A Zero Field Monte Carlo Algorithm Accounting for the Pauli Exclusion Principle

Sergey Smirnov, Hans Kosina, Mihail Nedjalkov, and Siegfried Selberherr

Institute for Microelectronics, TU-Vienna, Gusshausstrasse 27-29/E360, A-1040 Wien, Austria, smirnov@iue.tuwien.ac.at

Abstract. A Monte Carlo method for calculation of the carrier mobility in degenerate bulk semiconductors at zero electric field is presented. The method is obtained as a limiting case of an existing small-signal approach replacing the distribution function by the Fermi-Dirac distribution which is valid at zero electric field. The general form of the Boltzmann equation which takes into account the Pauli exclusion principle in the scattering term is used to derive the integral representation of a Boltzmann-like equation for a small perturbation of the distribution function. The method allows calculation of the whole mobility tensor in comparison with the one particle Monte Carlo algorithm which is traditionally used to compute low field carrier mobility.

1 Introduction

The low field carrier mobility of a bulk semiconductor is an important kinetic property of a semiconductor. It is used to analyze the carrier transport in semiconductor devices at low applied voltages and enters expressions for high field mobility models as an additional parameter. Thus the knowledge of the low field carrier mobility and its correct dependence on the material properties such as the doping concentration are necessary to adequately simulate carrier transport in semiconductor devices.

The standard approach for obtaining the low field carrier mobility is a single particle Monte Carlo method. Within this method in order to calculate the low field mobility along the direction of the electric field one should carefully choose the magnitude of the applied electric field. On the one hand side, the magnitude of the electric field must be as low as possible. In principle it is desirable to have zero electric field. However, there exist limitations related to the increase of the variance of standard Monte Carlo methods. On the other hand side, the field must not be too high to avoid a mobility reduction due to carrier heating.

In addition to these disadvantages, the standard approaches only give one component of the carrier mobility namely the component along the direction of the electric field. For isotropic conditions it does not make any difference since the mobility tensor is diagonal and all diagonal values are equal. However when anisotropy is present, for example in strained semiconductors, the mobility tensor elements may be different and several Monte Carlo simulations are required to obtain all the components of the tensor.

To overcome these problems associated with standard Monte Carlo methods a new Monte Carlo algorithm has been suggested recently [1], which solves the Boltzmann equation at exact zero field and represents a limiting case of a small signal algorithm obtained in [2]. One of the most remarkable properties of the algorithm is the absence of self-scattering that allows to significantly reduce calculation time and achieve very good accuracy of the results. This method is restricted to the simulation of low doped semiconductors. The quantum mechanical Pauli exclusion principle is not included in the scattering term of the Boltzmann equation used for the derivation of the algorithm. As a result there are limitations on the doping level of materials analyzed by this technique. It allows to obtain excellent results at low and intermediate doping levels while results obtained for higher doping levels where the effects of degenerate statistics are more pronounced are incorrect. As the standard Monte Carlo methods exhibit a very high variance especially in degenerate case, it is thus desirable to have a powerful technique to analyze the carrier mobility at high doping levels.

In this work we present a zero field algorithm to account for degenerate statistics. The Pauli exclusion principle is taken into consideration in the scattering term of the Boltzmann equation. As a result the Boltzmann equation becomes nonlinear. Using this nonlinear equation we derive a generalized zero field algorithm applicable for the analysis of highly doped materials.

2 Nonlinear Boltzmann Equation and its Linearized Form

Let us consider a bulk homogenous semiconductor. Then we can neglect the space dependence of the distribution function and the differential scattering rate. We also assume the differential scattering rate to be time invariant. With these conditions the time dependent Boltzmann equation taking into account the Pauli exclusion principle takes the following form:

$$\frac{\partial f(\mathbf{k}, t)}{\partial t} + \frac{q\mathbf{E}(t)}{\hbar} \nabla f(\mathbf{k}, t) = Q[f](\mathbf{k}, t), \tag{1}$$

where E(t) is an electric field and q is the particle charge. Q[f](k,t) represents the scattering operator which is given by the following expression:

$$Q[f](\mathbf{k},t) = \int f(\mathbf{k}',t)[1 - f(\mathbf{k},t)]S(\mathbf{k}',\mathbf{k}) d\mathbf{k}' - \int f(\mathbf{k},t)[1 - f(\mathbf{k}',t)]S(\mathbf{k},\mathbf{k}') d\mathbf{k}',$$
(2)

where $S(\mathbf{k}', \mathbf{k})$ stands for the differential scattering rate. Thus $S(\mathbf{k}', \mathbf{k})d\mathbf{k}$ is the scattering rate from a state with wave vector \mathbf{k}' to states in $d\mathbf{k}$ around \mathbf{k} , $f(\mathbf{k}, t)$ is the distribution function and the factors $[1 - f(\mathbf{k}, t)]$ mean that the final state must not be occupied according to the Pauli exclusion principle. As can be seen from (2), there are terms $f(\mathbf{k}, t)f(\mathbf{k}', t)$ which render the equation nonlinear.

