SIMULATION OF SEMICONDUCTOR DEVICES AND PROCESSES Vol. 4
Edited by W. Fichtner, D. Aemmer - Zurich (Switzerland) September 12-14, 1991 - Hartung-Gorre

Analysis of Filter Techniques for Monte-Carlo Device Simulation

Hans Kosina and Siegfried Selberherr
Institute for Microelectronics, Technical University of Vienna
Gußhausstraße 27-29, A-1040 Vienna, Austria

Abstract

A method for assigning Monte-Carlo calculated quantities to nonuniform grids is presented. Assignment is done in a computationally efficient way by convolution allowing both even and odd weighting functions. The problem of data normalization and differentiation of data with statistical origin is analyzed. The method is applied to study nonlocal transport effects in submicron semiconductor structures.

1 Introduction

Monte–Carlo (MC) simulation of semiconductor devices is based upon the Particle–Mesh method. Statistical quantities are assigned to a grid on which discretized differential equations are to be solved subsequently. In its original formulation the Particle–Mesh method consists of four steps [1]:

(1) assign charge to the grid
(2) solve the equation for electrostatic potential on the grid
(3) calculate force field at the grid points
(4) interpolate the forces at the particle positions

Low order charge assignment methods such as Nearest-Grid-Point scheme (NGP) and Cloud-in-Cell scheme (CIC) are widely used [1][2]. These schemes employ piecewise constant and piecewise linear assignment functions. One major drawback of using higher order assignment functions would be the drastically increased interpolation effort.

We present a method which separates the recording of data during MC-runtime, and their assignment to a nonuniform grid. Assignment by convolution is done only once after the MC-simulation. Thus the complexity of the weighting function has negligible influence on the total computation time. Choosing odd weighting functions this method further provides a means for consistent differentiation of the MC-quantities. The assignment of both the quantity and its derivative to the nonuniform grid is fully independent of the grid spacing.

2 Method

In the cloud shape interpretation of charge assignment the charge of a particle is considered to be spatially distributed according to a cloud shape function. This shape depends on the assignment scheme. The area of overlap of the cloud with a certain grid cell determines the amount of charge assigned to this cell. Since the areas of overlap with the surrounding grid cells have to be evaluated for each particle and since the number of particles is very large,

the complexity of the function describing the cloud shape has much impact on the required computation time.

It should be noted that in the stationary case the motion of just one particle needs to be calculated. Since we now have an ensemble of scattering events instead of an ensemble of real particles, each scattering event, as it has a well determined position in x-space, can be imagined to carry a cloud and the above described method can be applied.

In the assignment function interpretation, each grid point carries a replica of the assignment function. The value of this function at the position of a particle determines the fraction of charge assigned to the grid point.

This interpretation is the starting point of our method. We introduce an intermediate step by firstly assigning the quantities of interest to a very fine grid. The computational cheapest NGP assignment scheme is sufficient. In this way during the MC-simulation the MC raw-data is recorded. This fine grid is fully independent of the nonuniform grid. The raw data is assumed to be represented by a step-shaped function R(x) which is constant inside the intervals of the very fine grid. Assignment of a quantity to a certain point x inside the simulation domain is done by convolution of the raw data function R(x) with an assignment function $W(x)$. $W(x)$ is different from zero just in the interval $[-H/2, H/2]$, where H is the filter width.

$$Q(x) = \int_{x-H/2}^{x+H/2} R(x')\, W(x'-x)\, dx' \tag{1}$$

This convolution integral has to be evaluated at the discrete points of the nonuniform grid only once after the MC-simulation. The number of gridpoints is several orders of magnitude lower than the number of particles. Therefore evaluating (1) at the gridpoints is much less expensive than letting each particle carry its own cloud shape. Two dimensional weighting functions are built up by the product of two one-dimensional ones.

The derivative of the quantity $Q(x)$ is obtained by convolving the raw data function with the negative derivative of the weighting function, as easily can be seen by differentiating (1) with respect to x.

In our implementation we have taken a cosine filter with the weighting function

$$W(x) = \frac{1}{H}\left(1 + \cos\left(\frac{2\pi x}{H}\right)\right).$$

In principle any even function which has a maximum at $x = 0$ and which vanishes outside the interval $[-H/2, H/2]$ would be useful. For differentiation we consequently have to use a sine filter. In general the weighting function for differentiation has to be an odd function.

