Solid-State Electronics Vol. 33, No. 6, pp. 591–623, 1990
Printed in Great Britain

0038-1101/90 $3.00 + 0.00
Pergamon Press plc

THE USE OF SIMULATION IN SEMICONDUCTOR TECHNOLOGY DEVELOPMENT

D. C. COLE[1], E. M. BUTURLA[1], S. S. FURKAY[1], K. VARAHRAMYAN[1], J. SLINKMAN[1], J. A. MANDELMAN[1], D. P. FOTY[1], O. BULA[1], A. W. STRONG[1], J. W. PARK[1], T. D. LINTON JR[1], J. B. JOHNSON[1], M. V. FISCHETTI[2], S. E. LAUX[2], P. E. COTTRELL[1], H. G. LUSTIG[1], F. PILEGGI[1] and D. KATCOFF[1]

[1]IBM General Technology Division, Essex Junction, Vermont 05452, U.S.A. and [2]IBM Research Division, Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, NY 10598, U.S.A.

(Received 28 November 1989; in revised form 5 February 1990)

Abstract—An overview is presented on the types of problems encountered in semiconductor technology development that are actively studied today via simulation methods. Most of the simulation examples presented here are ones that have been explicitly used in actual industrial semiconductor device design cycles to aid in the optimization of device structures. The examples described here include *process* simulations, such as the diffusion of dopant atoms, oxidation, etching, deposition, and epitaxial growth, as well as *device* simulations, which predict the flow of charge carriers and field distribution within a semiconductor device, given its material structure and operating conditions. The main aim here is to illustrate, by example, some of the capabilities of state-of-the-art simulators used in characterizing and predicting semiconductor process and device-related phenomena. We will attempt to outline the degree of sophistication of the physics incorporated in such simulation programs, and provide some contrast to the fundamental physics required for a complete physical description. As will be indicated, simulation development necessarily involves molding the appropriate physical models and numerical algorithms into a package that can be handled in a reasonable length of time by modern computing systems. We briefly outline some of the advances that have been made, and some concerns that remain, in such simulation development.

1. INTRODUCTION

In this article, we will present a number of pictorial examples that demonstrate various uses of simulation in semiconductor technology development. By means of these examples, we will survey various capabilities of simulators we have ourselves used on a regular basis to aid in the development of advanced semiconductor devices. Our belief is that such examples will quickly portray to the reader of this special issue of *Solid-State Electronics* how the applied physical theory often discussed here has been used for simulation experiments within industry, as well as academia.

To accomplish this task, yet limit the overview to a reasonable length, we will need to de-emphasize the details of any one physical model, as well as the numerical solution techniques involved in the simulation, and simply refer the reader to the appropriate literature on that subject. Our intent is to give the reader a strong feel for the wide range of problems that can be solved with today's simulators. This work is meant to complement the expected detailed discussion on physical models in the remainder of this issue, which, due to the complicated physical behavior involved, generally needs to be embodied in a simulation program to extract quantitative predictive results.

The simulation examples we will discuss here involve predictions of *process* behavior related to the fabrication of semiconductor devices, such as the diffusion of dopant atoms in crystalline materials, oxidation, ion implantation, etching, and deposition of thin films. Given the structure of a semiconductor device via such process simulation, the resulting *device* behavior can then be simulated by calculating the transport of carriers and associated field distributions due to applied, external fields. For both of these cases, we will examine real problems that have been studied at IBM during the course of semiconductor technology device development over the past 15 years. Our author list reflects our intention to gather a number of simulation examples that have been used as active research tools and aids for characterizing and predicting process and device behavior of semiconductor products under development.

To begin, we note that there are already a number of review papers related to process and device simulation[1–13], as well as books[14–21], and numerous articles dealing with specific models and applications involving semiconductor-related simulation programs. The field is still evolving, as more detailed physics is incorporated into such programs, as more powerful numerical algorithms are investigated, and as the capabilities of computing systems continue to improve. Our paper will present, by example, what we believe to be a fairly up-to-date snapshot in time of this evolving field of semiconductor-simulation capability.

However, we note here that our examples are almost exclusively drawn from simulation programs developed within IBM, such as FEDSS[22–24], TRIM[25,26], FIELDAY[27–39], OYSTER[40–42], and DAMOCLES[43–46], with an example on lithography taken from the University of California at Berkeley program, SAMPLE[47–52]. Our "survey" then should not be regarded as a review on the details of simulators in existence today, but rather as an overview on the types of problems such simulators are capable of handling, using IBM simulators for illustration.

Nevertheless, the reader should be aware that a number of other comprehensive simulation programs are in existence today. Indeed, several universities have created programs and released them for public use, such as the device programs PISCES[53], SEDAN[54], HFIELDS[55–57], and MINIMOS [58–61], and the process programs SUPREM I[62], II[63], III[64] and IV[65–67]. Also, other development groups in academia and industry have developed their own set of simulation programs. Here we mention, as a partial list, the programs PREDICT[68,69], IMPACT[70], SMART-P[71], and COMPOSITE[72], as well as the set of programs by AT&T and TMA[73]. Each of these programs has certain advantages over one another regarding specific physical models, the precise mathematical equations solved, the numerical algorithms used, and the pre- and post-processing handling of data. (For specific comparisons between simulation programs, see such yearly conference proceedings as NASECODE and NUPAD, which often contain papers that emphasize the dissemination of this type of information.)

The outline of the remainder of our paper consists, first of all, of a brief background in Section 2 on *modeling* activities in semiconductor technologies. Here, we will contrast various approaches that have been used for simulation-related studies, and attempt to convey what we mean by modeling in general. Next, in Section 3, we will describe a brief history of process simulation, and then move on to some specific examples: namely, the construction of a trench isolation region, the exposure and development of a photoresist masking region, and the formation of a BiCMOS structure. Section 4 presents a similar treatment on device behavior simulations. Here, the described examples include: (1) the effect, due to hot electrons, of accumulated charge trapped in the oxide region of a MOSFET device, (2) leakage current in a 3-D cell structure, (3) the "turn-on" characteristics of the corner region of a MOSFET device, (4) simulations of device behavior at room and liquid nitrogen temperatures, and (5) a Monte Carlo electron transport simulation of a MOSFET. Section 5 then turns to the problem of handling the large amount of data generated by these programs, as well as the detailed data required as input to the programs. Here we mention the emerging field of *scientific visualization*, which is proving to be very helpful in analyzing the generated data, as well as recent improvements in the geometrical modeling of 3-D structures, which is needed for characterizing complicated device structures. Also, in this section we note the importance of establishing an environment in which the different process and device simulators, and their pre- and post-processing programs, interact with each other and use a common data repository. Finally, Section 6 contains a few concluding remarks.

2. BACKGROUND ON SIMULATION FOR SEMICONDUCTOR RELATED PHENOMENA

2.A. *Level of physical detail in simulation programs*

Our present understanding of the most fundamental microscopic physics involved in semiconductor-related phenomena is generally considered to be quite accurate. A quantum mechanical description of semiconductor phenomena usually starts with an appropriate "many-body" Hamiltonian. Aside from initial conditions, this Hamiltonian is regarded as containing all of the required information for describing the correlated motion of the *electrons* and *nuclei* in a semiconductor material, while under the influence of applied external fields, thermal radiation, and other possible electromagnetic sources. In principle, this information is sufficient for an accurate description of all semiconductor phenomena that enter into semiconductor device fabrication and operation, whether the physical process of concern is, for example, the diffusion of dopant atoms, or the flow of electrons and holes.

