

3D Simulations of Ultra-small MOSFETs with Real-space Treatment of the Electron–Electron and Electron-ion Interactions

W. J. GROSS^a, D. VASILESKA^{b,*} and D. K. FERRY^b

^a*Intel Corp., Chandler, AZ 85226;* ^b*Department of Electrical Engineering,
Arizona State University, Tempe, AZ 85287-5706*

(Received 16 December 1998; In final form 14 December 1999)

We present a 3D Ensemble Monte Carlo particle-based simulator with a novel real-space treatment of the short-range electron–electron and electron-ion interactions. By using a corrected Coulomb force in conjunction with a proper cutoff range, the short-range portion of the force is properly accounted for, and the ‘double counting’ of the long-range interaction is eliminated. The proposed method naturally incorporates the multi-ion contributions, local distortions in the scattering potential due to the movement of the free charges, and carrier-density fluctuations. The doping dependence of the low-field mobility obtained from 3D resistor simulations closely follows experimental results, thus supporting the appropriateness of the proposed scheme. Simulations of ultra-small MOSFETs demonstrate that the short-range electron–electron and electron-ion interactions are responsible for the fast thermalization of the carriers at the drain end of the device, which occurs over distances that are on the order of few nanometers. The omission of the short-range portions of these two interaction terms leads to significant overestimation of the distance over which carriers thermalize.

Keywords: 3D device simulations, discrete impurity effects, electron–electron and electron-ion interactions

1. INTRODUCTION

Transport in semiconductors has traditionally been described with the help of the Boltzmann Transport Equation (BTE). Powerful techniques have been developed to solve the BTE, and combined with Poisson equation solvers, these

methods form the basis for the theoretical analysis and design of modern semiconductor devices. In its most common formulation for semiconductors, the BTE is given as [1]

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{r}} f + \frac{d\mathbf{k}}{dt} \cdot \nabla_{\mathbf{k}} f = \left. \frac{\partial f}{\partial t} \right|_{\text{scatt}}, \quad (1)$$

*Corresponding author. Tel.: (480) 965-6651, e-mail: vasilesk@imap2.asu.edu

where \mathbf{v} is the carrier velocity, \mathbf{k} is the crystal momentum and $f(\mathbf{r}, \mathbf{k}, t)$ is the carrier distribution function, which gives the density of particles with momentum \mathbf{k} at the point \mathbf{r} at time t . The terms on the left-hand side describe the change in the distribution function with respect to time, concentration gradient, and applied field. The right-hand side represents the dissipation terms in the system, which include various scattering mechanisms that balance the driving terms on the left.

The simplest approximation to the BTE is the drift-diffusion (DD) model, in which the current density for each carrier (electron, hole) is the sum of a drift and a diffusion component [2, 3]. One of the main deficiencies of the DD model is its failure to account for carrier heating and velocity overshoot effects. Another rather popular method, which overcomes some of the limitations of the DD model, is the hydrodynamic model (HD). The hydrodynamic equations are easily formulated by calculating various moments of the BTE. The zeroth, first, and second-order moments of the BTE, which describe particle, momentum and energy conservation, are discussed in detail in [4], although the basic ideas are really nearly 100 years old. Since the HD equations take electron energy into account, they can produce better results in high-field conditions. Quantum corrections to the HD model also have been developed [5]. It is important to note that one of the difficulties in using the HD approach is in the numerical nature of the equations. For example, when the average carrier velocity exceeds certain limiting values, the conservation laws become hyperbolic in nature, which can lead to formation of numerical shock waves, that require very careful algorithmic implementation. Even if no shock forms, the same problem arises if space charge domains arise due to, for example, the Gunn effect. A way to avoid this problem is to use the essentially non-oscillatory (ENO) [6], or the weighted ENO scheme [7], that allow for adaptive discretization stencils.

The limitations of the continuum models (DD and HD) lie in the fact that the distribution function is not known *a priori*. Calculations that rely

on the knowledge of the distribution function, such as electron–electron interactions, will produce erroneous results based upon the actual form of the distribution function used. The accuracy of the models can be enhanced by increasing the order/moment used at the expense of a more complex set of equations to solve. Another major drawback of the continuum models is that electron–electron and electron-ion interactions can only be treated as an additional scattering mechanism in the mobility model [8], or the momentum and the energy relaxation times. Therefore, by using continuum modeling, it is impossible to fully account for carrier-ion interactions that lead to multiple scattering events and cause local variations in the electron and ion densities (a “derivation” of the BTE neglects multi-particle correlations).

The Ensemble Monte Carlo (EMC) method, which solves the Boltzmann transport equation [1, 9–21], overcomes most of the difficulties related to the carrier heating effects, since it does not make any assumption upon the form on the distribution function. Within the EMC scheme, point-like particles are used to represent electrons (or holes) within the device. The momentum and energy of each particle within the ensemble are continuously updated, as is the real space position of the particle and any other relevant dynamic parameter. The updated momentum (and corresponding energy) results from the various forces applied to the particle, whereas the real-space movement of the electrons through the device depends upon the local velocity that results from the electric field obtained through the self-consistent solution of the Poisson equation.

Due to the real-space tracking of the particle position, the EMC method allows for more exact treatment of the Coulomb interaction between charged particles (particle-ion and particle–particle interactions) through the addition of a molecular Dynamics (MD) loop [22–25]. This coupled EMC-MD scheme has been shown to give simulation mobility results in excellent agreement for multiple scattering situations [25] and with the

experimental data for bulk samples with high substrate doping levels [26]. It has been corrected for both degeneracy [26] and the many-body exchange interaction [27]. However, it has proven to be quite difficult to incorporate the coupled EMC-MD approach when inhomogeneous charge densities, characteristic of semiconductor devices, are encountered [28]. An additional problem with the use of the coupled EMC-MD approach, in a typical particle-based device simulation, arises from the fact that both the electron-electron ($e-e$) and electron-ion ($e-i$) interactions are already included in the self-consistent potential, at least within the Hartree approximation (the long-range carrier-carrier interaction enters the Poisson equation through the real part of the Hartree energy). The magnitude of the resulting so-called mesh force depends upon the volume of the cell and, for commonly employed mesh sizes in device simulations, usually leads to ‘double counting’ of the force if a separate Coulomb interaction is added to the EMC transport kernel [29]. The proper treatment of the Coulomb interactions is very important for modeling ultra-small devices, in which, as shown experimentally [30, 31] and confirmed with numerical simulations [32–36], fluctuations in the number and the position of the dopant atoms within the active region of the device can lead to significant fluctuations in what should be identical devices fabricated on the same chip.

It is important to note that, while traditionally the direct solution of the BTE through EMC simulation has been considered computationally inefficient when compared to approximate moment equation methods, in 3D (which must be considered for sub-70 nm devices) this advantage is no longer clear, and indeed, full 3D simulation of FET structures using EMC techniques indicate that the performance bottleneck arises from the 3D Poisson equation solver, rather than the transport kernel [37].

This paper is organized as follows: In Section 2, we briefly describe the general features of our 3D particle-based simulator. The technique used to account for the short-range portions of the $e-e$

and the $e-i$ interactions and its inclusion is described in Section 3. Here, we also give some representative results for the low-field electron mobility found in resistor simulations, and from simulations of ultra-small MOSFETs. Conclusions derived from the work presented here and some highlights of future directions of research are discussed in Section 4.

