Journal of Computational Electronics 3: 305–309, 2004

A Self-Consistent Event Biasing Scheme for Statistical Enhancement

M. NEDJALKOV, S. AHMED AND D. VASILESKA
Department of Electrical Engineering, Arizona State University, Tempe, AZ 85287-5706
mixi@asu.edu

Abstract. The event biasing approach for statistical enhancement is generalized for self-consistent device simulations, posed by mixed boundary and initial conditions transport problems. It is shown that the weight of the particles, as obtained by biasing of the Boltzmann equation, survives between the successive steps of solving the Poisson equation. Particular biasing techniques are applied to the simulation of a 15 nm MOSFET and the convergence of the terminal and channel currents is analyzed.

Keywords: Monte Carlo, statistical enhancement, event biasing

1. Introduction

Statistical enhancement aims at reduction of the time necessary for computation of the desired device characteristics. Enhancement algorithms are especially useful when the device behavior is governed by rare events in the transport process. Such events are inherent for sub-threshold regime of device operation, simulations of effects due to discrete dopant distribution as well as tunneling phenomena. Virtually all Monte Carlo device simulators with statistical enhancement use population control techniques [1]. They are based on the heuristic idea for splitting of the particles entering given phase space region Ω of interest. The alternative idea—to enrich the statistic in Ω by biasing the probabilities associated with the transport of classical carriers—gives rise to the event-biasing approach. The approach, first proposed for the Ensemble Monte Carlo technique (time-dependent problem) [2], has been recently extended for the Single Particle Monte Carlo technique (stationary problem) [3]. In the next section we introduce the basic steps of derivation of the approach in the presence of both initial and boundary conditions. The linearity of the transport problem is utilized, where Coulomb forces between the carriers are initially neglected. The generalization of the approach for Hartree carriers has been found in the iterative procedure of coupling with the Poisson equation. Self-consistent simulation results are presented and discussed in the last section.

2. Event Biasing

The Ensemble Monte Carlo (EMC) technique is designed to evaluate averaged values $\langle A \rangle$ of generic physical quantities a such as carrier density and velocity:

$$\langle A \rangle(\tau) = \int dQ A(Q) f(Q) = \int dQ f_0(Q) g(Q) \quad (1)$$

Here $Q = (\mathbf{k}, \mathbf{r}, t)$ defines the integration over the phase space and time $t \in (0, \infty)$, and $A = a\theta_\Omega \delta(t - \tau)$ introduces the indicator θ_Ω of the phase space domain, where the mean value is evaluated at time τ. Equation (1) is the usual expression for a statistical mean value, augmented by a time integral with the purpose to be conveniently approached with the formal theory of integral equations. It has been shown that the Boltzmann equation can be formulated as a Fredholm integral equation of a second kind with a free term f_0. The latter is determined by the initial condition in evolution problems [2,4] or, in the case of stationary transport, by the boundary conditions [3]. The second equality in (1) follows from the relationship between an integral equation and its adjoint equation. It shows that the mean value $\langle A \rangle$ is determined by f_0 and by the solution of the adjoint Boltzmann

equation:

$$g(Q') = \int dQ K(Q', Q) g(Q) + A(Q') \quad (2)$$

$$K = S(\mathbf{k}', \mathbf{k}, \mathbf{r}) e^{-\int_{t'}^{t} \lambda(\mathbf{K}(y), \mathbf{R}(y)) dy} \times \theta_D(\mathbf{r}) \delta(\mathbf{r}' - \mathbf{r}) \theta(t - t') \quad (3)$$

where S is the usual scattering rate from lattice imperfections, λ is the total out-scattering rate, θ_D is the device domain indicator, which is discussed later, θ is the Heaviside function and the trajectories, initialized by $\mathbf{k}$, $\mathbf{r}$, t', are formulated with the help of the electrical force $\mathbf{F}$ and the velocity $\mathbf{v}$:

$$\mathbf{K}(t) = \mathbf{k} + \int_{t'}^{t} \mathbf{F}(\mathbf{R}(y))\, dy \quad \mathbf{R}(t) = \mathbf{r} + \int_{t'}^{t} \mathbf{v}(\mathbf{K}(y))\, dy$$

If both initial f_i and boundary f_b conditions are taken into account, it is shown that f_0 becomes:

$$f_0(Q) = f_i(\mathbf{k}, \mathbf{r}) e^{-\int_0^t \lambda(\mathbf{K}(y), \mathbf{R}(y)) dy} + \int_0^t \mathbf{v}_\perp(\mathbf{k}) f_b(\mathbf{k}, \mathbf{r}, t_b) e^{-\int_{t_b}^t \lambda(\mathbf{K}(y), \mathbf{R}(y)) dy} dt_b \quad (4)$$

While f_i is defined only at the initial time $t = 0$, the function f_b is defined only at the device boundary Γ and for values of $\mathbf{k}$ such that the corresponding velocity inwards D. $v_\perp$ is the velocity component normal to Γ so that a velocity-weighted distribution drives the particle flux, injected into the device at times $t_b \leq t$. f_0 in (4) governs both the transient and the stationary behavior of a device. The latter is established in the long time limit, provided that f_b is time independent: usually f_b is assumed to be the equilibrium distribution function.

