Accurate impact ionization model which accounts for hot and cold carrier populations

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

(Received 12 October 2001; accepted for publication 16 November 2001)

Conventional macroscopic impact ionization models which use the average carrier energy as a main parameter can not accurately describe the phenomenon in modern miniaturized devices. Here, we present a model which is based on an analytic expression for the distribution function. In particular, the distribution function model accounts explicitly for a hot and a cold carrier population in the drain region of metal–oxide–semiconductor transistors. The parameters are determined by three-even moments obtained from a solution of a six-moments transport model. Together with a nonparabolic description of the density of states, accurate closed form macroscopic impact ionization models can be derived based on familiar microscopic descriptions. © *2002 American Institute of Physics.* [DOI: 10.1063/1.1445273]

Accurate calculation of impact ionization rates in macroscopic transport models is becoming more and more important due to the ongoing feature size reduction of modern semiconductor devices. Conventional models which use the average carrier energy as a main parameter fail because impact ionization is very sensitive to the shape of the distribution function, in particular to the high-energy tail. The average energy is not sufficient for obtaining this information and higher order moments of the distribution function have to be considered. We favor a local description because the scattering operator in Boltzmann's equation is a functional of the local distribution function which should be reflected by the model.

Here, we present a refined version of a previously published model[1] developed for the use with a six-moments transport model[2] which also accounts for the kurtosis of the distribution function, $\beta_n=(3/5)\langle\mathcal{E}^2\rangle/\langle\mathcal{E}\rangle^2$, in addition to the carrier temperature. So far, we used

$$f(\mathcal{E})=A\exp\left[-\left(\frac{\mathcal{E}}{\mathcal{E}_{\mathrm{ref}}}\right)^b\right], \tag{1}$$

$$g(\mathcal{E})=g_0\sqrt{\mathcal{E}}(1+(\eta\mathcal{E})^\zeta) \tag{2}$$

for the symmetric part of the distribution function and for the nonparabolic density of states, respectively. The parameters $\mathcal{E}_{\mathrm{ref}}=\mathcal{E}_{\mathrm{ref}}(T_n,\beta_n)$ and $b=b(T_n,\beta_n)$ are functions of the carrier temperature and kurtosis and are determined in such a way that $f(\mathcal{E})$ reproduces the given moments T_n and β_n.[1] The parameters η and ζ of Eq. (2) are determined by a fit to either Kane's dispersion relation[3] or to pseudopotential data.[1] Due to the form of the fit expression for the density of states Eq. (2) it is possible to give algebraic expressions for the moments of Eq. (1) using Gamma functions. With the parameter values $\eta=1.4\ \mathrm{eV}^{-1}$ and $\zeta=1.08$ the error in the first three-even moments was found to be smaller than 1% when compared to the results obtained from Kane's relation.

Expression (1) is accurate inside the channel of metal–oxide–semiconductor (MOS) transistors and gives values of $b>1$. Although Eq. (1) gives reasonable approximations for the distribution function inside the drain region ($b<1$), there are two problems: Firstly, the high-energy tail is overestimated and secondly, during the transition from the channel to the drain region, the exponent b assumes the value 1 which corresponds to a Maxwellian distribution function, a result not confirmed by Monte Carlo simulations. This error is amplified when, for instance, impact ionization rates are calculated where only the high-energy tail of the distribution function is required.[1]

To improve the model, we note that at the drain junction, the hot carriers from the channel meet a large pool of cold carriers and two populations coexist. We account for this fact by using a superposition of two distributions, similar to the work of Sonoda *et al.*[4]

$$f(\mathcal{E})=A\Bigg\{\underbrace{\exp\left[-\left(\frac{\mathcal{E}}{\mathcal{E}_{\mathrm{ref}}}\right)^b\right]}_{f_1(\mathcal{E})}+c\ \underbrace{\exp\left[-\frac{\mathcal{E}}{k_BT_2}\right]}_{f_2(\mathcal{E})}\Bigg\}. \tag{3}$$