Nevertheless, despite what most people regard as a highly accurate understanding of the microscopic physics involved in semiconductor phenomena, considerable limitations are imposed on us when attempting to use this knowledge for predicting more macroscopic physical behavior. Indeed, *quantitative* descriptions of the complex, correlated motion of the *electrons* and *nuclei* are generally quite intractable, without introducing a large number of approximations. Here is where physical *models* enter: specifically, in order to make approximations, empirical relationships ("laws") and parameters must be introduced. The domain over which these approximations hold, and the range over which the empirical parameters have been measured, then constitute the bulk of the basis for accurate physical descriptions.

Speaking rather loosely, we may view most semiconductor related processes as consisting of the transport of *ions*, *electrons*, and *holes*, as well as their *mixing* behavior, such as occurs in chemical reactions, and in the recombination and generation of electrons and holes. In this way, we may group together much of the following diverse phenomena, such as the conduction of electrons and holes, ion implantation, and the diffusion of dopant atoms within a crystalline structure. Even the oxidation of, for example, silicon,

can be similarly characterized, since it involves the chemical reaction of oxygen combining with silicon, and the transport of the chemical reactants and products involved in this process.

Of course, in some semiconductor processes the act of transport is of lesser importance than certain chemical processes, as in the case of the exposure of photoresist materials with light. Also, some problems of interest only involve equilibrium conditions without the flow of particles, as occurs in the stress distribution in a semiconductor structure after oxide has formed, or in the electron concentration and the electrostatic potential distribution in a device when no current is flowing.

Nevertheless, for our present discussion we can use the calculation of transport-related properties to indicate the types of computations, and the various levels of physical approximations, that are typically involved in describing semiconductor-related phenomena. Indeed, let us be even more specific and briefly survey numerical calculations that have been carried out for one particular transport process. By doing so, a clearer picture should form on what we mean by different levels of physical approximations used within the context of numerical simulation. The difference between these approaches essentially lies in the extent to which phenomenological models are incorporated within the calculations.

Although essentially any example will do, let us consider here the case of diffusion of dopant atoms in silicon. We choose this example partly because similar information on other examples, such as electron transport, appears to be more widely known, and partly because diffusion is, indeed, a very important part of semiconductor device fabrication. We emphasize, however, that the general scenario of physical descriptions discussed below for diffusion also holds, more or less, for the other relevant semiconductor phenomena of interest, such as electron transport and ion implantation. Indeed, many different levels of physical accuracy are presently pursued by various researchers attempting to describe and understand the observed behavior of these physical processes. (This point will again be noted in Section 4.A when discussing carrier flow in semiconductors.)

The diffusion of impurities in silicon has received considerable attention during the past 30 years or so, due to its technological importance in fabricating silicon-based integrated circuits. The underlying physics is quite complicated, since diffusion necessarily involves the role of defects in the crystalline structure. Consequently, the overall behavior includes the collective interactions of the silicon and dopant atoms, whereby each atom is influenced by the others. In turn, the electronic charge configuration is modified by, and also modifies, the atomic motion.

Some very fundamental computations have been carried out that identify the important mechanisms involved in diffusion processes within silicon. Using first-principles quantum mechanical calculations involving pseudopotentials[74] and a local-density-functional approximation[75], the total energy has been calculated for different configurations of core electrons and nuclei. Assuming a particular quasi-static change in position of these nuclei, then the total change in energy of the configuration of particles can be found for the particular "path of diffusion" chosen. Detailed calculations were first reported in Refs [76] and [77], while more recent ones are found in Ref. [78].

Calculations such as these are quite interesting and have helped to shed considerable light on mechanisms of diffusion, such as the role played by interstitials and vacancies in diffusion processes[78,79], as well as other possible contributing mechanisms[80]. Nevertheless, such computations are still quite far from actually *simulating* the process of diffusion. More specifically, what we mean here is the capability to calculate the distribution of dopant atoms within silicon at any particular time, and for a variety of physically imposed conditions, such as temperature and initial distribution of dopant impurities.

In order to address this problem, approximations presently need to be made in the physical description of the problem, with empirical relationships and parameters included in the calculations. To obtain time-dependent dynamical properties for the distribution of dopant atoms, "molecular dynamics" calculations[81] have been carried out in such exploratory work as Refs [82] and [83]. Here, the trajectory of the atoms are treated classically, with semi-empirically based interaction potentials between particles. These calculations have shed additional light on the mechanisms of diffusion, such as the dynamical behavior of "jump events" as atoms move through a crystal lattice[84]. Also, this method has provided a classical calculation for atomic relaxation and relaxation energies[85]. Nevertheless, such a simulation approach is still too computationally intensive to give results for use in device fabrication simulation.

Here, we will briefly mention a few other approaches that have been used to simplify the problem still further before turning to the method presently used, nearly to the exclusion of all other methods, by semiconductor device technologists. An interesting Langevin equation approach can be found in Ref. [86] for studying the Brownian diffusion of interstitials in a lattice. More widely pursued simulation methods, with alternative, and additional, physical approximations, involve *hopping* models[87–91]. These computations generally treat the atoms in a crystal as lying at only specific locations in the lattice, and then describe their motion as a random walk of vacancies and interstitials. Still, despite the simplifying assumptions in this approach, such simulations demand large computational resources, as well as requiring a more accurate characterization than presently exists for the empirical parameters involved in the random walk description.

Consequently, the method most widely used by semiconductor technologists for diffusion process simulation directly entails the numerical solution of a macroscopic equation, i.e. the diffusion equation, which contains empirically based diffusion coefficients. More sophisticated treatments tie together the diffusion of impurities with the diffusion of point defects, by solving a set of coupled partial differential equations[67,92–94] for dopant atoms and point defects. The empirical coefficients here include the diffusion coefficients for each diffusing species and the generation and recombination rates for point defects.

Thus, we see a wide range of approaches to calculating characteristic properties for the single process of diffusion of dopant atoms in a crystalline structure. These approaches do *not* all yield the same information for the diffusion process. Indeed, the first-principles, total-energy-calculation approach yields values for parameters that might be used within a diffusion-equation-solving program, which in turn would be used to predict actual changes in concentration of dopant atoms within a crystalline material.

A major thrust for improving the state-of-the-art simulation of semiconductor-related phenomena is the investigation of how the different simulation approaches properly relate to each other: specifically, under what conditions does one approach become invalid or require modification and, at the other extreme, under what conditions does the extra detail of more first-principles calculations become largely unnecessary?

As emphasized earlier, the situation we have described for simulating the diffusion of dopant atoms in a crystal is very similar to the situation that exists for simulating other semiconductor-related processes, such as ion implantation and the flow of electrons and holes. More specifically, what we mean here is that just as many different levels of physical description are actively being pursued via computational means for studying the diffusion of dopant atoms, so also one can find many levels of physical accuracy that have been investigated via simulation-related methods for the other processes just mentioned. A major dividing line in these computations is the degree to which quantum mechanical effects need to be taken into account in the simulation, and to what extent only classical effects, or semi-classical ones, are adequate. In general, the complexity of computer programs increases as more detailed quantum mechanical effects are taken into account. Of course, as conventional FET and BJT semiconductor devices become smaller, more accurate accounts of quantum mechanical effects become increasingly essential. Indeed, for such novel device structures as described in Ref. [95], quantum mechanical principles are already an essential part of their operation.

