Moment expansion approach to calculate impact ionization rate in submicron silicon devices

Ken-ichiro Sonoda, Mitsuru Yamaji, Kenji Taniguchi, Chihiro Hamaguchi, and Scott T. Dunham

Citation: Journal of Applied Physics **80**, 5444 (1996); doi: 10.1063/1.362732
View online: http://dx.doi.org/10.1063/1.362732
View Table of Contents: http://scitation.aip.org/content/aip/journal/jap/80/9?ver=pdfcov
Published by the AIP Publishing

Moment expansion approach to calculate impact ionization rate in submicron silicon devices

Ken-ichiro Sonoda,[a] Mitsuru Yamaji, Kenji Taniguchi, and Chihiro Hamaguchi
Department of Electronic Engineering, Osaka University, Yamada-oka, Suita, Osaka 565, Japan

Scott T. Dunham
Department of Electrical, Computer, and Systems Engineering, Boston University, 44 Cummington Street, Boston, Massachusetts 02215

(Received 9 November 1995; accepted for publication 1 August 1996)

A method to calculate the impact ionization rate in submicron silicon devices is developed using both an average energy and an average square energy of electrons. The method consists of an impact ionization model formulated with the average energy and conservation equations for the average square energy in the framework of an energy transport model. Parameters for the transport equations are extracted in such a way that calculated moments based on these equations match Monte Carlo simulation results. The impact ionization generation rate in an n^+nn^+ structure calculated with this method agrees well with the results obtained from Monte Carlo simulation. The new method is also applied to a submicron *n*-MOSFET. The calculated distribution of the generation rate is found to be quite different from the results based on a conventional method. © *1996 American Institute of Physics.* [S0021-8979(96)08421-6]

I. INTRODUCTION

Much effort has been devoted to device miniaturization into a deep submicron region to accomplish higher density of integration and better circuit performance. Device degradation caused by hot carrier effects has been a primary concern in recent years because the power supply voltage has not been scaled down along with the physical dimensions of MOSFETs. A device simulator is a powerful tool to optimize device structure by satisfying both current drive capability and the suppression of device degradation. Although it is mandatory to utilize accurate physical models to obtain precise simulation results, the presence of a large field gradient and multi-dimensional nature of the physical phenomena make it difficult to simulate the hot carrier effects in deep submicron MOSFETs.

The most popular approach for the calculation of impact ionization (II) is a local field (LF) model in the framework of a drift-diffusion model. In the LF model, an II coefficient, α, which is the number of ionization events per unit length along a current path, is determined[1] by the local electric field E with the characteristic field strength E_c as $\alpha \propto \exp(-E_c/E)$. Strictly speaking, the LF model holds only in a homogeneous electric field because the energy distribution function of carriers is not described using a local electric field in the case where the field varies within the energy relaxation length, which is the product of the saturation velocity and the energy relaxation time for electrons, $v_{\mathrm{sat}}\tau_{\langle\epsilon\rangle}(\sim 0.05\ \mu\mathrm{m})$. The value $\tau_{\langle A\rangle}$ is the relaxation time of $\langle A\rangle \equiv \int\int A d\boldsymbol{k} / \int\int d\boldsymbol{k}$, which are used hereafter. The LF model tends to overestimate the II rate in an increasing-field region; a simulated substrate current is normally much larger than its experimental counterpart.

In order to overcome these shortcomings, a nonlocal field (NLF) model[2–4] was proposed. In this model, carrier trajectories are traced back along an electric field path to a point where the electrostatic potential difference is equal to the effective threshold energy for impact ionization, V_I. The ionization coefficient is expressed[4] with the path length d as $\alpha \propto \exp(-d/\lambda_p)$, where λ_p is the effective mean free path without being scattered by the optical phonons. In a homogeneous field, the model reduces to the local field model with the relation of $\lambda_p = V_I/E_c$. In the nonlocal field model, the impact ionization phenomenon is treated as a one-dimensional problem. In reality, high energy electrons move around in three-dimensional space so that it is questionable whether this method can be adopted to multi-dimensional device structure.

