MONTE CARLO ANALYSIS

As the name suggests, the Monte Carlo (MC) method is based on the selection of random numbers (1–4). In its present form, the method is attributed to Fermi, Von Neumann, and Ulam, who developed it for the solution of problems related to neutron transport. In principle, the MC method can be considered as a very general mathematical tool for the solution of a great variety of problems.

Among the various applications of the method, the following are probably the most important:

- Integrodifferential equations
- Matrix inversion
- Transport of nuclear particles
- Transport in semiconductors
- Modeling of semiconductor devices
- Process simulation

An important feature of the MC technique is that more precise results can be obtained by generating larger numbers of points. More generally, being based on random numbers, the results obtained with an MC procedure are never exact, but are rigorous in a statistical sense: The exact result lies in given intervals with given probabilities. The uncertainty of the results is strictly related to the variance of the possible outcomes, and it is smaller if the size of the sample (i.e., the amount of computations devoted to the solution of the problem) is larger. One basic element of the numerical procedure is the possibility to generate random numbers with given distributions starting from pseudorandom numbers uniformly distributed between 0 and 1. Modern computers provide sequences of numbers obtained with precise mathematical algorithms, starting from a given element (seed). For each seed, the sequence is perfectly predictable. However, it satisfies a large number of statistical test of randomness. Those pseudorandom numbers offer two great advantages: They can be generated in a very fast way, and they are reproducible (which is essential, for example, in debugging a code).

The applications of MC methods can be divided into two major groups. One consists of the direct reproduction on a computer of the microscopic dynamics of the physical process in a system which is already statistical in its nature. We use in this case the term "MC simulation." The second group consists of MC methods devised for the solution of well-defined mathematical equations or problems. In such cases, the methods provide a way of sampling a given statistical distribution. The majority of real cases are a mixture of the two extreme limits indicated above. The application to the study of semiconductor devices and processes is a good example. In fact, transport processess are statistical in nature, but are also accurately described by well-defined equations that may or may not correspond to the direct simulation of the physical system under examination. For instance, the MC solution of the Boltzmann transport equation (BTE) not only gives the distribution function that verifies the equation, but also yields information that is lost in the BTE itself. On the other hand, the direct simulation is at times very inefficient, as for example in the analysis of situations that are rare in the actual physical system. In such cases it is necessary to distort the simulation by applying some more sophisticated MC techniques that reduce the variance of the quantity of interest, giving up partially the advantages offered by the direct simulation.

The applications of the MC methods that we will focus on are particularly important in light of the growth in the field of microelectronics achieved in recent years. Semiconductor devices are nowadays built with their active dimensions well below 1 μm. Metal-oxide-semiconductor field-effect transistor (MOSFET) technology has moved already into the quarter-micron size, while high electron mobility transistors (HEMTs) are commercially available with 0.15 μm gate length. The reduction in size leads to a higher integration level as more devices can be put into a single chip. Such a push toward smaller and more powerful devices (which immediately translate into higher levels of integration and enhanced performance of the single devices as well as of the overall circuit) has been sustained by enormous advances in the area of fabrication and processing. A very precise control is nowadays possible on the device geometry and doping profile through techniques such as ion implantation, reactive ion etching, and electron and X-ray lithography. Furthermore, new possibilities for novel devices are offered by the capability to grow nanometer layered structures with extremely high quality by molecular beam epitaxy (MBE) and metal organic chemical vapor phase epitaxy (MOCVD). As we will see, MC approaches can be of great help in understanding and overcoming the limits of several technological or in attempting to improve the yield of integrated circuits. Moving into the submicron scale, many new physical phenomena become important that require a sophisticated theoretical treatment. There exists, therefore, a new challenge toward the understanding of the principles of operation of those novel devices. The MC method offers great advantages also in this direction.

Computer programs are extremely important for technology development. Computer-aided design (CAD) has become one of the keystones in microelectronics. The importance of such a field can be greatly appreciated focusing on the steps required for the fabrication of integrated circuits (ICs). The development of new technologies is traditionally driven by an experimental approach. A useful alternative is offered by software tools, which can lead to a speed up of the development cycle and a reduction of the development costs. In fact, those calculations can be considered as simulated experiments, which can be much faster and less expensive than real experiments. Furthermore, computer experiments allow a deep physical interpretation of the final results that leads to a better understanding of the problem at hand. This is particularly true for the MC simulations. The characteristic links between the different aspects of CAD can be summarized as follows (5). The output of the process simulation is fed directly into a device simulation program, which determines the electrical characteristics and the performance of the device. At this stage, the interplay between process and device simulation can suggest improvements on the processing steps deduced

from the simulated device performance. The output of the device simulator is then compacted in order to be inserted in a circuit simulation program, which determines the characteristics of the overall circuit. As this article will show, MC approaches are finding more widespread use as CAD tools, both at the level of device and at the level of process simulation.

Although it will not be possible to exhaust the complexity of the MC methods in such a short review, the present contribution is intended to give a critical overview of the MC algorithms used for device and process modeling and for circuit yield analysis. A more thorough investigation can be found in the references. For the MC simulation of semiconductor devices, a complete review can be found in Ref. 6.

MONTE CARLO DEVICE SIMULATION

The Monte Carlo technique is a fairly new tool in the area of device modeling, traditionally dominated by simulators based on drift-diffusion models or on balance-equation models (for an overview of the subject, see Refs. 7 and 8). The first MC application to the study of electronic transport in semiconductors is due to Kurosawa in 1966 (9). Shortly afterwards the Malvern group, in UK (Boardman, Fawcett, Hilsum, Swain, among others), provided the first wide application of the method to device simulation, focusing mainly on GaAs devices (10). Applications to Si and Ge boomed in the 1970s, thanks to work performed at the University of Modena, Italy, and at IBM Yorktown, USA. The reviews in Refs. 4 and 11 provide a deeper historical and technical perspective. The great attention reserved in recent years to the MC analysis of devices (12) is strictly connected to the availability of powerful and relatively inexpensive computers and workstations, which guarantee the necessary numerical resources for the computationally quite heavy simulations. Furthermore, with the recent advances in material growth, contact deposition, and impurity control, devices have become more transparent from a physical point of view. Incidentally, this has provided physical systems of extreme interest. At the same time, MC algorithms have gained in sophistication and are now able to handle phenomena and systems of great complexity. These are two fundamental steps since the necessary input for an MC simulation of semiconductor materials and devices is an accurate microscopic description of the physical system under investigation. Many semiconductor devices can be nowadays simulated with the MC method, which is becoming more and more a very useful modeling tool.

The most common (and also the most interesting) simulation of a semiconductor device is performed for many particles in parallel [Ensemble Monte Carlo (EMC), procedure] and coupled to Poisson's equation in order to obtain the self-consistent potential related to the charge distribution given directly by the MC procedure. For systems of great complexity, a one-particle Monte Carlo (OPMC) simulation can be performed on a given fixed potential previously determined.

Since no *a priori* assumptions are needed on the form of the real- and k-space carrier distributions, an MC simulator is the only reliable tool for the investigation of those physical phenomena that critically depend on the shape of the distribution or on the details of its tail (such as electron injection over potential barriers). Furthermore, the MC technique allows us to focus on particular physical mechanisms that might be of importance with regard to the device performance (e.g., intercarrier scattering, impact ionization, generation-recombination, etc.). The prices one has to pay are a very time-consuming algorithm and the requirement of a complete knowledge of the physical system under investigation. Often many assumptions have to be made in order to reduce the complexity of the model describing a given device.

The Monte Carlo Algorithm

In recent years, EMC simulations have been widely used to study the properties of semiconductor devices. Particular emphasis has been lately attributed to submicron structures, because of their performances in switching and high frequency operations (13). Once the basic physics involved in the transport of such devices is known, EMC simulation provides a formidable tool to determine their limits and characteristics and can be very helpful in modeling. Together with the determination of the macroscopic properties of a device, EMC also gives a microscopic description of the local electric field, charge density, velocity distribution, and so on.