Finally, we should note that at whatever level of physical sophistication that is used for a simulation, there exists a wide range of sophistication for the numerical methods by which the computational algorithm can be constructed. Indeed, enormous effort is required in creating algorithms that are as computationally robust and efficient as possible, so as to produce results that numerically converge over a wide range of initial conditions, and yield answers in reasonable amounts of time. The complexity of appropriate numerical algorithms, as well as the ever present limitations imposed by computer resources (i.e. memory and speed), constitute two of the major constraints for obtaining accurate physical descriptions in advanced semiconductor process and device simulation programs.

2.B. Criteria for successful simulation development

Having briefly outlined the different levels of physical detail that one can expect to find in different simulation programs, we will now turn to mention some of the key points that enable successful simulation development within a semiconductor device technology environment. Indeed, over the last decade, improved process and device simulation capability has become increasingly important in the design and development of complex integrated circuits. This increased importance has been primarily due to the growing complexity of semiconductor fabrication processes, ever shrinking device dimensions, and the need to reduce to competitive levels the time and cost associated with VLSI and ULSI design and development.

With the rapid progress in the IC technology, maintaining an advanced and effective simulation capability requires a sustained commitment of resources. There are three main areas that need to be addressed in the development of state-of-the-art simulation: namely, physical model development, software development, and experimental verification.

Indeed, advanced multidimensional process and device simulation calls for models that are comprehensive and physically based. However, this task is usually difficult to achieve, since the physics of many of the processes of interest is quite complicated. In such cases, an alternative is to use empirically based macroscopic models, as discussed in the previous section, with parameters chosen that are related either to experiment, or else to computations that have been carried out with more first-principle microscopic calculations. Such models, while limited in scope, are attractive because they are less computationally intensive than first-principle models. Presently, both complex and computationally intensive first-principle models, as well as simple and empirically based models, play important roles in the simulation of semiconductor related phenomena.

The degree of model complexity forces the simulation algorithms and associated software to be complex. At the same time, advanced software development calls for programs that are numerically robust, modular in nature, optimized in their execution, and easy to learn and use. Particularly from the users' point of view, this latter criteria is one of

the most desirable features that can greatly influence the extent to which a program is used for various applications. As part of this focus, much attention has to be given to pre- and post-processors, data storage and retrieval, and other components that are needed for a complete simulation system.

Finally, a strong experimentation and characterization program is vital to successful process and device model development and verification. Clearly, the predictive capability of any model is a strong function of its foundation in reliable data. Both data generated by modeling experiments as well as technology development efforts need to be considered. From the practical point of view, such data should be readily available, and in a format compatible with each other. An ideal solution is to have a computer-based system where a variety of data and information can be collected and stored in a systematic fashion, and at any point accessed and used with great ease. As an example, a system called BEST[96] has been implemented at IBM and is aiding in both technology as well as process model development and verification efforts. Through this system, data and pertinent information is readily stored and retrieved, and there are also options for their processing and analysis.

The realization of better physical models greatly depends on accurate knowledge of materials and process parameters. The limitations in direct measurement methods for determination of some key parameters is one of the challenges that need to be addressed. For example, measurement techniques for dopant profiling have led to a better understanding and development of models for diffusion and ion implantation. However, all currently used techniques are capable of only measuring 1-D profiles. Nevertheless, there have been recent reports[97,98] that seem promising for quantitative determination of 2-D and 3-D profiles. Such techniques may greatly aid in the development of multidimensional models for ion implantation and diffusion, which will then provide higher accuracy in predicting device structures and their material properties. Using this information in device simulation programs will then yield more accurate predictions for device behavior.

3. SIMULATION OF PROCESSES FOR FABRICATION OF SEMICONDUCTOR DEVICES

Here we will first present a brief history on the progress that has been made in simulating semiconductor device fabrication processes, and then indicate some of the key processes that need to be simulated in order to accurately describe device structures. Finally, we will turn to three advanced process simulation examples.

3.A. Brief history of process simulation

The origins of process simulation can be traced to the 1960s. During this period some of the basic process models were developed that in later years were implemented and expanded upon for process simulation. Examples of such early work can be found in Refs [99–104]. These references include work related to modeling of thermal oxidation by the linear-parabolic model[99], ion implantation by the Lindhard–Scharff–Schiott (LSS) theory[100], and dopant diffusion[101–104]. In the area of diffusion, such effects as the built-in electric field[101], segregation during thermal oxidation[102–103], and interactions in multiple-dopant diffusion[104] were considered.

In the 1970s, as in the decade before, the evolution of process simulation was primarily driven by technological needs. Models for processes of interest were extended and refined. In the area of ion implantation, the LSS theory became widely used for describing the ion range distributions in various materials. According to this theory, an ion implant profile may be described to first order by a Gaussian distribution. Range statistics for a variety of ion and target combinations were determined and made available[105,106]. In the area of diffusion, some of the progress dealt with dopant diffusion under extrinsic conditions. Examples of models reported for arsenic, boron, and phosphorus are provided in Refs [107–109] respectively.

With the trend toward smaller devices and more complex processes, a strong need developed to account for what had originally been second-order effects within certain processes. Early models relied primarily on analytical equations, which became inadequate under more complex circumstances. Consequently, numerically based models were necessary for general applicability. The computer implementation of such models led to the development of process simulation programs containing models for a variety of processes. With the introduction of the 1-D simulator SUPREM I[62] in 1977, followed by enhanced versions II[63] and III[64], the concept of a general purpose process simulator for modeling of semiconductor fabrication steps was established.

Over the last decade, as the shrinking of device dimensions has continued, 2-D process simulation has become a necessity. The origins of 2-D process simulation may be traced to some of the work that was conducted in the 1960s for the analysis of lateral diffusion effects[110,111]. This work represented early attempts at modeling dopant redistribution in 2-D, and led to a better understanding of lateral diffusion near a mask opening for constant-source and instantaneous-source diffusing conditions. While such early 2-D process simulators were relatively simple programs and dealt with only one type of process, they paved the way for modern 2-D process simulation programs.

In the early 1980s, a number of 2-D process simulators were reported[22,112,113]. These simulators primarily focused on modeling of ion implantation, diffusion, and oxidation, and contained fairly basic capabilities for other processes. There were also

some efforts in simulating lithography, etching, and deposition types of processes[47,48]. More recently, there has been an increasing demand for simulators that provide the capability of modeling in a comprehensive manner as many semiconductor processes as possible. This goal can either be achieved by incorporating new features in a simulator, or by establishing links between simulators with different capabilities. An example where a combination of these two options has been adopted is with the process simulator FEDSS[22]. In this case, as the models for ion implantation, diffusion, and oxidation were further developed[23,24], more realistic models for etching and deposition have been added. Moreover, a link has been established with the SAMPLE program[47–52] for lithography simulation. Since SAMPLE also contains models for a variety of etching and deposition processes, these additional models may now be accessed from FEDSS.

