Influence of the distribution function shape and the band structure on impact ionization modeling

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

(Received 3 July 2001; accepted for publication 6 September 2001)

Accurate modeling of impact ionization is a critical issue for deep submicron devices. All established analytical models for the distribution function based on the local electric field or on the local average carrier energy give rather poor results. Therefore, theoretically sound microscopic scattering rates cannot be properly transformed into accurate macroscopic models. We show that by accounting for the average square energy an accurate analytical description of the distribution function can be given. Together with a proper band structure model the analytical distribution function can then be used to evaluate microscopic models in a macroscopic device simulator. The present model is accurate for bulk and sub 100 nm feature size devices and involves only local quantities, which makes it a good choice for inclusion in conventional device simulators. © *2001 American Institute of Physics.* [DOI: 10.1063/1.1415366]

I. INTRODUCTION

For fast and accurate evaluation of complementary metal–oxide–semiconductor reliability issues in sub-100 nm technology very precise high field transport models are needed. In particular, hot-carrier effects like impact ionization are poorly described by models using the local electric field or the average carrier energy as a parameter. Several nonlocal models have been proposed which are, however, both difficult to implement in a device simulator and difficult to justify on a theoretical basis, especially for multidimensional problems. Other more promising approaches rely on the extension of the method of moments and aim at obtaining some information about the distribution function (DF) in addition to the average energy. One approach is to split the energy range at some characteristic energy and handle both energy ranges with a two population and two temperature model.[1,2] Although promising results have been obtained, this approach has some disadvantages: First, because of the splitting of the energy range the complete transport model has to be reformulated including the modeling of such fundamental parameters as mobility and relaxation times. Second, there is no theoretical basis for the splitting energy in terms of carrier transport. In Refs. 1 and 2 the splitting energy has been chosen to provide additional information about impact ionization and has been set equal to the band gap energy. However, many other scattering terms could benefit from a more detailed knowledge of the DF and would probably require a different splitting energy. In addition, the transition of carriers between the two energy ranges requires careful modeling and causes numerical problems. Third, two additional equations have to be solved: one for the high-energy tail concentration and one for its temperature.

Another possibility is to extend the moment hierarchy. In particular, equation sets based on six moments of Boltzmann's equation have been considered, where in addition to the carrier concentration and temperature the next higher even moment, the average square energy of the DF was taken into account. Promising results were obtained by Sonoda *et al.*[3] who proposed an analytical expression for the distribution function. Unfortunately, their expression is difficult to handle analytically. Therefore, they proposed an analytical impact ionization model as a function of both the average energy and the average square energy by fitting Monte Carlo (MC) results. Being a fit formula, the model gives only limited additional insight into the role of the average square energy.

Highly accurate impact ionization rates have been obtained by the solution of Boltzmann's transport equation, for example by applying the MC method. Based on the notion that the collision operator in Boltzmann's transport equation is a local functional of the DF, in our model we rely on a more accurate description of the local DF. In particular, we have developed an analytical model for the symmetric part of the DF, which goes beyond the Maxwellian approximation.[4] Together with a proper description of the band structure this model is then used to derive a macroscopic impact ionization model based on microscopic descriptions. This model fits nicely into the concept of conventional device simulators as it requires only one additional equation for the average square energy and evaluates only local quantities. Furthermore, it contains no free parameters and allows for a direct study of the influence of both the shape of the DF and the band structure.

II. SIX MOMENTS METHOD

Several authors have given higher order moment equations to obtain additional information about the DF, see e.g., Refs. 5 and 6. These equations were based on an Ansatz for the DF which was taken to be some expansion around a Maxwellian distribution. Sonoda *et al.*[3] added two equations for the fourth and fifth order moments of the Boltzmann's transport equation (BTE) to a standard energy-transport model taken from Ref. 7. In a more systematic approach the

[a]Author to whom correspondence should be addressed; electronic mail: grasser@iue.tuwien.ac.at

kurtosis β_n of the DF has been included.[8] In addition to the carrier temperature T_n, which represents the second order moment of $\mathbf{k}$, the model accounts for β_n, the normalized moment of fourth order in $\mathbf{k}$

$$T_n = \frac{2}{3}\frac{\langle \mathscr{E} \rangle}{k_B} \tag{1}$$

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

For a heated Maxwell–Boltzmann (MB) 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. However, although the value of β_{MB} depends on the energy it stays close to one. Compared to the models based on splitting of the energy domain, this approach has the advantage that it does not depend on any splitting energy. Furthermore, it requires only one additional second order partial differential equation for β_n, which includes two additional relaxation times.

III. REVIEW OF SOME LOCAL IMPACT IONIZATION MODELS

A commonly used assumption[9] for modeling impact ionization is that the ionization rates depend on the local carrier temperature T_n in the following way:[10]

$$G_{\mathrm{II}} = n g_{\mathrm{II}} \exp\left(-\frac{\mathscr{E}_c}{k_B T_n} \right), \tag{3}$$

with g_{II} and $\mathscr{E}_{\mathrm{c}}$ being fit factors, which strongly depend on the technology and device geometry. Although these local energy (LE) models are capable of approximately reproducing measured integral quantities like contact currents, the predicted ionization rates *inside* the devices have been shown to deviate significantly from MC simulations.[11] This limits the usefulness of the simulations and makes predictive device simulations difficult. To overcome these limitations Sonoda *et al.*[3] derived an expression for the impact ionization coefficients using six moments of the DF by fitting the following analytical expression to their MC simulations:

$$G_{\mathrm{II}} = n\ g_{\mathrm{II}} \exp\left(-\frac{\mathscr{E}_c}{\langle \mathscr{E} \rangle_{\mathrm{eff}}} \right), \tag{4}$$

$$\langle \mathscr{E} \rangle_{\mathrm{eff}} = \langle \mathscr{E} \rangle\ \exp[\,\gamma(\sqrt{\beta_n} - \xi_h)\,], \tag{5}$$

with g_{II}, $\mathscr{E}_{\mathrm{c}}$, γ, and ξ_h being fit factors. In their original article[12] Sonoda *et al.* gave a unique set of parameters which they refined later.[3] They noted that in the bulk case β_n saturates at high energies and introduced the saturated value of $\xi_h = \lim_{T_n \to \infty} \sqrt{\beta_n} = 0.88$. In addition γ was modeled with two different values for $\sqrt{\beta_n} > \xi_h$ and $\sqrt{\beta_n} < \xi_h$, respectively. However, even for bulk, β_n is a function of the carrier temperature and the doping concentration with $\beta_n(300\ \mathrm{K}) = 1$. When Kane's nonparabolicity approximation[13] is used, a value of $\sqrt{\beta_n} = 0.88$ is obtained at approximately 800 K as long as the doping concentration is low ($N_D < 10^{16}\ \mathrm{cm}^{-3}$).

IV. A NEW IMPACT IONIZATION MODEL

In the following we derive an impact ionization model based on an analytical description of the DF, on microscopic scattering rates, and on a proper model for the density of states. We restrict our discussions in the following to electrons in silicon where we assume Kane's expression[13] to be sufficiently accurate.

It has been frequently shown that the shape of the DF depends on whether the gradient of the field is positive or negative with respect to the current density. Several expressions have been considered and evaluated and we found that

$$f(\mathscr{E}) = A \exp\left[-\left(\frac{\mathscr{E}}{\mathscr{E}_{\mathrm{ref}}} \right)^b \right], \tag{6}$$

which is a generalization of Refs. 14 and 15, neatly captures the main features of the DF throughout the whole device. The parameters $\mathscr{E}_{\mathrm{ref}} = \mathscr{E}_{\mathrm{ref}}(T_n, \beta_n)$ and $b(T_n, \beta_n)$ are functions of the local temperature and kurtosis and are determined in such a way that Eq. (6) reproduces the given moments T_n and β_n.

A. The density of states

To evaluate the moments of the DF given by Eq. (6) an expression $g(\mathscr{E})$ for the density of states (DOS) is needed. Beside the simple parabolic band approximation, Kane's dispersion relation is widely used to incorporate nonparabolicity effects to a first order. Kane's relation gives

$$g(\mathscr{E}) = g_0 \sqrt{\mathscr{E}} \sqrt{1 + \alpha \mathscr{E}} (1 + 2\alpha \mathscr{E}) \tag{7}$$

for the DOS. The nonparabolicity factor α is generally considered a fitting parameter, with $\alpha \approx 0.5\ \mathrm{eV}^{-1}$ in silicon. Unfortunately, Eq. (7) cannot be handled analytically when evaluating the moments. A commonly used simplification is the power-law approximation

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

given by Cassi and Riccò.[14] However, this expression cannot capture the nonparabolic nature of the bands, because only the low- or the high-energy part can be fit accurately, but not both simultaneously as shown in Fig. 1. Therefore, we have proposed the expression[4]

$$g(\mathscr{E}) = g_0 \sqrt{\mathscr{E}} (1 + (\eta \mathscr{E})^{\zeta}), \tag{9}$$

which can capture both the parabolic nature at low energies and the nonparabolic nature at higher energies. Furthermore, Eq. (9) can be easily integrated analytically. The parameters $\eta = 1.401\ 32\ \mathrm{eV}^{-1}$ and $\xi = 1.081\ 28$ are obtained by a fit to Eq. (7) with $\alpha = 0.5\ \mathrm{eV}^{-1}$. In addition, Eq. (9) can be nicely fit to the pseudopotential data of Ref. 16 using the parameter values $\eta = 0.9\ \mathrm{eV}^{-1}$ and $\xi = 1.4$ (cf. Fig. 1).

B. Evaluation of the parameters

With Eq. (9) as an approximation for the DOS the following expressions for $T_n(\mathscr{E}_{\mathrm{ref}}, b)$ and $\beta_n(\mathscr{E}_{\mathrm{ref}}, b)$ are found:

$$T_n = \frac{2}{3k_B}\frac{m_1}{m_0}\ , \tag{10}$$

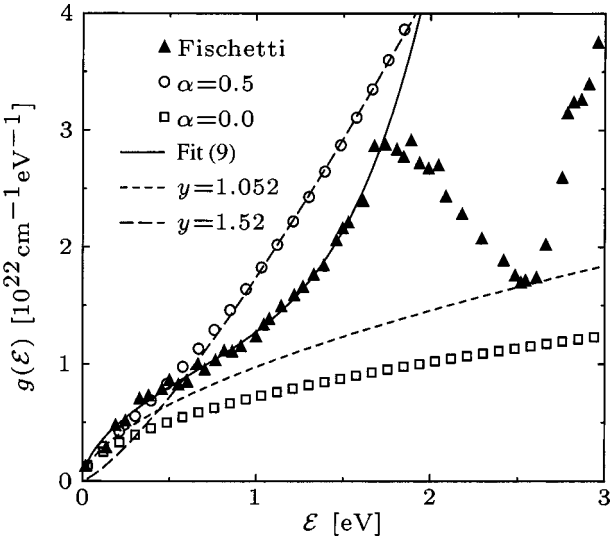

FIG. 1. Comparison of different expressions for the density of states.

