Variance Reduction in Monte Carlo Device Simulation by Means of Event Biasing

H. Kosina, M. Nedjalkov, and S. Selberherr

Institute for Microelectronics, TU Wien, Gusshausstr. 27–29
A-1040 Vienna, Austria, Hans.Kosina@iue.tuwien.ac.at

ABSTRACT

A theoretical analysis of the Monte Carlo (MC) method for the solution of the stationary boundary value problem defined by the Boltzmann equation is briefly outlined. This analysis clearly shows how event biasing can be used within the well-known One-Particle MC method. To enhance statistics in sparsely populated regions of interest, artificial carrier heating is introduced by increasing the probability for phonon absorption at the expense of phonon emission. The distribution of the scattering angle is biased to induce artificial carrier diffusion. Having identified the random variable whose realizations are statistically independent new variance estimates are proposed.

Keywords: event biasing, variance reduction, Monte Carlo method, Boltzmann equation, device simulation.

1 INTRODUCTION

In the MC simulation of semiconductor devices measures have to be taken to enhance the statistics in interesting phase space regions that are sparsely populated. The purpose of statistical enhancement is to reduce the variance of the MC estimates in those regions. Such a reduction comes at the expense of increased variance in other, less interesting regions [1]. There are two classes of statistical enhancement techniques, namely population control techniques and event biasing techniques. To date virtually all MC device simulation codes utilize population control techniques. On the other hand, the event biasing technique has been introduced in the field of semiconductor transport only one decade ago, when the *Weighted Ensemble* MC method has been proposed [2][3]. However, this method has not been used in MC device simulation yet. The present work focuses on the steady state, and consequently a *Weighted One-Particle* MC method is presented. It is applied to npn-structures and the suitability for variance reduction is studied.

2 THE MC METHOD

The kinetic equation considered is the stationary Boltzmann equation (BE), supplemented by boundary conditions. MC algorithms for the solution of the stationary boundary value problem can be constructed formally as outlined in the following.

2.1 Theory

In a first step the stationary BE is transformed into an integral equation of the second kind,

$$f(x) = \int f(x')K(x',x)\,\mathrm{d}x' + f_0(x) \qquad (1)$$

where the six-dimensional variable x stands for $(\mathbf{k},\mathbf{r})$. The kernel K describes the propagation of the particles, while the free term f_0 contains the boundary condition. Since the integral form of the BE is a backward equation, the corresponding forward equation, that is the conjugate equation, needs to be derived. Its kernel is given by $K^\dagger(x,x') = K(x',x)$.

$$g(x') = \int g(x)K^\dagger(x,x')\,\mathrm{d}x + A(x') \qquad (2)$$

The two integral equations are stated explicitly in [4]. With a forward method only mean values of some arbitrary quantity $A(x)$ can be evaluated due to the following equality.

$$\int A(x)f(x)\,\mathrm{d}x = \int f_0(x)g(x)\,\mathrm{d}x \qquad (3)$$

Note that usage of this equation precludes a point-wise evaluation of the distribution function f, because the choice $A(x) = \delta(x)$ can not be treated by the MC method.

Substituting the conjugate equation recursively into the right hand side of (3) yields an iteration series, the elements of which are finally evaluated by means of MC integration. Using this procedure, the One-Particle MC algorithm is obtained in a formal way. In particular, recovered are the well-known probability densities for trajectory construction, and both the *time averaging* and the *before-scattering* methods for mean value calculation.

In addition, the random variable can be identified whose realizations are statistically independent. A realization of this random variable, say X, is a complete numerical trajectory that starts and terminates at the domain boundary. The i-th realization consists of all generated random variables for the i-th trajectory, such as the initial state at the domain boundary, $\mathbf{k}_0, \mathbf{r}_0$, the free flight times t_j, and the after-scattering states, $\mathbf{k}_j^a$,

$$x_i = \{\mathbf{k}_0, \mathbf{r}_0, t_0, \mathbf{k}_1^a, t_1, \ldots \mathbf{k}_j^a, t_j, \ldots\} \qquad j \le N_i \quad (4)$$

where $N_i + 1$ is the number of free flight segments for the considered trajectory. Another random variable $Y(X)$ needed

below is defined by its realizations

$$y_i = \{\mathbf{k}_1^b, \mathbf{r}_1, \ldots \mathbf{k}_j^b, \mathbf{r}_j, \ldots\} \qquad j \le N_i, \tag{5}$$

which contain all before-scattering states $\mathbf{k}_b^j$ and the particle locations at the times of scattering, $\mathbf{r}_j$. Knowledge of the independent random variables forms the basis for determination of the variance of the MC estimators and thus for stating error estimates.