3.B. Essential phenomena needed to be described

3.B.1. Ion implantation. Ion implantation is the principal technique for controlled doping of semiconductor devices. For submicron VLSI technologies, where low "thermal budgets" are required to minimize profile broadening by diffusion, correct prediction of ion implanted profiles has become crucial. Analytical models based on a Gaussian or other distribution functions may be used for simulation of ion-implanted profiles. For example, in FEDSS, models based on Gaussian, joined half-Gaussian, and Pearson IV distributions are available for simulation of any element into silicon or into films such as oxide or nitride on silicon. A Gaussian distribution may be used to account for the lateral scattering effects. Specific models for As, P, and B are available. For all cases, angled implants and irregular surface geometries, including undercuts, are allowed.

In addition to analytic implant models, Monte Carlo programs exist for accurately simulating the implantation of ions into semiconductor materials. For example, the TRIM Monte Carlo ion implantation program[25,26] is used within the FEDSS process simulator for 1-D and 2-D implantation in multilayered structures. Such a capability can greatly aid in the understanding of ion implantation (especially lateral scattering in 2-D) and can be used to generate new and improved parameters for analytic models, or to extend their range. Another powerful Monte Carlo program for ion implantation in solids is MARLOWE[114,115]. A desirable feature of MARLOWE is its capability of treating crystalline materials, rather than simply assuming the semiconductor material to be amorphous. Consequently, this program can simulate the channeling of dopant atoms in crystalline materials.

3.B.2. Diffusion. Correct modeling of dopant diffusion is one of the key elements in process simulation for VLSI. Point defects are responsible for most nonequilibrium effects in diffusion, and a complete diffusion model must be able to account for defect–dopant interactions. Presently, most process simulators model diffusion by a numerical solution of the diffusion equation for dopants, and incorporate nonlinear effects through use of a concentration-dependent dopant diffusivity relationship. Nonequilibrium effects, typically due to thin film/substrate interfacial reactions, are incorporated via boundary conditions. The examples we show later, involving the diffusion of dopant atoms, were obtained via this means with FEDSS.

To directly account for the influence of point defects, the relation for dopant diffusivity may be defined with terms that are dependent on the local concentration of point defects. These in turn can be obtained from the solution of diffusion equations for defects. This approach may be useful for modeling such effects as 2-D oxidation enhanced diffusion of dopants by supersaturation of silicon interstitials[67] (see Ref. [116] for a review on the diffusion of impurities in silicon). Some recent reports also present alternative models with improved degrees of generality in terms of an explicit accounting for the interactions of dopants with point defects[92–94]. The most general form requires the simultaneous solution of a set of nonlinear coupled differential equations for dopants and defects, where the point defect distributions are explicitly calculated. This approach is under development for extensions to FEDSS. Such a simulator will be applicable for modeling a variety of explicit point-defect mediated diffusion effects, such as oxidation enhanced diffusion and nitridation-retarded diffusion. The simulation of rapid thermal annealing (RTA) processes will also require, at the very least, such a formulation, since the short times and high temperatures associated with RTA constitute a highly nonequilibrium state where impurity and point defect "reactions" tend to dominate "diffusion."

3.B.3. Oxidation. Historically, modeling of thermal oxidation of silicon has been based upon the well-known linear-parabolic growth model due to Deal and Grove in 1965[99]. This model is inaccurate in the "very thin" oxide regime (i.e. oxides less than about 20 nm thick). Since this regime has become very important in present day technology, many models have been proposed to account for the growth of thin oxides. In FEDSS, we have chosen two models for the thin-oxide regime: one is an "extension" by J. Slinkman and C. Schaeffer of the work by Deal and Grove that accounts for stress[117], while the other is due to Hu[118].

The problem of modeling oxidation is compounded considerably in 2-D, where one can no longer ignore the mechanical behavior of the growing oxide. Specifically, oxide is known to exhibit fluid flow at process temperatures, but there is currently no consensus as to whether it is viscous, viscoelastic, or even plastic flow[119]. The examples we present shortly are results of FEDSS simulations where the

following models are invoked: the "extended" Deal–Grove growth model[117] and the viscous flow model of Ref. [120] that was modified in Ref. [24]. The oxidation of both silicon and polysilicon are simulated and there is some capability for (empirical) "oxinitridation."

3.B.4. Etching and deposition. The complex nature of material etching and deposition processes has limited the development of physically based models for such processes. Consequently, many simulators, such as FEDSS, include user-defined models for material removal and addition. Presently FEDSS also has a basic etch model that is material-selective and etch-rate sensitive, as well as models for isotropic and anisotropic deposition.

3.C. Trench process simulation

Now let us turn to the first of our examples of semiconductor process related simulations. As will be seen, the example we have chosen involves a large number of process steps (i.e. $N \gtrsim 50$). Moreover, the structure described here represents a technologically important component in the advanced CMOS development of high density DRAMs: namely, the creation of trench structures in silicon[121,122]. Filling these trenches with highly doped polysilicon enables them to be used as storage-capacitor nodes (also see Section 4.C), while an oxide fill provides space conserving isolation regions to be formed between adjacent semiconductor devices (also see Section 4.D).

One of the difficulties involved with fabricating these trench structures is that stress-induced dislocations can form around the trenches. These defects can cause poor device performance, or possibly even an inoperable device. Such dislocations often arise during thermal oxidation, since volume expansion from the reaction at the oxide–silicon interface can create large pressures around the trench structure. To fully understand and prevent such dislocations requires costly and time consuming experimentation. Simulation, as a complementary tool, can greatly aid in reducing the required experimental efforts. An accurate simulation will pinpoint the regions of high stress, as well as enabling the technologist to perform simulation experiments of alternative process steps that might reduce, or even eliminate, these effects.

Indeed, using a comprehensive process simulation program, one can simulate a fairly complex structure and analyze stresses that may arise during thermal oxidation, as is shown in the following example. This simulation was carried out using FEDSS. Figure 1(a) shows the mesh generated for a structure of interest that was subsequently used to form a semi-recessed oxide (SROX) region. The structure in Fig. 1(a) consists of a trench formed in silicon, which was first covered by a triple layer of insulators, then filled with polysilicon, and finally covered, as shown, by pad nitride. Figure 1(b) illustrates the simulated structure after oxidation. Here, the computational algorithm used for simulating the oxidation growth required automatic mesh generation capability[23,24]. A corresponding SEM picture for the actual processed structure is shown in Fig. 1(c). The shape of the simulated SROX compares well with the SEM cross section. Based on simulation, the total SROX thickness is 0.59 μm, the lateral "bird's beak" structure is 0.45 μm, the distance of the pad nitride wrapage is 0.26 μm, and the oxide thickness on silicon is 0.45 μm. These results are in good agreement with the corresponding values of 0.57, 0.45, 0.24, and 0.44 μm from the SEM measurement in Fig. 1(c). For the structure in Fig. 1(b), a stress distribution can be calculated. For illustrative purposes, the hydrostatic pressure as calculated by FEDSS is shown in Fig. 1(d) (shear stress components are not shown). The bright green region in the nitride corresponds to high compressive stresses due to the growing oxide. In the vertical bird's beak region, we see that a tensile (bright red) hydrostatic pressure is predominent in the growing oxide due to the oxide volume expansion. However, we note that shear forces can contribute significantly as the oxide flows in this constricted volume region of the bird's beak. The FEDSS simulation predicts high compressive stresses on the surrounding silicon and polysilicon regions. These stresses can contribute to the formation of dislocations in the silicon, which are observed under certain process conditions.

