Convergence Estimation for Stationary Ensemble Monte Carlo Simulations

C. Jungemann[1], S. Yamaguchi[2], and H. Goto[2]
[1]Institut für Theoretische Elektrotechnik und Mikroelektronik
Universität Bremen, Postfach 33 04 40, 28334 Bremen, Germany
[2]Fujitsu Limited, Kawasaki, 211-88, Japan (Company Mail No.: C832)

***Abstract*—A criterion for the convergence of Monte Carlo simulations is necessary to ensure the reliability of the results and to guarantee efficiency. Due to the finite scattering rate in Monte Carlo simulations all quantities are in general correlated in time. This makes the estimation of the stochastic error of the sampled statistics difficult. In this work the theoretical basis of a method found in literature is explored which allows to calculate the stochastic error of stationary Ensemble Monte Carlo simulations and which requires only a rough estimate of the magnitude of the largest correlation time of the sampled quantities. The feasibility of the method is demonstrated by application to substrate current calculations for nMOSFETs.**

I. Introduction

Due to the shrinking device dimensions in MOS technologies hot carrier effects and nonlocal transport become more and more important for device simulation [1], [2]. A method capable of accounting for these effects is the Ensemble Monte Carlo (EMC) device simulation based on the semi classical Boltzmann transport equation and Poisson equation [3], [4], [2]. However, this method exhibits not only a numerical error caused by the discretization of the phase space, but also stochastic noise inherent to the MC method [5]–[7], [3]. In general this error decreases with simulation time and MC simulations are performed until a certain degree of convergence is achieved. Thus a reliable convergence criterion is necessary to avoid on the one hand a waste of CPU time in the case that the error is overestimated and on the other hand erroneous results in the case that the error is underestimated.

For the convergence estimation the standard deviation of the sampled quantity is required [5], [8], [6], [7], [9] which in general is difficult to calculate for EMC simulations, because the particles are correlated among each other and in time. The particle-particle correlation is due to the self-consistent solution of the EMC method and the Poisson equation, carrier-carrier scattering and other multi-particle effects or statistical enhancement methods [10]–[12]. This unknown correlation makes it difficult to divide the particle ensemble into independent subensembles necessary for an evaluation of the standard deviation. Due to the finite scattering rate quantities calculated with the MC method are also in general correlated in time. This problem can be circumvented with a method described in [3]. The simulation history is divided into subhistories which are much longer than the correlation time of the quantity and the sampled data are averaged over those subhistories. In this case the resulting samples of the different subhistories are nearly independent (uncorrelated) and the usual textbook formulas for the calculation of the stochastic error can be used [5], [8], [7].

In the next section a brief review of the statistics of correlated quantities is given which requires the precise knowledge of the correlation function [8]. Since the correlation function is normally not known, the above mentioned method [3] is applied. The procedure is explained in detail and its theoretical basis is explored. In the third section this method is then applied to the case of stationary substrate current calculations for nMOSFETs and discussed.

II. Theory

In stationary EMC simulations quantities are estimated by averaging not only over all particles of the ensemble but also by averaging the quantities over the simulation time assuming an ergodic system [8], [3]. This averaging can be achieved by integrating the simulated quantity over time (Other methods like fixed frequency sampling or before scattering statistics [3] can be used as well.) An estimation of the mean value $\overline{\eta}$ of an arbitrary quantity $X(t)$ (a single particle variable (e.g. velocity) or an ensemble average (e.g. substrate current)) is given by [8]:

$$\overline{\eta} = \frac{1}{T_{sim}} \int_0^{T_{sim}} X(t)dt \ , \tag{1}$$

where T_{sim} is the amount of simulated time. This estimator is unbiased [8], since its expected value equals the expected value of the random variable:

$$E\{\overline{\eta}\} = \frac{1}{T_{sim}} \int_0^{T_{sim}} E\{X(t)\}dt = \eta \ . \tag{2}$$

$E\{\}$ is the operation of the calculation of the expected value and η the expected value of the random variable ($\eta = E\{X\}$) [8].