2.2 Event Biasing

The kernel of the conjugate equation $K^\dagger$ yields the natural probability distributions, which are then used for the construction of the particle trajectory. However, it is possible to choose other than the natural probabilities for the MC integration of the terms of the iteration series. In that case one constructs numerical trajectories that are different from the physical ones.

With any quantity of interest, $A(\mathbf{k}, \mathbf{r})$, a random variable $\Psi_A(Y)$ is associated. Using the common One-Particle MC algorithm and the before-scattering method, a realization of Ψ_A is of the form

$$\psi_{Ai} = \sum_{j=1}^{N_i} \frac{A(\mathbf{k}_j^b, \mathbf{r}_j)}{\lambda(\mathbf{k}_j^b, \mathbf{r}_j)} \tag{6}$$

with λ being the total scattering rate. Summation is again over one complete trajectory starting and terminating at the domain boundary.

Changing probability distributions requires compensatory changes of random variables. The formal treatment can be summarized in a simple rule. Whenever in the course of numerical trajectory construction a random variable, for example, a free flight time or an after scattering state, is selected from a numerical density rather than from a physical density, the weight of the test particle changes by the ratio of the physical over the numerical density, evaluated at the selected value. Introducing the weight of the particle w_j constructed by this rule, the realization of Ψ_A becomes

$$\psi_{Ai} = \sum_{j=1}^{N_i} w_j \frac{A(\mathbf{k}_j^b, \mathbf{r}_j)}{\lambda(\mathbf{k}_j^b, \mathbf{r}_j)} \tag{7}$$

The weight of each injected particle is set to $w_1 = 1$, the subsequent weights evolve randomly. This gives rise to increasing variance in the MC estimates with increasing time cumulated by a trajectory. However, since the particle weight is reset whenever a particle leaves and reenters the domain, the variance stays bounded.

The motivation for using arbitrary probabilities is the possibility to guide particles towards a region of interest. We increase carrier diffusion against a retarding field by introducing artificial carrier heating. Controlled by a parameter $M_1 \ge 1$, the probability for phonon absorption is increased at the expense of phonon emission,

$$\lambda_a' = \lambda_a + \lambda_e \left(1 - \frac{1}{M_1}\right), \quad \lambda_e' = \frac{\lambda_e}{M_1}. \tag{8}$$

If in the MC simulation phonon absorption is selected, the particle weight is to be multiplied by λ_a/λ_a', otherwise by $\lambda_e/\lambda_e' = M_1$. The distribution of the flight time is not affected, because the sum of emission and absorption rate stays unchanged.

Carrier diffusion can be enhanced by modifying the distribution of the scattering angle. Only isotropic processes are affected. For these the distribution of $\chi = \cos\theta$ is constant: $p(\chi) = 1/2$ for $\chi \in (-1, 1)$. Here θ is defined as the angle between the after-scattering wave vector and the field direction. The following modified probability density is assumed:

$$p'(\chi) = \begin{cases} \dfrac{1}{2M_2} & -1 \le \chi < \chi_0 \\ \dfrac{M_2}{2} & \chi_0 \le \chi < 1 \end{cases} \tag{9}$$

$M_2 \ge 1$ is a given parameter, χ_0 is determined from the normalization. The cumulative probability at this point evaluates to $P'(\chi_0) = (1 + M_2)^{-1}$. With a random number r, evenly distributed between 0 and 1, one obtains for $r < P'(\chi_0)$

$$\chi_r = 2M_2 r - 1, \quad \frac{p}{p'} = M_2,$$

and otherwise

$$\chi_r = 1 - \frac{2(r-1)}{M_2}, \quad \frac{p}{p'} = \frac{1}{M_2}.$$

This means that the particle weight is either reduced or increased by the factor M_2 whenever χ is generated from the density (9).