The exact calculation of the II coefficient demands the use of an energy distribution function, because II events are caused by energetic carriers. The average energy of the distribution function is a basic physical parameter to characterize the energy distribution. It is more appropriate to use the average energy rather than the electric field to express the impact ionization coefficient. In a homogeneous high electric-field, the average energy, $\langle\epsilon\rangle$, is given by the electric field E as $\langle\epsilon\rangle = qEv_{\mathrm{sat}}\tau_{\langle\epsilon\rangle}$. This relation is combined with the local field model to give

$$\alpha \propto \exp\left(-\frac{\epsilon_c}{\langle\epsilon\rangle}\right). \tag{1}$$

This model is called a local energy (LE) model because the ionization coefficient is expressed as a function of the local average energy. The characteristic energy, ϵ_c, is obtained from $\epsilon_c = qE_cv_{\mathrm{sat}}\tau_{\langle\epsilon\rangle}$. This model is often used in energy transport models.[5–7] Other expressions for the II generation rate have been proposed[8–10] using the average energy of carriers.

Recent studies[11–14] have shown that the average energy is not a good measure to predict II generation rate in an inho-

[a]On leave from ULSI Laboratory, Mitsubishi Electric Corporation, Mizuhara, Itami, Hyogo 664, Japan.

 0021-8979/96/80(9)/5444/5/$10.00

TABLE I. The impact ionization models investigated and their physical parameters.

Models	Formulas and parameters
Local field (LF) model	$\alpha=\alpha_0 \exp\left(-\frac{E_c}{E}\right)$, $\alpha_0=1.40\times10^6$ cm^{-1}, $E_c=1.32\times10^6$ V/cm
Nonlocal field (NLF) model	$\alpha=\alpha_0 \exp\left(-\frac{d}{\lambda_p}\right)$, $V_I=\int_{x-d}^{x} \boldsymbol{E}\cdot d\boldsymbol{l}$, $\lambda_p=V_I/E_c$, $V_I=1.1$ eV
Local energy (LE) model	$G_{ii}=ng_{ii\,0} \exp\left(-\frac{\epsilon_c}{\langle\epsilon\rangle}\right)$, $g_{ii\,0}=1.39\times10^{14}$ s^{-1}, $\epsilon_c=8.78$ eV
New model	$G_{ii}=ng_{ii\,0} \exp\left(-\frac{\epsilon_c}{\langle\epsilon\rangle_{\rm eff}}\right)$, $\langle\epsilon\rangle_{\rm eff}=\langle\epsilon\rangle\exp(\gamma(\xi-\xi_h))$, $\gamma=\begin{cases}4.2 & (\xi<\xi_h)\\ 1.6 & (\xi>\xi_h)\end{cases}$, $\xi_h=0.88$

mogeneous field. The authors have proposed[14] the impact ionization model using the average square energy of electrons $\langle\epsilon^2\rangle$ in addition to the average energy $\langle\epsilon\rangle$. The pair creation rate by II in an inhomogeneous field is expressed as

$$G_{\rm ii}=ng_{\rm ii0}\exp\left(-\frac{\epsilon_c}{\langle\epsilon\rangle_{\rm eff}}\right), \tag{2}$$

$$\langle\epsilon\rangle_{\rm eff}=\langle\epsilon\rangle\exp(\gamma(\xi-\xi_h)), \tag{3}$$

and

$$\xi\equiv\frac{\sqrt{(3/5)\langle\epsilon^2\rangle}}{\langle\epsilon\rangle} \tag{4}$$

which gives higher accuracy compared with previous models. The values of the parameters in the above expressions are shown in Table I.

In this paper, we propose a new method to calculate II generation rate in deep submicron MOSFETs. It is based on the moment expansion of distribution functions. The transport equations for the average square energy as well as the carrier concentration and the average energy are derived from the Boltzmann transport equation (BTE) in Sec. II. Simulation results for an n^+nn^+ structure and an n-MOSFET are shown in Sec. II.

