ELSEVIER

Mathematics and Computers in Simulation 55 (2001) 191–198

MATHEMATICS AND COMPUTERS IN SIMULATION

www.elsevier.nl/locate/matcom

Variance of the ensemble Monte Carlo algorithm for semiconductor transport modeling☆

M. Nedjalkov*, H. Kosina

Institute for Microelectronics, TU-Wien, Gusshausstr. 27-29, A1040 Vienna, Austria

Abstract

The ensemble Monte Carlo algorithm (EMC) is the most frequently used tool for simulation of the transient transport in semiconductors and structures. The common definition of the algorithm is of a procedure based on imitation of the real transport phenomena. Often EMC is accepted as a simulated experiment rather than as a numerical method. Recently it has been shown that the EMC can be obtained by an application of the numerical Monte Carlo (MC) theory to the integral form of the Boltzmann equation (BE) [1–3]. The approach has been further used to prove under a general condition the convergence of the algorithm [4]. In this work we utilize the approach to investigate the variance of the EMC. It is proved that the algorithm has a finite variance and an analytical result is derived. This allows to assign the precision estimates of the numerical MC method to the EMC.

Keywords: Variance; Ensemble Monte Carlo algorithm (EMC); Semiconductors; Simulation

1. The ensemble Monte Carlo algorithm

Within the semi-classical concepts, the carrier transport in semiconductors is an alteration of elementary events taking place in the six-dimensional phase space $\boldsymbol{K} \times \boldsymbol{R}$. These are the drift and scattering events characterized below. The point-like particles drift along trajectories, defined by first order differential equations $\dot{\boldsymbol{k}} = \boldsymbol{F}(\boldsymbol{r})$, $\dot{\boldsymbol{r}} = \boldsymbol{v}(\boldsymbol{k})$ (Newton's law) where $\boldsymbol{F}(\boldsymbol{r}) = q\boldsymbol{E}(\boldsymbol{r})/\hbar$, $\boldsymbol{E}$ is the electric field, q is the particle charge and $\boldsymbol{v}$ the velocity. The time dependence of the electric field is not written explicitly. A phase space point $\boldsymbol{k}, \boldsymbol{r}$ at time t uniquely determines the solution of the Newton's equations $\boldsymbol{K}(\tau)$, $\boldsymbol{R}(\tau)$ at some evolution time τ. Assume that a trajectory takes the values $\boldsymbol{k}_a, \boldsymbol{r}_a$ and $\boldsymbol{k}_b, \boldsymbol{r}_b$ at the times $t_a < t_b$. There are two ways to initialize the trajectory

$$\boldsymbol{K}(\tau) = k_a + \int_{t_a}^{\tau} \boldsymbol{F}(\boldsymbol{R}(y))\,\mathrm{d}y = \boldsymbol{k}_b - \int_{\tau}^{t_b} \boldsymbol{F}(\boldsymbol{R}(y))\,\mathrm{d}y \tag{1}$$

☆ This work has been supported by the 'Fonds zur Förderung der wissenschaftlichen Forschung', project P13333-TEC, Austria.
* Corresponding author. Present address: CLPP, BAS, G. Bontchev Str. Bl25A, 1113 Sofia, Bulgaria..

0378-4754/01/$20.00
PII: S0378-4754(00)00261-5

$$\boldsymbol{R}(\tau) = \boldsymbol{r}_a + \int_{t_a}^{\tau} \boldsymbol{v}(\boldsymbol{K}(y))\,\mathrm{d}y = \boldsymbol{r}_b - \int_{\tau}^{t_b} \boldsymbol{v}(\boldsymbol{K}(y))\,\mathrm{d}y$$

That is, to choose $\boldsymbol{k}_a, \boldsymbol{r}_a, t_a$ so that $\tau > t_a$ or to choose $\boldsymbol{k}_b, \boldsymbol{r}_b, t_b$ so that $\tau < t_b$. At evolution time τ the reached phase space point is the same for both ways of initialization. We refer to the first initialization as to a forward one and to the second as to a backward one.