For the calculation of the standard deviation of the estimated mean the correlation function is required [8], [13]:

$$E\{X(t)X(t')\} = \eta^2 + \sigma^2 C(t,t') \ , \tag{3}$$

where $C(t,t')$ is the correlation coefficient and σ the standard deviation of the random variable X [8]:

$$\sigma^2 = E\{(X-\eta)^2\} \ . \tag{4}$$

0-7803-3775-1/97/$10.00 © 1997 IEEE.

The expected value of the standard deviation of the estimated mean $\overline{\sigma}^2$ is [8]:

$$\begin{aligned} E\{\overline{\sigma}^2\} &= E\{(\overline{\eta}-\eta)^2\} = E\{\overline{\eta}^2\} - \eta^2 \\ &= \frac{1}{T_{sim}^2}\int_0^{T_{sim}}\int_0^{T_{sim}} E\{X(t)X(t')\}dt'dt - \eta^2 \\ &= \frac{\sigma^2}{T_{sim}^2}\int_0^{T_{sim}}\int_0^{T_{sim}} C(t,t')dt'dt \ . \end{aligned} \tag{5}$$

In the case that the simulated time T_{sim} is much larger than the correlation time τ_C and the system is stationary [8]:

$$\tau_C = \int_0^{\infty} C(t, t+\tau)d\tau \tag{6}$$

the following approximation holds under quite general conditions [8]:

$$\frac{1}{T_{sim}^2}\int_0^{T_{sim}}\int_0^{T_{sim}} C(t,t')dt'dt \approx \frac{2\tau_C}{T_{sim}} \ . \tag{7}$$

Thus for large T_{sim} the expected value of the standard deviation $\overline{\sigma}$ reads:

$$E\{\overline{\sigma}^2\} \approx \frac{2\tau_C\sigma^2}{T_{sim}} \ . \tag{8}$$

An unbiased estimator for the standard deviation of the estimated mean is given by:

$$\begin{aligned} \overline{\sigma}^2 &= \frac{\int_0^{T_{sim}}\int_0^{T_{sim}} C(t,t')dt'dt}{T_{sim}^2 - \int_0^{T_{sim}}\int_0^{T_{sim}} C(t,t')dt'dt} \\ &\quad \times \left[\frac{1}{T_{sim}}\int_0^{T_{sim}} X^2(t)dt - \overline{\eta}^2\right] \ . \end{aligned} \tag{9}$$

The evaluation of this formula requires in general the exact knowledge of the correlation coefficient which is normally not known. Applying approximation (7) yields:

$$\overline{\sigma}^2 \approx \frac{2\tau_C}{T_{sim}}\left[\frac{1}{T_{sim}}\int_0^{T_{sim}} X^2(t)dt - \overline{\eta}^2\right] \ . \tag{10}$$

For the evaluation of this formula the precise value of the correlation time is still necessary and the simulated time T_{sim} must be much larger than the correlation time τ_C.

To circumvent the problem of the unknown correlation coefficient the scheme reported in [3] is utilized. The mean value is still estimated with (1) apart from a splitting of the integral into N small time steps T:

$$\overline{\eta} = \frac{1}{N}\sum_{i=1}^{N} X_i \tag{11}$$

with:

$$X_i = \frac{1}{T}\int_{(i-1)T}^{iT} X(t)dt \quad \wedge \quad T \gg \tau_C \tag{12}$$

and:

$$T_{sim} = NT \quad \wedge \quad N \gg 1 \ . \tag{13}$$

The estimator for the standard deviation of the estimated mean $\tilde{\sigma}$ now reads:

$$\tilde{\sigma}^2 = \frac{1}{N-1}\left[\frac{1}{N}\sum_{i=1}^{N} X_i^2 - \overline{\eta}^2\right] \ . \tag{14}$$

Equations (11) and (14) resemble the well-known formulas for sampling of independent (uncorrelated) random events [8]. The difference lies in the definition of the sample. Averaging the random variable over the time T which is much larger than the correlation time τ_C not only causes the samples X_i to be (nearly) mutually independent, but also ensures that all available information is sampled.