II. MOMENT EXPANSION METHOD

There are two ways to derive the energy transport model from the BTE: the energy balance (EB) model[15] and the hydro-dynamic (HD) model.[16] Both of them are derived by multiplying the BTE by a group velocity $\boldsymbol{u}$ and/or an energy ϵ and integrating them over $\boldsymbol{k}$-space. A major difference between these models lies in the treatment of the collision term in the BTE. In the HD model, the relaxation times for each moment (average velocity, average energy, etc.) are introduced to express the collision terms after the integration over $\boldsymbol{k}$-space. In the EB model, on the other hand, the collision term in the BTE is expressed as $-(f-f_0)/\tau$, in which only a microscopic relaxation time τ is included. The resulting equations of these models take different forms. From the practical point of view, however, they are regarded as similar, except for the definition of the parameters used in the equations.[6,17]

Several attempts have been made to clarify which model is more suitable for semiconductor device simulation,[6,18] and it has been concluded that neither the EB nor the HD model correctly describes the moments in an inhomogeneous field. We treat the parameters in the equations only as fitting variables to reproduce the moments which are close to those obtained from Monte Carlo simulation.

The energy transport model used in this study consists of the following equations:[18]

$$\frac{\partial n}{\partial t}+\nabla\cdot(n\langle\boldsymbol{u}\rangle)=\left(\frac{\partial n}{\partial t}\right), \tag{5}$$

$$\frac{\partial n\langle\epsilon\rangle}{\partial t}+q\boldsymbol{E}\cdot(n\langle\boldsymbol{u}\rangle)+\nabla\cdot(n\langle\epsilon\boldsymbol{u}\rangle) =-n\frac{\langle\epsilon\rangle-\langle\epsilon\rangle_0}{\tau_{\langle\epsilon\rangle}}+\left(\frac{\partial n\langle\epsilon\rangle}{\partial t}\right)_{\rm GR}, \tag{6}$$

with the consecutive equations

$$n\langle\boldsymbol{u}\rangle=-n\mu\boldsymbol{E}-\mu\frac{kT_n}{q}\nabla n-\mu\frac{k}{q}(1+\nu_0)n\nabla T_n \tag{7}$$

and

$$n\langle\epsilon\boldsymbol{u}\rangle=\frac{\tau\langle\epsilon\boldsymbol{u}\rangle}{\tau\langle\boldsymbol{u}\rangle}\frac{5}{3}\left(kT_n(n\langle\boldsymbol{u}\rangle)-qn\mu\left(\frac{k}{q}\right)^2(1+\nu_1)T_n\nabla T_n\right). \tag{8}$$

The electron temperature T_n is defined by $\langle\epsilon\rangle=(3/2)kT_n$. The terms $(\partial n/\partial t)_{\rm GR}$ and $(\partial n\langle\epsilon\rangle/\partial t)_{\rm GR}$ mean the relaxation rates for the electron density and the average energy due to the generation-recombination (GR) process, respectively. In Eqs. (7) and (8), $n\langle\boldsymbol{u}\rangle$ and $n\langle\epsilon\boldsymbol{u}\rangle$ are particle- and energy-fluxes, respectively.

For the relaxation rate caused by the GR process, only the II is taken into account in this study, i.e., $(\partial n/\partial t)_{\rm GR}=G_{\rm ii}$. The relaxation rate of the average energy due to II is formulated as $(\partial n\langle\epsilon\rangle/\partial t)_{\rm GR}=\epsilon_{\rm loss}G_{\rm ii}$, where $\epsilon_{\rm loss}$ is the net average energy loss by impact ionization.