2.3 Variance Estimation

The statistical average of A is estimated by a sample mean,

$$\langle \mathrm{A} \rangle \simeq \frac{\overline{A}}{T} = \frac{1}{TN} \sum_{i=1}^{N} \psi_{Ai} \tag{10}$$

with T being a proper normalization constant. N denotes the number of independent realizations of Ψ_A, that is the number of constructed trajectories. From the sample ψ_{Ai} the sample variance $\sigma^2\{A\}$ is computed, which gives a variance estimate for $\overline{A}$:

$$\sigma^2\{\overline{A}\} = \frac{\sigma^2\{A\}}{N} \tag{11}$$

The new method avoids the commonly assumed separation of the particle's history into sub-histories of some artificially predefined duration.

As an example the One-Particle MC method is applied to an npn-structure shown in Fig.1. The computed variance estimate for the current density is plotted in Fig.2. In this case, A is given by $A(\mathbf{k}, \mathbf{r}) = -e\ v_x(\mathbf{k})\ W_p(\mathbf{r})$, with the group velocity v_x and the charge assignment function W_p for grid point p [5].

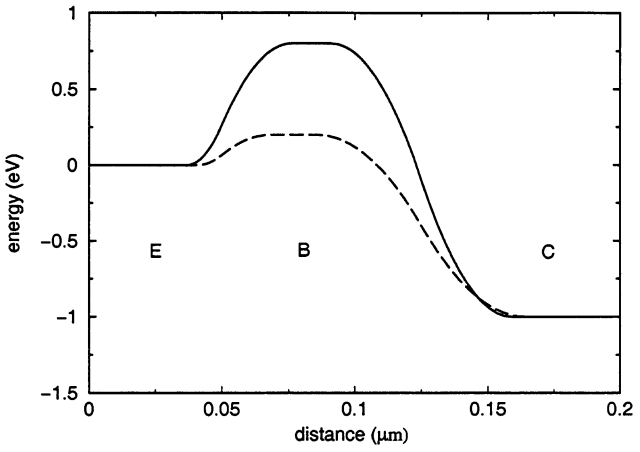

Figure 1: Conduction band edges of the simulated Si npn-structures. E-B barrier heights are 0.2 eV and 0.8 eV.

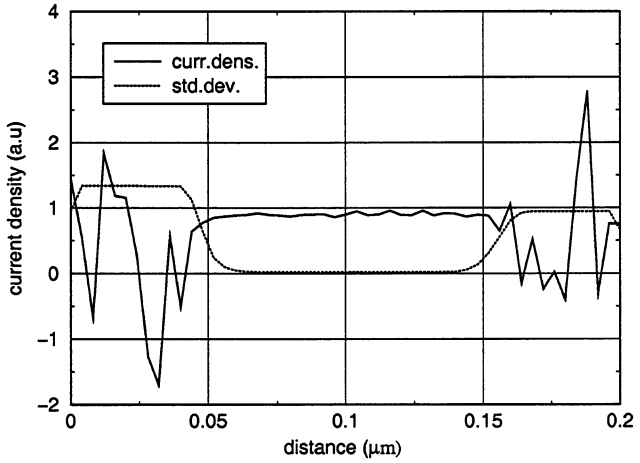

Figure 2: Current density and standard deviation in an npn-structure with 0.2 eV barrier height. No statistical enhancement method is used. $N_{\text{scatt}} = 5 \cdot 10^8$.

3 RESULTS AND DISCUSSION

Using the modified probabilities described in Section 2.2 electron transport through the npn-structure with a 0.8 eV energy barrier has been simulated (Fig.1). To enhance statistics in the emitter-base barrier region artificial carrier heating is introduced. In the barrier and the base region the distribution of the scattering angle is biased to induce artificial carrier diffusion towards the collector.