Figures 1(e)–(g) illustrate how a 3-D analysis can be carried out for the previous structure. Figure 1(e) shows the mesh of Fig. 1(a) extruded into the third dimension, with some of the pad nitride layer cut away on the sides. Thus, Fig. 1(a) represents the cross section along points $A \rightarrow B$ in Fig. 1(e). Taking the cross section along points $C \rightarrow D \rightarrow E$, a different oxide growth structure and stress distribution can be examined. Figure 1(f) shows the oxide growth, plus a blowup of the bird's beak on the left-hand side of the trench. The triple layer of oxide–nitride–oxide is visible in the blowup. Figure 1(g) shows the associated hydrostatic pressure plot.

3.D. Lithography simulation

Our second process simulation example deals with lithography. Since the wavelength of visible light is comparable to device dimensions for submicron device structures, then for conventional photoresist exposure methods, the image obtained from a lithographic mask will not match the intended masking design. Moreover, other complicating factors often arise due to topographical variations in the photoresist and in the structures underneath the photoresist. Developing the photoresist introduces another factor that typically offsets the final shape from the intended shape, as well as giving the edge of the photoresist a nonvertical profile.

In order to obtain accurate process simulation capabilities, these factors need to be taken into account since they govern the size and shape of the

actual physical mask: namely, the final photoresist structure. The developed photoresist "mask" controls the sections of underlying structures that are eventually oxidized, or subjected to ion implantation, or other processes.

The simulation program SAMPLE [47–52] can be used to predict developed photoresist shapes under various conditions. Here we show an example obtained by linking the SAMPLE and FEDSS simulators together. As can be seen in Fig. 2(a), simulated resist contours were obtained from SAMPLE that corresponds to developments at 10 s intervals. In Fig. 2(b), these contours have been mapped and reflected onto a structure of silicon, oxide, and nitride, which in turn was generated from FEDSS. The meshes associated with these structures are not shown in this figure, but they are shown in Fig. 3. Here, various other options within FEDSS are illustrated for mapping the SAMPLE profiles onto a FEDSS structure, such as positioning the two pieces together, or perhaps extending, or truncating, the end points of the photoresist structure. Control over the generated mesh within the photoresist region, as well as the other regions, is, of course, essential in obtaining accurate simulation results for subsequent process steps.

3.E. BiCMOS structure

Our third process simulation example consists of the construction of a field-effect transistor and a bipolar junction transistor that are placed in close proximity to each other. The inherent complexity associated with the development of the fabrication process for a structure, such as this one, stems in part from the myriad independent parameters involved in each process step. Moreover, fabrication is compounded by the inadequacy of experimental methods to properly characterize impurity distributions in 2 and 3 dimensions. Carefully planned experiments, even when based on statistical arguments, cannot hope to *economically* provide a complete characterization of device performance as a function of all possible process variations. Simulation can be a very effective complement to experimental work, particularly in the examination of very localized phenomena. Figures 4(a) and (b), for example, depict the results of a partial sequence of *simulated* process steps that define a local cross section of a typical BiCMOS device structure. The simulated structure shown was constructed using the FEDSS process simulation program.

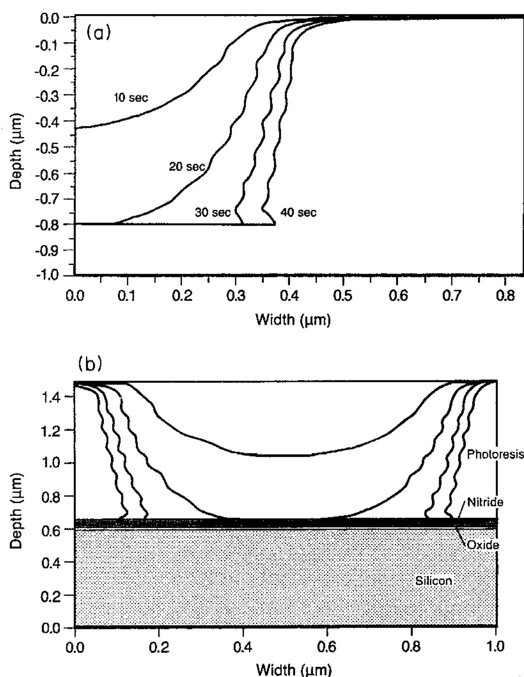


Fig. 2. Example of FEDSS-SAMPLE lithography simulation. (a) Simulated resist contours from SAMPLE after development at 10-second intervals. (b) Mapped and reflected SAMPLE resist contours on a silicon, oxide and nitride structure generated by FEDSS.

The initial process simulation step consisted of the growth of an epitaxial layer on a substrate of bare silicon. Implantation of a subcollector then followed. This region is clearly visible as the large red cloud in Fig. 4. An alignment signature was preserved at the appropriate point in the process by exposing the structure to a furnace anneal in an oxidizing ambient. These oxidation steps were accomplished as follows: first, a pad oxide growth was simulated, then a pad nitride deposition, and finally a furnace anneal in the appropriate ambient conditions. For this last step, parameters such as ambient gas (*wet or dry* O_2), ambient pressure, HCl content, gas flow rates, time, and temperature, were specified to characterize the process.

Several features of the resulting SROX structures are noteworthy that are predicted by the simulation. For example, a noticeable slope exists between points A and B in Fig. 4. This effect is due to the earlier subcollector oxidation process, which initially caused a slight topographical variation. The close-up of Fig. 4(b) highlights this phenomenon as well as providing a more detailed view of the MOSFET.

A sequence of ion-implantation steps was simulated. The purpose of these steps was to optimize device performance by (1) achieving a low-resistance collector contact, (2) tailoring the FET threshold voltage, and (3) creating lightly-doped source (LDS) and drain (LDD) structures, with halos. Most of the implants were done through oxide layers—sacrificial, gate, or base, as appropriate. The impurity source for the emitter region of the BJT was prepared by first depositing, and then doping, a layer of polysilicon.

Although only a few features of the BiCMOS structure have been highlighted here, this process simulation contains well over 100 distinct process steps that are necessary to accurately characterize the state-of-the-art BiCMOS device structure. Consequently, upon coupling this structure with a device simulation program, enough detail exists within the