The drift is interrupted by scattering events due to the lattice imperfections considered local in position and instantaneous in time. They are accounted for by a function $S(\boldsymbol{k}, \boldsymbol{k}', \boldsymbol{r})\,\mathrm{d}\boldsymbol{k}'$ giving the scattering frequency from $\boldsymbol{k}$ to $\mathrm{d}\mathbf{k}'$ around $\boldsymbol{k}'$. The elementary events are described by theoretical models for the energy dispersion $\epsilon(\boldsymbol{k})$, the variety of scattering mechanisms included in S and parameters typical for any semiconductor. For a concrete semiconductor there could be a number of theoretical models reflecting on different level of complexity the dispersion and interaction relations. The main quantity of interest is the distribution function $f(\boldsymbol{k}, \boldsymbol{r}, t)$ giving the particle density at time t. It provides the complete information about the mean values of the single-particle physical characteristics. The distribution function is a solution of the BE determined by the initial and/or boundary condition. In this work we consider evolution problems determined by initial conditions only. The phenomenological derivation of the BE can be found in most textbooks and includes two steps: the particle number in a small phase space domain is expressed as a balance of the positive and negative contributions of the particles entering and leaving the domain due to the drift and scattering processes. Finally the domain volume is let to zero (which means that the number of all carriers is let to infinity), thus providing the common integro-differential form of the BE.

The idea of the EMC is very natural: it is to imitate the elementary transport events happening to the real carriers. An ensemble of test particles is selected according the initial condition. A sequence of drift and scattering events is simulated according to the theoretical model for any test particle. The particle number at any time in any phase space volume unit is built up in a natural way and gives an estimate for the solution of the BE. The solution corresponds to the theoretical model which is supposed to describe adequately the physical problem. In this sense the imitation EMC is another way of stating the BE. Moreover, MC provides also information, related to the finite carrier number of the real processes. This information is lost in the BE [5]. EMC estimates the mean $f_{\Omega} = \int f(\boldsymbol{k}, \boldsymbol{r}, t)\theta_{\Omega}(\boldsymbol{k}, \boldsymbol{r})\,\mathrm{d}\boldsymbol{k}\,\mathrm{d}\boldsymbol{r}$ of the BE solution in a given phase space sub-domain Ω (i.e. the realtive number of the carriers inside). The domain indicator θ_{Ω} is unity if the arguments belong to Ω and zero otherwise. A test particle is followed by a generation of the probabilities for the drift duration and the after scattering state which are obtained from the scattering frequency function S. The probability for a carrier to scatter from $\boldsymbol{k}$ into $\mathrm{d}\boldsymbol{k}'$ around $\boldsymbol{k}'$ is $S(\boldsymbol{k}, \boldsymbol{k}', \boldsymbol{r})\,\mathrm{d}\boldsymbol{k}'/\lambda(\boldsymbol{k}, \boldsymbol{r})$ where $\lambda(\boldsymbol{k}, \boldsymbol{r}) = \int \mathrm{d}\boldsymbol{k}' S(\boldsymbol{k}, \boldsymbol{k}', \boldsymbol{r})$. The probability for drift without scattering during the time interval $(0, t)$ is given by $\exp\left(-\int_0^t (\boldsymbol{K}(\tau), \boldsymbol{R}(\tau))\,\mathrm{d}\tau\right)$. When multiplied by $\lambda(\boldsymbol{K}(t), \boldsymbol{R}(t))\,\mathrm{d}t$ it becomes the probability for scattering in time interval $\mathrm{d}t$ after a succesful drift $(0, t)$. The main steps in the algorithm are the following

1. choose the evolution time t and place N carriers at points $\boldsymbol{k}_0, \boldsymbol{r}_0$, selected according to the initial carrier distribution $f_0(\boldsymbol{k}_0, \boldsymbol{r}_0)$, assign to Ω a counter ν, begin the trajectory of every particular carrier at $\boldsymbol{k}_{\mathrm{s}} = \boldsymbol{k}_0, \boldsymbol{r}_{\mathrm{s}} = \boldsymbol{r}_0$ and $t_{\mathrm{s}} = 0$;
2. use the point $\boldsymbol{k}_s, \boldsymbol{r}_s, t_s$ for a forward initialization of the Newton trajectory $\boldsymbol{K}(t'), \boldsymbol{R}(t')$, with a probability density $\lambda(\boldsymbol{K}(t'), \boldsymbol{R}(t'))\exp\left(-\int_{t_s}^{t'} \lambda(\boldsymbol{K}(y), \boldsymbol{R}(y))\,\mathrm{d}y\right)$ generate the next scattering time t';
3. if $t' < t$ use the probability density $S(\boldsymbol{K}(t'), \boldsymbol{k}', \boldsymbol{R}(t'))/\lambda(\boldsymbol{K}(t'), \boldsymbol{R}(t'))$ to generate the after-scattering state $\boldsymbol{k}'$, set the values $\boldsymbol{k}_{\mathrm{s}} = \boldsymbol{k}', \boldsymbol{r}_{\mathrm{s}} = \boldsymbol{R}(t'), t_{\mathrm{s}} = t'$ and repeat 2;
4. if $t' > t$ increment ν by $\theta_{\Omega}(\boldsymbol{K}(t), \boldsymbol{R}(t))$ and begin with the next carrier from 1.