Optimal values for M_1 and M_2 are not known a priori. For instance, if M_1 is chosen too small, not enough particles will surmount the barrier, rendering statistical enhancement inefficient. On the other hand, choosing M_1 and M_2 too large, plenty of numerical trajectories will pass through the low concentration region. However, due to the aggressive biasing the individual particle weights will evolve to extremely different values, predominantly to extreme small ones. Because of the large spreading of the particle weights the recorded averages will again show a large variance. Reasonable values found for the considered structure are $M_1 = 2$ and $M_2 = 2$.

The described behavior of the event biasing scheme suggests the usage of additional variance reduction techniques [6]. The general goal must be a reduction of the spreading of the weights. Such techniques are not used in this study. Instead, the evolution of the particle weight is governed predominantly by the weighted MC algorithm. Explicit measures are taken only to prevent weights from getting extremely high or low.

The event biasing scheme has been compared with a simple particle split method. To first order such comparison is fair since the light-weight particles generated with either method are not further recycled.

Fig.3 demonstrates for the mean energy that with event biasing the correct physical mean values are reproduced. Also shown is the mean energy of the simulated particles, which is considerably higher the physical mean energy.

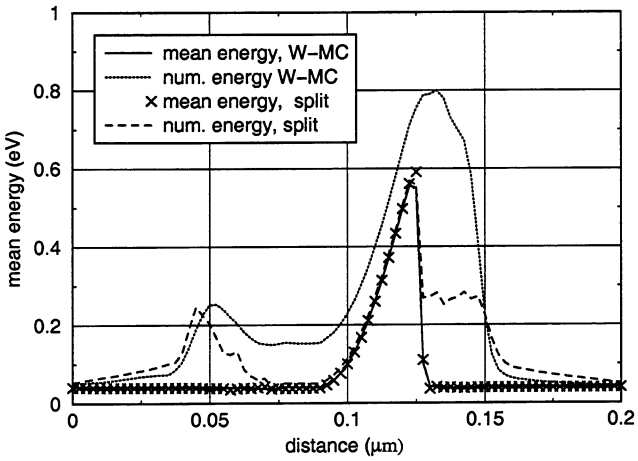

Figure 3: Mean energy of the physical system (mean) and of the simulated carriers (num.) in the npn-structure with 0.8 eV barrier height. Comparison of the event biasing (W-MC) method and a particle split method is shown.

In the simulation shown in Fig.4 a biased boundary distribution is assumed. Electrons are injected from the emitter contact with a Maxwellian distribution at 5 times the lattice temperature. Again the correct physical mean energy is obtained.

In Fig.5 the electron concentration and the standard deviations of the two MC methods are depicted. In the quasi-neutral base region (75–90nm) event biasing gives $\simeq$ 13 times less standard deviation. Due to the $1/\sqrt{N}$ dependence of σ (see (11)) this amounts to an estimated performance gain of $13^2 \simeq 170$ for the given example. Fig.6 shows the superior convergence of W-MC. Because of the poor convergence of

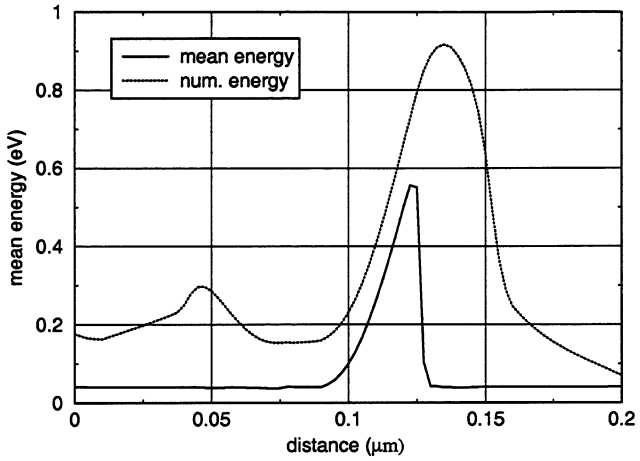

Figure 4: Mean energy in the 0.8 eV structure. In addition to Fig.3, electrons are injected at the left contact at 1500K, whereas at the right contact a Maxwellian at 300K is assumed.

the split method $1.5 \cdot 10^{10}$ scattering events needed to be processed to permit realistic comparisons in Figures 3–5.