We now have to determine the five parameters A, $\mathcal{E}_{\mathrm{ref}}$, b, c, and T_2 which describe the distribution function, that is, we need two heuristic relations in addition 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 distribution function in the drain region, we look at the second-order moment T_1 and the fourth-order moment β_1 of the hot distribution $f_1(\mathcal{E})$ only (Fig. 1) which was extracted from the total distribution function obtained by a Monte Carlo simulation in a postprocessing step. Additional simulations show that the temperature T_2 of the cold Maxwellian distribution function $f_2(\mathcal{E})$ rapidly relaxes to the lattice temperature T_L and will be modeled as $T_2=T_L$ in this work. The kurtosis β_1, however, is crucial for an accurate description of the high-energy tail. Interestingly, β_1 can be modeled accurately via the bulk relation $\beta_{\mathrm{Bulk}}(T_n)$ which can be derived from the homogeneous six-moments model[2] as

$$\beta_{\mathrm{Bulk}}(T_n)=\frac{T_L^2}{T_n^2}+2\,\frac{\tau_\beta}{\tau_\mathcal{E}}\,\frac{\mu_S}{\mu_n}\left(1-\frac{T_L}{T_n}\right), \tag{4}$$

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

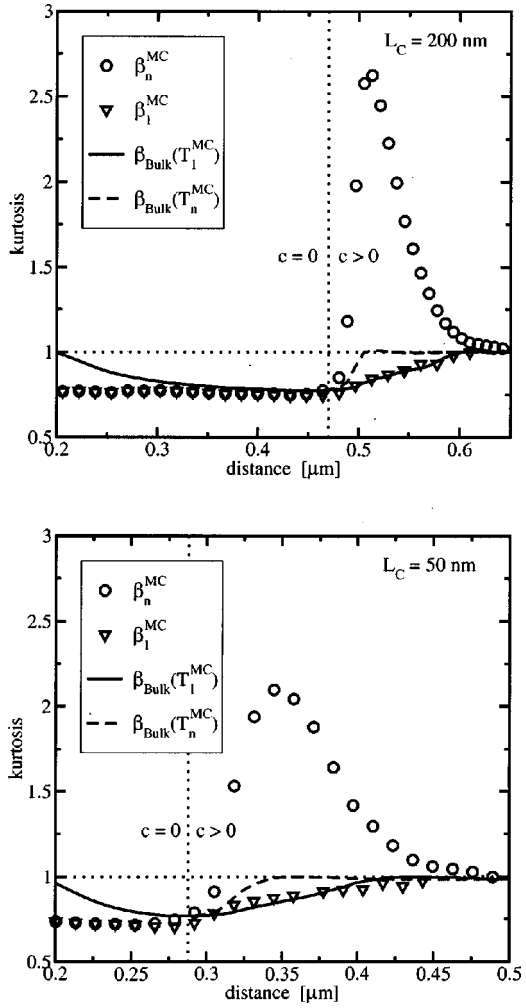

FIG. 1. Comparison of β_1^{MC} with two analytical models. When the temperature of the hot distribution function T_1 is used in the bulk characteristic, accurate results are obtained for n^+-n-n^+ test structures with L_C=200 nm and L_C=50 nm. Note that $\beta_{\mathrm{Bulk}}(T_n)$ does not properly describe the behavior of β_1. Also shown is the kurtosis of the total distribution function β_n.