2. DESCRIPTION OF THE SIMULATOR

The basic steps involved in a prototypical EMC device simulator are summarized in Figure 1. In steady-state, the drain current is calculated via the net number of particles crossing the drain contact per unit time. By performing several computer runs, the averaged current-voltage characteristics of the device under investigation are obtained. It is important to mention that there are several constraints that must be met when choosing the mesh size and the time step [28]. Small mesh sizes lead to large arrays that require extensive memory usage and long iteration times. Yet, the mesh

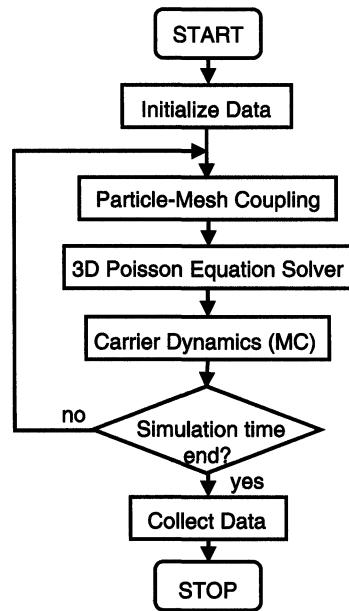


FIGURE 1 A flow-chart of a generic EMC particle-based simulator.

size must be sufficiently small to account for fluctuations in doping, carrier concentration, and electric fields. These two requirements imply that the mesh spacing be on the order of the extrinsic Debye length. Having too small a mesh spacing can create artificially large fluctuations in the carrier and dopant concentrations. The time-step should be small enough (on the order of the inverse plasma frequency) to avoid any nonphysical behavior, such as plasma oscillations. In addition, the force seen by an electron needs to be frequently updated to minimize errors due to carrier movement. For example, the time step should be smaller than the time required to cross a mesh boundary. The constraints described above lead to time steps in the femto-second range and mesh spacings in the 5–10 nm range.

2.1. Particle-mesh Coupling

In the case of DD or HD models, the Poisson equation solver and the transport kernel (either DD or HD) usually use different solution algorithms and are coupled by a mere subroutine call. The carrier concentrations at each node are variables that are passed between the two kernels. The solution is found when, for example, the potential update reaches a predetermined level. In the case of particle-based simulators, there are additional factors to be considered. These factors are described below.

2.1.1. Charge Assignment

The calculation of carrier concentration at the node points is essential to solving Poisson's equation on the mesh. Hockney and Eastwood [38] outline several methods for charge assignment, all of which have been implemented in our simulator (although only one is used at a time). The simplest one is the nearest-grid-point (NGP) method, in which the total charge of the particle is assigned to the nearest grid point. A more accurate method is the cloud-in-cell (CIC) scheme in which each grid point of the “cell”, in which the electron

is located, is assigned a charge weighted by the distance from the electron to that point. This produces a smoother electron concentration over the nearest mesh points. One artifact of this method is the creation of a “self-force” [39] if the grid spacing is nonuniform or the permittivity is not constant. Laux [40] proposed a nearest-element-center (NEC) scheme that puts the charge of the electron in the middle of the current mesh “cell”. This method eliminates the self-force, but causes fluctuations in the carrier concentration as the electrons move across cell boundaries. These fluctuations can increase the time required to solve Poisson's equation at each iteration step. Hybrid schemes using the NEC and CIC methods along different axes in multi-dimensional simulations have shown improved results. The CIC and NEC methods are first-order methods in that only the nearest node points are considered.

For a small mesh spacing, large fluctuations in carrier concentrations can occur from node to node, since the number of nodes may be greater than the actual number of simulated particles. For 1-D and 2-D simulations, the number of particles can be chosen such that the fluctuations are minimized, and the charge of the particles is determined from the device size along the “constant” axes [28]. A ratio of 10 particles or more per grid point helps reduce the fluctuations. However, the use of molecular dynamics requires that the particle charge is set to unity. In this case, the number of particles is limited by the device size and doping concentrations. For 3D simulations, with a mesh spacing of 10 nm in all dimensions, a source/drain doping of 10^{19} cm^{-3} results in 10 electrons per node point. For a uniform mesh spacing of 5 nm, this reduces to 1.25 electrons per node point. Thus, the desired “higher” accuracy of the low mesh spacing results in larger fluctuations from node to node (and iteration to iteration).

2.1.2. Electric Field Calculation

During the EMC iteration, electrons are accelerated by the Coulombic forces (due to other

electrons and impurities) and by the electric fields determined by the solution of Poisson's equation. The electric field seen by each electron must be calculated at the beginning of each EMC iteration. To minimize self-scattering, the interpolation scheme for electric field calculation should be the same as the one used for charge assignment [38, 40]. For example, for the CIC scheme, the electric field at each node of the cell is weighted by the distance between the electron and the node. The same weighting factors used to assign the charge from the electron to the node points must be used for field interpolation. As discussed in more detail in Section 3, the short-range portion of the Coulomb forces between an electron and surrounding electrons and ions must be added to the mesh force to properly account for the Coulomb interactions.

2.1.3. Boundary Conditions

As electrons move through the device, they will often encounter the device boundaries. Therefore, to obtain meaningful device measurements, the treatment of these boundaries must be physical in nature. A reflecting boundary condition is employed in non-critical areas of the device (artificial boundaries), such as the bottom and sides of the device away from the contacts. Electrons that cross these boundaries are reflected back into the structure, with an equal velocity and opposite direction normal to the boundary. For reflections at the semiconductor – oxide interface in the channel, the reflected angle can be randomized to take into account surface scattering. A combination of these two techniques was used with 50% probability of each occurring, similar to the method used by Fischetti and Laux [28]. This is an unsatisfactory approach, however, and is an area where further work is needed, especially as scattering at the interface is known to be an important mobility limiting process in ultra-small MOSFETs due to the closer confinement of the carriers to the interface [41].

2.1.4. Ohmic Contacts

The simulation of contacts requires a more careful consideration than the reflecting boundaries, because electrons crossing the source and drain contact regions are counted towards the corresponding terminal current. In order to conserve charge in the device, the electrons exiting the contact regions must be re-injected. Several common models for contacts include [42]:

- Electrons are injected at the opposite contact with the same energy and wavevector \mathbf{k} . If the source and drain contacts are in the same plane, as in the case of MOSFETs simulations, the sign of \mathbf{k} , normal to the contact will change. This is an unphysical model, however [43].
- Electrons are injected at the opposite contact with a wavevector randomly selected based upon a thermal distribution.
- Contact regions are considered to be in thermal equilibrium. The total number of electrons in a small region near the contact are kept constant, with the number of electrons equal to the number of dopant ions in the region.
- Another method uses ‘reservoirs’ of electrons adjacent to the contacts. Electrons naturally diffuse into the contacts from the reservoirs, which are not treated as part of the device during the solution of Poisson’s equation. This approach gives results similar to the velocity weighted Maxwellian [42], but at the expense of increased computational time due to the extra electrons simulated.

There are also several possibilities for the choice of the distribution function – Maxwellian, displaced Maxwellian, and velocity-weighted Maxwellian. Here, we use the charge-neutral method, which is simpler than the reservoir method, while allowing for rather accurate modeling of the contacts.

2.2. 3D Poisson Equation Solvers

In the numerical solution of the 3D Poisson equation, the application of a conventional

finite-difference scheme leads to algebraic equations having a well-defined structure. When using a seven-point scheme, a septa-diagonal coefficient matrix results. In general, the resulting system of equations can be represented by the matrix equation $\mathbf{Ax} = \mathbf{b}$ [2]. The most suitable methods for the solution of this matrix equation are direct methods, but the computational cost becomes prohibitive as the number of equations increases, which is normally the case in 2D and 3D device simulations. This has led to the development of iterative procedures that utilize the well-defined structure of the coefficient matrix. The simplest and most commonly used iterative procedures are the Successive Over-Relaxation (SOR) and the Alternating Direction Implicit (ADI) methods. Both methods lose their effectiveness when complex problems are encountered and when the equation set becomes large, as it is usually the case in 3D problems. The ILU methods, and the use of preconditioning together with conjugate gradients, provide a significant increase in the power of iterative methods.