Only when the condition $f(\mathbf{k},t) \ll 1$ is valid the factors $[1 - f(\mathbf{k},t)]$ can be replaced by unity and the equation takes the usual linear form.

To linearize (1) we write the electric field in the form:

$$\boldsymbol{E}(t) = \boldsymbol{E}_s + \boldsymbol{E}_1(t),\tag{3}$$

where E_s stands for a stationary field and $E_1(t)$ denotes a small perturbation which is superimposed on a stationary field. It is assumed that this small perturbation of the electric field causes a small perturbation of the distribution function which can be written as follows:

$$f(\mathbf{k},t) = f_s(\mathbf{k}) + f_1(\mathbf{k},t), \tag{4}$$

where $f_s(\mathbf{k})$ is a stationary distribution function and $f_1(\mathbf{k}, t)$ is a small deviation from a stationary distribution. Substituting (4) into (2) the scattering operator $Q[f](\mathbf{k}, t)$ takes the form:

$$Q[f](\mathbf{k},t) = \int (f_s(\mathbf{k}') + f_1(\mathbf{k}',t))[1 - f_s(\mathbf{k}) - f_1(\mathbf{k},t)]S(\mathbf{k}',\mathbf{k}) d\mathbf{k}' - \int (f_s(\mathbf{k}) + f_1(\mathbf{k},t))[1 - f_s(\mathbf{k}') - f_1(\mathbf{k}',t)]S(\mathbf{k},\mathbf{k}') d\mathbf{k}'.$$

$$(5)$$

It should be noted that in spite of the fact that $f_1(\mathbf{k},t) \ll 1$ one should take care when linearizing terms such as $1 - f_s(\mathbf{k}) - f_1(\mathbf{k},t)$ as especially in the degenerate case it may happen that the inequality $1 - f_s(\mathbf{k}) \ll f_1(\mathbf{k},t)$ becomes valid because of $[1 - f_s(\mathbf{k})] \to 0$.

Neglecting terms of second order we derive the zeroth order equation:

$$\frac{q}{\hbar} \mathbf{E}_{s} \nabla f_{s}(\mathbf{k}) = [1 - f_{s}(\mathbf{k})] \int f_{s}(\mathbf{k}') S(\mathbf{k}', \mathbf{k}) d\mathbf{k}' -
-f_{s}(\mathbf{k}) \int [1 - f_{s}(\mathbf{k}')] S(\mathbf{k}, \mathbf{k}') d\mathbf{k}',$$
(6)

and the first order equation:

$$\frac{\partial f_1(\mathbf{k},t)}{\partial t} + \frac{q}{\hbar} \mathbf{E}_s \nabla f_1(\mathbf{k},t) = -\frac{q}{\hbar} \mathbf{E}_1(t) \nabla f_s(\mathbf{k}) + Q^{(1)}[f](\mathbf{k},t), \tag{7}$$

where we introduced the notation $Q^{(1)}[f](\mathbf{k},t)$ for the first order scattering operator which has the form:

$$Q^{(1)}[f](\mathbf{k},t) = [1 - f_s(\mathbf{k})] \int f_1(\mathbf{k}',t) S(\mathbf{k}',\mathbf{k}) d\mathbf{k}' - f_1(\mathbf{k},t) \int [1 - f_s(\mathbf{k}')] S(\mathbf{k},\mathbf{k}') d\mathbf{k}' - f_1(\mathbf{k},t) \int f_s(\mathbf{k}') S(\mathbf{k}',\mathbf{k}) d\mathbf{k}' + f_s(\mathbf{k}) \int f_1(\mathbf{k}',t) S(\mathbf{k},\mathbf{k}') d\mathbf{k}'.$$

$$(8)$$

Equation (7) is linear with respect to $f_1(\mathbf{k},t)$, it is a kinetic equation which differs from the usual form of the Boltzmann equation. The first difference is related to the presence of the additional term on the right hand side being the term proportional to \mathbf{E}_1 which additionally depends on the stationary distribution. The second difference is in the expression for the scattering operator which now has a more complex form and also depends on the stationary distribution.