$$\beta_n = \frac{3}{5}\frac{m_0 m_2}{m_1^2}, \tag{11}$$

with the moments m_l given as

$$m_l = \int_0^\infty \mathscr{E}^l f(\mathscr{E}) g(\mathscr{E}) d\mathscr{E} \tag{12}$$

$$= A g_0 \frac{\mathscr{E}_{\mathrm{ref}}^{l+3/2}}{b}\left[\Gamma\left(\frac{2l+3}{2b}\right) + (\eta\mathscr{E}_{\mathrm{ref}})^{\zeta}\Gamma\left(\frac{2l+2\zeta+3}{2b}\right)\right] \tag{13}$$

$$= A g_0 \frac{\mathscr{E}_{\mathrm{ref}}^{l+3/2}}{b} F_l(\mathscr{E}_{\mathrm{ref}}, b). \tag{14}$$

As there is no analytical inversion of Eqs. (10) and (11) with respect to $\mathscr{E}_{\mathrm{ref}}$ and b, the solution is obtained by a two-dimensional Newton procedure. A comparison of DFs obtained by MC simulation and their analytical counterparts is given in Figs. 2 and 3 for the bulk case and for an n^+-n-n^+ test structure with channel length $L_c = 200$ nm, respectively. The analytical expression slightly overestimates

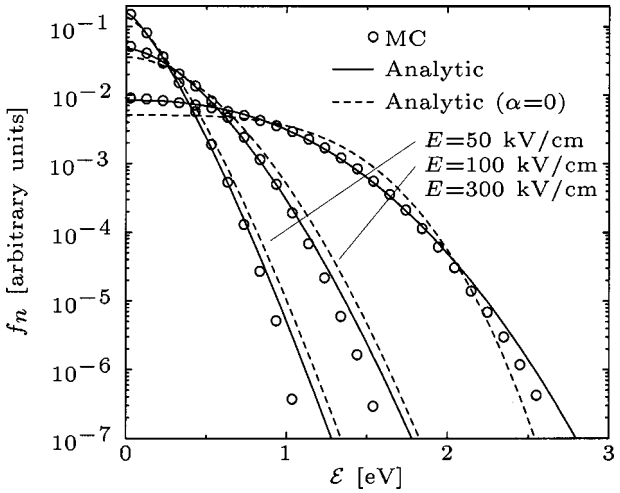

FIG. 2. Comparison of analytical expressions for the DF with MC results for the bulk case.

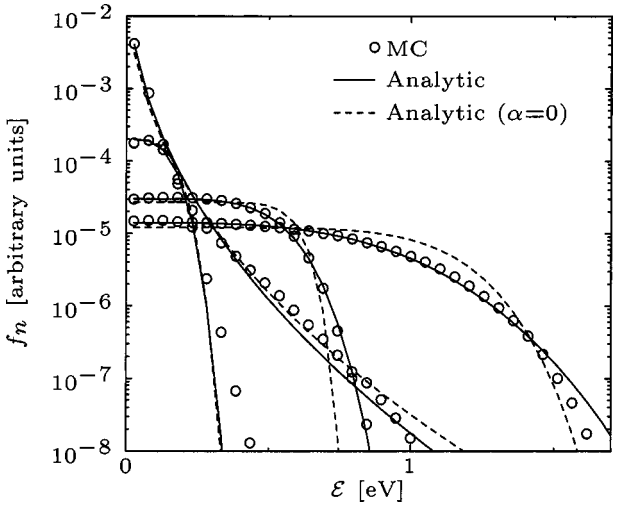

FIG. 3. Comparison of analytical expressions with MC results for the short channel device.

the high-energy tail when the curvature of the DF is too strong. Also shown are the results based on a parabolic band approximation ($\eta = 0$). The assumption of parabolic bands introduces errors of up to 50% when used to recalculate the temperature and the kurtosis.

C. The macroscopic model

With these analytical descriptions for the DF and the DOS, microscopic scattering rates can be evaluated to yield models suitable for macroscopic transport models. A popular model based on this idea has been published by Quade *et al.*[17,18] However, lacking a proper description of the DF, they used a Maxwellian shape and parabolic bands which had a significant impact on the carefully derived model: To reproduce measured values, the impact ionization threshold energy had to be set to $\mathscr{E}_{\mathrm{th}} = 4$ eV in Ref. 10 which is more than three times higher than the theoretical value $\mathscr{E}_{\mathrm{th}} \approx \mathscr{E}_{\mathrm{g}} = 1.12$ eV.

An extension of Quade's model to the six moments DF Eq. (6) is possible but gives rather complex expressions. As the inclusion of anisotropy effects into the modeling of impact ionization is avoided even in very elaborate descriptions,[19] we restrict ourselves to the isotropic expressions given by Keldysh,[20] and the extension used by Fischetti *et al.*[19] In the following, only the results based on Keldysh's formula

$$P_{\mathrm{II}}(\mathscr{E}) = P_0\left(\frac{\mathscr{E} - \mathscr{E}_{\mathrm{th}}}{\mathscr{E}_{\mathrm{th}}}\right)^2 \tag{15}$$

will be explicitly given because from a mathematical point of view Fischetti's expression is a mere superposition of three Keldysh expressions with different parameters.