Within incomplete factorization schemes [44], the matrix \mathbf{A} is decomposed into a product of lower (\mathbf{L}) and upper (\mathbf{U}) triangular matrices, each of which has four non-zero diagonals in the same locations as the ones of the original matrix \mathbf{A} . The unknown elements of the \mathbf{L} and \mathbf{U} matrices are selected in such a way that the seven diagonals common to both \mathbf{A} and $\mathbf{A}' = \mathbf{LU}$ are identical and the six superfluous diagonals represent the matrix \mathbf{N} , *i.e.*, $\mathbf{A}' = \mathbf{A} + \mathbf{N}$. Thus, rather than solving the original system of equations $\mathbf{Ax} = \mathbf{b}$, one solves the modified system $\mathbf{LUx} = \mathbf{b} + \mathbf{Nx}$, by solving successively the matrix equations $\mathbf{LV} = \mathbf{b} + \mathbf{Nx}$ and $\mathbf{V} = \mathbf{Ux}$, where \mathbf{V} is an auxiliary vector. It is important to note that the six superfluous terms of \mathbf{N} affect the rate of convergence of the ILU method. Stone [45, 46] suggested the introduction of partial cancellation, which minimizes the influence of these additional terms and accelerates the rate of convergence of the ILU method. By using a Taylor series expansion, the superfluous terms appearing in \mathbf{A}' are partially balanced by subtracting approximately equal terms. To provide

rapid convergence, in this so-called Stone's Strongly Implicit (SIP) procedure, the fraction that is canceled is varied on successive pairs of iterations.

Even though the ILU and the SIP methods show improved convergence behavior when compared to, for example, the SOR method, their convergence slows down when approaching the solution. The basic Conjugate Gradient (CG) algorithm is one of the best known iterative techniques for solving sparse Symmetric Positive Definite (SPD) systems, but it loses its applicability when the resulting system of equations is not SPD. In such circumstances, the best alternative are the Lanczos-type algorithms, which solve not only the original system $\mathbf{Ax} = \mathbf{b}$ but also the dual linear system $\mathbf{A}^T \mathbf{x}^* = \mathbf{b}^*$. In recent years, the Conjugate Gradient Squared (CGS) method due to Sonneveld [47] has been recognized as an attractive transpose-free variant of the Bi-Conjugate Gradient (Bi-CG) iterative method [48]. This method works quite well in many cases, but the very high variations in the residual vectors often cause the residual norms to become inaccurate, which can lead to substantial buildup of rounding errors and overflow. The Bi-CGSTAB method due to Van der Vorst [49] is a variant of the CGS algorithm, which avoids squaring of the residual polynomial. It has been demonstrated that the convergence behavior of this method is smoother because it produces more accurate residual vectors and, therefore, more accurate solutions. In conjunction with the Bi-CGSTAB method, a successful preconditioning matrix can be obtained by using ILU factorization [50]. If \mathbf{L} and \mathbf{U} are the strictly-lower and strictly-upper triangular parts of \mathbf{A} , then the preconditioning matrix is

$$\mathbf{K}_{\text{ILU}(k)} = (\mathbf{L} + \tilde{\mathbf{D}})\tilde{\mathbf{D}}^{-1}(\mathbf{U} + \tilde{\mathbf{D}}), \quad (2)$$

where $\text{diag}(\mathbf{K}_{\text{ILU}(k)}) = \text{diag}(\mathbf{A})$. The case $k=0$ is used here, and means no fill-ins are allowed. Once the diagonal $\tilde{\mathbf{D}}$ is computed, scaling of the original matrix \mathbf{A} is performed

$$\tilde{\mathbf{A}} = \tilde{\mathbf{D}}^{-1/2} \mathbf{A} \tilde{\mathbf{D}}^{-1/2} = \text{diag}(\tilde{\mathbf{A}}) + \tilde{\mathbf{L}} + \tilde{\mathbf{U}}. \quad (3)$$

The preconditioning matrix for this symmetrically scaled matrix is of the form

$$\tilde{\mathbf{K}} = (\tilde{\mathbf{L}} + \mathbf{I})(\mathbf{I} + \tilde{\mathbf{U}}), \quad (4)$$

where \mathbf{I} is the identity matrix. The Bi-CGSTAB method is now applied to the preconditioned system

$$(\tilde{\mathbf{L}} + \mathbf{I})^{-1}\tilde{\mathbf{A}}(\mathbf{I} + \tilde{\mathbf{U}})^{-1}\tilde{\mathbf{x}} = (\tilde{\mathbf{L}} + \mathbf{I})^{-1}\tilde{\mathbf{b}}, \quad (5)$$

where $\tilde{\mathbf{b}} = \tilde{\mathbf{D}}^{-1/2}\mathbf{b}$, and the solution of the original system of equations is obtained *via* $\mathbf{x} = \tilde{\mathbf{D}}^{-1/2}(\mathbf{I} + \tilde{\mathbf{U}})^{-1}\tilde{\mathbf{x}}$. In the calculation of the product $(\tilde{\mathbf{L}} + \mathbf{I})^{-1}\tilde{\mathbf{A}}(\mathbf{I} + \tilde{\mathbf{U}})^{-1}\tilde{\mathbf{x}}$, appearing in (5), extra work is avoided by using the Eisenstat's trick [51]

$$\begin{aligned} &(\tilde{\mathbf{L}} + \mathbf{I})^{-1}\tilde{\mathbf{A}}(\mathbf{I} + \tilde{\mathbf{U}})^{-1}\tilde{\mathbf{p}} \\ &= \tilde{\mathbf{t}} + (\tilde{\mathbf{L}} + \mathbf{I})^{-1}\{\tilde{\mathbf{p}} + [\text{diag}(\tilde{\mathbf{A}}) - 2\mathbf{I}]\tilde{\mathbf{t}}\}, \end{aligned} \quad (6)$$

where $\tilde{\mathbf{t}} = (\mathbf{I} + \tilde{\mathbf{U}})^{-1}\tilde{\mathbf{p}}$. More details on the convergence behavior of each of these methods and the number of operations required per iterations can be found in Ref. [52].

2.3. Ensemble Monte Carlo

The Monte Carlo model, used in the transport portion of the simulator, is based on the usual Si band-structure for three-dimensional electrons in a set of nonparabolic Δ valleys with energy-dependent effective masses. The six conduction band valleys are included through three pairs: valley pair 1 pointing in the (100) direction, valley pair 2 in the (010) direction, and valley pair 3 in the (001) direction. The explicit inclusion of the longitudinal and transverse masses is important and this is done in the program using a Herring-Vogt transformation [53]. Intravalley scattering is limited to acoustic phonons and the MD-derived forces. For the intervalley scattering, we include both *g*- and *f*-phonon processes. It is important to note that, by group symmetry considerations, the zeroth-order low-energy *f*- and *g*-phonon processes are forbidden. Nevertheless, three zeroth-order *f*-phonons and three zeroth-order *g*-phonons with various energies are usually assumed [17].

We have taken into account this selection rule, and considered two high-energy *f*- and *g*-phonons and two low-energy *f*- and *g*-phonons in our bulk EMC simulations. The high-energy phonon scattering processes are included *via* the usual zeroth-order interaction term, and the two low-energy phonons are treated *via* a first-order process [54]. The first-order process is not really important for low-energy electrons but gives a relevant contribution for high-energy electrons. The low-energy phonons are important in achieving a smooth velocity saturation curve, especially at low temperatures. The phonon energies and the coupling constants used in our model are determined so that the experimental temperature dependent mobility and velocity-field characteristic is consistently recovered. The choice of these parameters is summarized in Table I. We also include impact ionization *via* a soft impact-ionization model [55].

The Monte Carlo program was tested by simulating 10,000 electrons over a period of several picoseconds. Figure 2 shows the electron transient velocity for an instantaneously applied

TABLE I Phonon model for bulk electrons for Si

Phonon	Phonon-type	Coupling constant	Phonon energy
zeroth	<i>f</i> -phonon	$8 \times 10^8 \text{ eV/cm}$	59 meV
zeroth	<i>g</i> -phonon	$8 \times 10^8 \text{ eV/cm}$	63 meV
first	<i>f</i> -phonon	2.5 eV	23 meV
first	<i>g</i> -phonon	4.0 eV	18 meV

We use 9.0 eV for the acoustic phonon coupling constant, and treat it as elastic scattering process.

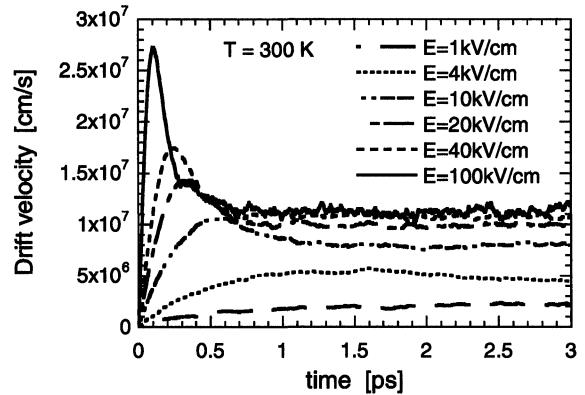


FIGURE 2 Electron transient velocity overshoot at $T = 300 \text{ K}$.