A recursive replacement of Eq. (2) into itself gives rise to the von-Neumann expansion, where the solution g is presented as a sum of the consecutive iterations of the kernel on A. If replaced in (1), the expansion gives rise to the following series for $\langle A \rangle$.

$$\langle A \rangle(\tau) = \sum_i \langle A \rangle_i(\tau) \quad (5)$$

Consider the second term in (5) [1] augmented with the help of two probabilities P_0 and P to become the expectation

$$\langle A \rangle_2 = \int dQ' dQ_1 dQ_2 P_0(Q') P(Q', Q_1) P(Q_1, Q_2) \times \frac{f_0(Q')}{P_0(Q')} \frac{K(Q', Q_1)}{P(Q', Q_1)} \frac{K(Q_1, Q_2)}{P(Q_1, Q_2)} A(Q_2) \quad (6)$$

value of a random variable (r.v.). It takes values determined by the second row with a probability given by the product in the first row. $\langle A \rangle_2$ is evaluated according to the numerical Monte Carlo theory as follows. P_0 and P are used to construct numerical trajectories: (i) $P_0(Q')$ selects the initial point Q' of the trajectory. (ii) $P(Q', Q)$ selects the next trajectory point Q provided that Q' is given. The fraction W_2 in front of A, called weight, is a product of weight factors $\frac{f_0}{P_0}$, and $\frac{K}{P}$ evaluated at the corresponding points $Q_0 \rightarrow Q_1 \rightarrow Q_2$, selected by application of $P_0 \rightarrow P \rightarrow P$. The sample mean of N realizations of the r.v., calculated over N trajectories $(Q' \rightarrow Q_1 \rightarrow Q_2)_n$, estimates the mean value $\langle A \rangle_2$:

$$\langle A \rangle_2 = \frac{1}{N} \sum_{n=1}^{N} (W_2 A)_n \quad \langle A \rangle = \frac{1}{N} \sum_{n=1}^{N} (WA)_n \quad (7)$$

The iterative character of the multiple integral (6) has been used to introduce a consecutive procedure for construction of the trajectories. It can be shown that a single trajectory, obtained by successive applications of P, contributes to the estimators of all terms in (5) simultaneously; i.e. the procedure is generalized by the second equation in (7) for a direct evaluation of $\langle A \rangle$. Next we establish the link between (7) and the EMC technique, which is due to particular choice of the initial, P_0^B, and transition, P^B, densities. P^B, which can be deduced from (3), is a product of the conditional probabilities for free-flight and scattering, associated with the evolution of the real carriers. The ratio K/P^B is then the domain indicator θ_D which takes values 1 if the trajectory belongs to D and 0 otherwise. The choice of P_0^B is complicated by the presence of both initial and boundary terms in (4). They decompose (7) into two terms which are evaluated separately: $\langle A \rangle = \frac{1}{N_1} \sum_{n=1}^{N_1} (WA)_n + \frac{1}{N_2} \sum_{n=1}^{N_2} (WA)_n$ The initial probability P_0^B for each estimator is obtained from f_i and $v_\perp f_b$ respectively, with the help of two normalization factors: the number of initial carriers N_i and the total number N_J of the injected into the device particles. The ratio f_0/P_0^B for each of the estimators becomes N_i and N_J respectively, and can be eliminated by the

choice $N_1 = N_i$ and $N_2 = N_J$. The two sums can be merged back to give

$$\langle A \rangle = \sum_{n=1}^{N_i+N_J} (WA)_n = \sum_{n=1}^{N_\tau} \theta_\Omega(n) a_n \qquad (8)$$

(8) accounts that only trajectories which belong to D give contributions. As only the endpoint of such trajectories matters for the estimator, we speak about particles inside the device. N_τ is the number of such particles at time τ, and $\theta_\Omega(n)$ is 1 or 0 if the n-th particle is inside or outside Ω. All particles have weight unity and evolve as real Boltzmann carriers: the EMC technique for transport problems posed by initial and boundary conditions is recovered. A choice of alternative probabilities is called event biasing. The probabilities for initial and/or boundary distributions, free flight duration, type of scattering and the selection of the after-scattering state direction can be biased. It can be shown that (8) is generalized to $\langle A \rangle = \sum_{n=1}^{N_\tau^b} W_n \theta_\Omega(n) a_n$ where the position of the N_τ^b biased particles is accounted in θ_Ω.