The physical parameters μ, $\tau\langle\epsilon\rangle$, $\tau\langle\boldsymbol{u}\epsilon\rangle/\tau\langle\boldsymbol{u}\rangle$, and $\epsilon_{\rm loss}$ are determined to fit the E-$\langle u\rangle$, E-$\langle\epsilon\rangle$ and E-$\langle\epsilon u\rangle$ curves obtained from MC simulations[14] in homogeneous fields as shown in Fig. 1. The values of these parameters are summarized in Table II. For the mobility μ, the empirical expression

$$\mu=\frac{\mu_0}{\left(1+\left(\frac{\mu_0E(\langle\epsilon\rangle)}{v_{\rm sat}}\right)^\beta\right)^{1/\beta}} \tag{9}$$

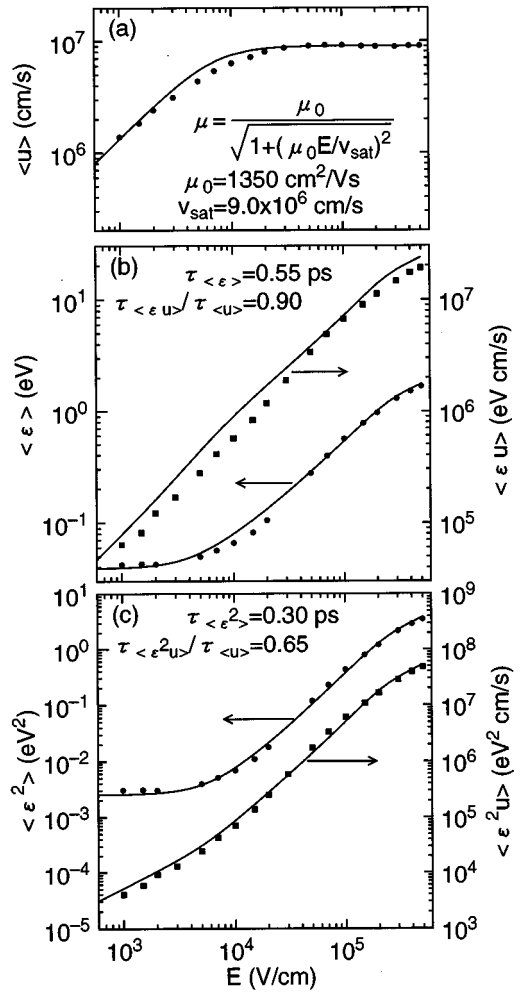

FIG. 1. Moments in a homogeneous electric field calculated using Monte Carlo (solid symbols) and energy balance equations (lines): (a) the average velocity, $\langle u\rangle$; (b) the average energy, $\langle\epsilon\rangle$, and the energy flux, $\langle\epsilon u\rangle$; and (c) the average square energy, $\langle\epsilon^2\rangle$, and the flux of square energy, $\langle\epsilon^2 u\rangle$.