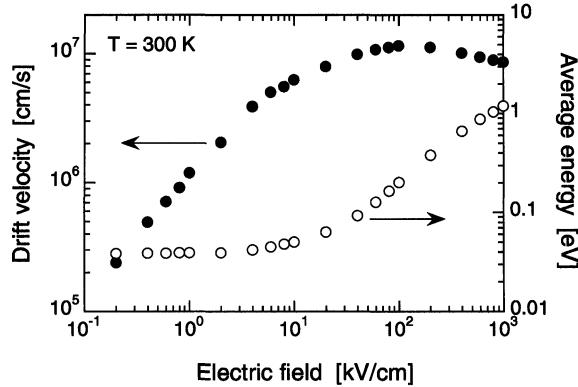


FIGURE 3 Steady-state velocity-field and energy-field characteristics at $T=300\text{ K}$.

electric field. At low fields, the electron velocity slowly approaches the steady state limit, which equals the product of the low field electron mobility μ_0 times the electric field. At higher fields, the electron velocity shows an overshoot effect as a result of the smaller momentum relaxation time when compared to the energy relaxation time. From the results shown in Figure 2, it is seen that steady-state is achieved in approximately 0.5–2 ps. The simulation was extended for 2–3 ps beyond this time in order to determine the steady-state velocity. At fields above 10 kV/cm, the optical phonon scattering rate limits the electron mean free path, resulting in a velocity saturation of $\sim 1.1 \times 10^7\text{ cm/s}$. Figure 3 shows the average steady-state velocity *versus* electric field. The simulated mobility at 1 kV/cm is $1500\text{ cm}^2/\text{V}\cdot\text{s}$, which is an accepted value for bulk silicon at 300 K. Also shown in this figure is the electric field dependence of the average electron energy. For low fields, the energy is $\sim 40\text{ meV}$, or $1.5k_B T$. As the field increases, the energy increases quadratically. Impact ionization limits the energy increase in the tail of the electron distribution.

3. SHORT-RANGE COULOMB INTERACTIONS

A commonly accepted approach in particle simulations is to include ionized-impurity scattering as

an additional scattering potential in the \mathbf{k} -space portion of the EMC transport kernel. Then, some type of screening model must be used. A similar approach is adopted for the electron–electron interaction. In both cases, the expressions for the scattering rates only take into account binary collisions between an electron and a single charge (electron or ion), and therefore second-order effects of multiple electron–electron and electron–ion interactions are neglected. In addition, local variation in electron and ion densities are not considered. To calculate, for example, the $e-e$ scattering rate, one must know the local distribution function. The CPU time it takes to re-calculate the distribution function for each iteration time-step makes this approach rather impractical to use. Nevertheless, many device simulations have been performed using a tabulated distribution function [56].

The use of a real-space treatment of the $e-e$ and $e-i$ interactions eliminates the deficiencies of the screened-scattering approach. In this method, the force upon each electron due to all other charges in the device (electrons and impurities) is added to the mesh force by the summation of all Coulomb forces. In three dimensions, the total Coulomb force, \mathbf{F}_i on each electron is

$$\mathbf{F}_i = \sum_{j=1, j \neq i}^{N_e} \frac{q \mathbf{a}_{ij}}{4\pi\epsilon \cdot |\mathbf{r}_i - \mathbf{r}_j|^2} + \sum_{k=1}^{N_i} \frac{q \mathbf{a}_{ik}}{4\pi\epsilon \cdot |\mathbf{r}_i - \mathbf{r}_k|^2} \quad (7)$$

where $|\mathbf{r}_i - \mathbf{r}_j|$ is the distance between the electron and other electrons, $|\mathbf{r}_i - \mathbf{r}_k|$ is the distance from the electron to an ion, \mathbf{a}_{ij} and \mathbf{a}_{ik} are unit vectors along the force direction, and N_e and N_i are the total number of electrons and ions, respectively.

The actual force seen by an electron is a combination of the molecular dynamics forces from other electrons and ions in the device and the applied field. The boundary conditions (both Dirichlet and Neumann) can be thought of as additional charges near the device boundaries that produce known potential or fields at the

boundaries. If the actual amount of charge and the placement of that charge were known, then the Coulomb force equation could be used to take into account these boundary conditions. However, the determination of the boundary charges is more difficult and less efficient (CPU time-wise) than solving for the mesh force. If infinitely small mesh is used, then the force calculated from the mesh potentials would be identical to the Coulomb force. However, even a practical solution (spacing in the 2.5 nm range) is very costly in terms of both CPU time and memory.

Figure 4 shows a comparison of the mesh and Coulomb forces *versus* separation range for a two-electron case, with a mesh spacing of 10 nm in all

three directions. It is easily seen that the mesh force (\mathbf{F}_{mesh}) and the Coulomb force (\mathbf{F}_{ee}) are identical when the electrons are separated by several mesh points (30–50 nm apart). Within 3–5 mesh points, the mesh force starts to deviate from the Coulomb force. When the electrons are within the same mesh cell, the mesh force approaches zero due to the smoothing of the electron charge when it is divided amongst the nearest node points. From the results shown in Figure 4, it is also clear that the mesh force alone cannot account for the short-range electron–electron and electron–ion interactions. One way of incorporating these interactions is to determine the Coulomb force on each electron due to all the other charges (electrons and ions) in the device using Eq. (7) and add the mesh force calculated from the solution of Poisson's equation. The mesh force is needed to take boundary conditions and applied biases into account. This method would result in ‘double counting’ of the long-range electron–electron and electron–ion interactions. Even accepting the ‘double counting’ error, this method still requires a large amount of CPU time for the Coulomb force calculations. 10,000 electrons and ions in a device leads roughly to about 50 million electron–electron collisions, and 100 million electron–ion collisions. To reduce the calculation time, while minimizing the loss of accuracy, a cut-off range is usually used, which limits the number of collision events drastically [38].

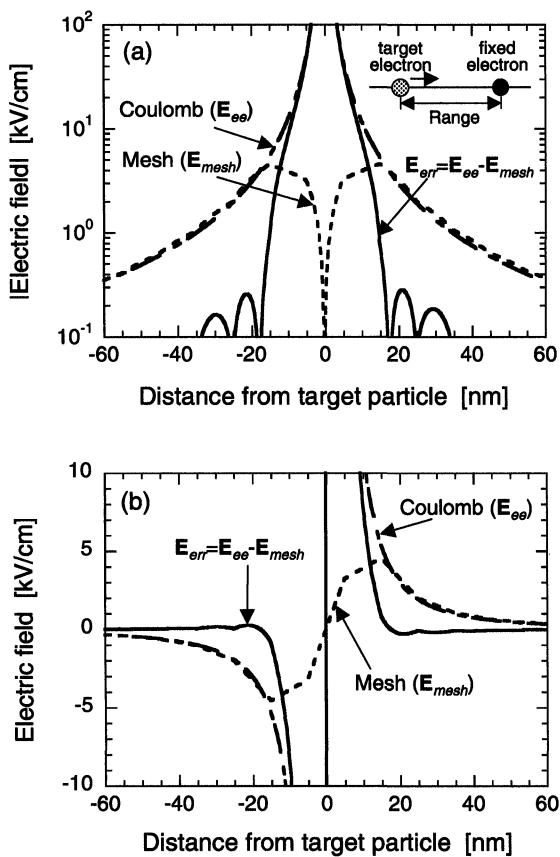


FIGURE 4 Mesh force *versus* Coulomb force on (a) log scale and (b) linear scale. The mesh is uniformly spaced at 10 nm in all three dimensions. Also shown is the corrected force, \mathbf{F}_{err} , which is the difference of the two forces.