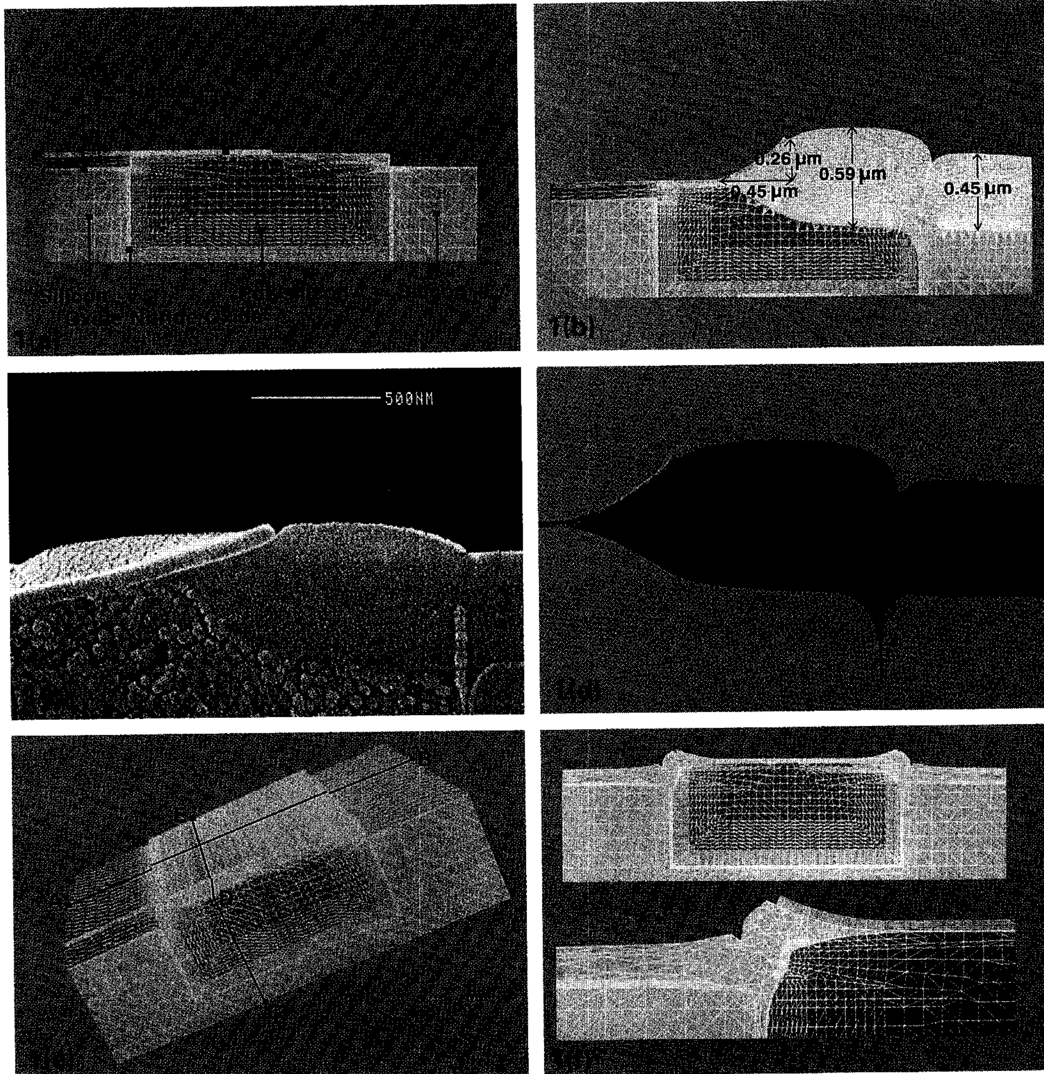
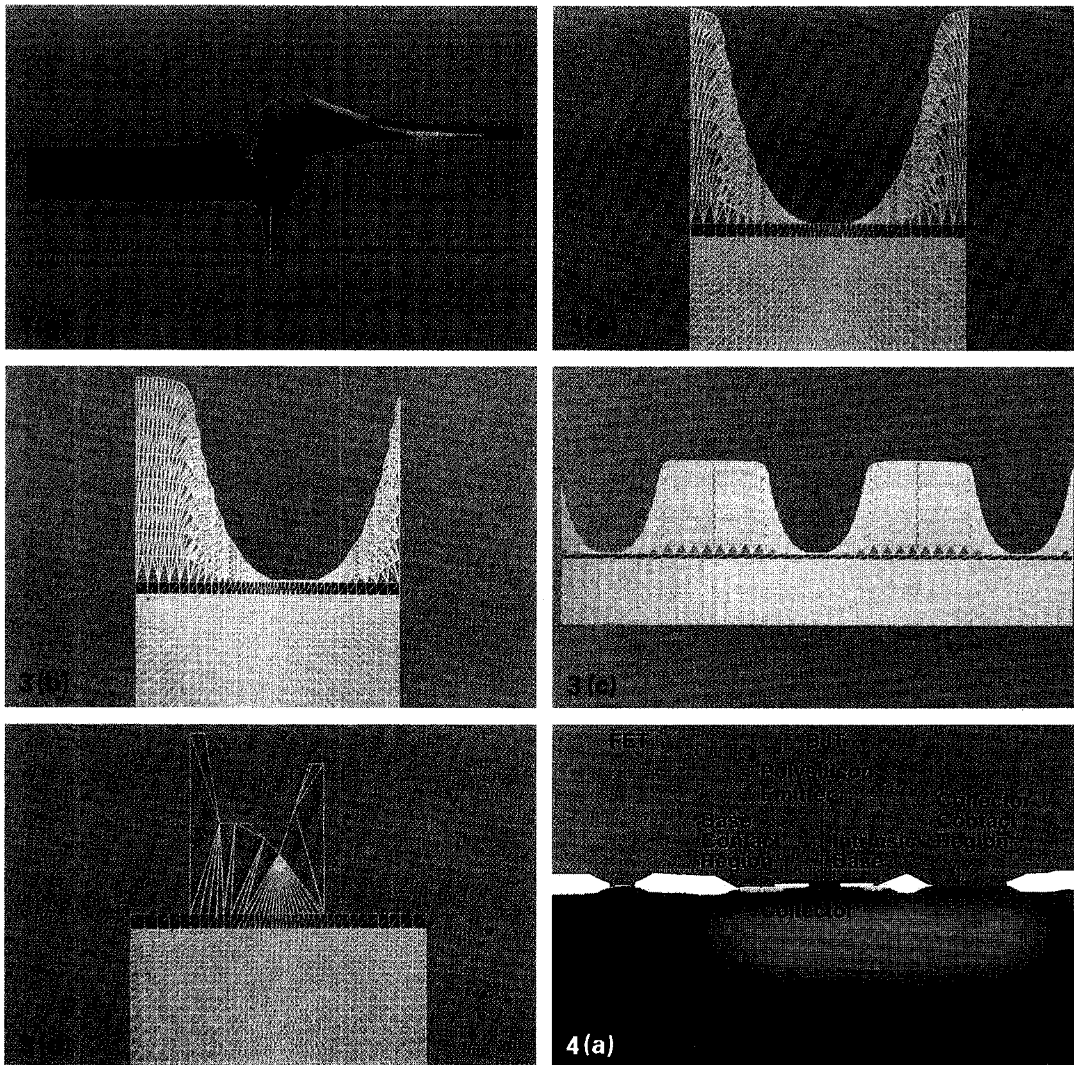


Fig. 1. (a) FEDSS mesh for a structure consisting of a trench formed in silicon, covered by a triple layer of insulators, and filled with polysilicon. The top surface is partially covered by pad nitride. (b) FEDSS mesh for the structure shown in Fig. 1(a) after thermal oxidation to form a semi-recessed oxide region. (c) SEM picture for actual semi-recessed oxide region simulated in Fig. 1(b). (d) Simulated hydrostatic pressure distributions in the semi-recessed oxide region at the end of oxidation. High compressive (green) hydrostatic pressure exists in the nitride pad region on top of the lateral bird's beak area, and high tensile (red) hydrostatic pressure exists around the vertical bird's beak region. The maximum compressive and tensile regions are approximately 10^9 dynes/cm^2. These quantities are shown here on a logarithm scale, with 64 separate contours. (e) Here, the 2-D mesh in Fig. 1(a) was extruded into the third dimension, with some of the pad nitride cut away. (f) The top part represents the cross section indicated by points $C \rightarrow D \rightarrow E$ in Fig. 1(e) after growing oxide for 10 min. The bottom part shows a blowup of the area of interest. (g) The hydrostatic pressure distribution for the bottom part of Fig. 1(f). (See overleaf for Fig. 1(g).)