4 CONCLUSION

For the One-Particle MC method it has been demonstrated that event biasing is a competitive statistical enhancement technique. It can be used on its own or in combination with other, presently used variance reduction techniques. Implementation of the method does not require structural changes of an existing code. A new variance estimate is proposed, based on the observation that one numerical trajectory represents one independent realization of a random variable.

5 ACKNOWLEDGMENT

This work has been supported by the FWF, Austria, project P13333-TEC, and the European Commission, project NANOTCAD, IST-1999-10828.

REFERENCES

[1] M. Gray, T. Booth, T. Kwan, and C. Snell, "A Multicomb Variance Reduction Scheme for Monte Carlo Semiconductor Simulators," *IEEE Trans.Electron Devices*, vol. 45, no. 4, pp. 918–924, 1998.

[2] C. Jacoboni, "A New Approach to Monte Carlo Simulation," in *Int.Electron Devices Meeting*, (Washington, D.C.), pp. 469–472, IEEE Electron Devices Society, 1989.

[3] M. Nedjalkov and P. Vitanov, "Application of the Iteration Approach to the Ensemble Monte Carlo Technique," *Solid-State Electron.*, vol. 33, no. 4, pp. 407–410, 1990.

[4] H. Kosina, M. Nedjalkov, and S. Selberherr, "Theory of the Monte Carlo Method for Semiconductor Device Simulation," *IEEE Trans.Electron Devices, Special Issue on Computational Electronics*, vol. 47, no. 10, pp. 1898–1908, 2000.

[5] R. Hockney and J. Eastwood, *Computer Simulation Using Particles*. Bristol and Philadelphia: Adam Hilger, 1988.

[6] C. Wordelman, T. Kwan, and C. Snell, "Comparison of Statistical Enhancement Methods for Monte Carlo Semiconductor Simulation," *IEEE Trans.Computer-Aided Design*, vol. 17, no. 12, pp. 1230–1235, 1998.

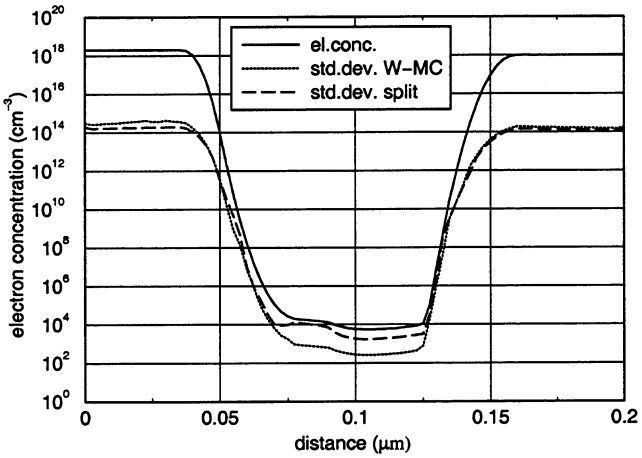

Figure 5: The electron concentration varies by more than 14 orders. In the base region W-MC gives significantly less variance than the split method.

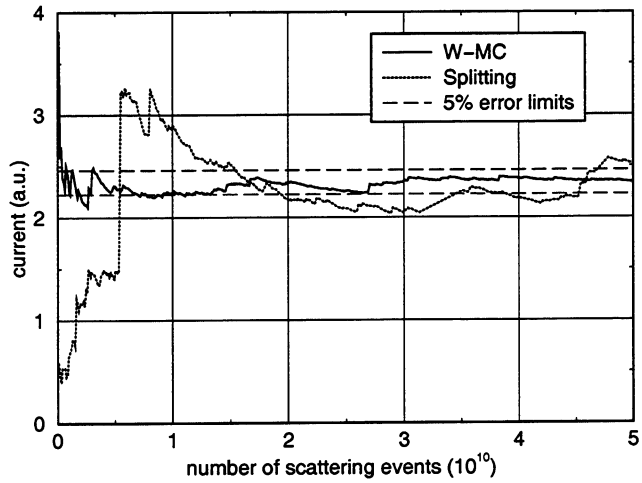

Figure 6: Evolution of the device current for extremely long simulation times, evaluated every 10^8 scattering events. W-MC converges faster and shows better stability than the split method.