Although the cutoff range alleviates the problem of CPU time needed for the molecular dynamic force calculations, it does not completely solve the problem of the ‘double counting’ of the long-range force. For electrons and ions outside of the cutoff range, the mesh force is the sole force calculated, and thus no ‘double counting’ occurs. However, the cutoff creates a discontinuity in the force at the cutoff range, within which the force will be ‘double counted’. Making the cutoff range too small can result in underestimating the short-range force, which would basically negate the effect of the added MD routine. Therefore,

the cutoff range must be set to a value beyond which the molecular dynamics force is equal to the mesh force. Table II shows the average electron–electron distance (range) *versus* doping concentration. Setting the cutoff range too large would result in excess CPU time needed to calculate the molecular dynamics forces. For a typical MOSFET with source/drain doping of greater than 10^{19} cm^{-3} , the average range is less than 5 nm. For a 5 nm mesh, which is a practical mesh size for a three-dimensional device simulation, the mesh force alone cannot provide accurate results even for the average interaction, let alone those of a close range. Hence, it is imperative that the short-range interactions be included separately.

We utilize a scheme that allows us to calculate the corrected Coulomb force [29]. Briefly, an adjusted Coulomb force is calculated using

$$\mathbf{F}_{\text{err}} = \mathbf{F}_{ee} - \mathbf{F}_{\text{mesh}}. \quad (8)$$

Figure 5 shows the corrected Coulomb force for the case of a single $e-e$ collision. One limitation of this scheme is that the mesh must be uniform in all directions. Another limitation arises from the fact that mesh spacing less than the extrinsic Debye length is needed in critical device regions (such as junctions and beneath a MOS gate). However, the small mesh spacing is only required to resolve high fields determined by the solution of Poisson's equation. As was shown in Figure 4, the Coulomb force is orders of magnitude higher than the mesh force for small separations. Therefore, the high fields in critical regions will be resolved through use of the molecular dynamics force. The application of this scheme to both resistor simulations and simulations of ultra-small MOSFETs is shown below.

TABLE II Doping dependence of the $e-e$ range.

Doping [cm^{-3}]	$e-e$ range [nm]
10^{17}	21.5
5×10^{17}	12.6
10^{18}	10
5×10^{18}	5.9
10^{19}	4.6

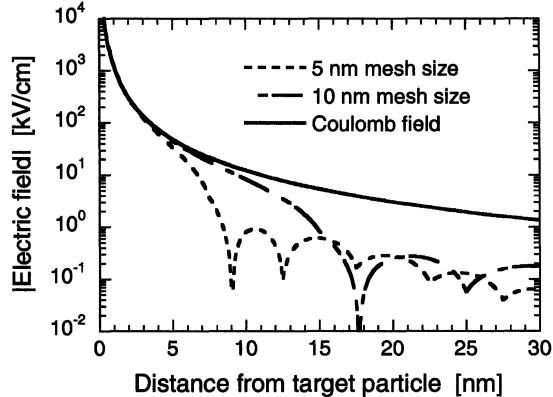


FIGURE 5 Corrected Coulomb force (F_{err}) vs. electron spacing for a single collision with mesh spacings of 10 and 5 nm. The Coulomb force is shown for comparison.

3.1. Resistor Simulations

To investigate the impact of adding the short-range $e-i$ and $e-e$ forces on the low-field electron mobility derived from resistor simulations, several *n*-type resistors with various sizes and doping levels were simulated using our 3D EMC particle-based simulator. Under our modified scheme, each electron is allowed to interact simultaneously with all the charges present in the system that fall within the range determined from our look-up table *via* the tabulated correction force. To account for the dynamical nature of the temporal changes in the charge configuration, a time step of 0.2 fs is used in the simulations. The electron-phonon interactions and the impact ionization process have been taken into account through a standard EMC procedure. A small electric field, on the order of 10^2 V/cm , was used for computing the steady-state velocities to ensure operation in the linear-response regime. The simulation results for the mobility are shown in Figure 6. Also shown in this figure are the mobility results reported in [21], calculated with a bulk EMC technique using the Brooks-Herring approach [57] for the $e-i$ interaction, and the measured data [58, 59] for the case when the applied electric field is parallel to the $\langle 100 \rangle$ crystallographic direction. From these results, it is obvious that adding the correction force to the mesh

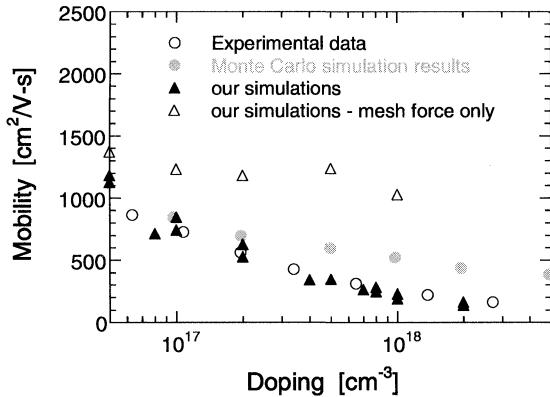


FIGURE 6 Doping-dependence of the low-field electron mobility obtained from 3D resistor simulations.

force leads to mobility values that are in very good agreement with the experimental data. It is important to note that if only the mesh force is used in the free-flight portion of the simulator, the simulation mobility data points are significantly higher than the experimental values. While resistor tests are not a true estimate of the power of any new method, we remark that previous studies of such structures have established the role of multi-ion interactions [25], and the need for proper dynamic screening with Fermi-Dirac statistics [60].

3.2. Modeling of Ultra-small MOSFETs

We have used the 3D particle-based simulator described in Section 2, in which we have also incorporated the modified MD scheme (described in the beginning of this section) to account for the short-range portions of the $e-e$ and $e-i$ interactions, to model ultra-small n -MOSFETs shown schematically in Figure 7. The average substrate doping of the device is $3 \times 10^{18} \text{ cm}^{-3}$, whereas the average doping of the source and drain regions is 10^{19} cm^{-3} . The impurity atoms, corresponding to these average doping densities, were placed within the device using the following procedure. We draw a random number k from the Poisson distribution with mean value of the process $a = N_{\text{imp}} V_{\text{disc}}$, where N_{imp} is the average donor or acceptor concentration, and V_{disc} is the corresponding volume

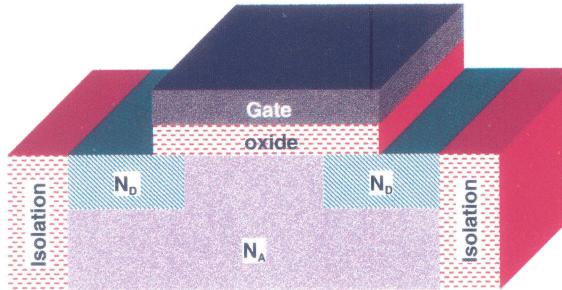


FIGURE 7 Simulated device. (See Color Plate I).

of the region. Each of these k impurity atoms is then uniformly distributed within V_{disc} using triplets of independent uniformly distributed random numbers, and its position coordinates are saved for later usage in the MD routine. We use either the nearest-element-cell (NEC) or the cloud-in-cell (CIC) scheme when assigning doping density to the node points. This procedure allows us to mimic the physical ion implantation process and study fluctuations in the device parameters in devices with different number and different distribution of the impurity atoms in their active and source, drain and substrate regions.

We point out that the average doping of the source and drain regions in any of the devices that we have studied is significantly lower than what a fabricated device would have. This allows us to use larger mesh sizes and reduce the total number of particles in the simulation, thus reducing the computation time. The higher source and drain resistances of our devices affect only the magnitude of the current, but does not significantly influence the velocity overshoot in the channel. Channel length of the simulated device is 80 nm, the junction depth is 5.5 nm, the extensions of the source and drain regions are 50 nm, and the channel width of the device is 80 nm. The oxide thickness is 3 nm.