After all N trajectories are followed, f_Ω is given by the estimator value divided by N. The above considerations introduce the EMC from a physical point of view. An analysis of the numerical properties of the algorithm, such as convergence conditions, variance and probable error require to reformulate the EMC in terms of the numerical MC methods. Relevant for this purpose will be the MC method for solving integrals.

2. Monte Carlo integration

We consider the task of evaluation of an integral

$$I = \int_{-\infty}^{\infty} \phi(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} p(x)\psi(x)\,\mathrm{d}x \tag{2}$$

where the function ϕ is absolute integrable. Suppose that p is non-negative and $\int_{-\infty}^{\infty} p(x)\,\mathrm{d}x = 1$, i.e. p is a density function. Then the function ψ can be interpreted as a random variable with a mean value $E\{\psi\} = I$ and a variance $\sigma_\psi^2 = E\{\psi^2\} - E^2\{\psi\}$, supposed to be finite. We consider an experiment of N independent, identically distributed (IID) (by p) random variables ψ_i [6]. The random variable $\bar{\psi} = \sum_{i=1}^{N} \psi_i/N$ is called sample mean and has the following properties: $E\{\bar{\psi}\} = E\{\psi\} = I$ and $\sigma_{\bar{\psi}}^2 = \sigma_\psi^2/N$. After the CLT, under a certain general conditions, the probability $P\{\bar{\psi} \le x\}$ for $\bar{\psi}$ to have an value less or equal to x approaches the normal distribution with mean I and variance $\sigma_{\bar{\psi}}^2$ when N tends to infinity. This allows to estimate the precision by which the sample mean approaches the mean value. We give two estimates, called 'the rule of 3σ' and probable error r_N

$$P\{|\bar{\psi} - I| \le 3\sigma_{\bar{\psi}}\} \simeq 0.997; \quad r_N = 0.6745\sigma_{\bar{\psi}} \tag{3}$$

The concept of the Monte Carlo method for evaluating integrals is to generate numerically such IID random variables. Note that the existence of the mean value is sufficient for the method to be applied [7]. But the method can be supplied by the above numerical characteristics only if the variance is finite.

A generalization to a multidimensional integral offers the following cases. Now x is a multidimensional point in given domain A.

- C1: The discussed case of $\phi(x) = \psi(x)p(x)$, $p(x) > 0$ and $\int_A p(x)\,\mathrm{d}x = 1$. Then the simplest algorithm for calculating I is very short: (1) with probability $p(x)$ generate N trials points x_i; (2) calculate the estimator $\mu_i = \psi(x_i)$ and (3) calculate the sample mean $\sum_{i=1}^{N} \mu_i/N$ According to the CLT it is an estimate for the value of I.
- C2: The function ϕ is a product of two functions $\phi(x) = \psi(x)p(x)$, $p(x) \ge 0$ but $\int_A p(x)\,\mathrm{d}x \ne 1$.
  (a): If $J = \int_A p(x)\,\mathrm{d}x$ is known the problem is reduced to the first case by using $\phi(x) = (J\psi(x))$ $(p(x)/J)$. Now the probability function $p(x)/J$ is normalized to unity in A.
  (b): When $\int_A p(x)\,\mathrm{d}x < 1$, another possibility is to extrapolate p to a domain B, where $A \in B$ such that $J = \int_B p(x)\,\mathrm{d}x = 1$ holds. The problem is related to C1 by setting $\psi(x) = 0$ when $x \notin A$ and generating points in B.
- C3: We can choose an arbitrary density $p(x)$ such that $p(x) \ne 0$ if $g(x) \ne 0$. Then we can write $\phi(x) = (\phi(x)/p(x))p(x)$ and refer to C1.

To investigate the variance of the EMC, it is necessary to determine the density function and the random variable of the carrier evolution process. For this purpose the algorithm must be associated to an integral

according to C1. $f_\Omega = \int f(\boldsymbol{k}, \boldsymbol{r}, t)\theta_\Omega(\boldsymbol{k}, \boldsymbol{r})\,\mathrm{d}\boldsymbol{k}\,\mathrm{d}\boldsymbol{r}$ provides a starting expression which must be further processed since f is unknown.