3 Normalization and Differentiation

The MC raw-data recorded on the fine grid is in its nature a density. In order to obtain a quantity per particle (energy, velocity), this density has to be normalized by the particle density. In that context it is important first to assign both the density under consideration and the particle density to the nonuniform grid and then to perform normalization. Ordering the two operations in this way maintains the statistical significance of the different grid cells, as they are populated differently. In the alternative method — first normalizing the raw-data and then doing assignment — the contents of the grid cells of the fine grid would be weighted by the assignment function independently of the particle number contained by each cell.

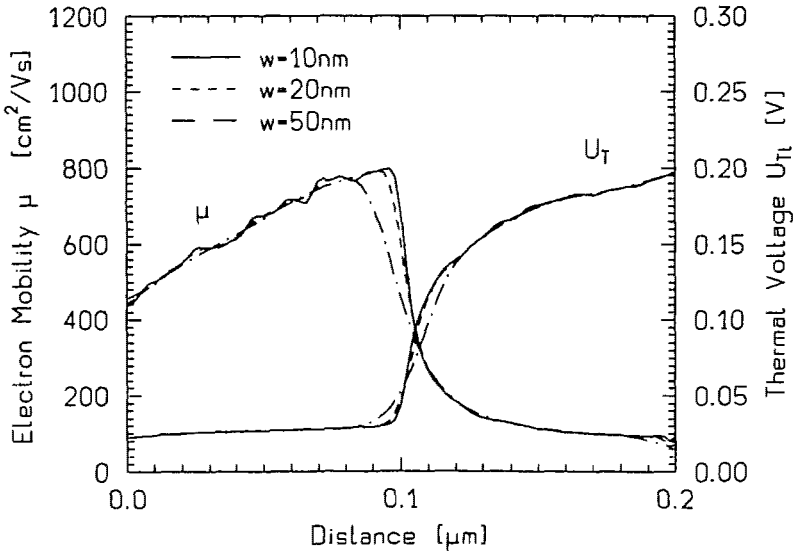

Figure 1: Electron mobility and thermal voltage for three different values of the filter width. A field step at $x = 0.1\mu m$ with $E1 = 10^3 V/cm, E2 = 10^5 V/cm$ is assumed.

In the following $CF\{\}$ and $SF\{\}$ denote application of cosine filter and sine filter respectively. Then normalization can be written as

$$Q(x) = \frac{CF\{R\}}{CF\{n\}}. \tag{2}$$

For derivation of a normalized quantity one has to differentiate (2). Inserting the convolution integral (1) yields:

$$\frac{dQ}{dx} = \frac{d}{dx}\frac{CF\{R\}}{CF\{n\}} = \frac{SF\{R\}\,CF\{n\} - CF\{R\}\,SF\{n\}}{CF\{n\}^2} \tag{3}$$

With this method the derivative can be calculated directly from the MC raw-data, independently of the spacing of the nonuniform grid.

4 Discussion

The method presented decouples several tasks. Firstly the design of the nonuniform grid can be done by solely considering the discretization constraints of the partial differential equations. The grid is allowed to be adapted during the solution process.

Secondly, the choice of the filter width depends on the required accuracy of the MC-simulation. The smaller the filter width, the more accurate is the MC-simulation and hence the more particles have to be used. In some cases less accuracy may be sufficient. Then the amount of particles can be reduced and in turn a larger filter width has to be chosen.

Thirdly the spacing of the fine grid can be chosen freely in a wide range. A lower limit for the grid spacing is not implied by the method. However consumption of memory would prevent too fine grids. The upper limit is related to the filter width, since the weighting function should cover a sufficient large number of grid cells. A minimum number of ten seems useful.