For the six moments model[8] the additional terms to the right hand side of the balance equations are required, which can be obtained via

$$G_{\mathrm{II},l} = \int_{\mathscr{E}_{\mathrm{th}}}^\infty \mathscr{E}^l P_{\mathrm{II}}(\mathscr{E}) f(\mathscr{E}) g(\mathscr{E}) d\mathscr{E}. \tag{16}$$

Here $l=0,1,2$ denote the entries for the continuity, energy balance, and kurtosis balance equations, respectively. For the calculation of the moments, an accurate expression for the DOS is needed which captures both the low- and the high-energy region. For the evaluation of Eq. (16), however, we are only interested in the high-energy part of the density of states ($\mathscr{E}\geqslant\mathscr{E}_{\text{th}}$) where Cassi's expression can be used to simplify the integration. To avoid the odd unit of x we rewrite Eq. (8) as

$$g(\mathscr{E})=g_0\sqrt{\mathscr{E}_C}\left(\frac{\mathscr{E}}{\mathscr{E}_C}\right)^{\lambda}, \tag{17}$$

where the new parameters $\mathscr{E}_C$ and λ relate to x and y as

$$\lambda=\tfrac{3}{2}y-1 \tag{18}$$

$$\mathscr{E}_C=(x^{3/2}y)^{2/(3-3y)}. \tag{19}$$

A least-square fit of Eq. (17) in the energy range 1.1–2 eV to Eq. (7) gives $\mathscr{E}_C=0.35$ eV and $\lambda=1.326$. Approximation Eq. (17) introduces an error less than 0.5% for all the impact ionization rates calculated in this article. Noting that $A\,g_0$ can be expressed via the carrier density $n=m_0$ we obtain

$$G_{\text{II},l}=P_0\frac{Ag_0}{b}\mathscr{E}_C^{\lambda+1/2}\mathscr{E}_{\text{ref}}^{l+\lambda+1}H_l(\mathscr{E}_{\text{ref}},b) \tag{20}$$

$$=nP_0\mathscr{E}_{\text{ref}}^{l}\left(\frac{\mathscr{E}_{\text{ref}}}{\mathscr{E}_C}\right)^{\lambda-1/2}\frac{H_l(\mathscr{E}_{\text{ref},b})}{F_l(\mathscr{E}_{\text{ref},b})}, \tag{21}$$

where

$$H_l(\mathscr{E}_{\text{ref}},b)=\Gamma_{1,l}-2z_{\text{th}}{}^{-1/b}\Gamma_{2,l}+z_{\text{th}}{}^{-2/b}\Gamma_{3,l}, \tag{22}$$

$$\Gamma_{j,l}=\Gamma\left(\frac{j+l+\lambda}{b},z_{\text{th}}\right), \tag{23}$$

$$z_{\text{th}}=\left(\frac{\mathscr{E}_{\text{th}}}{\mathscr{E}_{\text{ref}}}\right)^{b}, \tag{24}$$

$$\Gamma(a,z)=\int_z^{\infty}e^{-t}t^{a-1}dt\ \text{(incomplete Gamma function)} \tag{25}$$

have been used. We used $P_0=4.18\times10^{12}$ s^{-1} and $\mathscr{E}_{\text{th}}=1.12$ eV consistently with the MC simulation, values that fit available experimental data as shown in Fig. 4.

V. COMPARISON WITH MONTE CARLO SIMULATIONS

In the following we give a detailed comparison of the analytically calculated impact ionization rates with results obtained by MC simulations. The analytical expression Eq. (21) was evaluated using T_n and β_n directly from the MC simulation. We calibrated the parameters of Sonoda's model to give a better fit to our MC data (see Table I for the parameter values used in this article). 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 $\alpha=0.5$ eV^{-1} and impact ionization was modeled using Keldysh's expression Eq. (15).

In addition, Eq. (21) can also be used to demonstrate the inaccuracies introduced by assuming a heated Maxwellian distribution when deriving macroscopic impact ionization rates as done for example by Quade *et al.*[17] For a heated Maxwellian distribution function, which is characterized by $\mathscr{E}_{\text{ref}}=k_BT_n$ and $b=1$, Eq. (21) delivers results close to the local energy model Eq. (3).

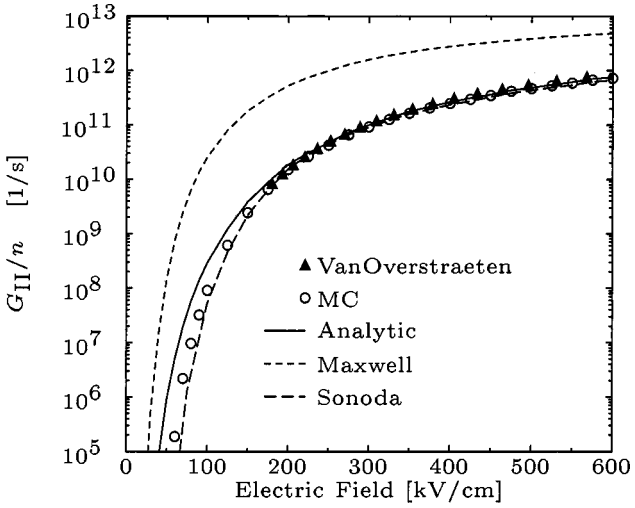

FIG. 4. Comparison of the various expressions for the bulk case.