3. Integral equations related to the Boltzmann equation

The distribution function f can be expressed as a solution Fredholm integral equation of second kind [8,9]

$$f(\boldsymbol{k}, \boldsymbol{r}, t) = \int_0^\infty \mathrm{d}t' \int \mathrm{d}\boldsymbol{k}' \int \mathrm{d}\boldsymbol{r}' f(\boldsymbol{k}', \boldsymbol{r}', t') S(\boldsymbol{k}', \boldsymbol{K}(t'), \boldsymbol{r}') \exp\left(-\int_{t'}^{t} \lambda(\boldsymbol{K}(y), \boldsymbol{R}(y))\,\mathrm{d}y\right)$$
$$\times \delta(\boldsymbol{r}' - \boldsymbol{R}(t'))\theta(t - t') + f_0(\boldsymbol{K}(0), \boldsymbol{R}(0)) \exp\left(\int_0^t \lambda(\boldsymbol{K}(y), \boldsymbol{R}(y))\,\mathrm{d}y\right) \tag{4}$$

here a backward initialization is used, determined by $\boldsymbol{k}$, $\boldsymbol{r}$, t. The common integro-differential form of the BE is recovered if the δ and θ functions are accounted by the integration, the identities $\boldsymbol{K}(t) = \boldsymbol{k}$, $\boldsymbol{R}(t) = \boldsymbol{r}$ are used on the left and a differentiation with respect to the time t is performed.

The EMC simulates f_Ω by using a forward initialization so that our further steps are directed to reverse the initialization used in (4). We write (4) in a short notation as $f(Q) = \int \mathrm{d}Q' f(Q') K(Q', Q) + f_0(Q)$ and introduce an conjugate equation $g(Q') = \int \mathrm{d}Q g(Q) K(Q', Q) + g_0(Q')$, [3]. If the equation for f is multiplied by $g(Q)$ and integrated over Q, the equation for g is multiplied by $f(Q')$ and integrated over Q' and the two equations are subtracted, the result is

$$\int f(Q) g_0(Q)\,\mathrm{d}Q = \int g(Q) f_0(Q)\,\mathrm{d}Q \tag{5}$$

In this way the task to calculate a functional of f is transformed into a task to calculate a functional of the solution of the conjugate equation. The free term of the conjugate equation is the function g_0 which determines the concrete functional of f.

The conjugate equation of (4) is obtained by a replacement of the primed with the unprimed integration variables

$$g(\boldsymbol{k}', \boldsymbol{r}', t') = \int_0^\infty \mathrm{d}t \int \mathrm{d}\boldsymbol{k} \int \mathrm{d}\boldsymbol{r}\, g(\boldsymbol{k}, \boldsymbol{r}, t) S(\boldsymbol{k}', \boldsymbol{K}(t'), \boldsymbol{r}') \exp\left(-\int_{t'}^{t} \lambda(\boldsymbol{K}(y), \boldsymbol{R}(y))\,\mathrm{d}y\right)$$
$$\times \delta(\boldsymbol{r}' - \boldsymbol{R}(t'))\theta(t - t') + g_0(\boldsymbol{k}', \boldsymbol{r}', t') \tag{6}$$

By using a Liouville transform, the integration variables $\boldsymbol{k}$, $\boldsymbol{r}$ can be changed to $\boldsymbol{K}(t')$, $\boldsymbol{R}(t')$. Further $\boldsymbol{K}(t')$ is denoted by $\boldsymbol{k}^a$ and $\boldsymbol{R}(t')$ with $\boldsymbol{r}''$. This allows to obtain $\boldsymbol{k} = \boldsymbol{K}(t)$ and $\boldsymbol{r} = \boldsymbol{R}(t)$ in a forward initialization with $\boldsymbol{k}^a$, $\boldsymbol{r}''$ at t. Further the δ and θ functions are taken into account in the integration. Additionally, the variable of time integration is renamed from t to τ.