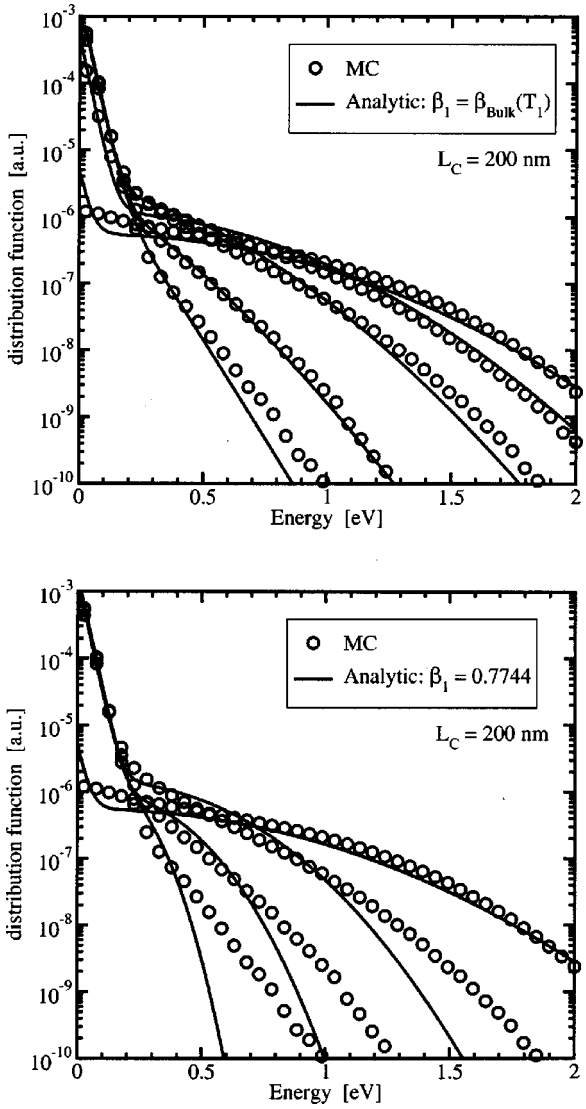

FIG. 2. Comparison of analytical expressions for the distribution function with Monte Carlo results at different positions inside an n^+-n-n^+ test-structure with L_C=200 nm. Note the error in the tail when a constant value for β_1 is assumed (bottom figure).

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.

A comparison of the model $\beta_1=\beta_{\mathrm{Bulk}}(T_1)$ with Monte Carlo data is shown in Fig. 1, where the temperature of the high-energy tail T_1 has been taken as the argument. Note that $\beta_{\mathrm{Bulk}}(T_n)$ approaches unity too quickly as also shown in Fig. 1 which underlines the idea of modeling the hot and cold electrons as separate populations.

For the calculation of the parameters $\mathcal{E}_{\mathrm{ref}}$, b, and c, we have to detect the regions where a cold population exists. Monte Carlo simulations show that inside the channel, the tail of the distribution function is always less populated than in the bulk case. Therefore, we detect the drain region when $\beta_n>\beta_{\mathrm{Bulk}}(T_n)$ is fulfilled. Inside the channel, we assume $c=0$ because no cold subpopulation exists. Inside the drain region, however, we need to determine c and T_2 to explicitly allow for two separate populations.

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

$$\begin{pmatrix} T_n(\mathcal{E}_{\mathrm{ref}},b,c) \\ \beta_n(\mathcal{E}_{\mathrm{ref}},b,c) \\ \beta_1(\mathcal{E}_{\mathrm{ref}},b,c) \end{pmatrix} = \begin{pmatrix} T_n^{\mathrm{MC}} \\ \beta_n^{\mathrm{MC}} \\ \beta_{\mathrm{Bulk}}[T_1(\mathcal{E}_{\mathrm{ref}},b,c)] \end{pmatrix}. \tag{5}$$

Note that T_n, β_n, and β_1 are analytic expressions derived from the moments of Eq. (3) and T_n^{MC} and β_n^{MC} were taken from Monte Carlo simulations. As stated, in the channel, $c=0$ is assumed and the last row of Eq. (5) is dropped. Since there is no cold population, $T_1=T_n$ holds inside the channel. This also holds at the transition point which guarantees a continuous transition between the two regions. A comparison with distribution functions obtained by Monte Carlo simulations is shown in Fig. 2. Note that a constant value for $\beta_1=\beta_h$ as used in Ref. 4 underestimates the tail of the distribution function and thus the associated impact ionization rate. Furthermore, an approach based on a constant β_1 works only for the high field case because otherwise β_n will never reach β_h and the model erroneously creates a cold population throughout the whole device. For intermediate bias conditions, a spurious cold population would be predicted in the larger part of channel.