A. Bulk case

A comparison of the various models to measured data from van Overstraeten[21] for the bulk case is given in Fig. 4. Even for the bulk case the DF deviates significantly from the Maxwellian shape, which manifests itself in too large ionization rates in the Maxwellian approximation. This is due to the overestimation of the high-energy tail where the Maxwellian approximation appears as a straight line in a semilogarithmic plot which is too crude an approximation as shown in Fig. 2. Sonoda's model with refitted parameters shows good accuracy in the whole energy range. Note, however, that not all parameters can be calibrated with bulk data only because $\sqrt{\beta_n}$ is smaller than ξ_h in the whole region of significant impact ionization. The parameters for $\sqrt{\beta_n}>\xi_h$ have to be adjusted using inhomogeneous data which is quite tedious. For the new expression no calibration is required as the same physical parameters as in the MC simulation are used. Good agreement is obtained in the whole energy range.

B. Inhomogeneous case

For the inhomogeneous case we considered the n^+-n-n^+ test structure used in Ref. 3 with channel lengths $L_c=1000$ nm, $L_c=200$ nm, and $L_c=50$ nm. The bias conditions were chosen in such a way that the maximum electric field inside the devices reaches approximately 300 kV/cm in all structures. In contrast to our previous work,[4] where we

TABLE I. Refitted coefficients for the Sonoda model.

g_{II}	$\mathscr{E}_c$	ξ_h	γ
0.8×10^{14} s^{-1}	5.3 eV	0.88	4.2 ($\sqrt{\beta_n}<\xi\chi_h$) 1.6 ($\sqrt{\beta_n}\geqslant\xi\chi_h$)

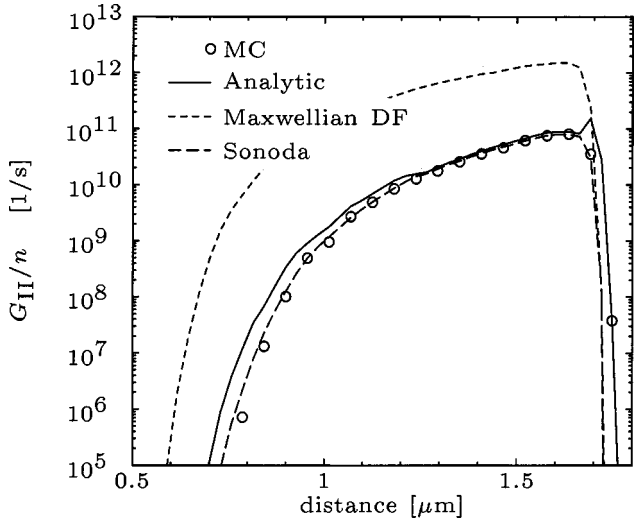

FIG. 5. Comparison of the new expressions with other models for an n^+-n-n^+ test structure with $L_c=1000$ nm.

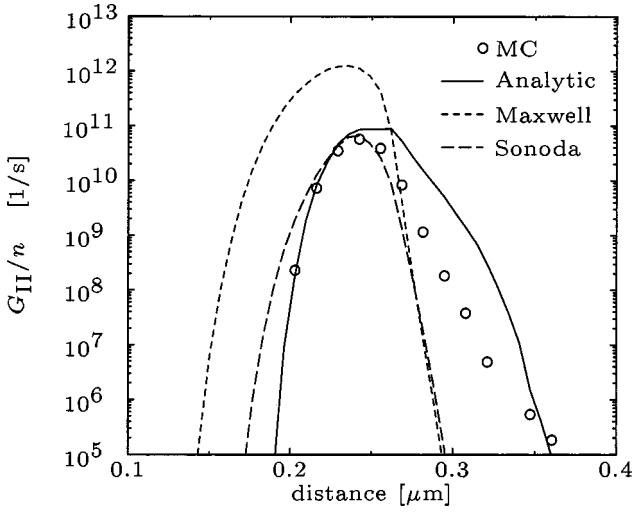

FIG. 7. Comparison of the present expressions with other models for an n^+-n-n^+ test structure with $L_c=50$ nm.

employed abrupt doping profiles, a somewhat smoother transition has been used in this article. A comparison of Eq. (21) with the MC results and the model proposed by Sonoda *et al.* is given in Figs. 5, 6, and 7 for the three n^+-n-n^+ test structures.

Both the recalibrated Sonoda model and Eq. (21) deliver excellent results for all three test cases. In the drain region, however, only the present model predicts the ionization caused by the hot-electron tail, which is however slightly overestimated. The contribution of the tail electrons to the total ionization rate becomes more important when the device size is reduced. These results are in accordance with Fig. 3 of Ref. 3 where the ionization rates predicted by the Sonoda model fall off too quickly and a calibration in this region failed. Problematic is the transition region from $\beta_n < \beta_{\mathrm{MB}}$ to $\beta_n > \beta_{\mathrm{MB}}$: When $\beta_n \approx \beta_{\mathrm{MB}}$, the analytical DF assumes a close-to-Maxwellian shape, which results in the small spikes in the figures. MC simulations show that the DF is s-shaped. There is a point of inflection, which, however, cannot be reproduced with the chosen Ansatz. A refined Ansatz is currently under investigation.

The weakness of assuming a heated Maxwellians distribution is obvious from Figs. 5 to 7. Note that the resulting profiles are similar to the profiles predicted by LE models. In general, ionization rates based on the local energy start rising too early, fall off too sharply, and considerably overestimate the ionization rates if not calibrated for the investigated device. In particular, these models cannot capture impact ionization caused by hot electrons in the drain because there the cold carriers dominate the average energy which is close to the equilibrium value.