$$g(\boldsymbol{k}', \boldsymbol{r}', t') = \int_{t'}^\infty \mathrm{d}\tau \int \mathrm{d}\boldsymbol{k}^a S(\boldsymbol{k}', \boldsymbol{k}^a, \boldsymbol{r}', t') \exp\left(-\int_{t'}^{\tau} \lambda(\boldsymbol{K}(y), \boldsymbol{R}(y))\,\mathrm{d}y\right)$$
$$\times g(\boldsymbol{K}(\tau), \boldsymbol{R}(\tau), \tau) + g_0(\boldsymbol{k}', \boldsymbol{r}', t') \tag{7}$$

The trajectory $\boldsymbol{K}(y)$, $\boldsymbol{R}(y)$ appears in a forward initialization, given by $\boldsymbol{k}^a$, $\boldsymbol{r}'$ at t'. The explicit form of g_0 is determined by our concrete task to calculate f_Ω. The expression for f_Ω must be augmented in order

to include all relevant integration variables $f_\Omega = \int_0^\infty \mathrm{d}t' \int \mathrm{d}\boldsymbol{k}' \int \mathrm{d}\boldsymbol{r}' f(\boldsymbol{k}', \boldsymbol{r}', t')\delta(t - t')\theta_\Omega(\boldsymbol{k}', \boldsymbol{r}')$. This defines $g_0(\boldsymbol{k}', \boldsymbol{r}', t') = \delta(t - t')\theta_\Omega(\boldsymbol{k}', \boldsymbol{r}')$. According to (5) f_Ω becomes

$$f_\Omega = \int_0^\infty \mathrm{d}t' \int \mathrm{d}\boldsymbol{k}' \int \mathrm{d}\boldsymbol{r}' f_0(\boldsymbol{K}(0), \boldsymbol{R}(0)) \exp\left(-\int_0^{t'} \lambda(\boldsymbol{K}(y), \boldsymbol{R}(y))\,\mathrm{d}y\right) g(\boldsymbol{k}', \boldsymbol{r}', t')$$

Using again the Liouville transform we change the backward initialization at $(\boldsymbol{k}', \boldsymbol{r}', t')$ to a forward initialization at $(\boldsymbol{k}_0, \boldsymbol{r}_0, 0)$.

$$f_\Omega = \int_0^\infty \mathrm{d}t' \int \mathrm{d}\boldsymbol{k}_0 \int \mathrm{d}\boldsymbol{r}_0 f_0(\boldsymbol{k}_0, \boldsymbol{r}_0) \exp\left(-\int_0^{t'} \lambda(\boldsymbol{K}(y), \boldsymbol{R}(y), y)\,\mathrm{d}y\right) g(\boldsymbol{K}(t'), \boldsymbol{R}(t'), t') \tag{8}$$

4. Variance of the EMC

The resolvent series of (7) gives the solution $g = \sum_{i=0}^\infty g^{(i)}$ as a sum of the successive iterations $g^{(i+1)}(Q') = \int \mathrm{d}Q \boldsymbol{K}(Q', Q) g^{(i)}(Q)$ with $g^{(0)} = g_0$. Accordingly, (8) becomes $f_\Omega = \sum_{i=0}^\infty f_\Omega^{(i)}$. It has been shown that the resolvent series of the conjugate equation converges [4]. Hence for any evolution time t we can choose the order n large enough, such that $\sum_{i=n}^\infty f_\Omega^{(i)}$ becomes less than a desired small ϵ. The finite sum $f_\Omega^{[n]} = \sum_{i=0}^n f_\Omega^{(i)}$ which approaches f_Ω for large n will be further analyzed. We consider $f_\Omega^{(2)}$ which is obtained as

$$\begin{aligned} f_\Omega^{(2)} = & \int_0^\infty \mathrm{d}t_1 \int_{t_1}^\infty \mathrm{d}t_2 \int_{t_2}^\infty \mathrm{d}t_3 \int \mathrm{d}\boldsymbol{k}_1 \int \mathrm{d}\boldsymbol{k}_2 \int \mathrm{d}\boldsymbol{k}_0 \int \mathrm{d}\boldsymbol{r}_0 \{f_0(\boldsymbol{k}_0, \boldsymbol{r}_0)\} \\ & \times \left\{\exp\left(-\int_0^{t_1} \lambda(\boldsymbol{K}_0(y), \boldsymbol{R}_0(y))\,\mathrm{d}y\right) S(\boldsymbol{K}_0(t_1), \boldsymbol{k}_1, \boldsymbol{R}_0(t_1))\right\} \\ & \times \left\{\exp\left(-\int_{t_1}^{t_2} \lambda(\boldsymbol{K}_1(y), \boldsymbol{R}_1(y))\,\mathrm{d}y\right) S(\boldsymbol{K}_1(t_2), \boldsymbol{k}_2, \boldsymbol{R}_1(t_2))\right\} \\ & \exp\left(-\int_{t_2}^{t_3} \lambda(\boldsymbol{K}_2(y), \boldsymbol{R}_2(y))\,\mathrm{d}y\right) \delta(t - t_3)\theta_\Omega(\boldsymbol{K}(t_3), \boldsymbol{R}(t_3)) \end{aligned} \tag{9}$$