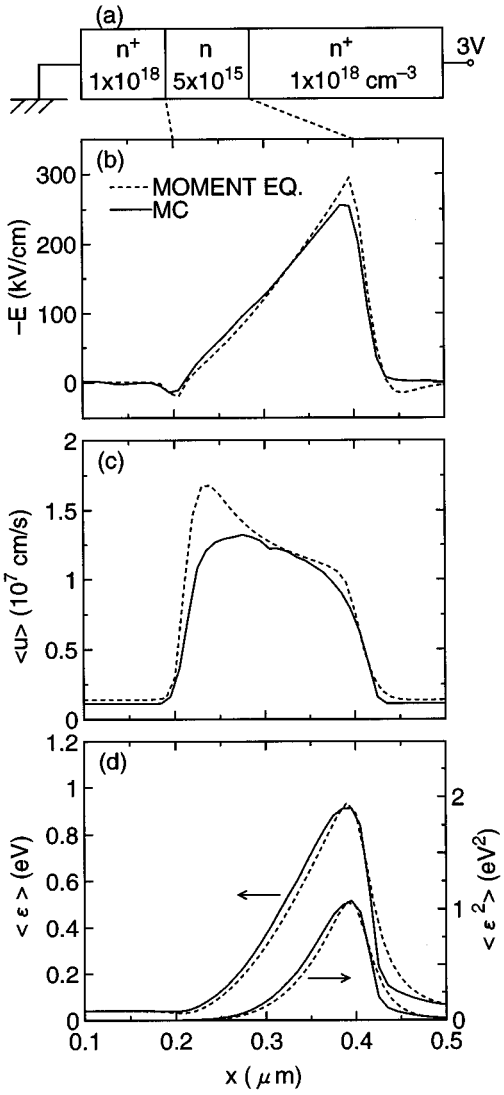

FIG. 2. (a) An n^+nn^+ structure used to determine the nonlocal transport parameters, ν_0, ν_1, and ν_2. Calculated results from MC and energy balance model: (b) the electric field, (c) the average velocity, and (d) the average energy and average square energy. Solid and dashed lines indicate the calculated results of MC and energy balance model, respectively.

is used, where $E(\langle\epsilon\rangle)$ is the electric field as a function of the average energy in a homogeneous field.

The remaining parameters ν_0 and ν_1, which relate to the nonlocal transport effect, cannot be determined uniquely because they depend on the field profile.[19] In order to derive these parameters, we employ the one-dimensional test structure as shown in Fig. 2 which has a similar field profile along the channel of the deep submicron MOSFETs. The parameters ν_0 and ν_1 are calibrated to fit the average velocity and the average energy to MC results.

TABLE II. Values of parameters used in the energy balance model.

Parameters	Values
μ_0	1350 cm^2/Vs
v_{sat}	9.0×10^6 cm/s
β	2.0
$\tau_{\langle\epsilon\rangle}$	0.55 ps
$\tau_{\langle\epsilon^2\rangle}$	0.30 ps
$\tau_{\langle\epsilon u\rangle}/\tau_{\langle u\rangle}$	0.90
$\tau_{\langle\epsilon^2 u\rangle}/\tau_{\langle u\rangle}$	0.65
ϵ_{loss}	1.5 eV
ϵ^2_{loss}	8.0 eV^2
ν_0	−0.2
ν_1	−2.0
ν_2	−1.5

The average square energy, $\langle\epsilon^2\rangle$, is calculated using the particle concentration and the average energy, which are obtained from the energy transport model. The conservation equations for $\langle\epsilon^2\rangle$ are derived from the BTE as follows:

$$\frac{\partial n\langle\epsilon^2\rangle}{\partial t} + 2q\boldsymbol{E}\cdot(n\langle\epsilon\boldsymbol{u}\rangle) + \nabla\cdot(n\langle\epsilon^2\boldsymbol{u}\rangle) = -n\,\frac{\langle\epsilon^2\rangle-\langle\epsilon^2\rangle_0}{\tau\langle\epsilon^2\rangle} + \left(\frac{\partial n\langle\epsilon^2\rangle}{\partial t}\right)_{GR}, \tag{10}$$

and

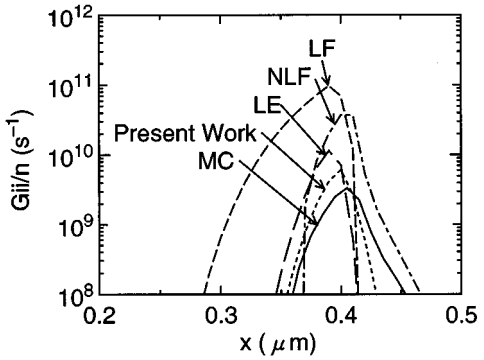

FIG. 3. Calculated impact ionization generation rates in the n^+nn^+ structure shown in Fig. 2. The parameters in the impact ionization models are set so as to give the same generation rate in a homogeneous electric field.

$$n\langle \epsilon^2 \boldsymbol{u}\rangle = \frac{\tau\langle \epsilon^2 \boldsymbol{u}\rangle}{\tau\langle \boldsymbol{u}\rangle}\frac{7}{3}\left(\langle \epsilon^2\rangle(n\langle \boldsymbol{u}\rangle) - n\mu\frac{kT_n}{q}(1+\nu_2)\nabla\langle \epsilon^2\rangle\right). \tag{11}$$