The basic steps of a standard EMC self-consistent device simulation are (see Fig. 1) as follows:

1. *Set Up Geometry and Discretization Scheme.* Two parameters that play an important role in the choice of the time step and the grid size are the plasma frequency and the Debye length. For simple device geometry a one-dimensional description can be sufficient. Normally, MC simulations are performed using a two-dimensional grid (that is, assuming homogeneity along the third direction). Since the simulation is inherently three-dimensional, there are no principal reasons that prevent a fully three-dimensional analysis.
2. *Charge Assignment.* The charge of each particle is assigned to a particular mesh point. Since it is not possible to simulate all the electrons present in a real device, each simulated particle (also referred to as "superparticle") represents a cloud of electrons for the purpose of estimating currents, charge, and field distributions. For all other purposes, each individual particle carries its elementary charge e. The doping charge is also added to the mesh according to its distribution. A sufficiently large number of simulated particles is needed to guarantee the statistical validity of the results and to reduce the level of numerical noise. Such a number depends on the dimensionality of the spatial grid, ranging typically from thousands of particles for one-dimensional descriptions, up to several tens of thousand particles for two- or three-dimensional systems.
3. *Potential Solution.* Poisson's equation is solved to determine the electrostatic potential at the mesh points. In connection to EMC simulations, a finite-difference scheme is generally used, although some attempts to use a finite-element approach have been presented (14). The solution can be obtained in several ways, among which we can list the Fourier analysis cyclic reduction, the conjugate gradient, the multigrid, and the direct matrix inversion methods. The first three methods provide effective algorithms that allows the inclusion of special boundaries. The latter requires a matrix inversion at the beginning of the simulation. The new poten-

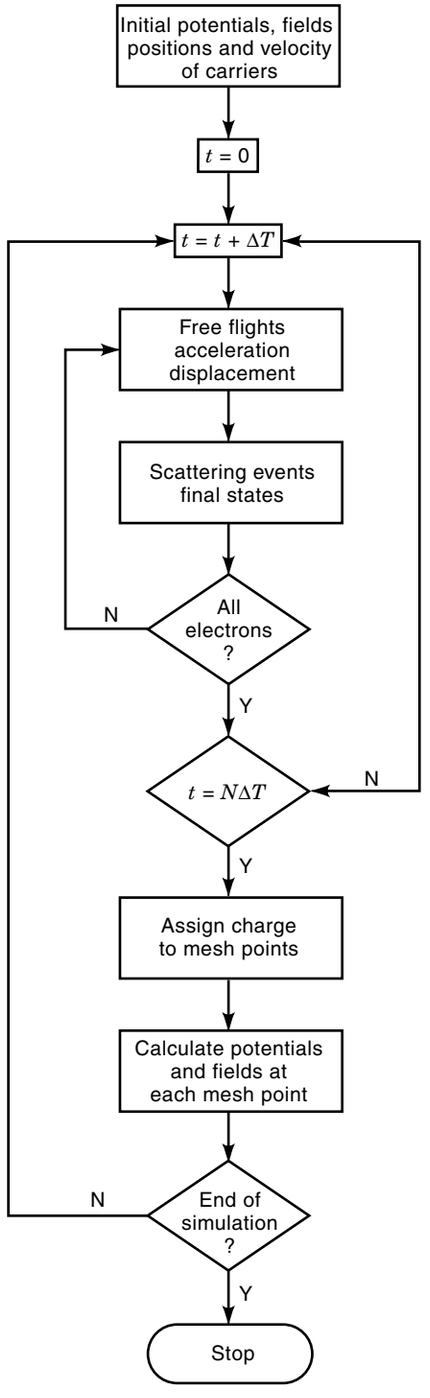

Figure 1. Flow chart of a typical self-consistent MC device simulation.

tial is calculated with a simple matrix multiplication at fixed times during the simulation. The electrostatic field is then obtained from the potential with a finite-difference algorithm.

4. *Flights.* Each particle, now treated as an individual electron, undergoes the standard MC sequence of scattering and free flights, subject to the local field previously determined from the solution of the Poisson equation. The MC sequence is stopped at fixed times, at which the field is adjusted following the steps described above. The carrier dynamics is computed by solving Newton's equations of motion in the crystal, described by its band structure. The length of a free flight is generated randomly according to the total scattering rate (i.e., the number of collisions per unit time). At the end of each flight, carriers scatter (with impurities, lattice vibrations, other carriers, etc.), with the probability of each event being weighted by its scattering rate (relative to the total one). Finally, a new state (i.e., the new energy, wavevector, occupied band) following the collision is determined according to the differential cross section of the process that has terminated the free flight. The scattering probability and the probability distribution of the final states are computed using quantum mechanics (starting from the so-called Fermi Golden Rule). A particular event (collision or no collision, which type of collision, which final state) is selected randomly, by comparing the probability of occurrence of that event to a random number.

The description of the problem is completed by setting initial and boundary conditions. The initial conditions are not so important, since only the self-consistent steady-state result is usually retained. Boundary conditions are instead crucial, in particular in submicron devices, where contact properties drastically influence the whole behavior of the device.

The steady-state current is given directly by the net number of particles crossing one contact per unit time. More effectively, an extension of Ramo's theorem to two-dimensional situations allows the calculation of the current in three terminal devices simply by summing the velocities of all particles found inside given portions of the device (15). By performing several computer runs, it is possible to construct the I–V characteristics of the device. Important device parameters can also be extracted from the simulation. For example, the output resistance $R_{\text{out}} = \Delta V_{\text{ds}}/\Delta I_{\text{ds}}$ for constant gate bias V_{gs} and the transconductance $g_{\text{m}} = \Delta I_{\text{ds}}/\Delta V_{\text{gs}}$ for constant drain bias V_{ds} can be obtained from a series of runs, starting from normal operating conditions and varying the drain-to-source and gate-to-source bias, respectively. Parasitic elements can also be calculated, referring to an equivalent circuit description of the device. For example, the source-to-gate C_{sg} and source-to-drain C_{gd} capacitances are obtained by applying a step change respectively to the gate (ΔV_{gs}) and drain voltage (ΔV_{gd}) as $C_{\text{gs}} = \Delta Q_{\text{gs}}/\Delta V_{\text{gs}}$ and $C_{\text{gd}} = \Delta Q_{\text{gd}}/\Delta V_{\text{gd}}$, where ΔQ is the total equivalent charge flowing from the gate in response to the step potential. From the time dependence of the charge flow, it is also possible to estimate the parasitic source and drain resistance.

In an MC simulation, ohmic contacts are usually treated as ideal contacts, by maintaining a neutrality condition near the metal boundary. Due to such an assumption, the simulation results are those of an ideal intrinsic device. The comparison of the simulated characteristics with the one measured on real devices requires the consideration of a finite resistance for each ohmic contact. This can be done quite simply by scaling the simulated I–V characteristics using independently determined values of the contact resistances (16).

Traditionally, device simulators have been based on drift diffusion (DD) or hydrodynamical (HD) models. The basis of the two methods (as also the MC) is the Boltzmann transport

equation. By taking the first three moments of BTE, three coupled equations are obtained which describe the spatial and temporal evolution of the average carrier concentration, velocity, and energy. Within the HD approach, the three equations are solved numerically, with some simplifying assumptions (such as the introduction of momentum and energy relaxation times for the equations of first and second moments). The DD approach assumes that carriers are always in equilibrium with the lattice temperature. Thus the energy equation drops out, and only the continuity equation is left, provided that the current density is expressed in a phenomenological way in terms of the carrier mobility and diffusivity. Clearly, the HD scheme is far superior to the DD one, since it can account (when all terms are considered) for carrier heating and nonhomogeneous distributions of the carrier temperature. The MC procedure stands on an even higher level, since it provides (also in nonhomogeneous, nonstationary conditions) an exact solution of the Boltzmann equation. It correctly describes nonlocal effects (in space $\boldsymbol{r}$ or time t) typical of situations where the field inside the device varies appreciably over lengths comparable with the electron mean free path, and the electrons at a given position carry information about the field value at another position. Unfortunately, the complexity and cost of each approach is inversely proportional to the refinement of the physical model it is based on. Therefore, the use of one approach or another depends on the specific device under investigation. The MC is the best technique to study situations where nonstationary effects are important (as for example in submicron devices). It is safe to anticipate that as the tendency toward the miniaturization of devices will continue in the future, MC simulators will progressively extend their applicability. The success of MC in device modeling will ultimately depend on the compromise between two tendencies, one to use very sophisticated physical models (which lead to very costly but extremely accurate simulations), the other to rely on simplified models, sacrificing a bit of accuracy for a reduced complexity and cost of the algorithms. Along this line, it will be extremely useful to be able to combine different methods in order to fully exploit the potential of each approach. Examples of MC simulators present in the literature can be found in Refs. 17–43.

Special Features

In the following section, we focus on special aspects of the MC simulation that are not generally considered because of their difficulty, although they can be of great importance in the device performance.

Full Band Simulation. A physical description of the semiconductor band structure is needed as input to the MC simulation. Traditionally, bands have been described via nonparabolic dispersion relations, because this allows the analytical calculation of the differential and total scattering rates for a variety of scattering mechanisms. Furthermore, the calculation of the field-induced acceleration during a free flight is trivial. A typical example is the model used for GaAs (44), which includes a three-valley description of the conduction band (T, L, and X valley) and a description of the three valence bands (heavy, light, and split-off holes). The nonparabolic dispersions are introduced with nonparabolicity factors treated as fitting paramenters and adjusted as to reproduce a variety of experimental results, including those provided by time-resolved spectroscopy (6). The carrier interaction with polar optical, acoustic, equivalent and nonequivalent intervalley, intraband and interband phonons, and ionized impurities can be easily considered. Unfortunately, nonparabolic dispersions provide an adequate band description only up to energies not too far from band edge (typically 1 eV or less). Thus, higher band states need to be described more accurately any time high-energy effects are important in a device (as for instance in the case of carrier injection above energy barriers, or when impact ionization occurs). The most popular approach is that based on a full band description, achieved via pseudopotential methods (45). The full $\epsilon(\boldsymbol{k})$ dispersion is calculated numerically, either on the entire Brillouin zone or only on the irreducible wedge. The scattering rates are then calculated using directly the electronic states coming out of the numerical routine. The improvement in the band description level is paid in terms of computational time and memory required with respect to simpler band models.