3 Integral Form of the First Order Equation

To construct a new Monte Carlo algorithm we reformulate the Boltzmann equation of the first order into integral form. For this purpose we introduce a new differential scattering rate and new total scattering rate defined by the following expressions:

$$\widetilde{S}(\mathbf{k}', \mathbf{k}) = [1 - f_s(\mathbf{k})]S(\mathbf{k}', \mathbf{k}) + f_s(\mathbf{k})S(\mathbf{k}, \mathbf{k}'), \tag{9}$$

$$\widetilde{\lambda}(\mathbf{k}) = \int ([1 - f_s(\mathbf{k})]S(\mathbf{k}, \mathbf{k}') + f_s(\mathbf{k}')S(\mathbf{k}', \mathbf{k})) d\mathbf{k}' = \int \widetilde{S}(\mathbf{k}, \mathbf{k}') d\mathbf{k}'. \quad (10)$$

It is worth noting that the similarity with the standard Boltzmann equation is only formal as both, differential scattering rate and total scattering rate, are now functionals of the stationary distribution function which is the solution of the equation of zero order (6).

With these definitions the scattering operator of the first order $Q^{(1)}[f](\mathbf{k},t)$ formally takes the conventional form:

$$Q^{(1)}[f](\boldsymbol{k},t) = \int f_1(\boldsymbol{k}',t)\widetilde{S}(\boldsymbol{k}',\boldsymbol{k}) d\boldsymbol{k}' - f_1(\boldsymbol{k},t)\widetilde{\lambda}(\boldsymbol{k}), \qquad (11)$$

and the Boltzmann-like equation can be rewritten as follows:

$$\frac{\partial f_{1}(\boldsymbol{k},t)}{\partial t} + \frac{q}{\hbar} \boldsymbol{E}_{s} \nabla f_{1}(\boldsymbol{k},t) = \int f_{1}(\boldsymbol{k}',t) \widetilde{S}(\boldsymbol{k}',\boldsymbol{k}) d\boldsymbol{k}' - f_{1}(\boldsymbol{k},t) \widetilde{\lambda}(\boldsymbol{k}) - \frac{q}{\hbar} \boldsymbol{E}_{1}(t) \nabla f_{s}(\boldsymbol{k}). \tag{12}$$

We derive the integral form of this equation using techniques described in [3]. Introducing a phase space trajectory $\mathbf{K}(t') = \mathbf{k} - \frac{q}{\hbar} \mathbf{E}_s(t - t')$ which is the solution of Newton's equation, and taking into account that $f_1(\mathbf{K}(t_0), t_0) = 0$ for $t_0 < 0$ we obtain the following integral form:

$$f_{1}(\mathbf{K}(t),t) =$$

$$= \int_{0}^{t} dt' \int d\mathbf{k}' f_{1}(\mathbf{k}',t) \widetilde{S}(\mathbf{k}',\mathbf{K}(t')) \cdot \exp\left(-\int_{t'}^{t} \widetilde{\lambda}[\mathbf{K}(y)] dy\right) -$$

$$-\frac{q}{\hbar} \int_{0}^{t} \mathbf{E}_{1}(t') [\nabla f_{s}](\mathbf{K}(t')) \cdot \exp\left(-\int_{t'}^{t} \widetilde{\lambda}[\mathbf{K}(y)] dy\right) dt'.$$
(13)

Finally we assume an impulse like excitation of the electric field, $E_1(t) = \delta(t)E_{im}$, and obtain:

$$f_{1}(\mathbf{K}(t),t) =$$

$$= \int_{0}^{t} dt' \int d\mathbf{k}' f_{1}(\mathbf{k}',t) \widetilde{S}(\mathbf{k}',\mathbf{K}(t')) \cdot \exp\left(-\int_{t'}^{t} \widetilde{\lambda}[\mathbf{K}(y)] dy\right) +$$

$$+ G(\mathbf{K}(0)) \exp\left(-\int_{0}^{t} \widetilde{\lambda}[\mathbf{K}(y)] dy\right), \tag{14}$$

where

$$G(\mathbf{k}) = -\frac{q}{\hbar} \mathbf{E}_{im} \nabla f_s(\mathbf{k}). \tag{15}$$

The essential difference of this integral representation from the one of the non-degenerate approach consists in the appearance of the new differential scattering rate $\widetilde{S}(\boldsymbol{k}',\boldsymbol{k})$ and of the total scattering rate $\widetilde{\lambda}(\boldsymbol{k})$. Another difference which is common for both approaches is the additional free term on the right hand side which in general cannot be treated as an initial distribution because it also takes negative values.