The key simulation parameters used in our simulations are summarized in Table III. A uniformly spaced 5 nm grid was used in all three dimensions. The cutoff range for the MD calculations was set to 10 nm, or twice the grid spacing, which is just beyond the point where the mesh

TABLE III Key simulator parameters used for MOSFET device simulations

Parameter	Value/Method used
Mesh Spacing	5 nm
Cutoff range for $e-e$ and $e-i$ interaction	10 nm
Poisson solver	SIP method
Time-step	0.2 fs
Absolute Error Level for the Poisson solver	26 μ V
Relative Error Level for the Poisson solver	0.1%
Charge assignment scheme	CIC
Iterations per bias point	10,000

calculated force and the Coulomb force converge. The chosen combination of the mesh spacing, mesh voltage tolerance level, and Coulomb force cutoff range led to acceptable simulation time. Increasing the mesh spacing by a factor of two would result in an 8x reduction in the total number of nodes. However, the force cutoff range must be doubled, resulting in a more than 4x increase in the number of $e-e$ and $e-i$ collisions.

In Figure 8, we show the averaged potential energy profile of this device, for applied biases $V_G = 1.0$ V and $V_D = 1.0$ V. The averaging was done over a 2 ps time period. As a result of the averaging procedure, any observable fluctuation

in the potential profile can be associated with the fluctuations in the doping density because of the atomistic treatment of the impurity atoms. Fluctuations in the potential that would naturally arise from the movement of the discrete electrons are thus smeared out. The use of larger mesh sizes leads to further smoothing of the potential profile, because of the larger number of impurity atoms that would fall within one unit cell, thus leading to smaller relative fluctuations in the number of dopant atoms between neighboring unit cells. For the mesh sizes and the doping densities that we use, there are on the average 1.25 donor-type impurity atoms per unit cell in the source and drain regions and 0.375 acceptor-type impurity atoms per unit cell in the substrate region of the device.

The average electron velocity and the average electron energy along the channel of this device are shown in Figure 9. The short-range $e-e$ and $e-i$ interactions have been included in these simulations *via* our MD procedure. As expected, both the electron velocity and energy increase when going from the source to the drain and they reach a peak value at the end of the channel, decreasing towards their equilibrium values in the

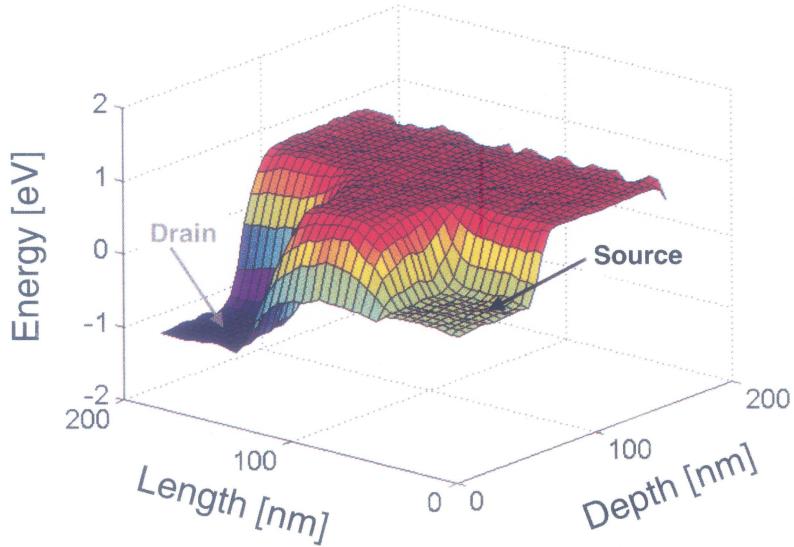


FIGURE 8 Potential energy profile of the device for $V_G = V_D = 1$ V. (See Color Plate II).

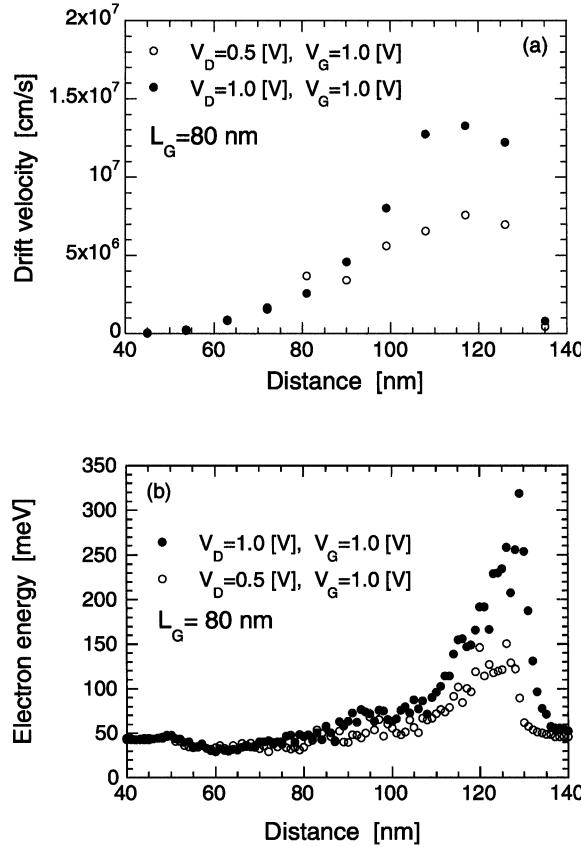


FIGURE 9 (a) Average electron velocity along the channel; (b) Average electron energy along the channel. On both figures, filled circles correspond to the case $V_D = 1$ V and $V_G = 1$ V. Open circles represent our simulation results for $V_D = 0.5$ V and $V_G = 1$ V. The short-range $e-e$ and $e-i$ interactions are included in these simulations via the MD routine. The channel extends from 50 to 130 nm.

drain end. It should be noted, however, that the velocity actually peaks well before the drain, due to the Coulomb scattering from the drain. Comparing the results shown in Figure 9a with the steady-state results for bulk samples shown in Figure 3, it is also clear that when the drain bias V_D equals 1.0 V, the peak carrier velocity is higher than its steady-state counterpart, thus suggesting presence of slight velocity overshoot.

To demonstrate the influence of the $e-e$ and $e-i$ interactions, in Figure 10a and 10b, we plot the average electron velocity and average electron energy of the carriers for the case when the short-range

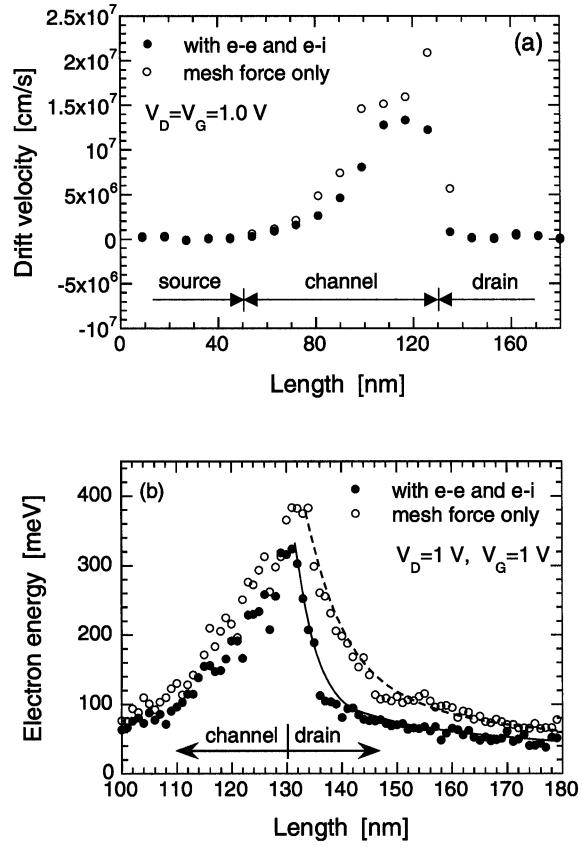


FIGURE 10 (a) Average velocity of the electrons along the channel, with and without the inclusion of the $e-e$ and $e-i$ interactions; (b) Average energy of the electrons coming to the drain from the channel. For both figures, we use $V_D = 1$ V and $V_G = 1$ V. Filled (open) circles correspond to the case when the short-range $e-e$ and $e-i$ interactions are included (omitted) in the simulations. The channel extends from 50 to 130 nm.