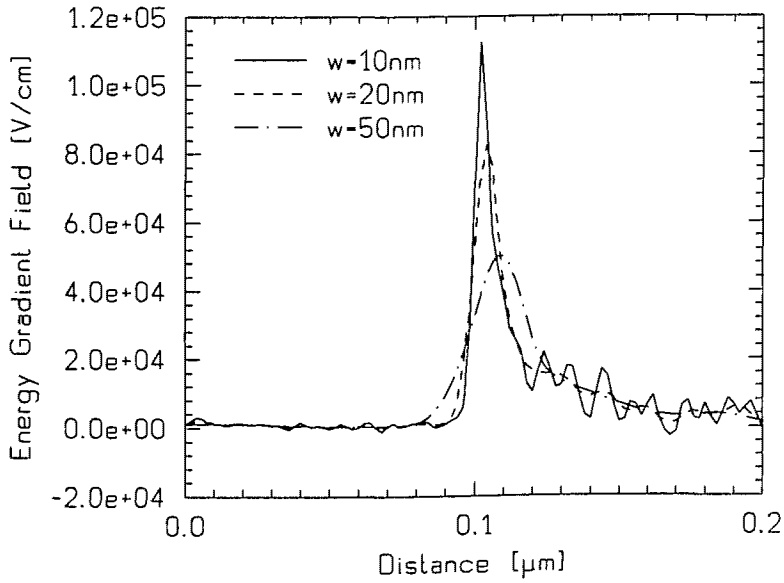

Figure 2: The energy gradient field is more sensitive with respect to the filter width. Although the peak value changes by more than a factor of two, the integral over the different curves is conserved.

5 Examples

To study the transport process in submicron structures we have calculated the following quantities:

$$U_T = q^{-1} < \hbar k_x v_x >$$

$$E_{eg} = \frac{dU_T}{dx}$$

$$(\frac{dp}{dt})_c = < \tau_m^{-1} \hbar k >$$

$$\mu = q v_d \; / \; (\frac{dp}{dt})_c$$

Here U_T denotes the thermal voltage, E_{eg} the energy gradient field and $(\frac{dp}{dt})_c$ the momentum loss rate. Furthermore v_d is the electron drift velocity and μ the mobility. The role of the energy gradient field is discussed in [3]. The quantities $(\mu(x), U_T(x), E_{eg}(x))$ are parameters in a general current relation including nonlocal effects which reads for electrons

$$J = q \; n \; \mu \; (E + E_{eg} + \frac{U_T}{n} \frac{dn}{dx}).$$

In a first example we have simulated a $0.2\mu m$ structure, assuming a field step at $x = 0.1\mu m$ $(E1 = 1kV/cm, E2 = 100kV/cm)$. The MC-simulator is based on the physical model of Si mainly as it is described in [4].

Fig. 1 depicts the nonlocal mobility which is the ratio of drift velocity and momentum loss rate and the thermal voltage for different filter widths. A filter width in the range of $20nm$

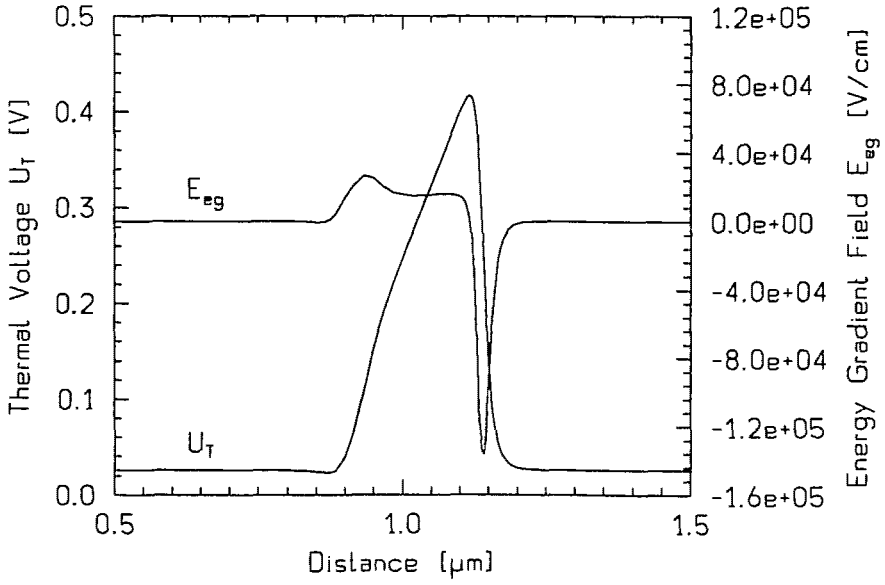

Figure 3: Thermal voltage and energy gradient field in an n^+-n-n^+ structure. The nonsymmetric profiles indicate electron heating over most of the n^- area and rapid cooling at the $n-n^+$ junction.

to $50nm$ seems to be a good compromise between smoothing the raw data and maintaining steep gradients. The energy gradient field (EGF) calculated by (3) is shown in Fig. 2. The field peak approaches values of $80kV/cm$ and $50kV/cm$ for a filter width of $20nm$ and $50nm$ respectively. Due to consistency of (2) and (3) the integral over the EGF is simply determined by the difference of the boundary values of U_T. As can be seen in Fig. 1, this difference is nearly independent of the filter width. Thus the area under the different field peaks is conserved.