4 Zero Field Approach

When the electric field tends to zero, the distribution function approaches its equilibrium which is in case of particles with fractional spin represented by the Fermi-Dirac distribution function:

$$f_{FD}(\epsilon) = \frac{1}{\exp\left[-\frac{E_f - \epsilon}{k_B T_0} + 1\right]},\tag{16}$$

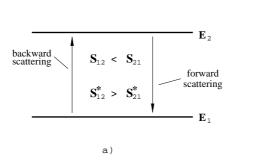
where E_f denotes the Fermi energy, ϵ stands for an electron energy and T_0 is an equilibrium temperature equal to the lattice temperature. Since the stationary distribution is known, there is no necessity to solve the zeroth order equation (6). As can be seen from (16), in equilibrium the distribution function depends only on the carrier energy and not on the wave vector. This fact allows one to significantly simplify (10) using the Fermi golden rule [4]:

$$S(\mathbf{k}, \mathbf{k}') = \frac{V}{2\pi^{2}\hbar} |V_{fi}|^{2} \delta[\epsilon(\mathbf{k}') - \epsilon(\mathbf{k}) \pm \Delta\epsilon]. \tag{17}$$

Making use of the delta function in the last expression we can rewrite (10) in the following manner:

$$\widetilde{\lambda}(\mathbf{k}) = [1 - f_{FD}(\epsilon_f)]\lambda(\mathbf{k}) + f_{FD}(\epsilon_f)\lambda^*(\mathbf{k}), \tag{18}$$

where ϵ_f denotes the final carrier energy and we introduced the backward scattering rate $S^*(\mathbf{k}, \mathbf{k}') = S(\mathbf{k}', \mathbf{k})$ and $\lambda^*(\mathbf{k}) = \int S^*(\mathbf{k}, \mathbf{k}') d\mathbf{k}'$. (18) represents a linear combination of the forward and backward total scattering rates. In the



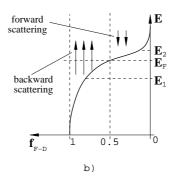


Fig. 1. Schematic illustration of the scattering processes at high degeneracy.

non-degenerate case, $f_{F-D}(\epsilon) \ll 1$, we obtain $\lambda(\mathbf{k}) = \lambda(\mathbf{k})$ that means that scattering processes are mostly determined by the forward scattering rate and thus the algorithm developed in [1] for nondegenerate statistics is restored. On the other hand side for highly doped semiconductors, $f_{F-D}(\epsilon) \sim 1$, scattering processes are dominantly backward $\lambda(\mathbf{k}) = \lambda^*(\mathbf{k})$. In case of intermediate doping levels both forward and backward scattering contribute to the kinetics. It should be also mentioned that as at high doping levels the backward scattering rate is dominant, the probability to scatter to higher energy states is larger than to lower energy states as schematically shown in Fig.1(a). This means that lower energy levels are already occupied by particles i.e. $f_{F-D}(\epsilon) \approx 1$ (see Fig.1(b)) and, due to the Pauli exclusion principle, scattering to these energy levels is quantum mechanically forbidden.

The additional free term in (14) cannot be considered as an initial distribution because function $G(\mathbf{k})$ may take negative values. However, in case of zero electric field the stationary distribution is known analytically and we can calculate $G(\mathbf{k})$ explicitly:

$$G(\mathbf{k}) = \frac{q}{k_B T_0} \mathbf{E}_{im} \cdot \mathbf{v} \frac{\exp\left[-\frac{E_f - \epsilon}{k_B T_0}\right]}{\left(\exp\left[-\frac{E_f - \epsilon}{k_B T_0}\right] + 1\right)^2},$$
(19)

where v denotes the group velocity. This expression can be rewritten in the following manner:

$$G(\mathbf{k}) = \frac{q\mathbf{E}_{im}\langle \widetilde{\lambda} \rangle}{k_B T_0} \frac{\mathbf{v}(\mathbf{k})[1 - f_{FD}(\epsilon)]}{\widetilde{\lambda}(\mathbf{k})} \left\{ \frac{\widetilde{\lambda}(\mathbf{k}) f_{FD}(\epsilon)}{\langle \widetilde{\lambda} \rangle} \right\}$$
(20)