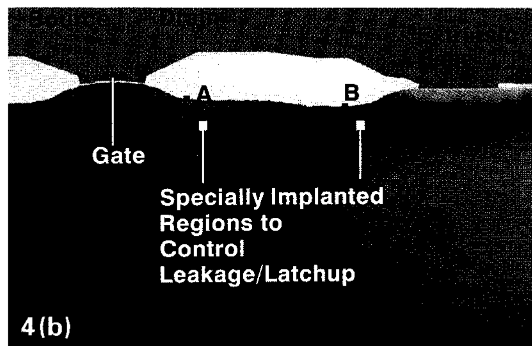


Fig. 3. Example of FEDSS-SAMPLE mapping options. (a) A mapped, reflected, and meshed structure. (b) Translation of an isolated resist structure. (c) Periodic resist structure generated from translation. (d) This unusual looking structure is shown here to indicate that the user can define his own structure, and can also specify the mesh density in the photoresist structure. The mesh in the photoresist is automatically matched up with the mesh in the semiconductor structure, thereby enabling subsequent process calculations for the two regions.

Fig. 4. BiCMOS cross section generated by FEDSS. (a) An n-MOSFET and a vertical *npn* BJT are shown. (b) Blowup of the MOSFET to show the source and drain, which contain LDD and LDS/halo structures.

simulated structure to yield the following types of device design information: (1) sensitivity analysis of process ground rules, (2) leakage, (3) breakdown, (4) latchup, and (5) general device performance.

4. SIMULATION OF SEMICONDUCTOR DEVICE BEHAVIOR

Our treatment of semiconductor device behavior simulation is similar to the outline of Section 3 on process simulation. First Section 4.A contains background on advances that have been made in device simulation. Sections 4.B–4.G then turn to the parts we mainly want to emphasize in this article: namely, specific pictorial examples of simulations of semiconductor device behavior.

4.A Background on device simulation

Until fairly recently, advances made in the actual simulation of conventional field-effect and bipolar junction transistors have largely been confined to the so-called *drift-diffusion* (DD) description of electron transport. The equations embodying this description for the flow of electrons and holes in a semiconductor material are given for the 1-D case in Shockley's classic paper in 1949[123], and for the general 3-D case in Shockley's monograph[124] and Van Roosbroeck's article[125], which were both published in 1950. Of course, clear discussions can be found in more recent standard references, such as Refs [126] and [127]. For a concise formulation of the initial DD approach for studying device behavior, see Section 2 of Ref. [128]. Here, the DD relations for the current densities, the current continuity equation, and Poisson's equation, are outlined for this "conventional model", as well as what Ref. [128] terms "auxiliary relations" between the carrier densities and their quasi-Fermi energies, and between the intrinsic Fermi level and the electrostatic potential.

Significant developments leading to, and extending, this DD approach were made by Bardeen and Shockley[129] for a dopant-dependent variable energy bandgap, and Kroemer[130] for recognizing that a position-dependent bandgap would affect the flow of electrons and holes as though "quasi-electric" fields were present. Other advances in accounting for additional physical effects within a DD approach have largely been concerned with heterostructures, more accurate recombination–generation models, Fermi–Dirac statistics for degenerate conditions, some additional high-doping effects, attempts at modeling changes in the effective local mobility at material interfaces, and contact models. Reference [14] describes how some of these effects have been incorporated into a numerical simulation scheme (also see, for example, Ref. [131] and references cited therein). The physical models used in a DD formulation are essentially phenomenological ones, such as mobility and recombination–generation models.

A substantial effort in DD simulation has involved the numerical aspects. Indeed, since carrier concentrations can typically vary by 10–20 orders of magnitude or more in standard semiconductor devices, then a simple linear or polynomial interpolation for discretization of the carrier densities is quite impractical. Instead, the discretization scheme that Scharfetter and Gummel proposed in 1969[132], with some modifications and extensions[131,133], has been the basic approach for simulators today. Even with this scheme, some regions of a device require a very fine discretization, such as in the channel region of an FET, where the carrier density changes dramatically within just 20 nm across the depth of channel.

Apparently the first suggested numerical scheme for solving the DD equations was made by Gummel in 1964[134] for a 1-D analysis of a bipolar transistor. Two-dimensional simulations were developed in the late 60s and early 70s[135–141]. As smaller devices have been made, geometry effects have become increasingly important, thereby necessitating the need for 3-D simulators. Development of such programs were first reported in the early 80s[29,142–143]. (For specific examples where 3-D simulation capability is essential for accurate device characterization of present day technology, see Sections 4.C and 4.D.)

The DD simulation approach has been the main tool for practical device engineering types of application. Indeed, nearly all of industry presently uses the DD method for device analysis of conventional FET and BJT structures. Even for MOSFET device structures with submicron channel lengths of 0.5–1.0 μm, the DD method can yield satisfactory analysis of many aspects of device behavior, such as the source-to-drain current as a function of applied biases, and estimates of leakage currents. For typical simulation examples of devices of these sizes, improvements beyond the DD method are not usually the most significant factor for obtaining sensibly accurate simulation results. Instead, other critical factors exist, such as a close characterization of doping profiles and of material interface properties.

Nevertheless, the physics embodied by the DD approach is limited. Developing the next generation of simulation capability demands considerable work. Hence, we presently see the situation where the literature contains numerous papers on the development of simulation methods that are physically more accurate than the DD method. Here, a major concern is making these new programs numerically robust, and in keeping the computational time down so as to remain practical for multiple simulation runs of semiconductor devices. Undoubtedly, as device sizes continue to shrink, these methods will eventually become essential for all of industry in general.

An improved physical description over the DD method becomes increasingly critical as silicon based device dimensions approach 0.25 μm, and smaller, in order to account for effects such as velocity

overshoot. Also, quasi-ballistic transport becomes important for MOSFETs with channel lengths below about 0.10 μm. Indeed, even for much longer channel-length devices, an accurate description of the energy and velocity distribution of electrons throughout the device certainly requires more than just the usual DD description coupled with the common assumption of a drift-Maxwellian velocity distribution. Detailed band-structure effects are required to properly characterize the energy distribution of carriers far away from the valleys and peaks in the conduction and valence bands[43]. Moreover, high electric fields can occur over fairly short distances in even conventional MOSFETs with channel lengths of 1 μm or more, as in the case when pinch-off conditions occur. Indeed, a distance of 20 nm is not an uncommon length for the width at half-maximum of the peak of the electric field component along the channel, just at the onset of pinch-off conditions. Also, the average current flow usually curves away from the Si/SiO_2 interface and then curves into the heavily doped drain pocket (see the illustration in Fig. 5a). Since the drain region in conventional MOSFETs often has effective junction widths around 50 nm, or smaller, large electric fields can exist in these regions that give rise to high energy carriers. (See, for example, Fig. 20a, which gives the energy distribution of electrons in a 0.35-μm channel length n-MOSFET. Here, much of the high-energy region of the electrons lies along the edge of the drain pocket.)

Thus, even for a MOSFET with a channel length of 1 μm or more, we can expect that there will be small, but critical, regions within the device in which the electron distribution will be altered from a simple drifted-Maxwellian velocity form. Within these regions, highly energetic carriers can result in carrier transport up into the gate oxide, or cause a sufficient number of impact ionization events to yield significant substrate current. As discussed by Barker[144], we expect that the full Boltzmann transport equation is necessary for accurate descriptions of the complete behavior of submicron devices (i.e. the accurate calculation of quantities like gate and substrate currents for MOSFETs and the calculation of general device behavior during fast switching conditions), and that even the Boltzmann equation may not be fully satisfactory for device scales below 0.25 μm.