The expected value of the estimator (14) is:

$$\begin{aligned} E\{\tilde{\sigma}^2\} &= \frac{\sigma^2}{N-1}\left(\frac{1}{T^2}\int_0^{T}\int_0^{T} C(t,t')dt'dt \right. \\ &\quad \left. - \frac{1}{T_{sim}^2}\int_0^{T_{sim}}\int_0^{T_{sim}} C(t,t')dt'dt\right) \end{aligned} \tag{15}$$

If the approximation (7) holds for T, the estimator gives the correct standard deviation (cf. (8)):

$$E\{\tilde{\sigma}^2\} \approx E\{\overline{\sigma}^2\} \approx \frac{2\tau_C\sigma^2}{T_{sim}} \ . \tag{16}$$

Thus the estimator (14) gives an (approximately) unbiased estimation of the standard deviation under the condition that T is much larger than the correlation time τ_C. Since no further information about the correlation coefficient is necessary, the estimator can be used as long as a rough estimate of the magnitude of the correlation time is available. Moreover the scheme works as well when the average over T (12) is evaluated with methods other than time integration, like before scattering statistics or fixed frequency sampling. To ensure efficiency the sampling frequency should be larger than the inverse correlation time.

With the standard deviation of the estimated mean the convergence of the MC simulation can be calculated. From the central limit theorem it follows that due to the averaging, the distribution function of the estimated mean approaches a Gaussian distribution [8]. Thus the probability (confidence level) that the estimated mean is within an interval $[\eta-\delta, \eta+\delta]$ (confidence interval) around the expected value is approximately [5], [8], [7]:

$$\begin{aligned} &P_{con}\{\eta-\delta \le \overline{\eta} \le \eta+\delta\} \approx \\ &\quad \frac{1}{\sqrt{2\pi\overline{\sigma}^2}}\int_{\eta-\delta}^{\eta+\delta} \exp\left(-\frac{(x-\eta)^2}{2\overline{\sigma}^2}\right) dx \ . \end{aligned} \tag{17}$$

If δ is chosen to be two times the standard deviation, the confidence level is 95.45% [5]. In the following the (estimated) relative error:

$$r_{err} \approx \frac{2\tilde{\sigma}}{\overline{\eta}} \approx \frac{2\overline{\sigma}}{\overline{\eta}} = \frac{\delta}{\overline{\eta}} \tag{18}$$

will be used [9]. MC simulations are stopped, after a given relative error has been achieved.

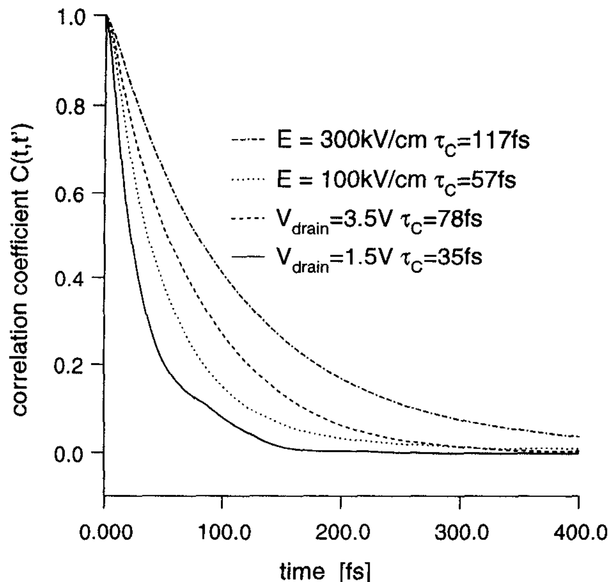

Fig. 1. Correlation coefficients of the substrate current in a $0.16\mu m$-nMOSFET ($V_{gate} = 1.5V$) and of the impact ionization coefficient in a homogeneous bulk system.