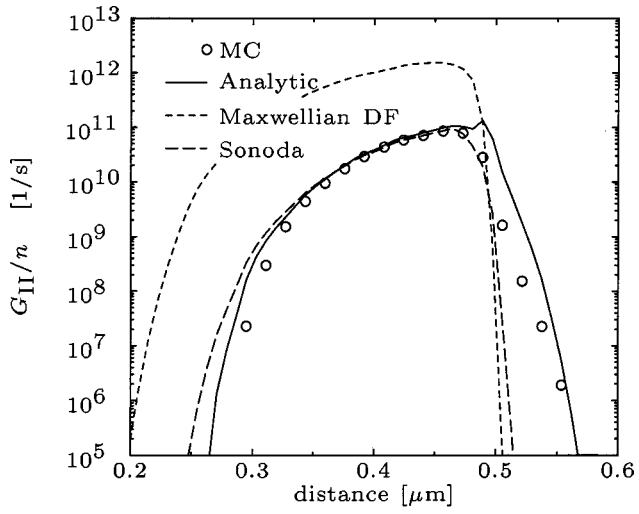

FIG. 6. Comparison of the new expressions with other models for an n^+-n-n^+ test structure with $L_c=200$ nm.

C. Higher order contributions

The higher order contribution to the source term of the energy balance equation is often modeled as[17]

$$G_{\mathrm{II},1} = \mathscr{E}_1 G_{\mathrm{II}}, \tag{26}$$

where $\mathscr{E}_1$ is frequently assumed to be the band gap energy ($\mathscr{E}_g = 1.12$ eV in silicon). In analogy, Sonoda *et al.* proposed a heuristic model for the kurtosis balance contribution[3]

$$G_{\mathrm{II},2} = \mathscr{E}_2 G_{\mathrm{II}}, \tag{27}$$

with $\mathscr{E}_2 = 8$ eV^2. With the new model Eq. (21) $\mathscr{E}_1$ and $\mathscr{E}_2$ can be directly calculated and turn out to be energy dependent. This is confirmed by MC calculations as shown in Figs. 8 and 9 for bulk and the long-channel device, respectively. The agreement with MC data of Eq. (21) for the higher order terms follows the same pattern as for G_{II} and is generally very good. The poor agreement in the contact regions will not influence simulation results as G_{II} is virtually zero. What is important, though, is the poor accuracy in the transition region from $\beta_n < 1$ to $\beta_n > 1$ where the spikes in $G_{\mathrm{II},l}$ are more pronounced due to the influence of the weight function $\mathscr{E}^l$. The approximation $\mathscr{E}_1 \approx \mathscr{E}_g$ seems to be quite reasonable, although the MC simulations indicate that a value of 1.5 eV (as in Ref. 3) or higher might be more appropriate. For $\mathscr{E}_2$ a value of 8 eV^2 seems to be much too large.

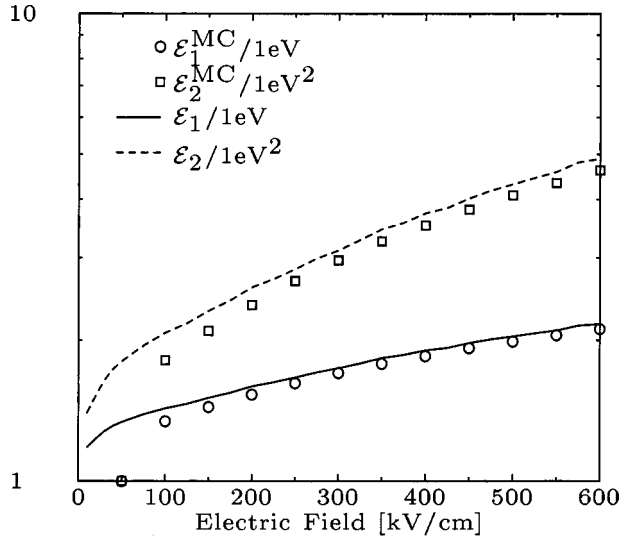

FIG. 8. Comparison of the higher order ratios $\mathscr{E}_1$ and $\mathscr{E}_2$ with MC data for the bulk case.

VI. PARABOLIC BAND APPROXIMATION

Inclusion of nonparabolicity effects considerably increases the complexity of the model. In particular, the calculation of $\mathscr{E}_{\text{ref}}$ and b requires a two-dimensional Newton method. Therefore, we investigate the influence of the parabolic band approximation on impact ionization calculation.

As pointed out before, the assumption of parabolic bands introduces errors of up to 50% when used to recalculate the temperature T_n and the kurtosis β_n. This is insofar of relevance as the parameters $\mathscr{E}_{\text{ref}}$ and b were derived under the constraint to *exactly* reproduce some given moments T_n and β_n. Furthermore, the maximum of the error coincides with the region of maximum impact ionization. Whereas a nonparabolic DOS results in DFs which fit the MC DFs over nearly the whole energy range, the parabolic approximation results in only an approximate fit (cf. Figs. 2 and 3).

