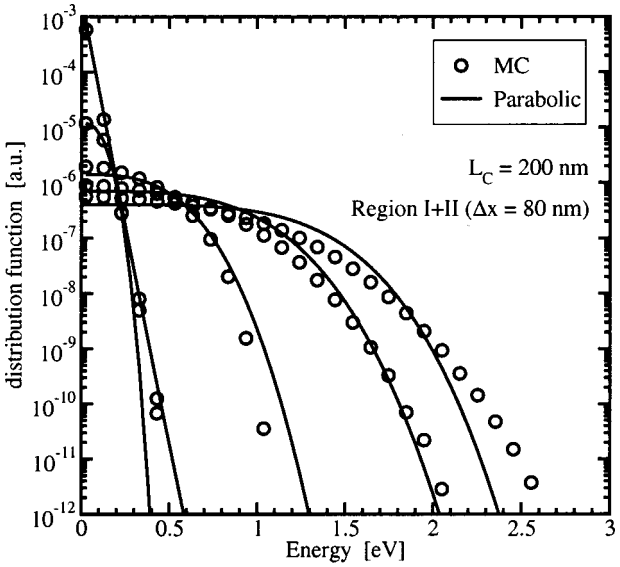


FIG. 11. The analytic distribution function inside Region II using the parabolic band approximation.

ACKNOWLEDGMENTS

This work has been partly supported by the ''Christian Doppler Forschungsgesellschaft,'' Vienna, Austria. Fruitful discussions with Dr. K. Sonoda, Mitsubishi Corp., are gratefully acknowledged.

APPENDIX A: EXTRACTION OF THE SUBMOMENTS

For the extraction of β_1 and T_1, which are the moments of the hot subpopulation, we used the following algorithm: First, MC simulations were performed which recorded the DF in the interval [0,3 eV] using 400 points. In a postprocessing step, a change in the curvature of the DF was detected, and the low-energy region of the DF, which contained the cold Maxwellian DF, was cut off. The low-energy region of the hot DF was then reconstructed by least-square fitting of Eq. (6) with $N=5$ to the remaining high-energy region. From the reconstructed DF, the second and fourth moments where taken to calculate β_1 and T_1. T_2 was determined in a similar fashion from the low-energy region.

[1] M. Lundstrom, *Fundamentals of Carrier Transport*, Modular Series on Solid State Devices, Vol. 10 (Addison-Wesley, Reading, MA 1990).
[2] R. Thoma, A. Emunds, B. Meinerzhagen, H. Peifer, and W. L. Engl, in *Intl. Electron Devices Meeting*, Washington, D.C. (IEEE Press, Piscataway, NJ, 1989), pp. 139–142.
[3] A. Abramo and C. Fiegna, J. Appl. Phys. **80**, 889 (1996).
[4] M. Chang, D. Dyke, C. Leung, and P. Childs, J. Appl. Phys. **82**, 2974 (1997).
[5] J. Bude, *Proceedings of the IEEE Symposium on VLSI Technology Digest of Technical Papers* Kyoto (Business Centers for Academic Societies, Japan, 1995), p. 101.
[6] C. Jungemann, S. Yamaguchi, and H. Goto, in *27th European Solid-State Device Research Conference*, edited by H. Grünbacher (Editions Frontieres, Stuttgart, Germany, 1997), pp. 336–339.
[7] G. Wolokin and J. Frey, in *Proceedings of the NASECODE VIII*, Vienna, 1992, pp. 107–108.
[8] T. Bordelon, V. M. Agostinelli, X.-L. Wang, C. M. Maziar, and A. F. Tasch, Electron. Lett. **28**, 1173 (1992).
[9] T. Grasser, H. Kosina, M. Gritsch, and S. Selberherr, in *31th European Solid-State Device Research Conference*, edited by H. Ryssel, G.

Wachutka, and H. Grünbacher (Frontier Group, Nuremberg, Germany, 2001), pp. 215–218.
[10] P. Scrobohaci and T.-W. Tang, IEEE Trans. Electron Devices **41**, 1197 (1994).
[11] T. Grasser, H. Kosina, M. Gritsch, and S. Selberherr, J. Appl. Phys. **90**, 2389 (2001).
[12] H. Reik and H. Risken, Phys. Rev. **124**, 777 (1961).
[13] D. Cassi and B. Riccò, IEEE Trans. Electron Devices **37**, 1514 (1990).
[14] N. Goldsman and J. Frey, Solid-State Electron. **31**, 1089 (1988).
[15] T. Vogelsang and W. Hänsch, J. Appl. Phys. **69**, 3592 (1991).
[16] S. Wang, N. Goldsman, and K. Hennacy, J. Appl. Phys. **71**, 1815 (1992).
[17] V. Arora, Jpn. J. Appl. Phys., Part 1 **24**, 537 (1985).
[18] M. Nekovee, B. Geurts, H. Boots, and M. Schuurmans, Phys. Rev. B **45**, 6643 (1992).
[19] S. Liotta and H. Struchtrup, Solid-State Electron. **44**, 95 (2000).
[20] T.-W. Tang, S. Ramaswamy, and J. Nam, IEEE Trans. Electron Devices **40**, 1469 (1993).
[21] A. Smith and K. Brennan, Solid-State Electron. **39**, 1659 (1996).
[22] 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.
[23] D. Matz, J. Phys. Chem. Solids **28**, 373 (1967).
[24] B. Geurts, M. Nekovee, H. M. J. Boots, and M. F. H. Schuurmans, J. Appl. Phys. **59**, 1743 (1991).
[25] H. Struchtrup, Physica A **275**, 229 (2000).
[26] D. Chen, E. C. Kan, U. Ravaioli, C.-W. Shu, and R. W. Dutton, IEEE Electron Device Lett. **13**, 26 (1992).
[27] K. Sonoda, S. T. Dunham, M. Yamaji, K. Taniguchi, and C. Hamaguchi, Jpn. J. Appl. Phys., Part 1 **35**, 818 (1996).
[28] A. Concannon, F. Piccinini, A. Mathewson, and C. Lombardi, in *International Electron Devices Meeting*, Washington, D.C. (IEEE Press, Piscataway, NJ, 1995), pp. 289–292.
[29] K. Hasnat, C.-F. Yeap, S. Jallepalli, S. A. Hareland, W.-K. Shih, V. M. Agostinelli A. F. Tasch, and C. M. Maziar, IEEE Trans. Electron Devices **44**, 129 (1997).
[30] C. Fiegna, F. Venturi, M. Melanotte, E. Sangiorgi, and B. Riccó, IEEE Trans. Electron Devices **38**, 603 (1991).
[31] T. Grasser, H. Kosina, and S. Selberherr, Appl. Phys. Lett. **80**, 613 (2002).
[32] E. Kane, J. Phys. Chem. Solids **1**, 249 (1957).
[33] M. Fischetti and S. Laux, Phys. Rev. B **38**, 9721 (1988).
[34] L. Keldysh, Sov. Phys. JETP **21**, 1135 (1965).