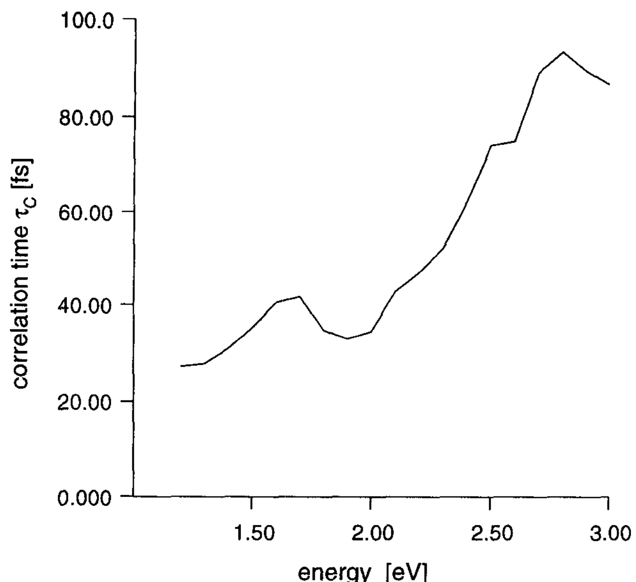

Fig. 2. Correlation time of the impact ionization coefficient for a field free homogeneous bulk system as a function of the injection energy.

III. Discussion and Results

Equations (11), (14) and (18) have been implemented into the in-house full-band EMC device simulation program FALCON for the evaluation of substrate currents in nMOSFETs [14]. With this program stationary nonself-consistent (regarding the electric field) simulations for silicon devices at room temperature have been performed [15], [14].

In Fig. 1 the correlation coefficient of the substrate current is shown for a $0.16\mu m$-nMOSFET ($V_{gate} = 1.5V$). Moreover correlation coefficients of the impact ionization coefficient in a homogeneous bulk system are shown. The corresponding correlation times are $35fs$ ($V_{drain} = 1.5V$), $78fs$ ($V_{drain} = 3.5V$), $57fs$ ($E = 100kV/cm$), and $117fs$ ($E = 300kV/cm$). The increase of the correlation time with the electric field or V_{drain} can be understood with the help of Fig. 2. In this graph the correlation time for a particle ensemble which is injected with a given energy into a field free bulk system (quantum yield experiment) as a function of this energy is shown. High energetic particles have a longer correlation time than low energetic particles. Since a high electric field results in a large fraction of high energetic particles, the correlation time of the substrate current increases with the electric field.

Since the correlation time exhibits a complicated dependency on the simulation conditions, it is not possible to obtain the correlation time before the actual MC simulation. On the other hand only a rough estimate of the correlation time is required to ensure that the sampling time T is much larger than the correlation time. With Figs. 1 and 2 a good guess of the magnitude of the correlation time of impact ionization in silicon can be made which is about $100fs$.

In Fig. 3 the relative error (18) of the impact ionization coefficient for a homogeneous bulk system is shown as a function of the time T. The total simulated time T_{sim} is $1000ps$, the number of particles 100, and the electric field $300kV/cm$. Assuming that the correlation coefficient in this case can be approximated by an exponential function ($C(t,t') = \exp(-|t-t'|/\tau_C)$) the relative error estimated with (18) can be calculated:

$$r_{err} = \frac{2\tilde{\sigma}}{\overline{\eta}} \approx \frac{2\overline{\sigma}}{\overline{\eta}}\sqrt{1 - \frac{1-\exp(-\frac{T}{\tau_C})}{\frac{T}{\tau_C}}}\,. \qquad (19)$$

As expected the relative error saturates for large times T and approaches the correct value of about 1.44% apart from statistical noise. For times T which are smaller than $1ps$ the condition that T must be much larger than the correlation time τ_C (here $117fs$) is violated. Thus $\tilde{\sigma}$ does not equal the standard deviation of the estimated mean $\overline{\sigma}$ and the relative error is considerably underestimated. On the other hand the time T should not be too long, since this reduces the total number of samples and increases therefore the statistical uncertainty of the estimated standard deviation.