$e-e$ and $e-i$ interaction terms were included (omitted) in our model. For clarity, in Figure 10b we plot the average energy of only those electrons which arrive at the drain end of the device from the channel. We use $V_D = V_G = 1.0$ V. From the results shown, it is quite clear that the inclusion of the short-range $e-e$ and $e-i$ interaction terms damps the previously discussed velocity overshoot effect, thus increasing the transit time of the carriers through the device, *i.e.*, reducing its cut-off frequency. On the other hand, the results shown in Figure 10b suggest that carrier thermalization occurs over distances that are on the order of few

nm when the $e-e$ and $e-i$ interactions are included *via* our modified MD loop. It is also important to note that using the mesh force alone does not even lead to complete thermalization of the carriers along the whole length of the drain extension.

4. CONCLUSIONS AND FUTURE DIRECTIONS OF RESEARCH

We have described the main features of our 3D EMC particle-based simulator, with emphasis on the proper inclusion of the $e-e$ and $e-i$ interactions *via* a MD approach. The proposed scheme, described in details in this paper, avoids the problem of the ‘double counting’ of the long-range Coulomb force, and at the same time leads to a more accurate description of the short-range components of the force. It naturally includes multiple-scattering events as well as fluctuations in the doping and electron densities, through the atomistic treatment of both impurity atoms and charged particles. The close agreement of our simulation results for the doping dependence of the low-field electron mobility, derived from resistor simulations, with the experimental measurements supports the suitability of our approach. This novel scheme has also been incorporated into our 3D EMC particle-based simulator, used to model n -MOSFETs with gate-length equal to 80 nm. The simulation results of this device suggest that the short-range components of the $e-e$ and $e-i$ interaction terms are responsible for the relaxation of the velocity overshoot effect at high drain biases and for the faster thermalization of the carriers when entering the heavily-doped drain region.

One limitation of our approach for proper inclusion of the $e-e$ and $e-i$ interactions is that the mesh must be uniform in all dimensions in regions where the Coulomb force is calculated. Another drawback of the present version of our simulator is that holes are accounted for using a zero hole current, also known as the quasi-equilibrium approximation. This means that we are

currently unable to gather information regarding substrate leakage currents.

We are currently using this simulator to address a number of important issues for ultra-small structures, including discrete impurity effects and fluctuations in device characteristics, such as threshold voltage, transconductance and on-state current. We will also investigate whether contacts will dominate the device behavior in ultra-small MOSFETs with gate-lengths less than 50 nm, including a realistic treatment of the contacts themselves as part of the simulation domain.

Acknowledgements

This work was supported in part by NSF under Contract No. ECS-9802596. One of the authors (DV) would also like to acknowledge the financial support from ONR under Contract No. N00014-99-1-0318 and from NSF under Contract No. ECS-9875051.

References

- [1] Ferry, D. K., *Semiconductors*, New York: Macmillan Publishing Company, First Ed. 1991.
- [2] Selberherr, S., *Analysis and Simulation of Semiconductor Devices*, New York-Wien: Springer-Verlag, 1984.
- [3] Selberherr, S. (1980). “MINIMOS-A Two-Dimensional MOS Transistor Analyzer”, *IEEE Trans. Electron Dev.*, **27**, 1540.
- [4] Bløtekjær, K. (1970). “Transport Equations for Electrons in Two-valley Model”, *IEEE Trans. Electron Dev.*, **17**, 38.
- [5] Zhou, J.-R. and Ferry, D. K. (1993). “Modeling of Quantum Effects in Ultrasmall HEMT Devices”, *IEEE Trans. Electron Dev.*, **40**, 421; Ferry D. K. and Zhou, J.-R. (1993). “On the Form of the Quantum Potential for Use in Hydrodynamic Equations for Semiconductor Device Modeling”, *Phys. Rev. B*, **48**, 7944.
- [6] Jerome, J. W. and Shu, C. W. (1995). “The Response of the Hydrodynamic Model to Heat Conduction, Mobility, and Relaxation Expressions”, *VLSI Design*, **3**, 131.
- [7] Hu, C. and Shu, C.-W., “Weighted Essentially Non-oscillatory Schemes on Triangular Meshes”, *Journal of Comp. Physics*, in press.
- [8] Fukuda, K. and Nishi, K. (1995). “A New Wide Applicable Mobility Model for Device Simulation Taking Physics-based Carrier Screening Effects Into Account”, *IEICE Trans. Electron.*, **78**, 281.
- [9] Kurosawa, T. (1996). “Monte Carlo Calculation of Hot Electron Problem”, In: *Proc. Intern. Conf. Phys. Semicond., J. Phys. Soc. Jpn.*, **21** (Suppl.), 424.