Pauli Exclusion Principle. Electrons obey the Fermi–Dirac statistics and must satisfy the Pauli Exclusion Principle (PEP). This means that not all the states are available because only two electrons of opposite spin can reside in each state. This aspect is not very important in the nondegenerate case, and electrons are distributed in a large interval of states; in the degenerate case the problem becomes more conspicuous. For instance, GaAs electrons are degenerate at room temperature at densities around 10^{19} cm^{-3}. This is the case for many devices of interest. Degeneracy is equivalent to a many-body interaction which reduces the phase space available for the electron final state in an induced transition. If $p(\boldsymbol{k})$ and $p(\boldsymbol{k}')$ are the probabilities that the initial and final state are, respectively, occupied, the total rate of transition $P(\boldsymbol{k}, \boldsymbol{k}')$ between two different states is given by $P(\boldsymbol{k}, \boldsymbol{k}') = p(\boldsymbol{k})\, S(\boldsymbol{k}, \boldsymbol{k}')\, [1 - p(\boldsymbol{k}')]$. Normally a semiclassical MC procedure works with the approximation $p(\boldsymbol{k}') = 0$ because all the states are considered available as final states. The inclusion of the PEP is then essentially the inclusion of this term in the total scattering rate. In the EMC procedure, this is obtained very easily because the particle distribution is known step by step. The algorithm generating the distribution function is set up by multiplying the scattering probability by the correction factor $1 - p(\boldsymbol{k}')$; $p(\boldsymbol{k}')$ is determined self-consistently, and a rejection technique is used after selecting the final state without the correcting Pauli factor (46).

Contact Simulation. The simulation of contacts is one of the most serious problems in MC device simulations, due in part to the limited knowledge of the physics of contacts. On the other hand, contacts are of great importance in a number of semiconductor devices, whose applications range from high-speed logic to microwaves. As the dimensions of these devices reach the submicron limit, contacts become the limiting factor for the performance in the ballistic or quasi-ballistic mode of operation. In general, semiconductor devices do not operate under charge neutrality conditions. The net charge inside the device is directly related through Gauss' law to the flux of the electric field on the boundaries, and consequently to the potential inside the device. Therefore, charge neutrality (that is, conservation of the number of particles) cannot be enforced in the simulation. Rather, an appropriate handling of the

boundaries is required to simulate a number of electrons that vary in time self-consistently with the potential distribution. Taking a field-effect transistor as an example, the most significant boundaries are at the source, drain, and gate electrodes. Source and drain contacts are usually treated as ideal ohmic contact by absorbing all the electrons that hit the electrodes and by injecting a number of electrons which maintain a neutral region in the adjacency of the electrodes. The Schottky barrier at the gate is treated as a perfectly absorbing electrode with a potential equal to the applied potential minus the barrier height. Although commonly assumed, the above conditions have never really been tested. One attempt to deal with the problem of contacts in a simulation of a one-dimensional metal–n–n^+ structure has been presented in Ref. 47.

Carrier–Carrier Scattering. Many devices are characterized by very high electron concentrations. In such a situation, one might have to worry about the possible influence of the interaction among the conducting electrons. A good example is provided by the heterojunction bipolar transistor (HBT), which will be examined later. In a standard device, such as a metal semiconductor field-effect transistor (MESFET) or HEMT, electrons are injected into the channel with a thermal energy distribution and a small initial velocity. In an HBT, electrons are injected from the emitter into the base, which they cross before being swept away from the high field at the base–collector junction. Due to the high doping (p-type) in the base region, electrons can be scattered by the collective excitations of the hole gas, as well as through normal binary collisions with the other electrons. Two main contributions to the carrier–carrier scattering can be identified:

- The individual carrier–carrier interaction via a screened Coulomb potential which accounts for two-body short-range interaction
- The electron–plasmon interaction, which accounts for the collective long-range behavior of the electron gas

In semiconductors, the plasmon energy at a reasonable electron density can be of the same order of magnitude as the characteristic phonon energies. In a device simulation, the scattering rates for electron–electron and electron–plasmon processes can be tabulated at the beginning of the simulation. Carrier–carrier scatterings are then treated as any other mechanisms in the MC algorithm, using appropriate rejection algorithms to account for the current carrier distribution function (48,49).

Optimization Procedures. An original, efficient algorithm has been implemented to calculate the appropriate duration of the free flights (depending on the actual carrier status). The method, which is based on a space-dependent definition of the scattering rate (50), leads to a drastic reduction in the number of self-scatterings, thus allowing a large savings in computation time (more than one order of magnitude compared with the conventional approaches). In areas where the electron population is very small, it is possible to extend a technique originally proposed by Phillips and Price (50), which allows one to obtain good statistics in rarely visited energy regions. Two situations are of particular interest. If a device presents regions with a high doping density N^+ connected to regions with low densities N^-, the carrier concentration will reflect (except at the interface between the different regions) the doping distribution. Therefore, most of the simulated carriers (roughly in the ratio N^+/N^-) will populate the highly doped, low-field regions, requiring an extensive amount of computation for the simulation of a quasi-thermal distribution. This is the case, for example, of the MOSFET or a MESFET with ion implanted source and drain contacts. A similar case is found in $\boldsymbol{k}$ space, when we are interested in the population of high-energy states, which are rarely touched by the carriers but might cause very important physical phenomena (a typical example is the carrier injection into SiO_2 for the channel of a MOSFET). The latter situation is the one examined by Phillips and Price. The population of the high-energy states can be enhanced by generating a fixed number N of carrier histories every time one of the simulated particles enters the rarely populated region. Each one of the N generated particles will have the same initial condition (equal to the state of the "parent" particle at the moment of the multiplication) and a weight $1/N$ for the calculation of the average quantities. The multiplication algorithm can be repeated several times at higher energies, originating an "avalanche" of carriers that fill up the tail of the distribution function at higher and higher energies (36). A similar multiplication technique has been also used in real space in Ref. 40. A peculiar situation is found when impact ionization phenomena are important. The knowledge of the high-energy tail of the carrier distribution function is then required. Furthermore, in the presence of carrier multiplication, the number of simulated particles would grow above the initially set value, diverging as breakdown is approached. A special multiplication technique for both energy and real space has been implemented (51), which is an extension of the approaches described above. Each particle is assigned a statistical weight which varies with its position in the device and its energy. With such approach, it is possible to account for regions with very different doping levels (as in bipolar transistors) and to obtain a reliable statistics of rare processes, keeping at the same time a constant number of particles.

Simulation Results

We show some results obtained by MC simulation of GaAs devices, leaving the interested reader to the variety of applications listed as references at the end of this article.

Heterojunction Bipolar Transistors. HBTs are receiving considerable attention because of their high-speed performance and high current-handling capability. The device shown in Fig. 2 has been simulated (44), and the results have been compared with measured data. By varying the base–collector voltage (in the common-base configuration), the electric field profile changes as depicted in Fig. 3. Very high values are reached in the collector region, with the maximum occuring at the base–collector interface. Electrons, injected from the emitter, cross the base where they strongly interact with the dense hole plasma. As they enter the collector, they are ballistically accelerated by the junction field, reaching velocities as high as 6×10^7 cm/s. The spatial extent of the velocity overshoot is limited to about 100 Å, as the electrons are rapidly scattered into the satelite valleys where they move at saturated velocity. Figure 4 indeed shows that for a collector voltage of 16 V, almost all electrons in the collector populate, in

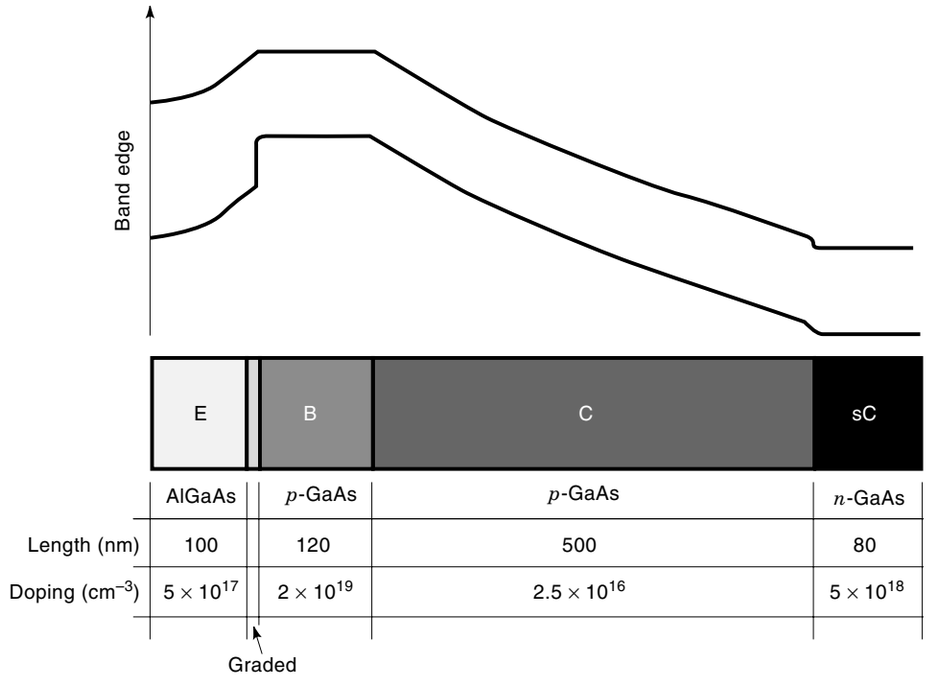

Figure 2. Energy band diagram and cross section of the simulated HBT.

equal number, the L and X valleys. There, they are strongly heated by the collector field, obtaining the high values of average energy illustrated in Fig. 5. Correspondingly, the calculated $(M - 1)$ factor, which measures the relative increase of collector current due to multiplication phenomena, is around unity at this voltage, in excellent agreement with the measured value (44). As V_{BC} is further increased, the holes created by primary ionization processes are in turn able to ionize, marking the onset of breakdown, which is predicted around 18 V.