The Boltzmann equation for Coulomb carriers becomes nonlinear via the interaction component $\mathbf{F}(f)(\mathbf{r}, t)$ of the electric force. As the results of the previous section are based on the linearity of the integral equations involved, it is no longer possible to apply the steps used to derive the event biasing. The solution is sought in the iterative procedure of coupling of the EMC technique with the Poisson equation: the latter is discretized by a decomposition of the device region into mesh cells Ψ_l. The particle system is evolved in time intervals $\Delta t \simeq 0.1 fs$. At the end of each time step, at say time τ, the charge density $qC(\mathbf{r}_l, \tau)$ is calculated and assigned to the corresponding grid points. We use the relation between C_l and the distribution function $f_{l,m} = f(\mathbf{r}_l, \mathbf{k}_m, \tau)$, which is estimated with the help of (8) by introducing a mesh Φ_m in the wave vector space, ($\Omega_{l,m} = \Psi_l \Phi_m$):

$$f_{l,m} = \frac{\sum_n \theta_{\Omega_{l,m}(n)}}{V_{\Omega_l} V_{\Phi_m}} C_l = \sum_m f_{l,m} V_{\Phi_m} N_\tau = \sum_l C_l V_{\Psi_l}$$

The charge density C_l is used to find the solution of the Poisson equation, which provides an update for the electric force $\mathbf{F}(\mathbf{r}, \tau)$. The latter governs the trajectories evolving the particles in the next time interval $\tau, \tau + \Delta t$. Between the steps of solving the Poisson equation the electric field is frozen so that event biasing can be applied. Assume that at time τ the particles emerge with weights W_n. Due to the event biasing the behavior of the biased particles differs from that of the EMC particles. The distribution function of the biased particles $f_{l,m}^{\text{num}}$ obtained from the above formula is entirely different from $f_{l,m}$. Nevertheless, as seen from (6), any biasing does not change the values of the physical averages. The Boltzmann distribution function is recovered by using the weights W_n: $f_{l,m} = \sum_n W_n \theta_{\Omega_{l,m}}(n) / V_{\Omega_l} V_{\Phi_m}$ Accordingly the correct $\mathbf{F}$ is provided by the Poisson equation. As the evolution is Markovian, $f_{l,m}$ presents the initial condition for the next time step. Numerical particles, having distribution $f_{l,m}^{\text{num}}$ and weights W_n present a biased initial condition for this step: the initial weight will be updated in the time interval $\tau, \tau + \Delta t$ by the weight factors according the chosen biased evolution. It follows that, the particle weights survive between the successive iteration steps, which completes the proof of the self-consistent biasing scheme.

3. Simulation Results

The MOSFET device chosen for the simulation experiments has gate length is 15 nm, channel doping 2×10^{19} cm^{-3}, and oxide thickness 0.8 nm. A similar device has already been fabricated by Intel [5].

The applied potentials $V_G = 0.375$ V, $V_D = 0.1$ V correspond to a sub-threshold regime, where the channel carrier density is three orders of magnitude less than the density in the S/D regions. The lattice temperature is $T = 300$ K. The chosen biasing techniques aim at increasing the number of numerical particles in the channel, but keep the total particle number in the device equal to 10^5, as used in the EMC technique. Particles which enrich the high energy domain of the distribution at the expense of obtaining weights $W < 1$ readily overcome the source potential barrier. The number of particles in the low energy domain is reduced, such particles have weights $W > 1$ and remain longer in the S/D regions. An alternative approach, which biases the scattering angle of the Coulomb interaction in order to create a flux of numerical particles towards the channel is currently under consideration. We first investigate the consistency of the biasing techniques in the thermodynamic limit of a very large number of simulated particles. Both Boltzmann and biased stochastic processes must give the same evolution of the physical averages. Indeed, the simulated physical characteristics, such as carrier, velocity and energy profiles, and in particular the electric field in X and Y directions, coincide with the results from the EMC approach at a given τ. Second for smaller particle number we seek

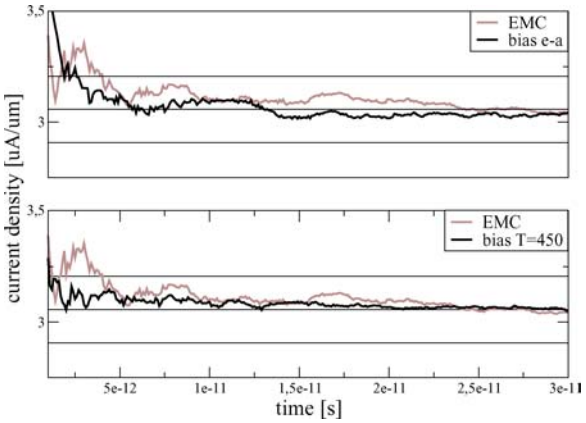

Figure 1. Comparison of the channel currents from biased e-a rates, boundary distribution ($T = 450$ K) and the EMC.