- [10] Boardman, A. D., Fawcett, W. and Rees, H. D. (1966). "Monte Carlo Calculation of the Velocity-field Relationship for Gallium Arsenide", *Sol. State Commun.*, **6**, 305.
- [11] Curby, R. C. and Ferry, D. K. (1971). "Monte Carlo Calculation of Impact Ionization Generation Rate", *Phys. Lett.*, **35A**, 64.
- [12] Bacchelli Montefusco, L. and Jacoboni, C. (1972). "Electron-Electron Interaction in Monte Carlo Transport Calculation", *Sol. State Commun.*, **10**, 71.
- [13] Lebwohl, P. A. and Price, P. J. (1970). "Hybrid Method for Hot Electron Calculations", *Sol. State Commun.*, **9**, 1221.
- [14] Ferry, D. K. and Barker, J. R. (1980). "On the Use of Monte Carlo Techniques for the Calculation of Transient Dynamic Response in Semiconductors", *Phys. Stat. Sol. (b)*, **100**, 683.
- [15] Ferry, D. K. (1980). "Long-time Tail of the Autocorrelation Function for Electron Drift in High Electric Fields in Silicon", *Phys. Rev. Lett.*, **45**, 758.
- [16] Ferry, D. K. (1980). "Diffusion in the Transient Dynamic Response Regime", *Phys. Lett.*, **78A**, 379.
- [17] Jacoboni, C. and Reggiani, L. (1983). "The Monte Carlo Method for the Solution of Charge Transport in Semiconductors with Applications to Covalent Materials", *Rev. Modern Phys.*, **55**, 645.
- [18] Fischetti, M. and Laux, S. (1988). "Monte Carlo Analysis of Electron Transport in Small Semiconductor Devices Including Band-structures and Space-charge Effects", *Phys. Rev. B*, **38**, 9721.
- [19] Lugli, P. (1990). "The Monte Carlo Method for Semiconductor Device and Process Modeling", *IEEE Trans. Electron Dev.*, **9**, 1564.
- [20] Patil, M. B. and Ravaioli, U. (1991). "Transient Simulation of Semiconductor Devices Using the Monte-Carlo Method", *Solid State Electron.*, **34**, 1029.
- [21] Tomizawa, K., *Numerical Simulation of Submicron Semiconductor Devices*. Boston: Artech House, 1993.
- [22] Jacoboni, C. (1974). "Recent Developments in the Hot Electron Problem", In: *Proc. Intern. Conf. Phys. Semicond.*, Rome: Tipograf. Marves, 1195.
- [23] Bosi, S. and Jacoboni, C. (1976). "Monte Carlo High-field Transport in Degenerate GaAs", *J. Phys. C: Sol. State Phys.*, **9**, 315.
- [24] Lugli, P. and Ferry, D. K. (1986). "Dynamical Screening of Hot Carriers in Semiconductors from a Coupled Molecular-dynamics and Ensemble Monte Carlo Simulation", *Phys. Rev. Lett.*, **56**, 1295.
- [25] Joshi, R. P. and Ferry, D. K. (1991). "Effect of Multi-ion Screening on the Electronic Transport in Doped Semiconductors: A Molecular-dynamics Analysis", *Phys. Rev. B*, **43**, 9734.
- [26] Lugli, P. and Ferry, D. K. (1986). "Degeneracy in the Ensemble Monte Carlo Method for High-field Transport in Semiconductors", *IEEE Trans. Electron Dev.*, **32**, 2431.
- [27] Kriman, A. M., Kann, M. J., Ferry, D. K. and Joshi, R. (1990). "Role of the Exchange Interaction in the Short-time Relaxation of a High-density Electron Plasma", *Phys. Rev. Lett.*, **65**, 1619.
- [28] Fischetti, M. V. and Laux, S. E. (1995). "Understanding Hot-electron Transport in Silicon Devices: Is there a Shortcut?", *J. Appl. Phys.*, **78**, 1058.
- [29] Gross, W. J., Vasileska, D. and Ferry, D. K. (1999). "A Novel Approach for Introducing the Electron-Electron and Electron-impurity Interactions in Particle-based Simulations", *Electron Dev. Lett.*, **20**(9), 463–465.
- [30] Mizuno, T., Okamura, J. and Toriumi, A. (1994). "Experimental Study of Threshold Voltage Fluctuation Due to Statistical Variation of Channel Dopant Number in MOSFET's", *IEEE Trans. Electron Dev.*, **41**, 2216.
- [31] Mizuno, T. (1996). "Influence of Statistical Spatial-non-uniformity of Dopant Atoms on Threshold Voltage in a System of Many MOSFET's", *Jpn. J. Appl. Phys.*, **35**, 842.
- [32] Wong, H.-S. and Taur, Y. (1993). "Three-dimensional atomistic simulation of discrete random dopant distribution effects in sub-0.1 μm MOSFET's", *IEDM Tech. Dig.*, p. 705.
- [33] Zhou, J. and Ferry, D. K. (1995). "3D Simulation of Deep-submicron Devices: How Impurity Atoms Affect Conductance", *IEEE Computational Science and Engineering*, **2**(2), 30.
- [34] Stok, P., Widdershoven, F. and Klassen, D. B. M. (1998). "Modeling Statistical Dopant Fluctuations in MOS Transistors", *IEEE Trans. Electron Dev.*, **41**, 1960.
- [35] Vasileska, D., Gross, W. J. and Ferry, D. K., "Modeling of Deep-submicrometer MOSFETs: Random Impurity Effects, Threshold Voltage Shifts and Gate Capacitance Attenuation", Extended Abstracts IWCE-6, Osaka 1998, *IEEE Cat. No. 98EX116*, p. 259.
- [36] Asenov, A. (1998). "Random Dopant Induced Threshold Voltage Lowering and Fluctuations in Sub-0.1 μm MOSFET's: A 3-D 'Atomistic' Simulation Study", *IEEE Trans. Electron Dev.*, **45**, 2505.
- [37] Ranawake, U. A., Huster, C., Lenders, P. M. and Goodnick, S. M. (1994). "PMC-3D: A Parallel Three-dimensional Monte Carlo Semiconductor Device Simulator", *IEEE Trans. on CAD*, **13**, 712.
- [38] Hockney, R. W. and Eastwood, J. W., *Computer Simulation Using Particles*, Maidenhead: McGraw-Hill, 1981.
- [39] "Self-force" is the force (electric field) seen by electron due to the electron itself.
- [40] Laux, S. E. (1996). "On Particle-mesh Coupling in Monte Carlo Semiconductor Device Simulation", *IEEE Trans. CAD Integr. Circ. Syst.*, **15**, 1266.
- [41] Vasileska, D. and Ferry, D. K. (1997). "Scaled Silicon MOSFET's: Part I - Universal Mobility Behavior", *IEEE Trans. Electron Devices*, **44**, 577.
- [42] Gonzalez, T. and Pardo, D. (1996). "Physical Models of Ohmic Contacts for Monte Carlo Device Simulation", *Solid State Electron.*, **39**, 555.
- [43] Blakey, P. A., Cherensky, S. S. and Sumer, P., "Self-consistent Particle-field Monte Carlo Simulation of Semiconductor Microstructures", In: *Physics of Submicron Structures*, New York: Plenum Press, 1984.
- [44] Gadiyak, G. V. and Obrecht, M. S., "The Use of Factorization Methods for Solving the Charge Transfer Equations in Semiconductor Devices", *Simulation of Semiconductor Devices and Processes: Proceedings of the Second International Conference*, pp. 147–156, Ed. by Board, K. and Owen, D. R. J., Swansea UK: Pineridge Press, 1986.
- [45] Stone, H. L. (1968). "Iterative Solution of Implicit Approximations of Multidimensional Partial Differential Equations", *SIAM J. Numer. Anal.*, **5**, 530.
- [46] Weinstein, H. G., Stone, H. L. and Kwan, T. V. (1969). "Iterative Procedure for Solution of Systems of Parabolic and Elliptic Equations in Three Dimensions", *I & EC Fundamentals*, **8**, 281.
- [47] Sonneveld, P. (1989). "CGS, a Fast Lanczos-type Solver for Nonsymmetric Linear Systems", *SIAM J. Sci. Stat. Comput.*, **10**, 36.

- [48] Saad, Y., *Iterative Methods for Sparse Linear Systems*, Boston: PWS Publishing Company, 1996.
- [49] Van der Vorst, H. A. (1992). "Bi-CGSTAB: A Fast and Smoothly Convergent Variant of Bi-CG for the Solution of Nonsymmetric Linear Systems", *SIAM J. Sci. Stat. Comput.*, **13**, 631.
- [50] Van der Vorst, H. A. (1989). "High Performance Preconditioning", *SIAM J. Sci. Stat. Comput.*, **10**, 1174.
- [51] Eisenstat, S. C. (1981). "Efficient Implementation of a Class of Preconditioned Conjugate Gradient Methods", *SIAM J. Sci. Stat. Comput.*, **2**, 1.
- [52] Vasileska, D., Gross, W. J., Kafedziski, V. and Ferry, D. K. (1998). "Convergence Properties of the Bi-CGSTAB Method for the Solution of the 3D Poisson and 3D Electron Current Continuity Equations for Scaled Si MOSFETs", *VLSI Design*, **8**, 301.
- [53] Herring, C. and Vogt, E. (1956). "Transport and Deformation-potential Theory for Many-valley Semiconductors with Anisotropic Scattering", *Phys. Rev.*, **101**, 944.
- [54] Ferry, D. K. (1976). "First-order Optical and Intervalley Scattering in Semiconductors", *Phys. Rev. B*, **14**, 1605.
- [55] Yamada, T., Zhou, J.-R., Miyata, H. and Ferry, D. K. (1994). "In-plane Transport Properties of Si/Si_{1-x}Ge_x Structure and its FET Performance by Computer Simulation", *Phys. Rev. B*, **49**, 1875.
- [56] Fischetti, M. V., Laux, S. E. and Crabbé, E., "Monte Carlo Simulation of High-energy Electron Transport in Silicon: Is there a Short-cut to Happiness?", In: *Hot Carriers in Semiconductors*, New York: Plenum Press, 1996, pp. 475–80.
- [57] Brooks, H. (1951). "Scattering by Ionized Impurities in Semiconductors", *Phys. Rev.*, **83**, 879.
- [58] Jacoboni, C., Canali, C., Ottaviani, G. and Alberigi-Quaranta, A. (1977). "A Review of Some Charge Transport Properties of Silicon", *Solid-state Electron.*, **20**, 77.
- [59] Canali, C., Ottaviani, G. and Alberigi-Quaranta, A. (1971). "Drift Velocity of Electrons and Holes and Associated Anisotropic Effects in Silicon", *J. Phys. Chem. Solids*, **32**, 1707.
- [60] Chung, W.-Y. and Ferry, D. K. (1988). "Dynamic Screening for Ionized Impurity Scattering in Degenerate Semiconductors", *Solid-state Electron.*, **31**, 1369, and references therein.

Authors' Biographies

William J. Gross received the B.S degree in microelectronic engineering from R.I.T. in 1990 and the M.S. degree in electrical engineering from Arizona State University in 1994. He is presently pursing his Ph.D in electrical engineering at ASU in the simulation of deep-submicrometer MOSFETs.

Dragica Vasileska is Assistant Professor at Arizona State University. She received her Ph.D. degree in Electrical Engineering from Arizona State University in 1995. Her research interests include semiconductor device physics, 3D device simulations and transport theory.

David K. Ferry is Regents' Professor of Engineering at Arizona State University. He is a Fellow of both the American Physical Society and the Institute of Electrical and Electronics Engineers. He pursues the physics and modeling of nanostructured devices.