The terms eclosed in curly brackets are those probability densities used in the EMC. f_0 is the initial carrier distribution, normalized to unity and can be used to generate initial points $\boldsymbol{k}_0, \boldsymbol{r}_0$ as required by step 1 of the EMC.

Following C2(a), we can rewrite the next term as a product of densities

$$\left\{\exp\left(-\int_0^{t_1} \lambda(\boldsymbol{K}_0(y), \boldsymbol{R}_0(y))\,\mathrm{d}y\right) \lambda(\boldsymbol{K}_0(t_1), \boldsymbol{R}_0(t_1)) \frac{S(\boldsymbol{K}_0(t_1), \boldsymbol{k}_1, \boldsymbol{R}_0(t_1))}{\lambda(\boldsymbol{K}_0(t_1), \boldsymbol{R}_0(t_1))}\right\}$$

They appear as the densities for the first free flight duration and after scattering state in the steps 2 and 3 of the EMC. The initialization of the free-flight trajectory is given by $\boldsymbol{k}_0, \boldsymbol{r}_0$ at 0 and the after-scattering state is chosen by the density S/λ, determined by the end flight coordinates $\boldsymbol{K}_0(t_1)$, $\boldsymbol{R}_0(t_1)$. This means that we deal with a conditional probabilities. The order they are applied is opposite to the order of the integrals, providing the normalization. Integration over $\boldsymbol{k}_1$ and then integration over t_1 of the above term

gives unity. The term has the meaning of a probability density for a transition from $\boldsymbol{k}_0$, $\boldsymbol{r}_0$ at 0 to $\boldsymbol{k}_1$, $\boldsymbol{R}_0(t_1)$ at t_1. We use the locality of the scattering in the real space $\boldsymbol{R}_0(t_1) = \boldsymbol{R}_1(t_1)$ to denote the term by: $P(\boldsymbol{k}_0, \boldsymbol{r}_0, 0|\boldsymbol{k}_1, \boldsymbol{R}_1(t_1), t_1)$. The next term in (9) in curly brackets has an equivalent interpretation and is accordingly written as $P(\boldsymbol{k}_1, \boldsymbol{R}_1(t_1), t_1|\boldsymbol{k}_2, \boldsymbol{R}_2(t_2), t_2)$. We conclude that the densities in (9) coincide with the EMC densities, generating the real carrier trajectories of the first two flight-scattering events.

The integration over t_3 of the last line term in (9) in the limits t_2 and ∞ gives $\exp(-\int_{t_2}^{t}\lambda(\boldsymbol{K}_2(y), \boldsymbol{R}_2(y))\,\mathrm{d}y)\theta(t-t_2)\theta_\Omega(\boldsymbol{K}(t), \boldsymbol{R}(t))$. The exponential function is a probability which can be expressed as an integral over a probability density

$$\exp\left(-\int_{t_2}^{t}\lambda(\boldsymbol{K}_2(y), \boldsymbol{R}_2(y))\,\mathrm{d}y\right) = \int_{t_2}^{\infty}\mathrm{d}t_3\left\{\exp\left(-\int_{t_2}^{t_3}\lambda(\boldsymbol{K}_2(y), \boldsymbol{R}_2(y))\,\mathrm{d}y\right)\lambda(\boldsymbol{K}_2(t_3), \boldsymbol{R}_2(t_3))\right\}\theta(t_3 - t)$$