High Electron Mobility Transistors. HEMTs are extremely appealing for microwave low-noise applications. Their superior performance is due to the spatial separation of the electrons (confined in a small-gap channel layer, typically GaAs or InGaAs) from the donors (fixed in the large-gap region or regions surrounding the channel). In the simulated device (52) a special type of doping is considered, denoted δ-doping, where two dopant planes are located in the AlGaAs layers respectively above and below an InGaAs channel. The substrate material of the device is GaAs. Due to the different lattice constant of GaAs and InGaAs, strain is present in the channel region. This type of HEMT is referred to as "pseudomorphic HEMT." Typical two-dimensional contour plots for a device with a gate length of 180 nm are shown in Figs. 6 to 9. The source-to-drain bias is 2 V. The potential distribution (Fig. 6) reveals that a high-field region exists in the channel between the end of the gate and the beginning of the cap layer (which is a doped GaAs region above the AlGaAs confining barrier, used in technology to reduce the capacitance and resistance at source and drain contact region). Entering this region, an electron becomes very hot, reaching average kinetic energies of few tenths of electronvolts, as indicated in the energy map of Fig. 7. Electron heating results in reduced confinement within the channel, since many of the electrons have sufficient energy to surmount the confining barrier. Furthermore, as a result of heating, the population of the upper L valley is very high in correspondence of the high field region, as visible in Figs. 8 and 9 which show respectively the total and L valley electron concentrations. Because of the hot-electron-induced real-space transfer, the current flows through both the bottom AlGaAs layer (and partially through the substrate) and the top AlGaAs barrier, which actually constitutes the access path to the drain cap region. The actual value of the electric field along the channel, along with the corresponding carrier drift velocity and average energy, is presented in Fig. 10. Due to the very short gate length, the field reaches a peak value of 200 kV/cm, resulting in the strong electron heating discussed above and a remarkable velocity

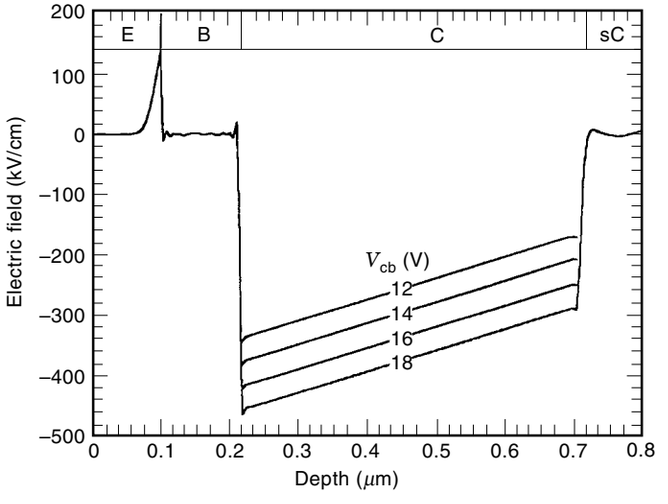

Figure 3. Electric field profile of the simulated HBT at various V_{BC}.

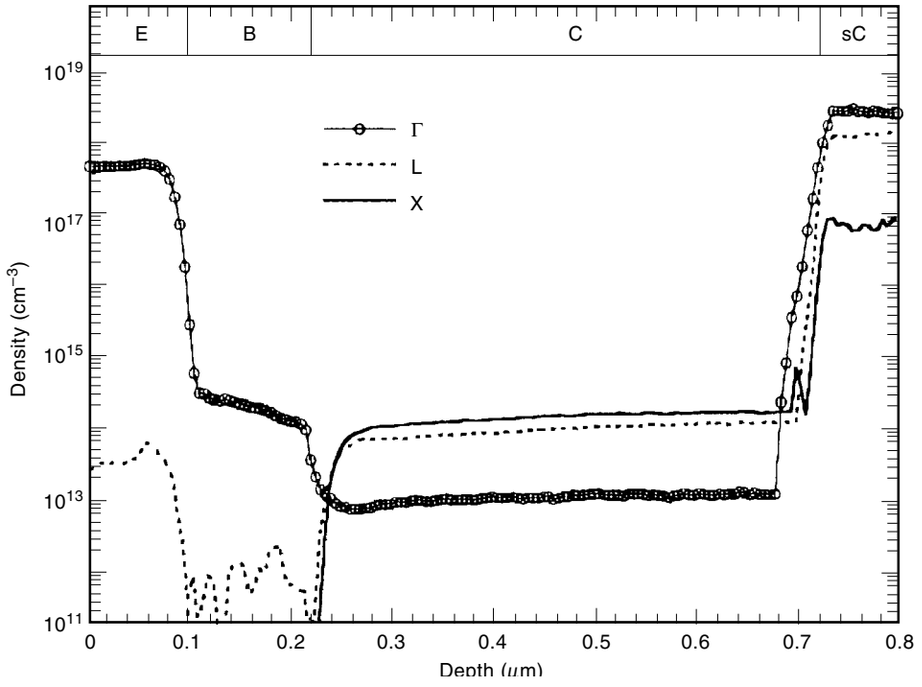

Figure 4. Electron concentration in the different valleys for $V_{BC} = 16$ V.

overshoot, which is in turn responsible for the excellent microwave performance of such devices.

MONTE CARLO PROCESS SIMULATION

We are going to discuss MC process simulation from the point of view of integrated circuit (IC) fabrication, since this is certainly the area where simulation plays the most important role. Many different interrelated steps contribute to the realization of the final product. The main processing steps can be classified into the following three categories:

1. Thermal processing and doping (ion implantation, predeposition, annealing, oxidation, epitaxial growth)
2. Pattern definition (reactive ion etching, deposition, evaporation, sputtering)

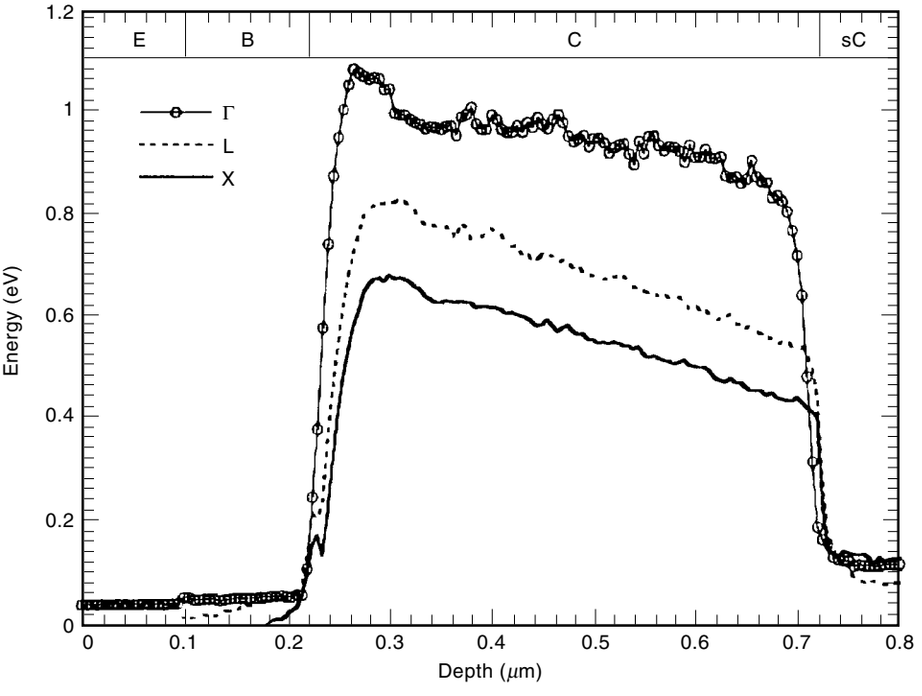

Figure 5. Electron average energy in the different valleys for $V_{BC} = 16$ V.

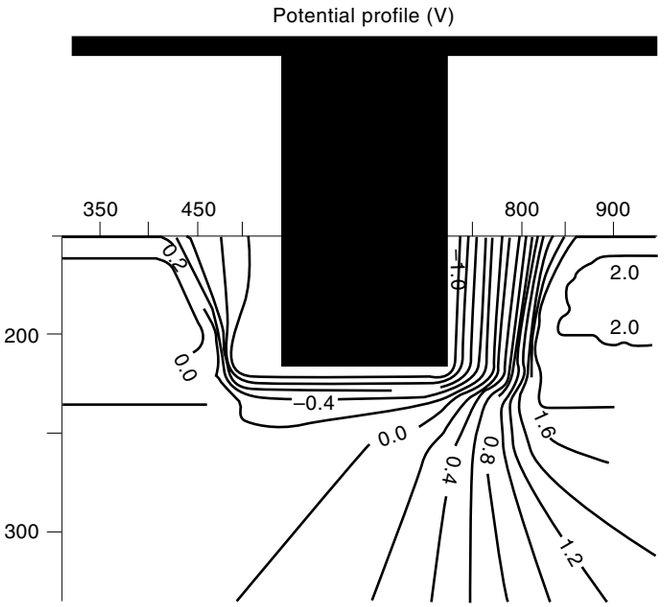

Figure 6. Two-dimensional plot of the equipotential lines for the simulated HEMT at a drain bias of 2 V.