A closed form macroscopic impact ionization rate is then obtained by integrating Keldysh's expression[5]

$$P_{\mathrm{II}}(\mathcal{E})=P_0\left(\frac{\mathcal{E}-\mathcal{E}_{\mathrm{th}}}{\mathcal{E}_{\mathrm{th}}}\right)^2 \tag{6}$$

with Eqs. (1) and (2) as

$$G_{\mathrm{II},l}=\int_{\mathcal{E}_{\mathrm{th}}}^{\infty}\mathcal{E}^{l}P_{\mathrm{II}}(\mathcal{E})f(\mathcal{E})g(\mathcal{E})d\mathcal{E} \tag{7}$$

$$\approx\int_{\mathcal{E}_{\mathrm{th}}}^{\infty}\mathcal{E}^{l}P_{\mathrm{II}}(\mathcal{E})f_1(\mathcal{E})g_c(\mathcal{E})d\mathcal{E} \tag{8}$$

$$=nP_0\mathcal{E}_{\mathrm{ref}}^{l}\left(\frac{\mathcal{E}_{\mathrm{ref}}}{\mathcal{E}_C}\right)^{\lambda-1/2}\times\frac{\Gamma_{1,l}-2z_{\mathrm{th}}^{-1/b}\Gamma_{2,l}+z_{\mathrm{th}}^{-2/b}\Gamma_{3,l}}{\Gamma\left(\frac{2l+3}{2b}\right)+(\eta\mathcal{E}_{\mathrm{ref}})^{\zeta}\Gamma\left(\frac{2l+2\zeta+3}{2b}\right)}, \tag{9}$$

where $\Gamma(a,z)$ is the incomplete Gamma function and

$$\Gamma_{j,l}=\Gamma\left(\frac{j+l+\lambda}{b},z_{\mathrm{th}}\right), \tag{10}$$

$$z_{\mathrm{th}}=\left(\frac{\mathcal{E}_{\mathrm{th}}}{\mathcal{E}_{\mathrm{ref}}}\right)^{b}, \tag{11}$$

$$g_c(\mathcal{E})=g_0\sqrt{\mathcal{E}_C}\left(\frac{\mathcal{E}}{\mathcal{E}_C}\right)^{\lambda}. \tag{12}$$

Equation (12) is an approximation of the high-energy region of the density of states ($\mathcal{E}_C=0.35$ eV and $\lambda=1.326$), based on Cassi and Riccò's[6] model, which has been introduced to simplify the final expression. Furthermore, it is assumed that only $f_1(\mathcal{E})$ contributes to the impact ionization rate. In Eq. (9) $l=0$, 1, 2 and denotes the entries for the continuity, energy balance, and kurtosis balance equations, respectively.

The final expression (9) is equivalent to the expression given in Ref. 1, but the parameters $\mathcal{E}_{\mathrm{ref}}$ and b are now calculated in a different way. We use the same parameter values as in the Monte Carlo simulation ($P_0=4.18\times10^{12}$ s^{-1}, $\mathcal{E}_{\mathrm{th}}=1.12$ eV), values which fit available experimental data for bulk. The accuracy for the bulk case is the same as in Ref. 1 as there is no cold population ($c=0$). A comparison with Monte Carlo simulations for n^+-n-n^+ test structures is shown in Fig. 3 where the improvement to the model of Ref. 1 is obvious. In addition, the results obtained by assuming a constant β_1 and the results obtained by a Maxwellian distribution function ($\mathcal{E}_{\mathrm{ref}}=k_BT_n$, $b=1$, and $c=0$) are shown which underline the importance of an accurate distribution function model for impact ionization rate modeling.