Equations (10) and (11) together with the energy transport model (5)–(8) are solved using the box integration method[20] to give the average square energy of the electrons. The relaxation rate of the average square energy due to II is formulated as $(\partial n\langle \epsilon^2\rangle/\partial t)_{\mathrm{GR}} = \epsilon_{\mathrm{loss}}^2 G_{\mathrm{ii}}$, where $\epsilon_{\mathrm{loss}}^2$ is the net loss of the average square energy by II.

The parameters in the equations, $\tau\langle \epsilon^2\rangle$, $\tau\langle \epsilon^2 \boldsymbol{u}\rangle/\tau\langle \boldsymbol{u}\rangle$, and $\epsilon_{\mathrm{loss}}^2$ are determined using Monte Carlo simulation in a homogeneous electric field, whereas the nonlocal transport parameter ν_2 is determined in such a way that the numerically calculated average square energy fits those obtained from MC simulations in the structure shown in Fig. 2. The set of physical parameters is summarized in Table II.

III. RESULTS AND DISCUSSION

II generation rates in the n^+nn^+ structure calculated with several II models described in Sec. I are shown in Fig. 3. The calculated results using the MC simulation are also shown. Parameters used in these models, which are shown in Table I, are adjusted so as to give the same generation rate in a homogeneous electric field.

Note that the LF model overestimates the generation rate by more than one order of magnitude. The NLF model yields a smaller generation rate in the up-field region and better accuracy in the down-field region than the LF model. This model, however, still shows an excessive generation rate in the up-field region. Unlike two previous models the LE model, which is often used in device simulation with an energy transport model, gives a reasonable maximum generation rate which is closer to the MC result. In the down-field region, this model gives poor agreement with the generation rate derived from MC simulation because the average energy in the region is lowered due to the existence of cold electrons diffused from the high concentration region.

In contrast to the models described above, the new II model [Eq. (2)], formulated with $\langle \epsilon\rangle$ and $\langle \epsilon^2\rangle$ well describes the II rate in both up- and down-field regions. In this model, the distribution functions at a given average energy in up- and down-field regions is differentiated by the use of the average square energy.

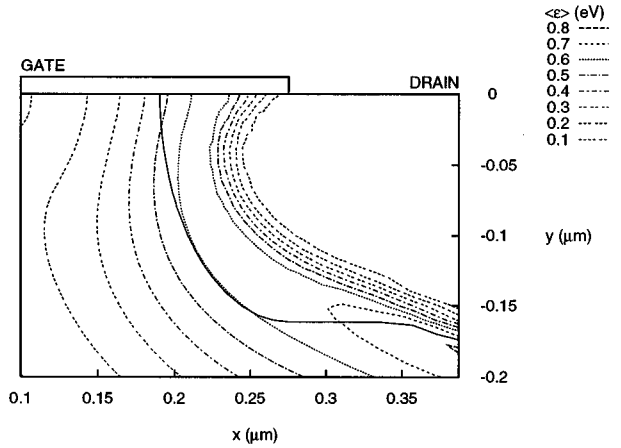

FIG. 4. The average energy of electrons at the drain junction of an n-channel MOSFET at V_d=2V and V_g=1V. The gate length is 0.55 μm ($-0.275 \leqslant x \leqslant 0.275$ μm). The Si/SiO_2 interface lies at y=0. The solid line indicates the metallurgical p-n junction.

We also calculate the II generation rate in an n-MOSFET with gate length of 0.55 μm, to study the applicability of the new method to a two-dimensional problem. Figure 4 shows the calculated average energy of electrons near the drain junction. The calculated II generation rates using the LE model and the new model are compared in Fig. 5. Although the calculated substrate currents are the same order of magnitude, the new model demonstrates that most of the II events occur in regions farther away from the Si/SiO_2 interface. In order to clarify the relation between the location of the II events and device degradation caused by carrier injection into gate oxide, further investigation will be necessary.