In Fig. 4 the substrate current of a $0.16\mu m$-nMOSFET is shown. The calculations are converged with a probability of 95% within $\pm$ 10%. The time T has been chosen to be $10ps$, considerably larger than the correlation time of the substrate current. The given CPU time values are for a Sun Sparc 20 workstation. The dependence of the CPU time on the drain voltage results from the fraction of high energetic carriers which increases with the drain voltage, thus improving the signal to noise ratio. Since no statistical enhancement method has been used for these simulations, the CPU time increases considerably when the threshold energy of impact ionization is approached. This demonstrates clearly that it is impossible to guess the necessary simulation time without a convergence es-

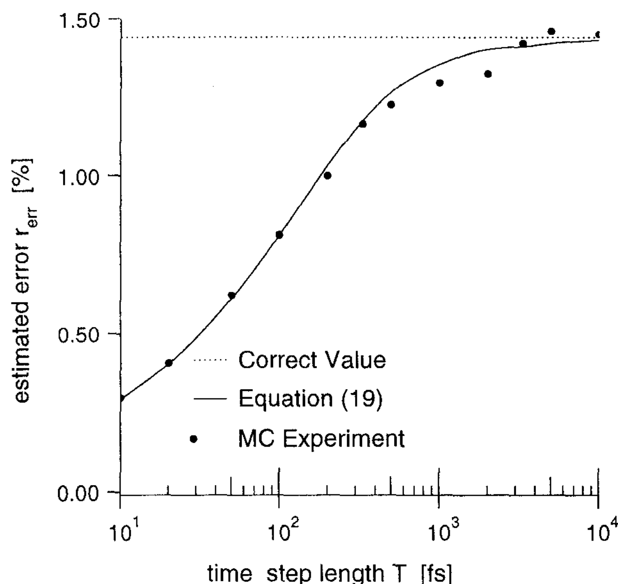

Fig. 3. Relative error of the impact ionization coefficient as a function of the time T for a homogeneous bulk system, an electric field of $300kV/cm$, 100 particles and a total simulated time of $1000ps$.

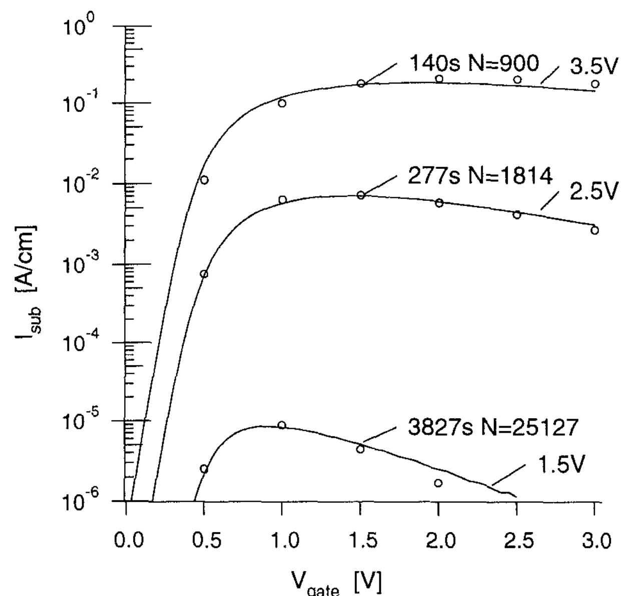

Fig. 4. Simulated substrate current (symbols) of a $0.16\mu m$-nMOSFET for different drain voltages and experimental results (solid lines) [14]. CPU times are for a Sun Sparc 20 workstation and N is the number of time steps T.

timation.