The macroscopic impact ionization model has been proven to deliver accurate results for the homogenous case[1] and inhomogeneous cases down to nanoscale devices. In particular, the tail of the impact ionization rate inside the drain area is correctly predicted which is not possible with models based on the average carrier energy only. The additional accuracy is provided by the kurtosis of the distribution function which can be obtained via a six-moments transport model, that is, a model which is one order higher than conventional energy-transport models. It is important to point out that the same parameters as in the Monte Carlo simulation were used to evaluate the model, so no fitting has been performed. Although we have taken a rather simple expression for the microscopic scattering rate which is known to be not very accurate for energies close to the threshold energy, an extension to more accurate models is straightforward. As the model involves only state variables of the equation system, it is well suited for the implementation into conventional numerical device simulators.

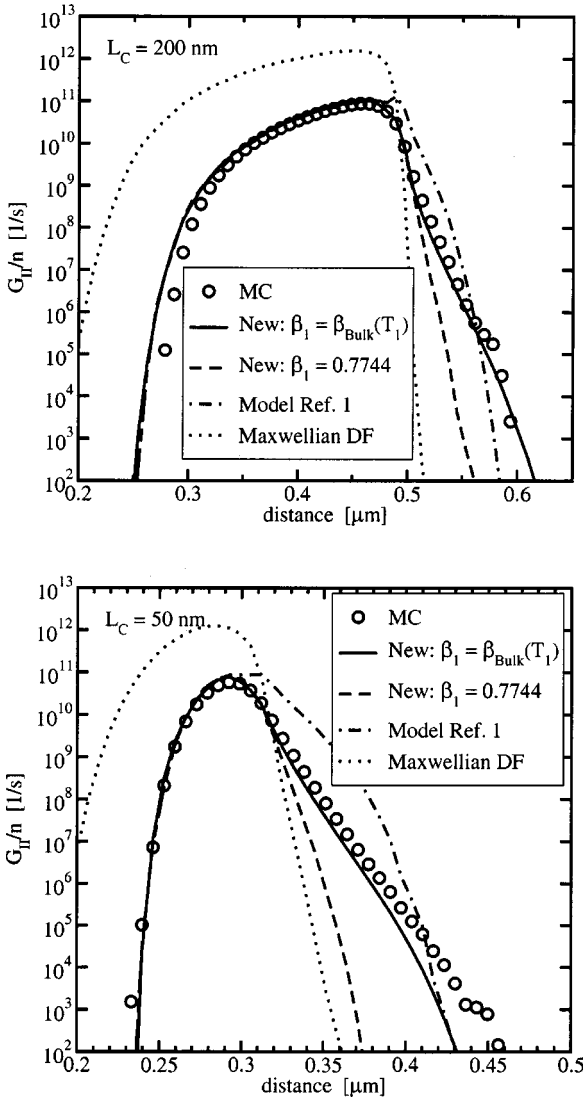

FIG. 3. Comparison of analytically obtained impact ionization rates with Monte Carlo results for two n^+-n-n^+ test structures. Note the improvement to the model of Ref. 1 and the large error when a Maxwellian distribution function is assumed ($E_{\mathrm{max}}=300$ kV/cm).

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

[1] T. Grasser, H. Kosina, and S. Selberherr, J. Appl. Phys. **90**, 6165 (2001).
[2] T. Grasser, H. Kosina, M. Gritsch, and S. Selberherr, J. Appl. Phys. **90**, 2389 (2001).
[3] E. Kane, J. Phys. Chem. Solids **1**, 249 (1957).
[4] K. Sonoda, S. T. Dunham, M. Yamaji, K. Taniguchi, and C. Hamaguchi, Jpn. J. Appl. Phys., Part 1 **35**, 818 (1996).
[5] L. Keldysh, Sov. Phys. JETP **21**, 1135 (1965).
[6] D. Cassi and B. Riccò, IEEE Trans. Electron Devices **37**, 1514 (1990).