In a second example a n^+-n-n^+ structure with an n-region of $0.25\mu m$ has been analyzed. The n region is doped with a concentration of $2\cdot10^{15}cm^{-3}$ and the n^+ regions with $5\cdot10^{17}cm^{-3}$.

Fig. 3 shows that the electrons in the n-region are heated up to $0.4eV$ and are cooled rapidly at the $n-n^+$ junction. This results in a peak energy gradient field of $-130kV/cm$. Derivation was done according to (3) with a filter width of $40nm$. In Fig. 4 the electric field and the momentum loss rate are compared. Due to the built-in potential of the n^+-n junction the electrons have to surmount a potential barrier, which has in the present example a value of $110meV$. Therefore to simulate this structure a particle split algorithm [5] has to be used.

6 Conclusion

A generalized method for the assignment of MC-calculated quantities and their derivatives to nonuniform grids has been proposed. The method is suited for problems of arbitrary dimension in real space and for abrbitrary grid types. The unique parameter filter width is an indicator for the accuracy of the MC-simulation.

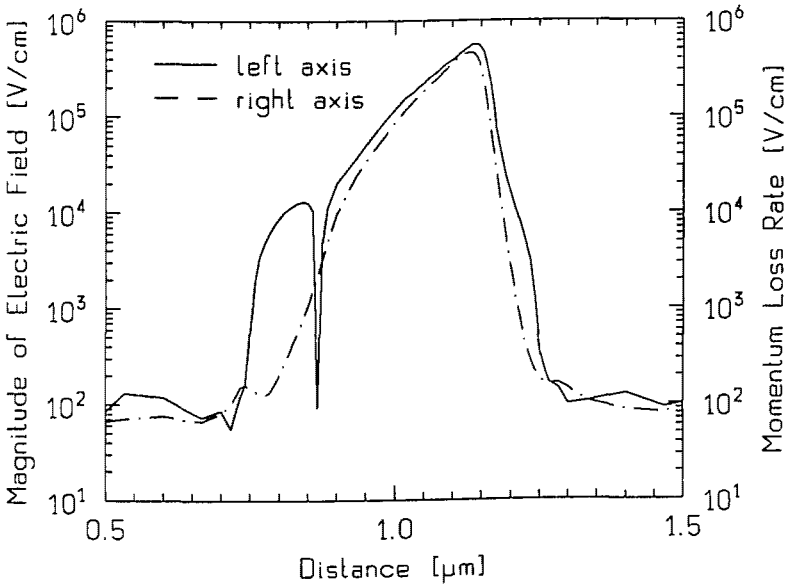

Figure 4: Magnitude of electric field and momentum loss rate in the same structure as in Fig. 3. The lobe in the electric field at about $0.8\mu m$ has opposite sign and forms a retarding barrier for electrons.

Acknowledgements

This work is supported by: AUSTRIAN INDUSTRIES – AMS International Unterpremstätten, Austria; DIGITAL EQUIPMENT Corporation Hudson, USA; SIEMENS Corporation Munich, FRG; and SONY Corporation Atsugi, Japan.

References

[1] R.W. Hockney, and J.W. Eastwood, *Computer Simulation using Particles*, New York: McGraw-Hill 1981.

[2] M. Tomizawa, K. Yokoyama and A. Yoshii, *Nonstationary Carrier Dynamics in Quarter-Micron Si MOSFET's*, IEEE Trans. on Computer Aided Design, CAD-7, No.2, pp.245-258, 1988.

[3] P.T. Landsberg, *D grad ν or grad(Dν) ?* , J. Appl. Phys., Vol.56, No.4, pp.1119-1122, 1984.

[4] 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., Vol.55, No.3, pp.698-705, 1983

[5] A. Phillips, and P.J. Price, *Monte Carlo calculations on hot electron energy tails*, Appl. Phys. Lett. , Vol.30, No.10, pp.528-530, 1977