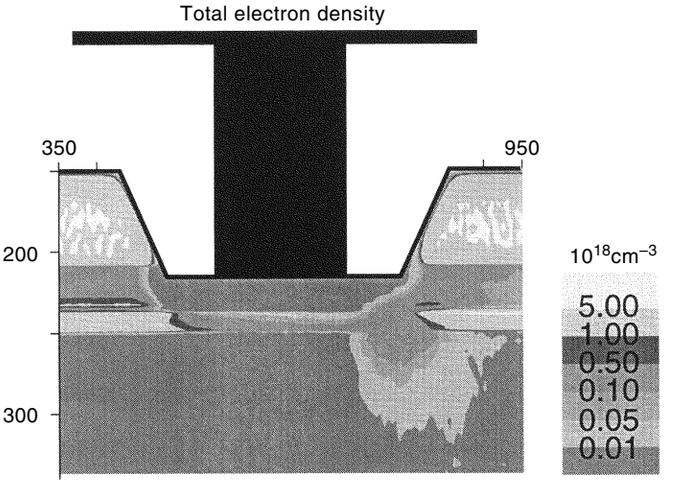

Figure 8. Two-dimensional representation of the total electron concentration for the simulated HEMT at a drain bias of 2 V. The darkest region corresponds to a density of 5×10^{18} cm^{-3}.

3. Pattern transfer (optical X-ray, electron beam lithography)

Setting up mathematical models for each step requires the knowledge of very complex physical and chemical phenomena, such as for instance the redistribution of atoms or impurities into a given material, or the energy exchange between fast projectiles and the substrate they interact with. As we will see in the examples, some drastic approximations are made in order to define a tractable model.

As for the case of MC device simulation, several analytical and numerical approaches exist in the literature that cover all the processing steps outlined above (5,53). The MC technique has been very successful in some applications, especially those where the simulation can be reduced to a series of uncorrelated events describing the trajectories of projectiles against target atoms. This is the case of the examples we will discuss below—that is, ion implantation and electron beam lithography. Other attempts have been made to use MC techniques to describe epitaxial growth (54,55).

As a general comment, we can say that once the mathematical model has been set up, the MC algorithm for process simulation presents fewer difficulties than the one for device simulation. This is because no self-consistency is required between internal potential and charge distribution and also because the transition probabilities are assumed to be constant between two successive stochastic events.

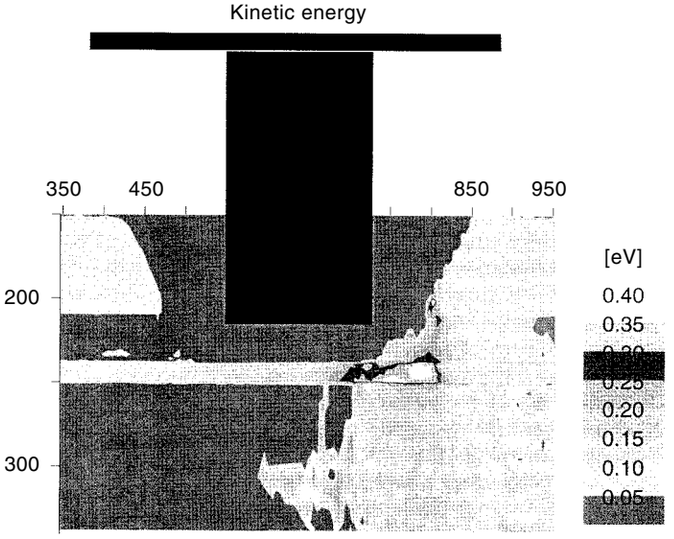

Figure 7. Two-dimensional representation of the electron average energy for the simulated HEMT at a drain bias of 2 V. The darkest region corresponds to an energy of 0.35 eV.

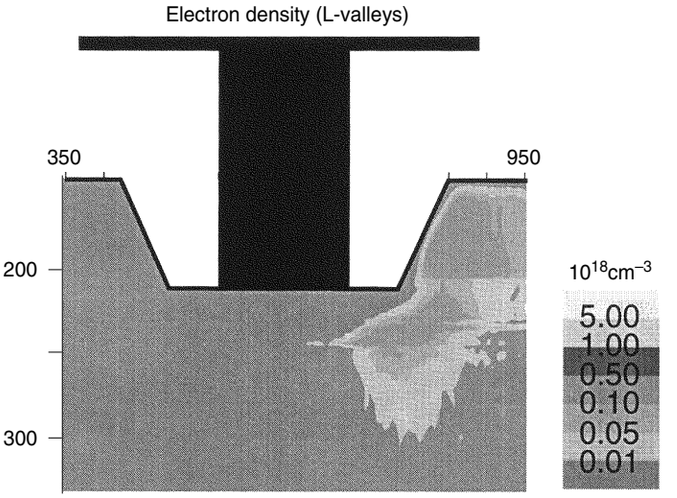

Figure 9. Two-dimensional representation of the L-valley electron concentration for the simulated HEMT at a drain bias of 2 V. The darkest region corresponds to a density of 5×10^{18} cm^{-3}.

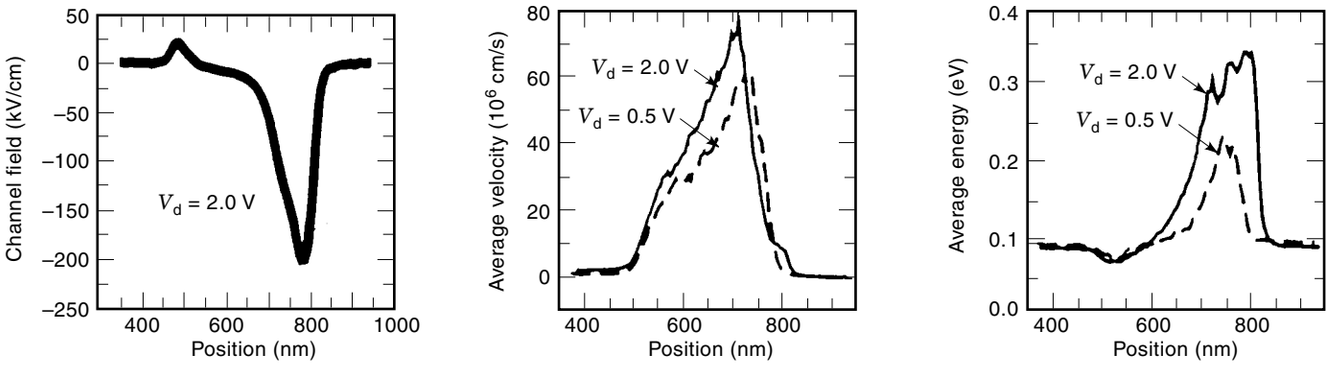

Figure 10. One-dimensional plot (along the channel direction) of the electric field (left), average velocity (center), and average energy (right) for two different drain voltages, respectively equal to 0.5 V (dashed lines) and 2 V (solid lines).

Ion Implantation

Ion implantation is one of the most important doping techniques for device fabrication, in particular for very large scale integration (VLSI) circuits. The successful application of this technique depends strongly on the ability to control the impurity profile for any implant condition. The three main processes involved in the penetration and slowing down of energetic ions into a material are:

1. The energy loss via collisions with the target atoms. Thousands of atoms (called "recoils") are displaced by each impinging ion. Local disorder and very high temperatures are reached in the region where the cascade occurs.
2. The thermalization of the excited zone, with possible diffusion of defects.
3. The long-range migration of defects.

Computer-aided design models for ion implantation fall into two broad categories: analytic distribution functions and MC methods. The former are computationally very inexpensive, but rely upon fits to experimental data to reproduce the observed profile of dopants ion and work well only for simple geometries in one dimension. In contrast, the MC approach attempts a first principle calculation based on two-body scattering theory. Although computationally expensive, it can handle the most complicated structures.

An MC simulation offers the following advantages (56,57):

- It accounts for implant profile discontinuities at the interface between different layers.
- It allows for a rigorous treatment of elastic scattering with the different types of atoms in a multiatomic target.
- It gives a full implant distribution rather than only a few of its moments.
- It can generate as-implanted profiles as a function of key parameters such as dose, energy, tilt, and rotation angles.
- It can include the recoil effect due to atoms that are knocked into deeper layers from an impinging ion.
- It allows the simulation of the defect generation due to ion implantation.

Furthermore, once a physically based model has been set up and verified, the MC method can be used to generate "experimental data" for semiempirical models which are highly computationally efficient and desirable in technology development and process optimization. Other important features of the MC simulation of ion implantation are its inherent three-dimensionality, the fact that ion backscattering is naturally accounted for, and the fact that both amorphous and crystalline targets can be considered. Those distinctive features make the MC simulation the most suitable approach to the study of ion implantation.