The θ function has been introduced according to C2(b) to augment the time interval so that the term in the curly brackets becomes equivalent to the density for the third free flight. Further we complete the term in the curly brackets by $S(\boldsymbol{K}_2(t_3), \boldsymbol{k}_3, \boldsymbol{R}_2(t_3))/\lambda(\boldsymbol{K}_2(t_3), \boldsymbol{R}_2(t_3))$ and obtain exactly $P(\boldsymbol{k}_2, \boldsymbol{R}_2(t_2), t_2|\boldsymbol{k}_3, \boldsymbol{R}_3(t_3), t_3)$. Additionally, an integration over $\boldsymbol{k}_3$ must be included in order to keep the value of $f^{(2)}$ unchanged

$$\begin{aligned} f_\Omega^{(2)} = {} & \int_0^\infty \mathrm{d}t_1 \int_{t_1}^\infty \mathrm{d}t_2 \int_{t_2}^\infty \mathrm{d}t_3 \int \mathrm{d}\boldsymbol{k}_1 \int \mathrm{d}\boldsymbol{k}_2 \int \mathrm{d}\boldsymbol{k}_3 \int \mathrm{d}\boldsymbol{k}_0 \int \mathrm{d}\boldsymbol{r}_0 \\ & \times \{f_0(\boldsymbol{k}_0, \boldsymbol{r}_0) P(\boldsymbol{k}_0, \boldsymbol{r}_0, 0|\boldsymbol{k}_1, \boldsymbol{R}_1(t_1), t_1) P(\boldsymbol{k}_1, \boldsymbol{R}_1(t_1), t_1|\boldsymbol{k}_2, \\ & \times \boldsymbol{R}_2((t_2), t_2) P(\boldsymbol{k}_2, \boldsymbol{R}_2(t_2), t_2|\boldsymbol{k}_3, \boldsymbol{R}_3(t_3), t_3)\}\theta(t - t_2)\theta_\Omega(\boldsymbol{K}(t), \boldsymbol{R}(t))\theta(t_3 - t) \end{aligned} \tag{10}$$

indeed, the θ functions do not depend on $\boldsymbol{k}_3$ and the $\boldsymbol{k}_3$-integration can be carried out explicitly to give unity.

In (10) $f^{(2)}$ is expressed as a mean value of the random variable $\theta(\cdot)\theta_\Omega(\cdot)\theta(\cdot)$, having a density function $p^{(2)}$ given by the term enclosed in curly brackets. We further extend $p^{(2)}$ to $p^{(n)}$ by a multiplication with $P(\boldsymbol{k}_3, \boldsymbol{R}_3(t_3), t_3|\boldsymbol{k}_4, \boldsymbol{R}_4(t_4), t_4) \ldots P(\boldsymbol{k}_n, \boldsymbol{R}_n(t_n), t_n|\boldsymbol{k}_{n+1}, \boldsymbol{R}_{n+1}(t_{n+1}), t_{n+1})$ and introducing integrations over $t_4 \ldots t_{n+1}$, $\boldsymbol{k}_3 \ldots \boldsymbol{k}_{n+1}$ in order to keep the value of $f_\Omega^{(2)}$ unchanged. Any term in the series of $f_\Omega^{[n]}$ can be processed in the same way. After summing up all terms we obtain the desired expression

$$\begin{aligned} f_\Omega^{[n]}(t) = {} & \int_0^\infty \mathrm{d}t_1 \ldots \int_{t_n}^\infty \mathrm{d}t_{n+1} \int \mathrm{d}\boldsymbol{k}_1 \ldots \int \mathrm{d}\boldsymbol{k}_n \int d\boldsymbol{k}_0 \int \mathrm{d}\boldsymbol{r}_0 p^{(n)}(\boldsymbol{k}_0, \boldsymbol{r}_0, \boldsymbol{k}_1, \ldots \boldsymbol{k}_n, t_1, \ldots t_{n+1}) \\ & \times \sum_{i=0}^{n} \theta_\Omega(\boldsymbol{K}_i(t), \boldsymbol{R}_i(t))\theta(t - t_i)\theta(t_{i+1} - t) \end{aligned}$$

Now we are ready to analyze the EMC for simulation of $f_\Omega^{[n]}$ in terms of the numerical Monte Carlo for evaluating integrals. The density function is $p^{(n)}$. One trial of $p^{(n)}$ consists of consecutive realizations of the conditional densities P. Any such realization requires a set of random numbers. The consecutive sets form a sequence of random numbers. We say that the process builds up a numerical trajectory, assigned

to this sequence which presents a possible real trajectory. The random variable is

$$\psi = \sum_{i=0}^{n} \theta_\Omega(\boldsymbol{K}_i(t), \boldsymbol{R}_i(t))\theta(t - t_i)\theta(t_{i+1} - t) \tag{11}$$

Only the coordinates at time t remain relevant. Due to the θ functions in time, the terms of the sum are mutually complementary. If during given trial the time $t \in (t_j, t_{j+1})$ is reached, the process can be interrupted since all terms with $i \neq j$ are zero. Then the value of the domain indicator is evaluated for the obtained phase space point at time t. N trials provide N independent values 0 or 1. According to C1, their sample mean gives the Monte Carlo estimate for $f_\Omega^{[n]}$.