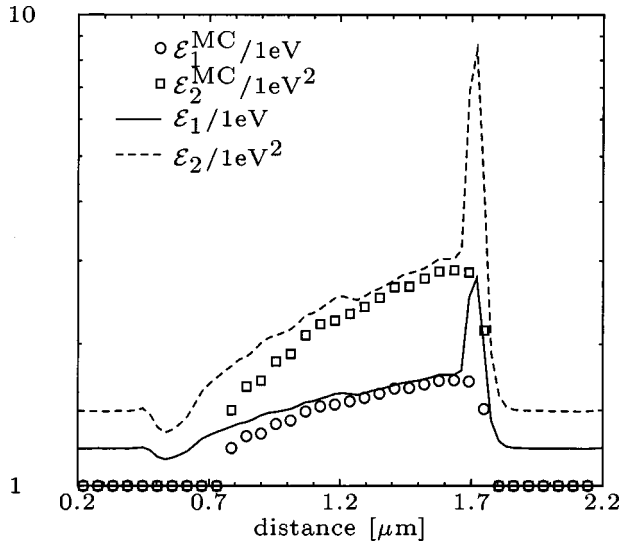

FIG. 9. Comparison of the higher order ratios $\mathscr{E}_1$ and $\mathscr{E}_2$ with MC data for the long-channel device.

TABLE II. Coefficients for (30).

i	0	1	2	3	4
b_i	0.265	0.05	−12.649	0.683	−2.659

For parabolic bands the temperature and kurtosis given by Eqs. (10) and (11) simplify to

$$T_n=\frac{2}{3}\frac{\Gamma\left(\frac{5}{2b}\right)}{\Gamma\left(\frac{3}{2b}\right)}\frac{\mathscr{E}_{\text{ref}}}{k_B}, \tag{28}$$

$$\beta_n=\frac{3}{5}\frac{\Gamma\left(\frac{3}{2b}\right)\Gamma\left(\frac{7}{2b}\right)}{\Gamma\left(\frac{5}{2b}\right)^2}. \tag{29}$$

In particular, b does not depend on T_n and neither does $\mathscr{E}_{\text{ref}}/(k_B T_n)$. A least square fit to the inversion of Eq. (29) has been obtained numerically

$$b_{\text{parabolic}}(\beta_n)=b_0+b_1\beta_n^{b_2}+b_3\beta_n^{b_4}. \tag{30}$$

The coefficients are given in Table II. Knowing $b, \mathscr{E}_{\text{ref}}$ can be obtained via Eq. (28). With $\eta=0$ and $\lambda=1/2$ we obtain for the impact ionization rates

$$G_{\text{II},l}=nP_0\mathscr{E}_{\text{ref}}^l\frac{H_l(\mathscr{E}_{\text{ref}},b)}{\Gamma\left(\frac{2+3}{2b}\right)}, \tag{31}$$

where

$$H_l(\mathscr{E}_{\text{ref}},b)=\Gamma_{1,l}-2z_{\text{th}}{}^{-1/b}\Gamma_{2,l}+z_{\text{th}}{}^{-2/b}\Gamma_{3,l}, \tag{32}$$

$$\Gamma_{j,l}=\Gamma\left(\frac{2j+2l+1}{2b},z_{\text{th}}\right). \tag{33}$$

Impact ionization rates calculated with Eq. (31) for bulk are shown in Fig. 10. The accuracy is astonishing, considering

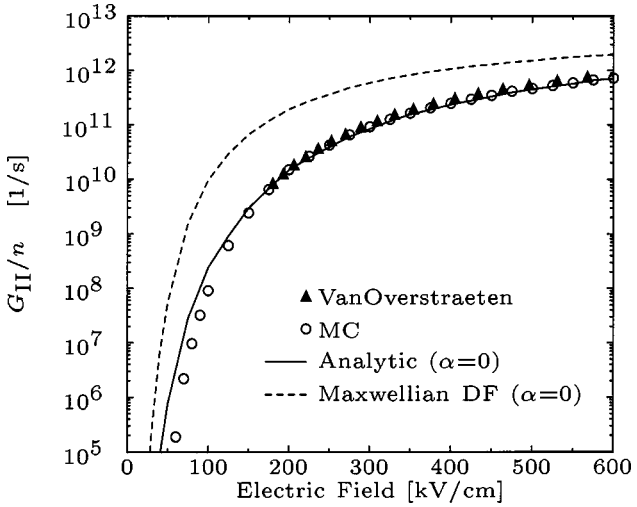

FIG. 10. Comparison of the present expressions based on parabolic bands for the bulk case.

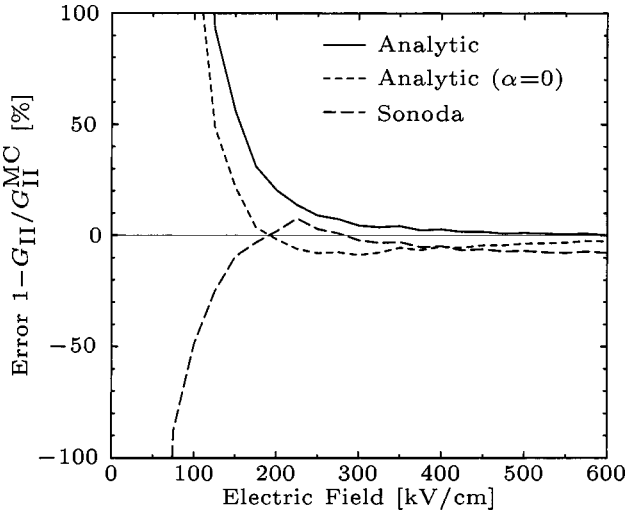

FIG. 11. Error of the expressions for G_{II} with nonparabolic and parabolic bands.

that neither the DF nor the DOS is correct. A detailed investigation shows that the errors in the DF and the errors in the DOS partially cancel. The error in the impact ionization rates is shown in Fig. 11 and results of similar accuracy are obtained for all the devices and bias conditions discussed in this article. With the same band structure as in the MC simulation, the analytical DF systematically overestimates the high-energy tail, which leads to a systematic overestimation of G_{II}, at least for the bulk case. Within the parabolic band approximation the analytic DF overestimates the high-energy tail for lower electric fields (cf. Figs. 2 and 3). As we calculate the ionization rate with a lower DOS compared to the Kane case, we obtain results closer to the MC data. For higher electric fields, however, the high-energy tail is underestimated and together with the lower DOS we obtain a systematic error of approximately 10% (cf. Figs. 2 and 3). Note that the discrepancy for all models at lower electric fields is of minor importance because of the low impact ionization rates.