The MC simulation is performed by following a large number of individual ion histories; each of them is made up of collisions with target atoms, along with straight flights between them. The model relies on two main assumptions. The first is that the projectile interacts with one atom at a time, and thus multiple collisions can be neglected. Such a "binary collision" model breaks down at low energies, when deflections can occur even at greater distances from a target atom. The second approximation involves the mechanisms for energy and momentum losses. The contributions coming from electronic losses (where the incident ion is excited or ejects atomic electrons, with large energy and small momentum transfer) and from nuclear losses (originated from nearly elastic collisions with the target atoms, characterized by large exchange of momentum) are considered independent of each other. The slowing down of the projectile results from the simple addition of the two effects. In the past, the basic theory to describe the penetration of charge particles into a solid was the one due to Lindhard, Scharff, and Schidøtt (LSS). The LSS theory has been successful in the prediction of primary ion range and damage distribution in amorphous semi-infinite substrates. Because of its assumption, it is not applicable to multilayers structures as often encountered in VLSI processing.

The simulation is made up of successive paths. After each of them, the energy of the ion is reduced by the amount of the electronic and nuclear energy losses, which are related to the momentum transfer to the target atom occurring during the collision. The history of each ion terminates either when its energy drops below a specified value or when the ion exits the target.

Different physical models can be used for the various phenomena involved in the energy loss process. An exhaustive

review is found in 58. A series of optimizations of MC procedures are discussed in Ref. 59. Such optimizations are crucial in order to overcome the major limitation of the MC approach—that is, the amount of computation needed to achieve an acceptable statistical accuracy.

A good example of a situation where the MC simulation works at its best is provided by channeling. If the target is crystalline in nature (and it is not damaged during the implantation process, as guaranteed by using low implant doses), then the stopping power of the medium is greatly reduced because of the large open spaces of the axial channels. The average ion penetration is then considerably increased with respect to amorphous targets. Although the number of ions that are well-channeled is considerably less than the number of the dechanneled ones, the well-channeled ions travel deep into the crystal and form a channeling tail in the resulting profile. It is the channeling tail that determines the junction depth. While the formulation of suitable transport equations for the channeling effect presents many problems, the direct MC simulation of the ion path through the target can be quite straightforward. The main obstacle is again the inadequacy of several available models for electronic stopping. Although still widely used, the LSS theory cannot properly describe channeling effects. Its shortcoming lies in the inability to account for the variation of electron densities between the atoms of the crystalline structure. In fact, an ion traveling along a crystal channel will experience less electronic stopping than another ion which more closely encounters the electron clouds of the target atoms. A proper model must therefore account for the electron distribution around each atom. For boron ions implanted in crystalline silicon, it has been shown that a combination of the two concepts of effective change and proton stopping power lead to reasonable models. There, the solid-state Hartree–Fock theory with a muffin tin structure has been employed, with the assumption of a spherically symmetric electron distribution around the Si atoms. While such an assumption is fairly reasonable for boron, which has most of the bound electrons in s orbitals (i.e., with spherically symmetric distribution), it has been pointed out that a further improvement is needed for the implantation of arsenic ions, where the electrons occupy p and d orbitals. One possibility is to calculate self-consistently the potential for the As projectile within the local density approximation, and then to determine the scattering phase shift and the scattering cross section needed in the simulation.

In the theoretical study of the implantation process, it is very important to be able to determine the cumulative damage imparted to the semiconductor crystal by the implanted ions. The MC simulation directly accounts for the effect of damage by explicitly simulating the formation of point defects, their recombination, and the effect that such defects have on each subsequently implanted ion (60,61). The procedure is as follows. As the simulated ion travels through the crystal, target atoms that receive sufficient energy (15 eV for Si) leave their lattice site, forming a vacancy or an interstitial. The interstitials travel until they lose the energy gained from the incident ion. The location of interstitials and vacancies are recorded for each collision cascade. At the end of each of these cascades, the interstitials and vacancies within the cascade are annihilated if they are located within a specified capture radius of one another. Furthermore, the recombination with damage caused by previous cascades is taken into account via a statistical recombination algorithm. The defects that survive recombination are weighted to reflect the fraction of the overall dose that the simulated projectile represents, and then they are summed to obtain the simulated point defect distribution as a function of depth. This distribution is then used to modify the crystal structure that subsequent ions will cross by determining the rate of statistical creation of (a) interstitials at one of the eight interstitial sites in the diamond lattice unit cell and (b) vacancies on lattice sites. Such a model is inherently homogeneous, and it appropriately describes the damage caused by lighter ions, such as for instance boron, where the net formation of cumulative damage occurs by the overlap of the damage caused by individual cascades. Heavier ions, such as arsenic, can lead to local phase transitions from the crystalline to the amorphous phase. This can be simulated by considering the presence of highly localized disordered regions, which are formed whenever the relative number of displaced atoms during an ion cascade exceeds a specified value (for instance, 10% in the region under consideration).

The traditional MC approach described above is based on the calculation of a large number of independent ion trajectories, where each ion is followed from its entrance in the target material down to its stopping or exit point. Peripheral areas with low exposure (that is, with a dopant concentration that is order of magnitudes lower than the maximum values) cannot be adequately represented, because insufficient statistics is provided by the simulation. This is a typical situation where rare events can play an important role. An approach similar to the one already used in electronic transport simulation has been suggested (62). The fundamental idea is to locally increase the number of calculated ion trajectories in areas with large statistical uncertainty. To this purpose, the simulated region is subdivided into layers, each characterized by a certain relative concentration level (with respect to the current maximum concentration achieved in the whole region). For each simulated ions, a number of "checkpoints" is set up during the flight path. At a checkpoint, the local dopant concentration is calculated together with the global maximum concentration. The corresponding layer of fixed relative concentration is determined, and a "trajectory split point" is defined at the checkpoint if the ion has moved into a layer with lower relative concentration than the one touched at the previous checkpoint. The position, energy, and velocity of the ion are stored, and they are used to generate a series of virtual branches of the ion trajectory that start at this split point. In this way, a tree of virtual trajectories is formed for each regular ion, so that the peripheral areas of the dopant concentration are represented by a much higher number of ion trajectories and the statistical noise is reduced. To obtain the correct concentration, each branch is assigned a weight. The virtual trajectories are generated with the same model and parameters as the regular ones, with their initial condition (energy, velocity, and position) being determined by the regular ion characteristic at the checkpoint.

Electron Beam Lithography

Electron beam lithography (EBL) is the standard way of fabricating masks for optical and X-ray lithography. Furthermore, direct electron beam writing on wafers is the only practical way to obtain ultrasmall linewidths. In EBL, finely

focused beams are used to expose polymeric resist layers. The ultimate resolution obtainable is not limited by the characteristics of the incident beam but rather by the electronic scattering with the resist and the underlying substrate. Such scattering leads to the so-called proximity effect and can be subdivided into three distinct contributions, namely the forward scattering within the resist, the backscattering within the resist, and the backscattering from the substrate.

The actual process of electron scattering in solids is very complex, and simplifying models are needed to achieve quantitative results via numerical techniques. As for the case of ion implantation, the best approach is the MC simulation. In fact, the simulation is very similar to the one described before in this section. In the simplest model (63,64), electrons undergo a series of elastic scatterings with the target nuclei. They also suffer energy losses because of the inelastic scatterings with the target electrons. The elastic scattering is modeled using the screened Rutherford cross section, while in the inelastic contributions the energy is assumed to be lost continuously between two successive collisions according to Bethe's formula. Between scattering events, the simulated electrons travel in a straight path, whose length is determined randomly according to the calculated cross sections. The sequence of free paths and scattering events is repeated until the electrons come to rest. Contrary to the ion implantation case, the quantity of interest here is the deposited energy rather than the position where the particle stops. It is, in fact, the energy passed by the incident electrons to the medium that creates the condition for the selective removal of the polymeric resist.

Several improvements have been suggested to the simple model described above. In particular, it was pointed out (65) that the production of secondary electrons as a result of an ionization process caused by the incident beam has to be accounted for. Since the energy deposition is inversely proportional to the electron energy, the contribution of the secondary electrons (which are slower and move in a direction almost perpendicular to the incoming flux) can be significant. An hybrid model has been set up which includes a discrete energy loss mechanism corresponding to ionization, in addition to the continuous energy loss mechanism described above. A further refinement of the model has led to the use of Mott cross section for the treatment of elastic collisions (66). Another improvement is needed in order to consider multilayered structures, as for instance those that serve as masks for X-ray lithography (67,68). Such masks are obtained by direct writing via EBL on a thin resist, followed by pattern transfer in a thick resist multilevel structure. An example is shown in Fig. 11, where the radial profile of the energy deposited at the resist–substrate interface is plotted for three different values of the beam energy, namely, 20, 30, and 40 keV. The simulated multilayer is constituted by 10 nm of Al, 1000 nm of resist [polymethyl methacrylate (PMMA)], 20 nm of Au as base plating, 200 nm of Cr for adhesion purposes, and 2000 nm of Si substrate. The length of the free path has to be determined by taking into account the details of the electron dynamics— that is, the possibility for an electron to cross one or more layers during its path. Forward and backward contributions are shown separately in the figure, each characterized by its own range. The parameter η is the ratio between the energy deposited by backscattered versus forward scattered electrons. The MC results show that the ultimate resolution