Comparisons with previous works are mentioned below. Concerning the numerical aspect to solve the BTE, it is not new to expand the distribution function with appropriate functions (Legendre polynomial, etc.) in order to translate the BTE into a more tractable form. The series expansion of the Legendre polynomials[21] has been extensively used to calculate II in high electric fields.[22–25] This method, however, requires the discretization in an energy- or momentum-space as well as in a real-space, which results in computational burden. On the other hand, the moment approach or a Taylor series expansion, which is used in this study, is the extension of the conventional drift-diffusion or energy transport device simulation technologies, and the discretization in the energy space is not necessary.

Several attempts have been made to improve the accuracy of high-energy population of the distribution function by solving conservation equations for the high-energy tail.[12,13] In these works, however, some ambiguity remains in defining the boundary conditions between the low- and high-energy regions. In our method, on the other hand, there is no need to draw the line between the low- and high-energy range of distribution functions.

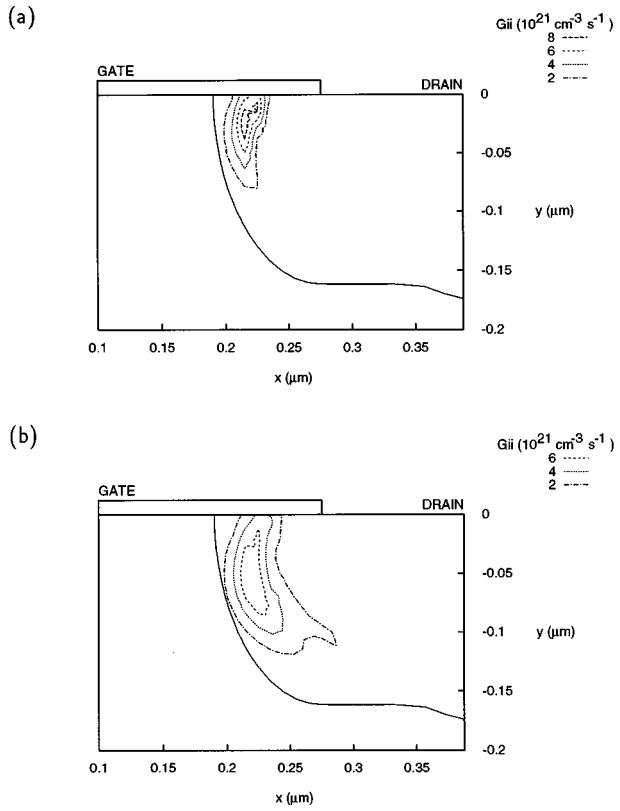

FIG. 5. Calculated electron-hole pair generation rates due to impact ionization (II) near the drain junction of the n-channel MOSFET shown in Fig. 4: (a) with the local energy (LE) II model, and (b) with the new II model (the average energy and the average square energy dependent). The solid lines indicate the metallurgical p-n junction.

We also mention the average energy of carriers obtained from energy transport equations. In the calculation of the energy transport equations reported by several authors,[7,26] the peak of the average energy shifts along the current direction because the cold electrons diffused from the high-density region are ignored. The resulting generation rate coincides with that from MC simulation. However, the ''average energy'' in this case has a different physical meaning from the conventional one.

IV. CONCLUSIONS

A method to calculate II in submicron Si devices has been developed, in which the II is formulated using the average energy and the average square energy. A set of transport equations to obtain the average square energy has been proposed in the framework of an energy transport model. Parameters used in the equations were extracted in such a way that the calculated results based on these equations agree well with Monte Carlo simulation results in both homogeneous and inhomogeneous fields. The new method was applied to investigate the II generation rate in a submicron n-MOSFET. The calculated results show that II events occur in deeper channel regions than the LE model's prediction.