VII. CONCLUSION

An impact ionization model based on an analytic description for the symmetric part of the DF has been derived. MC simulations show that the average square energy provides the necessary information for an analytical DF which goes beyond the heated Maxwellian shape assumption. Together with a proper model for the DOS an analytical description for impact ionization rates is derived based on the microscopic ionization rate. Although we have restricted ourselves to Keldysh's expression, the approach presented here can be easily extended to more complex models. The new model contains only physical parameters which are the same as in the MC simulation. Note that the fit parameters which have been introduced for the DOS are for mathematical convenience only: A third order Taylor expansion of Kane's dispersion relation delivers results of the same accuracy but with considerably more complicated expressions for the impact ionization rates. The present model allows a detailed investigation of the influence of the shape of the DF and the DOS. Assuming a heated Maxwellian shape results in large errors because the regions with strong impact ionization do not coincide with the regions where the average energy is high. Notably, the onset of impact ionization is not directly related to the beginning of the high average energy regions, a phenomenon frequently modeled by introducing artificial "dead lengths." Furthermore, impact ionization caused by hot carriers continues inside the drain regions where the average energy has already relaxed close to its equilibrium value. Both phenomena are directly related to the tail of the DF and do not require any artificial corrections, provided the necessary details of the DF are known. The DOS, on the other hand, is important to accurately calculate the parameters of the analytic DFs. For the particular case of parabolic bands the errors introduced in the parameter calculation seem to cancel out the errors introduced in the final integration of the microscopic impact ionization rate. As only the basic features of the real band structure are captured by Kane's dispersion relation, the accuracy of the parabolic approximation might be sufficient for the calculation of impact ionization rates. The new model fits nicely into the concept of macroscopic device simulators as it requires only local quantities.

[1] P. Scrobohaci and T.-W. Tang, IEEE Trans. Electron Devices **41**, 1197 (1994).
[2] J.-G. Ahn, C.-S. Yao, Y.-J. Park, H.-S. Min, and R. Dutton, IEEE Electron Device Lett. **15**, 348 (1994).
[3] K. Sonoda, M. Yamaji, K. Taniguchi, C. Hamaguchi, and S. Dunham, J. Appl. Phys. **80**, 5444 (1996).
[4] T. Grasser, H. Kosina, and S. Selberherr, *Conference Proceedings on Simulation of Semiconductor Devices and Processes*, Athens, Greece, September 2001 (Springer, Wien, 2001), pp. 46–49.
[5] B. Geurts, M. Nekovee, H. Boots, and M. F. H. Schuurmans, J. Appl. Phys. **59**, 1743 (1991).
[6] S. Liotta and H. Struchtrup, Solid-State Electron. **44**, 95 (2000).
[7] M. Vecchi and L. Reyna, Solid-State Electron. **37**, 1705 (1994).
[8] T. Grasser, H. Kosina, M, Gritsch, and S. Selberherr, J. Appl. Phys. **90**, 2389 (2001).
[9] Y. Apanovich, E. Lyumkis, B. Polsky, A. Shur, and P. Blakey, IEEE Trans. comput.-Aided Des. **13**, 702 (1994).
[10] K. Souissi, F. Odeh, H. Tang, A. Gnudi, and P. Lu, IEEE Trans. Electron Devices **40**, 1501 (1993).
[11] T. Grasser, H. Kosina, M. Gritsch, and S. Selberherr, in Proceedings of the 4th International Conference on Modeling and Simulations of Microsystems, Hilton Head Island, SC, March 2001, pp. 474–477.
[12] K. Sonoda, M. Yamaji, K. Taniguchi, and C. Hamaguchi, in *Simulation of Semiconductor Devices and Processes*, edited by H. Ryssel and P. Pichler, Wien, Austria (Springer, Wien 1995), Vol. 6, pp. 420–423.
[13] E. Kane, J. Phys. Chem. Solids **1**, 249 (1957).
[14] D. Cassi and B. Riccò, IEEE Trans. Electron Devices **37**, 1514 (1990).
[15] K. Hasnat, C.-F. Yeap, S. Jallepalli, S. Hareland, W.-K. Shih, V. Agostinelli, A. Tasch, and C. Maziar, IEEE Trans. Electron Devices **ED-44**, 129 (1997).
[16] M. Fischetti and S. Laux, Phys. Rev. B **38**, 9721 (1988).
[17] W. Quade, E. Schöll, and M. Rudan, Solid-State Electron. **36**, 1493 (1993).
[18] W. Quade, M. Rudan, and E. Schöll, IEEE Trans. Comput.-Aided Des. **10**, 1287 (1991).
[19] M. Fischetti, S. Laux, and E. Crabbé, J. Appl. Phys. **78**, 1058 (1995).
[20] L. Keldysh, Sov. Phys. JETP **21**, 1135 (1965).
[21] R. VanOverstraeten and H. DeMan, Solid-State Electron. **13**, 583 (1970).