where the term in curly brackets represents the normalized distribution function of the before-scattering states. The Monte Carlo algorithm contains the same steps as those in [1] except that the whole kinetics must be now considered in terms of $\widetilde{S}(\boldsymbol{k}, \boldsymbol{k}')$ and $\widetilde{\lambda}(\boldsymbol{k})$ instead of $S(\boldsymbol{k}, \boldsymbol{k}')$ and $\lambda(\boldsymbol{k})$. It may be seen by derivation of the second iteration term of the Neumann series for (14) using the forward formulation to obtain the ensemble Monte Carlo algorithm:

$$(f_{1}^{(2)})_{\Omega} = \int_{0}^{t} dt_{2} \int_{t_{2}}^{t} dt_{1} \int d\mathbf{k}_{2}^{a} \int d\mathbf{k}_{1}^{a} \int d\mathbf{k}_{1} \{G(\mathbf{k}_{i})\} \cdot \left\{ \exp\left(-\int_{0}^{t_{2}} \widetilde{\lambda}[\mathbf{K}_{2}(y) dy]\right) \widetilde{\lambda}[\mathbf{K}_{2}(t_{2})] \right\} \left\{ \frac{\widetilde{S}[\mathbf{K}_{2}(t_{2}), \mathbf{k}_{2}^{a}]}{\widetilde{\lambda}[\mathbf{K}_{2}(t_{2})]} \right\} \cdot \left\{ \exp\left(-\int_{t_{2}}^{t_{1}} \widetilde{\lambda}[\mathbf{K}_{1}(y) dy]\right) \widetilde{\lambda}[\mathbf{K}_{1}(t_{1})] \right\} \left\{ \frac{\widetilde{S}[\mathbf{K}_{1}(t_{1}), \mathbf{k}_{1}^{a}]}{\widetilde{\lambda}[\mathbf{K}_{1}(t_{1})]} \right\} \cdot \exp\left(-\int_{t_{1}}^{t_{1}} \widetilde{\lambda}[\mathbf{K}(y)] dy\right) \Theta_{\Omega}(\mathbf{K}(t)),$$

$$(21)$$

where k^a stands for an after-scattering wave vector and k_i denotes an initial wave vector. The quantity $\frac{\widetilde{S}[k,k']}{\widetilde{\lambda}[k]}$ represents a normalized after-scattering distribution. As can be seen from (9) and (10) it is normalized to unity. As it follows from (21), during Monte Carlo simulation a particle trajectory is constructed in terms of new quantities \widetilde{S} and $\widetilde{\lambda}$.

Another difference from the non-degenerate zero field algorithm is that the weight coefficient $\frac{v(k)}{\tilde{\lambda}(k)}$ must be multiplied by the factor $[1 - f_{FD}(\epsilon)]$.

With these modifications the steps of the algorithm are as follows:

- 1. Set $\nu = 0$, w = 0.
- 2. Select initial state k arbitrarily.
- 3. Compute a sum of weights: $w = w + [1 f_{FD}(\epsilon)][v_j(\mathbf{k})/\widetilde{\lambda}(\mathbf{k})].$
- 4. Select a free-flight time $\tilde{t}_f = -\ln(r)/\tilde{\lambda}(\mathbf{k})$ and add time integral to estimator: $\nu = \nu + wv_i\tilde{t}_f$ or use the expected value of the time integral: $\nu = \nu + w[v_i/\tilde{\lambda}(\mathbf{k})]$.
- 5. Perform scattering. If mechanism was isotropic, reset weight: w = 0.
- 6. Continue with step 3 until N k-points have been generated.
- 7. Calculate component of zero field mobility tensor as $\mu_{ij} = q\langle \widetilde{\lambda} \rangle \nu / (k_B T_0 N)$.

5 Results and Discussion

As the first example we calculate the doping dependence of the zero field mobility in silicon. The analytical band structure reported in [5] is adopted. The scattering processes included are acoustic deformation potential scattering, ionized impurity scattering [6] and plasmon scattering [7]. Ionized impurity scattering is treated as an isotropic process [8] which effectively reduces small-angle scattering.