ACKNOWLEDGMENT

One of the authors (K.S.) wishes to thank Dr. N. Kotani of ULSI Laboratory, Mitsubishi Electric Corporation for his encouragement and constant support.

[1] A. G. Chynoweth, Phys. Rev. **109**, 1537 (1958).
[2] T. Thurgate and N. Chan, IEEE Trans. Electron Devices **ED-32**, 400 (1985).
[3] R. Kuhnert, C. Werner, and A. Schütz, IEEE Trans. Electron Devices **ED-32**, 1057 (1985).
[4] H. J. Peifer, B. Meinerzhagen, R. Thoma, and W. L. Engl, IEDM Tech. Dig. 131 (1991).
[5] M. Fukuma, and W. W. Lui, IEEE Electron Device Lett. **EDL-8**, 214 (1987).
[6] Y. Apanovich, E. Lyumkis, B. Polsky, A. Shur, and P. Blakey, IEEE Trans. Comput.-Aided Des. **13**, 702 (1994).
[7] V. M. Agostinelli, Jr., T. J. Bordelon, X. Wang, K. Hasnat, C.-F. Yeap, D. B. Lemersal, Jr., A. F. Tasch, and C. M. Maziar, IEEE Trans. Electron Devices **41**, 1784 (1994).
[8] K. Rahmat, J. White, and D. A. Antoniadis, IEDM Tech. Dig. 115 (1991).
[9] W. Quade, E. Schöll, and M. Rudan, Solid-State Electron. **36**, 1493 (1993).
[10] J. Z. Peng, S. Longcor, and J. Frey, IEICE Trans. Electron. **E77-C**, 166 (1994).
[11] T. Iizuka and H. Kato, ''Transient Impact Ionization in Bulk Si,'' 1993 International Workshop on Numerical Modeling of Processes an Devices for Integrated Circuits (NUPAD V), 1994, p. 59.
[12] J.-G. Ahn, C.-S. Yao, Y.-J. Park, H.-S. Min, and R. W. Dutton, IEEE Electron Device Lett. **15**, 348 (1994).
[13] P. Scrobohaci and T.-W. Tang, IEICE Trans. Electron. **E77-C**, 134 (1994).
[14] K. Sonoda, S. T. Dunham, M. Yamaji, K. Taniguchi, and C. Hamaguchi, Jpn. J. Appl. Phys. 1 **35**, 225 (1996).
[15] R. Stratton, Phys. Rev. **126**, 2002 (1962).
[16] K. Bløtekjær, IEEE Trans. Electron Devices **ED-17**, 38 (1970).
[17] T.-W. Tang, S. Ramaswamy, and J. Nam, IEEE Trans. Electron Devices **40**, 1469 (1993).
[18] M. C. Vecchi and L. G. Reyna, Solid-State Electron. **37**, 1705 (1994).
[19] D. Chen, E. Sangiorgi, M. R. Pinto, E. C. Kan, U. Ravaioli, and R. W. Dutton, ''Analysis of Spurious Velocity Overshoot in Hydrodynamic Simulations,'' 1992 International Workshop on Numerical Modeling of Processes and Devices for Integrated Circuits (NUPAD IV), 1992, p. 109.
[20] A. Forghieri, R. Guerrieri, P. Ciampolini, A. Gnudi, M. Rudan, and G. Baccarani, IEEE Trans. Comput.-Aided Des. **7**, 231 (1988).
[21] G. H. Wannier, Bell Syst. Tech. J., 170 (1953).
[22] P. A. Wolff, Phys. Rev. **95**, 1415 (1954).
[23] N. Goldsman, Y.-J. Wu, and J. Frey, J. Appl. Phys. **68**, 1075 (1990).
[24] S.-L. Wang, N. Goldsman, and K. Hennacy, J. Appl. Phys. **71**, 1815 (1992).
[25] S. M. Cho and H. H. Lee, J. Appl. Phys. **71**, 1298 (1992).
[26] R. K. Cook and J. Frey, IEEE Trans. Electron Devices **ED-29**, 970 (1982).