There is a finite probability for $t_{n+1} < t$. This probability tends to zero with $n \to \infty$ because the resolvent series converges. Formally we may think n infinite, so that t is reached at any trial. This is consistent since the random variable is zero over the infinite part of numerical trajectory after the time t. Thus, the conclusions for $f_\Omega^{[n]}$ are valid also for f_Ω which leads to the EMC algorithm.

We note that C3 is given not only for completeness. It is the base for extension of the EMC to the weighted EMC algorithm [1,2].

The formulation of the EMC experiment is very short. The outcomes of our experiment are N independent, infinite sequences of random numbers. To any outcome the random variable being the phase space coordinates at time t is assigned by a rather complicated but well defined rule. Further, the function $\theta_\Omega(\boldsymbol{K}, \boldsymbol{R})$ defined in the phase space is given. This function determines another random variable $\psi = \theta_\Omega(\boldsymbol{K}(t), \boldsymbol{R}(t))$ assigned to the outcomes of the experiment. In this way the EMC provides the sample mean $\bar{\psi}$ of N IID variables ψ_i. They have mean value $E\{\psi_i\} = E\{\psi\} = f_\Omega$. They have equal variance $\sigma^2_{\psi_i} = \sigma^2_\psi$ which can be obtained analytically. Indeed, as seen from (11) $\psi^2 = \psi$, such that $E\{\psi^2\} = f_\Omega$. We obtain for the variance

$$\sigma^2_\psi = f_\Omega(1 - f_\Omega) \tag{12}$$

This equation proves that the random variable of the EMC has a finite variance and the precision estimate (3) holds with

$$\sigma^2_\psi(N) = \frac{f_\Omega(1 - f_\Omega)}{N} \tag{13}$$

5. Conclusions

The EMC has been formulated in terms of a combined experiment. f_Ω has been expressed as an multidimensional integral of the form $\int p(x)\psi(x)\,\mathrm{d}x$ where $p(x)$ appears to be the density function used in the EMC for the particle trajectory construction. ψ is the corresponding random variable, having expectation value f_Ω. It has been proved that the variance of ψ is finite and given by the expression (12). It further allows to express the variance of the EMC in terms of f_Ω and the number of the simulated particles N.

Since EMC is a simulated experiment, (13) can be applied to the real transport phenomena with $N = N_{\text{phys}}$ carriers. This can be a useful estimate in the case of small devices where the number of the carriers decreases to the limit where a statistical approach is possible.

References

[1] C. Jacoboni, A new approach to Monte Carlo simulation, in: International Electron Devices Meeting, IEEE Electron Devices Society, Washington, DC, 1989, pp. 469–472.
[2] L. Rota, C. Jacoboni, P. Poli, Weighted ensemble Monte Carlo, Solid-State Electron. 32 (12) (1989) 1417–1421.
[3] M. Nedjalkov, P. Vitanov, Application of the iteration approach to the ensemble Monte Carlo technique", Solid-State Electron. 33 (4) (1990) 407–410.
[4] M. Nedjalkov, I. Dimov, Convergency of the Monte Carlo algorithms for linear transport modeling, Mathematics and Computers in Simulations 47 (1998) 383–390.
[5] C. Jacoboni, P. Lugli, The Monte Carlo Method for Semiconductor Device Simulation, Springer, Wien, New York, 1989.
[6] A. Papoulis, Probabilities, random variables and stochastic processes, Series in Electrical Engineering, 3rd Edition, Mc Graw Hill, New York, 1991.
[7] I. Sobol, Numerical Methods Monte Carlo, Nauka, Moskow, 1973.
[8] F. Rossi, P. Poli, C. Jacoboni, The weighted Monte Carlo approach to electron transport in semiconductors, Semiconduct. Sci. Technol. 7 (1992) 1017.
[9] P. Vitanov, M. Nedjalkov, Iteration approach for solving the inhomogeneous Boltzmann equation, in: J. Miller (Ed.), Numerical Analysis of Semiconductor Devices and Integrated Circuits, Vol. VII, Copper Mountain, Front Range Press, Boulder, 1991, pp. 55–56.