Fig.2 shows two electron mobility curves obtained by the new zero field algorithm and the one particle Monte Carlo algorithm, respectively. The Pauli exclusion principle at low electric field has been included in the one particle Monte Carlo method using the Fermi-Dirac distribution. This leads to a prefactor $(1 - f_{FD}(\epsilon_f))$ for all scattering processes within the low field approach. The value of the electric field used for the one particle approach has been chosen to

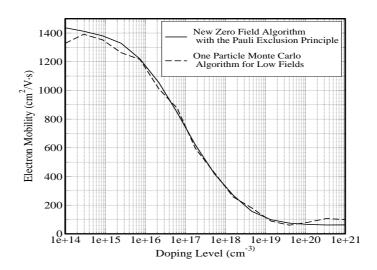


Fig. 2. Electron zero field and low field mobilities in Si.

be 1 kV/cm. The other physical parameters of both algorithms are the same. As can be seen from Figure 2, the curve obtained with the one particle Monte Carlo method has some variance at low and high doping levels while the zero field curve appears rather smooth. Moreover the calculation time for the new zero field algorithm is about 20% of the one particle Monte Carlo method.

Fig.3 shows the electron mobility as a function of doping concentration for GaAs obtained by both algorithms. The degeneracy effects are more pronounced in this semiconductor because of the smaller effective mass of electrons in the Γ valley. The absolute value of the electric field is the same as for silicon. As is seen from this figure, in addition to a high variance at low doping levels the one particle Monte Carlo gives an incorrect behavior of the mobility at high doping levels. This is related to the fact that the value 1 kV/cm of the electric field is still high from the viewpoint of using the Fermi-Dirac distribution with lattice temperature within the one particle algorithm. In order to obtain correct results at high doping levels by the one particle Monte Carlo method it is necessary to reduce further the magnitude of the electric field. However in this case the variance would increase considerably leading to an extremely long computation time.

6 Conclusion

A zero field Monte Carlo algorithm accounting for the quantum mechanical Pauli exclusion principle has been presented. The method has been derived from the integral representation of a linearized Boltzmann-like equation. It has been shown that particle trajectories are constructed in terms of a new scattering rate which in general represents a linear combination of the forward and backward scattering rates. It has been also pointed out that for energies below the Fermi

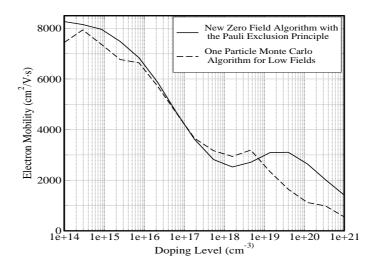


Fig. 3. Electron zero field and low field mobilities in GaAs.

level kinetic properties are predominantly determined by the backward scattering rate while for energy levels above the Fermi level the forward scattering rate is dominant. In the latter case the non-degenerate zero field algorithm is recovered.

References

- 1. H. Kosina, M. Nedjalkov, and S. Selberherr. Monte Carlo Analysis of the Small-Signal Response of Charge Carriers, Proc. 3rd Intl. Conf. on Large-Scale Scientific Computations, Sozopol, 175-182, 2001.
- 2. H. Kosina, M. Nedjalkov, and S. Selberherr. A Monte Carlo Method for Small Signal Analysis of the Boltzmann equation, *J. Appl. Phys.*, 87(9), 4308-4314, 2000.
- 3. H. Kosina, M. Nedjalkov, and S. Selberherr. Theory of the Monte Carlo Method for Semiconductor Device Simulation, *IEEE Trans. Electron Devices*, 47(10), 1898-1908, 2000.
- C. Jacoboni and P. Lugli. The Monte Carlo Method for Semiconductor Device Simulation, Wien-New York, Springer, 1989.
- C. Jacoboni and L. Reggiani. The Monte Carlo Method for the Solution of Charge Transport in Semiconductors with Applications to Covalent Materials, Rev. Mod. Phys., 55(3), 645-705, 1983.
- H. Kosina and G. Kaiblinger-Grujin. Ionized-Impurity Scattering of Majority Electrons in Silicon, Solid-State Electron., 42(3), 331-338, 1998.
- M. V. Fischetti. Effect of the electron-plasmon interaction on the electron mobility in silicon, *Phys.Rev.B*, 44(11), 5527-5534, 1991.
- 8. H. Kosina. A Method to Reduce Small-Angle Scattering in Monte Carlo Device Analysis, *IEEE Trans. Electron Devices*, 46(6), 1196-1200, 1999.