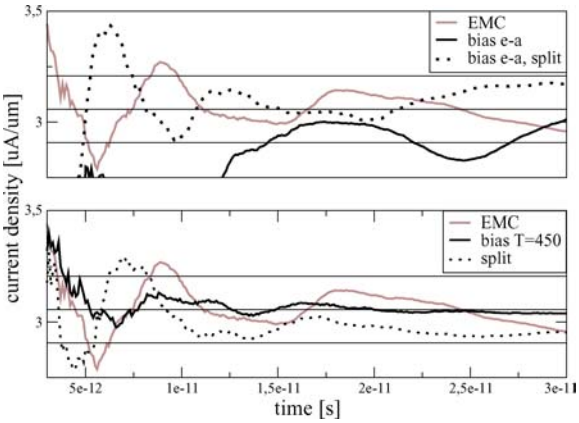

Figure 2. Terminal currents obtained by particle counting.

bias processes with smaller variance as compared to the EMC technique. The convergence of the cumulative averages for the channel and terminal currents (obtained from the velocity and particle counting respectively) is investigated.We first bias the phonon scattering by increasing the absorption and reducing the emission rates (e-a) in the half of the source region near the barrier in a 4 nm depth. Figure 1 (top) shows the biased channel current as compared to the EMC result for 30 ps evolution time. The 5% error region (straight lines) around the mean value is entered 2.5 ps earlier and the convergence is better. The second technique biases the boundary conditions by injection of particles with higher temperature $T = 450$ K (Fig. 1 bottom). The temperature-biased curve shows superior behavior.The corresponding terminal currents are much more noisy and show long time correlations due to the inter-particle interactions. The e-a biased curve is very unstable and enters the 5% error region in Fig. 2, top, after 15 ps evolution. We associate this behavior with the numerical error: the poor statistics are due to the appearance of very heavy ($W > 2$) particles in the source. To check this it is sufficient to apply the conventional particle splitting along with the e-a biasing. The result is presented by the dotted curve on the plot. The behavior is significantly improved at the expense of a 30% increase in the number of simulated particles. The terminal current corresponding to the biasing of the temperature of the injected particles again shows a superior behavior. This is due to an improved weight control: the weight determined during the injection remains constant in the evolution. Its maximal value for $T = 450$ K is 1.5. Furthermore the probability for interaction with the impurities, which dominates the S/D regions, drops for the majority of the particles due to their high energies. The conventional splitting technique cannot achieve such superiority. Experiments where the boundary particles at $T = 300$ K are split into sub-particles with weight $C_1 < 1$ above a given energy threshold ϵ_1 have been performed. In order to maintain the same total number of simulated particles (10^5), the weight below ϵ_1 becomes $C_2(C_1, \epsilon_1) > 1$. The best result obtained by varying ϵ_1 and C_1 is shown by the dotted curve on Fig. 2. The behavior of the curve resembles the EMC counterpart. An improvement is expected if the C_2 particles are split, which recovers the conventional split technique.

In conclusion, the event biasing approach has been derived in the presence of both initial and boundary conditions and generalized for self-consistent simulations. The approach is confirmed by the simulations presented. A bias technique, particularly useful for small devices, is obtained by injection of hot carriers from the boundaries. The coupling with the Poisson equation requires precise statistics in the S/D regions. It is shown that a combination of event biasing and population control approaches is advantageous for this purpose. We note that estimates for the variance and covariance can be introduced to measure the convergence [3].

Note

1. The δ function in each K takes away a spatial integral, while $\delta \in A$ sets the time $t_2 \in Q_2$ equal to the evaluation time τ.

References

1. C. Wordelman, T. Kwan, and C. Snell, "Comparison of statistical enhancement methods for Monte Carlo semiconductor simulation," *IEEE Trans.Computer-Aided Design*, **17**(12), 1230–1235 (1998).
2. L. Rota, C. Jacoboni, and P. Poli, "Weighted ensemble Monte Carlo," *Solid-State Electron.*, **32**(12), 1417–1421 (1989).
3. H. Kosina, M. Nedjalkov, and S. Selberherr, "The stationary Monte Carlo method for device simulation - Part I and Part II," *J. Appl. Phys.*, **93**(6), 3553–3571 (2003).
4. C. Jacoboni, "A new approach to Monte Carlo simulation," in *PRIEDM, IEEE Electron Devices Society*, (1989) p. 469.
5. R. Chau, B. Boyanov, B. Doyle, M. Doczy, S. Datta, S. Hareland, B. Jin, J. Kavalieros, and M. Metz, "Silicon nano-transistors for logic applications," *Physica E*, 1 (2003).