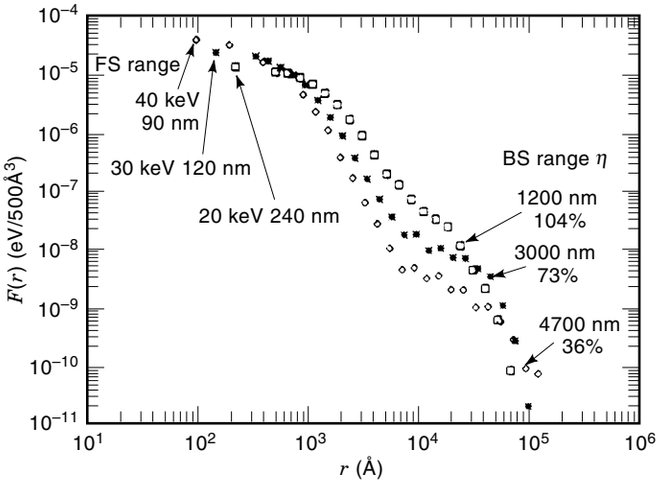

Figure 11. Radial distribution function of the total absorbed energy density for three different beam energies, with the relative contributions of forward scattering and backscattering.

limit is set by forward scattering, which has a width of 90 nm at the interface. The backscattering contribution from the metal is important despite the reduced thickness of the layer, but can be eliminated by increasing the beam energy. The simulation results can be parameterized in order to predict the EBL resolution under various exposure conditions. This is done via the so-called proximity function, defined from a Gaussian fit to the MC data. By a numerical convolution of the proximity function with the experimental pattern, it is possible to calculate the absorbed energy densities correspondent to the experimental situations. Figure 12 shows the absorbed energy density profile for the structure considered before and after an electron energy of 40 keV. The nominal line/space dimension is 100/400 nm. By assuming [according to the "Threshold solubility model" (69)] that the solvent devel-

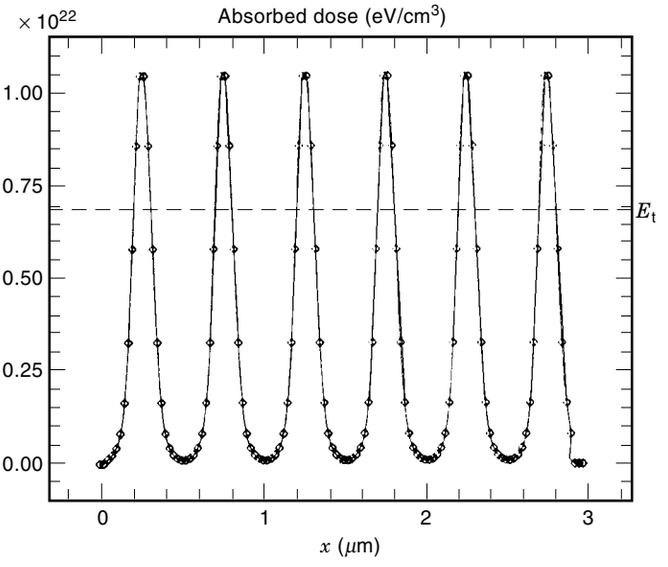

Figure 12. Profile of absorbed energy density for a nominal pattern of 5 lines/spaces of 0.1/0.4 μm.

ops all the PMMA irradiated above the threshold energy E_t, the developed lines would be around 150 nm. This value reproduces very well the experimental findings.

MC ANALYSIS OF CIRCUIT YIELD

The number of transistors on a single chip is approximately doubling every 2 years, as predicted by Moore's law. Such progress is sustained by the continuous drive toward smaller feature sizes, larger dies, and better packing efficiency. While technology allows a general reduction of the manufacturing costs per transistor, the level of investment to continuously upgrade such technology is becoming enormous. It is then vital for any manufacturer to improve the yield of the IC production. Yield can be defined as the ratio of the number of designs that pass the performance specifications to the total number of designs that are produced. Since the total number of designs produced might be large or unknown, yield is usually measured over a finite number of design samples or trials in the process known as yield estimation. As the number of trials become large, the yield estimate approaches the true design yield. Parameter values that have statistical variations are referred to as yield variables.

Statistical design techniques for ICs, including yield maximization and sensitivity minimization, are becoming increasingly important in IC technology development because cost effectiveness and competitiveness demand minimal experimental passes and short development times. A valuable statistical circuit simulation methodology must inherently predict device and circuit performance reliably and efficiently, relate their performance to fabrication-process parameters and account for the fluctuations of such parameters, and account for physical correlations among device model parameters. The use of numerical process and mixed-mode device/circuit simulators provides the needed capabilities for predicting the effects of technological variations on device/circuit performance. However, such an approach cannot be generally used for yield predictions or other statistical modeling due to the waste amount of computation time required. This is very efficiently done via MC techniques.

In general, there are two basic sources of yield loss in a fabrication process: local process faults and global process faults. The latter include for instance mask misalignment and linewidth variations. They can cause variations in speed and power consumption, thus affecting primarily parametric yield. The former include spot defects, such as oxide pinholes, extra metal, and extra or missing material defects. They affect functional yield, because they primarily affect circuit topology and can cause the chip to completely fail functionally. Local process faults are therefore called "catastrophic" faults. The most probable origin of such faults is the presence of dust particles, or in general some contaminant, on the mask or on the wafer surface. During the photolithographic processes, these particles lead to unexposed photoresist areas, or resist pinholes, thus causing unwanted material or unwanted etching of material on a layer. Although some crossover and coupling between the two types of faults occur, and manifest themselves at a circuit level, usually they are considered as uncorrelated and they are treated separately.

Catastrophic Yield

The effect of catastrophic faults on yield can be determined via a MC simulation made up of the following steps (70,71): (1) generation of a chip sample, according to specified layout design rules, (2) generation and placement of defects on the layout, and (3) analysis of the modified layout for circuit faults. A filtering step prevents uninteresting faults to be considered. Many chip samples are generated in a simulation. The defect diameter, type, and spatial distribution on each sample are selected randomly according to the defect statistics observed in the fabrication line. Once the defects have been placed on the layout, a series of fault analysis procedures are performed by looking at the defect neighborhood in order to determine which type of circuit fault has occurred (if any): short, open, open device, shorted device, new via, new device due to extra metal, or new device due to extra active material. Out of the resulting faults, only those affecting functional yield are kept, the others being filtered out. Some faults can also be combined together into a composite fault. Both fault analysis and filtering operation are guided by defect models, which specify which circuit faults can be caused by each defect type, which layers interact with the defect, and how layers are electrically connected together. The resulting output is a chip sample containing a list of the circuit faults that have occurred on it during the simulated fabrication. The chip sample fault lists are summarized to record the frequency of each unique fault combination and are then passed to an application postprocessor when the simulation is completed. In a typical simulation, the random number generator can be called millions of times. It is therefore essential to use fast generation algorithms.

Parametric Yield

Parametric yield analysis is the process of varying a set of parameter values, using specified probability distributions, to determine how many possible combinations result in satisfying predetermined performance specifications. The design is simulated over a given number of trials in which the yield variables have values that vary randomly about their nominal values with specified probability distribution functions. The number of passing and failing trials are recorded, and these numbers are used to compute an estimate of the yield.

Statistical design is the process of (1) accounting for the statistical variation in the parameters of a design, (2) measuring the effects of these variations, and (3) modifying the design to minimize these effects. This can be achieved by resorting to a purely numerical process and device simulators (72), to macro (response-surface polynomials) models (73), or to a seminumerical mixed-mode device/circuit simulator coupled to a parameter evaluator (74). In all three approaches, the first step is the specification of a set of measurable process parameters, which act as input variables for the simulation.

These variables have to be measurable, and they should be linked to more fundamental process parameters so that they can be used to control a fabrication process. For bipolar technology they can, for instance, be the base sheet resistivity, the epicollector doping density, and the effective surface recombination velocity at the emitter contact. The probability distribution for the input parameters has to be determined from measurements on different wafers and lots. In the second step, device model parameters are calculated, together with their variation following a given process fluctuation. In the example cited above for bipolar technology, model parameters can be the peak base doping density, the metallurgical

base width, and the epicollector width. Seminumerical approaches resorting to analytical formulation of device operation can be used, or, as an alternative, extensive process simulations coupled to physical device simulations can be performed. Clearly, the first approach is fast and inexpensive, while the second can be computationally very demanding but is free from the deficiency intrinsic to compact and simplified process and device models. The third step consists of a mixed-level device/circuit simulation that predicts the performance of a given circuit and its variation with fluctuation of process parameters.

In the MC-based yield analysis a series of trials is run in which random values are assigned to the design's statistical variable of interest, a simulation is performed, and the yield specifications are checked against the simulated measurement values. The number of passing and failing simulations is accumulated over the set of trials and used to compute the yield. The method is widely used and well-accepted as a way to estimate yield. The strength of the method is twofold: The accuracy of the MC estimates is independent of the number of statistical variables, and no simplifying assumptions are needed on the probability distribution of either component parameters values or performance responses. The weakness of the method is that a full network simulation is required for each trial and that a large number of trials is required to obtain high confidence and an accurate yield estimate.

Multiple yield analysis can be performed in order to adjust the nominal value of the yield variables to maximize the yield estimate. This process is called yield optimization or design centering.