Since the simulations are stopped when a certain level of convergence is achieved, the CPU time consumption is minimized. This resulted in the shortest CPU times reported so far for substrate current calculations [9], [14]. Of course the method is not limited to the case of substrate current calculations, but can be used for other quantities like drift velocity, relaxation times, gate currents as well.

IV. Conclusions

By applying the basic principles of estimation theory the theoretical basis of the standard-deviation-estimation method reported in [3] has been clarified. It has been shown that the scheme can be used in an easy and reliable way for the estimation of the stochastic simulation error of stationary Monte Carlo simulations. The estimator has the advantage that only a rough estimate of the magnitude of the correlation time is necessary which can be obtained easily. Furthermore the estimator is efficient because all information available in the Monte Carlo simulation is sampled. The method is universal and its feasibility has been demonstrated for the case of substrate current calculations.

Acknowledgments

The authors are grateful for discussions with S. Decker (University of Bremen).

References

[1] C. Jacoboni and P. Lugli, *The Monte Carlo Method for Semiconductor Device Simulation*, Springer Verlag, Wien, 1989.

[2] K. Hess, Ed., *Monte Carlo Device Simulation: Full Band and Beyond*, Kluwer Academic Publishers, Boston, 1991.

[3] C. Jacoboni and L. Reggiani, "The Monte Carlo method for the solution of charge transport in semiconductors with application to covalent materials", *Rev. Mod. Phys.*, vol. 55, pp. 645–705, 1983.

[4] R.W. Hockney and J.W. Eastwood, *Computer Simulation Using Particles*, Institute of Physics Publishing, Bristol and Philadelphia, 1988.

[5] J. M. Hammarsley and D. C. Handscomb, *Monte Carlo Methods*, Methuen/Chapman and Hall, London, 1964.

[6] S. M. Ermakow, *Die Monte-Carlo-Methode und verwandte Fragen*, R. Oldenbourg, München and Wien, 1975.

[7] R. Y. Rubinstein, *Simulation and the Monte Carlo method*, Wiley series in probability and mathematical statistics. John Wiley & Sons, New York, 1981.

[8] A. Papoulis, *Probability,Random Variables and Stochastic Processes*, Mc Graw-Hill, 2 edition, 1984.

[9] C. Jungemann, S. Yamaguchi, and H. Goto, "Efficient Full Band Monte Carlo hot-carrier simulation for silicon devices", in *Proc. ESSDERC*, 1996, pp. 821–824.

[10] A. Phillips and P. J. Price, "Monte Carlo calculations on hot electron energy tails", *Appl. Phys. Lett.*, vol. 30, pp. 528–530, 1977.

[11] M. G. Gray, T. E. Booth, T. J. T. Kwan, and C. M. Snell, "A multi-comb variance reduction scheme for Monte Carlo semiconductor simulators", *IEEE Trans. Electron Devices*, 1997.

[12] C. Jungemann, S. Decker, R. Thoma, W. L. Engl, and H. Goto, "Phase space multiple refresh: A general purpose statistical enhancement method for Monte Carlo device simulation", *IEEE Trans. Semi. Con. Mod. Sim.*, vol. http://www.ieee.org/journal/tcad, 2, 1997.

[13] C. W. Gardiner, *Handbook of Stochastic Methods*, Springer, Berlin, 1985.

[14] C. Jungemann, S. Yamaguchi, and H. Goto, "On the accuracy and efficiency of substrate current calculations for sub-μm nMOSFETs", *IEEE Electron Device Lett.*, vol. 17, pp. 464–466, 1996.

[15] J. M. Higman, K. Hess, C. G. Hwang, and R. W. Dutton, "Coupled Monte Carlo-drift diffusion analysis of hot-electron effects in MOSFET's", *IEEE Trans. Electron Devices*, vol. 36, pp. 930–937, 1989.