BIBLIOGRAPHY

1. J. M. Hammersley and D. C. Handscomb, *Monte Carlo Methods,* London: Methuen, 1964.
2. Yu. A. Shreider (ed.), *The Monte Carlo Method,* Oxford, UK: Pergamon, 1966.
3. J. Spanier and E. M. Gelbard, *Monte Carlo Principles and Neutron Transport Problems,* Reading, MA: Addison-Wesley, 1969.
4. C. Jacoboni and L. Reggiani, *Rev. Mod. Phys.,* **55**: 645–705, 1983.
5. B. R. Penumalli, in W. L. Engl (ed.), *Process and Device Modeling,* Amsterdam: North-Holland, 1986, p. 1.
6. C. Jacoboni and P. Lugli, *The Monte Carlo Method for Semiconductor Device Simulation,* Vienna: Springer-Verlag, 1989.
7. S. Selberherr, *Analysis and Simulation of Semiconductor Devices,* Vienna: Springer-Verlag, 1984.
8. G. Baccarani et al., in W. L. Engl (ed.), *Process and Device Modeling,* Amsterdam: North-Holland, 1986, p. 107.
9. K. Kurosawa, *J. Phys. Soc. Jpn., Suppl.,* **21**: 424, 1966.
10. R. W. Hockney and J. W. Eastwood, *Computer Simulation Using Particles,* New York: McGraw-Hill, 1981.
11. P. J. Price, *Semicond. Semimetals,* **14**: 249–308, 1979.
12. An idea of the groups involved worldwide on MC device simulation can be found on the WEB site: http://www.research.ibm.com/0.lum/laux/dam.html, together with details of the MC simulation program DAMOCLES developed at IBM Yorktown.
13. A. Yoshii and M. Tomizawa, in W. L. Engl (ed.), *Process and Device Modeling,* Amsterdam: North-Holland, 1986, p. 195.
14. S. Babiker et al., *Solid-State Electron.,* **39**: 629–635, 1996.
15. V. Gruzinskis, S. Kersulis, and A. Reklaitis, *Semicond. Sci. Technol.,* **6**: 602, 1991.
16. S. Babiker et al., *IEEE Trans. Electron Devices,* **ED-43**, 2032–2034, 1996.
17. J. Zimmermann and E. Constant, *Solid-State Electron.,* **23**: 915, 1980.
18. A. Yoshii, M. Tomizawa, and K. Yokoyama, *IEEE Trans. Electron Devices,* **ED-30**: 1376, 1983.
19. Y. Awano, K. Tomizawa, and N. Hashizume, *IEEE Trans. Electron Devices,* **ED-31**: 448, 1984.
20. M. Tomizawa, A. Yoshii, and K. Yokoyama, *IEEE Electron Devices Lett.,* **EDL-6**: 332, 1985.
21. P. Hesto, J. F. Pone, and R. Castagne, *Appl. Phys. Lett.,* **40**: 405, 1982.
22. Y. Park, T. Tang, and D. H. Navon, *IEEE Trans. Electron Devices,* **ED-30**: 1110, 1983.
23. C. Moglestue, *IEEE Trans. Comput.-Aided Des.,* **CAD-5**: 326, 1986.
24. Y. Park, D. H. Navon, and T. Tang, *IEEE Trans. Electron Devices,* **ED-31**: 1724, 1984.
25. M. A. Littlejohn et al., *J. Vac. Sci. Technol. B,* **1**: 449, 1983.
26. T. Wang, K. Hess, and G. J. Iafrate, *J. Appl. Phys.,* **59**: 2125, 1986.
27. K. Brennan, *IEEE Trans. Electron. Devices,* **ED-32**: 2197, 1985.
28. S. E. Laux and M. V. Fischetti, *IEEE Electron Devices Lett.,* **9** (9): 467–469, 1988.
29. M. V. Fischetti and S. E. Laux, *Phys. Rev. B,* **38**: 9721, 1988.
30. K. Yokoyama et al., *IEEE Trans. Electron Devices,* **ED-32**: 2008, 1985.
31. T. Wang and K. Hess, *J. Appl. Phys.,* **57**: 5336, 1985.
32. R. Fauquembergue et al., *Solid-State Electron.,* **31**: 595, 1988.
33. M. Tomizawa, A. Yoshii, and K. Yokoyama, *IEEE Electron Devices Lett.,* **EDL-5**: 362, 1984.
34. U. Ravaioli and D. K. Ferry, *IEEE Trans. Electron Devices,* **ED-33**: 677, 1986.
35. K. Throngnumchai, K. Asada, and T. Sugano, *IEEE Trans. Electron Devices,* **ED-33**: 1005, 1986.
36. E. Sangiorgi, B. Ricco, and F. Venturi, *IEEE Trans. Comput.-Aided Des.,* **CAD-7**: 259, 1988.
37. M. Tomizawa, K. Yokoyama, and A. Yoshii, *IEEE Trans. Comput.-Aided Des.,* **CAD-7**: 254, 1988.
38. S. Bandyopadhyay et al., *IEEE Trans. Electron Devices,* **ED-34**: 392, 1987.
39. M. V. Fischetti and S. E. Laux, *IEEE Trans. Electron Devices,* **38**: 650, 1991.
40. I. Kizilyalli et al., *IEEE Trans. Electron Devices,* **ED-40**: 234, 1993.
41. D. Park and K. Brennan, *IEEE Trans. Electron Devices,* **ED-37**: 618, 1990.
42. J. G. Adams, T. Tang, and L. E. Kay, *IEEE Trans. Electron Devices,* **ED-41**: 575, 1994.
43. P. Dollfus, *J. Appl. Phys.,* **82**: 3911, 1997.
44. C. Canali et al., *IEEE Trans. Electron Devices,* **ED-43**: 1769, 1996.
45. K. Hess (ed.), *Monte Carlo Device Simulation: Full Band and Beyond,* Boston: Kluwer, 1991.
46. P. Lugli and D. K. Ferry, *IEEE Trans. Electron Devices,* **ED-32**: 2431, 1985.
47. U. Ravaioli et al., *IEEE Trans. Electron Devices,* **ED-32**: 2097, 1985.
48. P. Lugli and D. K. Ferry, *Physica B,* **129**: 532, 1985.
49. P. Lugli and D. K. Ferry, *IEEE Electron Devices Lett.,* **EDL-6**: 25, 1985.
50. A. Phillips and P. J. Price, *Appl. Phys. Lett.,* **30**: 528, 1977.

51. D. Liebig et al., *Microelectron. Eng.,* **19**: 127, 1992.
52. P. Lugli et al., HEMT models and simulations, in R. Lee Ross, S. Swensson, and P. Lugli (eds.), *Pseudomorphic HEMTs: Technology and Applications,* Dordrecht, The Netherlands: Kluwer, 1996, pp. 141–163.
53. W. Fichtner et al., *Proc. IEEE,* **72**: 96, 1984.
54. J. Sigh, S. Dudley, and K. K. Bajaj, *J. Vac. Sci. Technol. B,* **4**: 878, 1986.
55. P. A. Maskym, *Semicond. Sci. Technol.,* **3**: 594, 1988.
56. M. T. Robinson and D. G. Besco, *Phys. Rev. B,* **9**: 5008, 1974.
57. A. M. Mazzone, *IEEE Trans. Comput.-Aided Des.,* **CAD-4**: 369, 1985.
58. H. Ryssel and J. P. Biersack, in W. L. Engl (ed.), *Process and Device Modeling,* Amsterdam: North-Holland, 1986, p. 31.
59. G. Hobler, E. Langer, and S. Selberherr, *Solid-State Electron.,* **30**: 445, 1987.
60. K. M. Klein, C. Park, and A. F. Tasch, *IEEE Trans. Electron Devices,* **ED-39**: 1614, 1992.
61. S. Yang et al., *IEEE Trans. Semicond. Manuf.,* **9**: 49, 1996.
62. W. Bohmayr et al., *IEEE Trans. Semicond. Manuf.,* **8**: 402, 1995.
63. K. Murata, D. F. Keiser, and C. H. Ting, *J. Appl. Phys.,* **52**: 4396, 1981.
64. D. C. Joy, *Microelectron. Eng.,* **1**: 103, 1983.
65. S. Horiguchi et al., *J. Appl. Phys.,* **39**: 512, 1981.
66. K. Murata et al., *J. Vac. Sci. Technol. B,* **5**: 124, 1987.
67. M. Gentili et al., *Microelectron. Eng.,* **9**: 147, 1989.
68. P. Lugli, *IEEE Trans. Comput.-Aided Des.,* **9**: 1164, 1990.
69. A. R. Neureuther, D. F. Keiser, and C. H. Ting, *IEEE Trans. Electron Devices,* **ED-26**: 686, 1979.
70. H. Walker and S. W. Director, *IEEE Trans. Comput.-Aided Des.,* **CAD-5**: 541, 1986.
71. J. Khare and W. Maly, *IEEE Trans. Semicond. Manuf.,* **9**: 518, 1996.
72. I. C. Kizilyalli et al., *IEEE Trans. Electron Devices,* **40**: 966, 1993.
73. M. O'Leary and C. Lyden, *IEEE J. Solid-State Circ.,* **30**: 279, 1995.
74. H. Cho and J. Fossum, *Solid-State Electron.,* **38**: 1065, 1995.

PAOLO LUGLI
University of Rome