Of course, most macroscopic device properties can certainly still be simulated with the DD method, or by modifications of this method, without necessarily solving the full Boltzmann equation. Parameterized phenomenological models, such as the mobility model[145–152], enable some shortcuts to be made. However, the details of the energy distribution, and even the precise carrier density (e.g. including an energy balance equation alters the calculated carrier density[57]), are simply not available without more sophisticated treatments.

In what follows, we outline some of the research to date on methods that go beyond the DD approach, in order to provide a flavor for the motivations, concerns, and focus in such activities today. Most work on improving the physical description for conventional device simulation has centered on the Boltzmann transport equation, which represents a semiclassical description of carrier transport. Quantum mechanical effects enter via the effective masses of the carriers, which depend on the carrier momentum, and via the scattering mechanisms. Otherwise, the carriers are treated within this description as classical charged particles following Newton's equation of motion.

In 1962, Stratton[153] made a significant contribution toward a better physical description of carrier transport, as well as an improved understanding of the relationship between the DD description and the more accurate Boltzmann equation description. Bløekjær[154] improved the theory further in 1970 and helped point the way toward what is now often referred to as the "hydrodynamical model" for carrier transport. This method involves multiplying the Boltzmann equation by successive powers of components of the velocity vector, and then calculating expectation values for all terms so as to yield relationships between moments of the velocity components. The resulting equations are essentially the Euler equations, as described in, for example, Refs [155] and [156]. The full set of velocity moments yields the complete electron velocity distribution. Dealing only with moments up to the second power, as presently done in most hydrodynamical models, means that approximations need to be made to produce phenomenological models for the remaining higher-order moments in the transport equations.

Cook and Frey[157] showed a clear set of assumptions that relate the DD approach to the Boltzmann equation. Further significant work on the hydrodynamical model can be found in Refs [158–165], with much of the attention focusing on improved numerical discretization schemes for handling the additional higher-moment terms.

Besides work on the hydrodynamical model, much has been accomplished by researchers on solving the full Boltzmann transport equation for device structures of interest. The main numerical approaches have been iterative and Monte Carlo methods. References [166–168] review numerical and analytical work on the Boltzmann equation up until the early 1980s. Since that time, the Monte Carlo method has clearly been the dominant focus in research activity. References [169–171] review this method.

The Monte Carlo method for solving the Boltzmann equation is similar to the molecular dynamics approach[81] described in the discussion on diffusion in Section 2.A. More specifically, however, the trajectories of electrons and holes are followed in space while under the influence of applied electric fields. Probabilities are calculated for scattering events due to phonons, carrier–carrier interactions, ionized and neutral impurities, and surface scattering at interfaces

between the different types of materials. These scattering events are treated as introducing instantaneous changes in the carrier trajectories. Impact ionization and the recombination–generation of electrons and holes can also be taken into account in a probabilistic manner. By summing over a large number of carrier trajectories, statistical results can be obtained that describe the energy and momentum distribution of carriers at each point in space within a device structure.

Fairly recently, the Monte Carlo simulation program called DAMOCLES was developed by M. Fischetti and S. Laux. For a detailed description of the physics embodied in this program, as well as new results for submicron device behavior, see Refs [43–46]. The program DAMOCLES has been used to simulate the behavior of 2-D device structures including both FETs and BJTs. Besides conventional silicon based devices, DAMOCLES can handle some heterostructure devices. Significant features of this program are that carrier–carrier interactions are taken into account, and Poisson's equation is solved self-consistently with the Monte Carlo part of the program. This program significantly advances the state-of-the-art in device modeling because it incorporates a full band structure for the carriers, along with statistical enhancement features for aiding the calculation of statistically rare events. Consequently, the program can simulate the behavior of carriers that are sufficiently energetic to lie far from the minimum and maximum in the first conduction and valence bands, as well as carriers that lie within higher-order energy bands. These regions are typically the important ones for giving rise to interesting hot-electron effects, as exhibited by gate and substrate currents in MOSFETs. Section 4.G describes an example of a simulation of a MOSFET using DAMOCLES.

Although much of the present interest in such Monte Carlo simulations lies with investigating the full behavior predicted by Boltzmann's transport equation, considerable interest also exists in reducing the enormous CPU times required by this method. Here we mention Refs [172] and [173], as well as a number of references cited in Ref. [173] that have explored more approximate, but faster, variations on the strict Monte Carlo approach.

Finally, we want to mention here the significant work that has been done by a large number of researchers on yet further improvements in the physical descriptions of charge-carrier transport in semiconductors than is provided by the semiclassical Boltzmann transport equation. For example, advances have been made in describing quantum-size effects, such as occur in the inversion and accumulation layers at semiconductor–insulator interfaces, heterojunctions, and quantum wells. Such problems have been treated in various simulation work by taking the one-electron Schrödinger equation into account[174]. Quantum interference effects and other related topics of interest are discussed by Datta in Ref. [175]. Finally, significant work has been done involving quantum transport theory; Refs [176] and [177] contain reviews of this field. A recent example of simulation results involving quantum transport theory, which was applied to different materials, is reported by Brunetti *et al.* in Ref. [178]. For an elementary introduction to computer simulation of quantum transport and tunneling, see Barker's discussion in Ref. [179].

Thus, as discussed in Section 2.A, many levels of physical accuracy are presently being researched for simulation-related work on carrier transport. We now turn to examples of simulation that we have used in our own work involving conventional FET and BJT device structures. Consequently, our focus is mainly on the more macroscopic DD description of carrier transport which, as discussed earlier, is adequate for predicting the general behavior of all but the most experimental of such structures in industry today. As will be seen, the 3-D aspects of conventional devices introduce many interesting issues that can be explored by DD simulation methods. We also report on some low-temperature simulation work via this method. Our final example in Section 4.G contains a simulation based on the more fundamental physical description of the semiclassical Boltzmann transport equation.

4.B. Some early examples of device simulation

The continuing efforts by device designers to shrink feature sizes has historically been marked by the discovery, and subsequent solution, of myriad problems. Simulation has helped to understand and characterize some of the problems, thereby enabling superior device designs to be realized. A typical problem solved by early 2-D device simulations using the DD method involved the so-called "short-channel" effect, so that threshold-voltage values and transconductance characteristics could be accurately calculated when long-channel analytic formulae were no longer applicable (see, for example, Ref. [140]).

Another set of problems that was more difficult to characterize was due to so-called "hot-electron" effects. Such problems increasingly became more of a concern in the late 70s as device dimensions were scaled without proportionate reductions in power-supply voltage. Observed performance degradation (e.g. substantial gate and substrate currents, threshold-voltage shifts, and decreased transconductance) in MOSFET structures of the kind shown in Fig. 5(a) was hypothesized at the time to be due to high-energy carriers resulting from strong electric fields in the vicinity of the drain, so that carriers could be injected into, and trapped within, the gate insulator. Figure 5(b) shows an example of typical experimental data that were found to result from operating devices under high bias conditions. The curve labeled *no charge* represents source-to-drain current vs gate potential, prior to charge

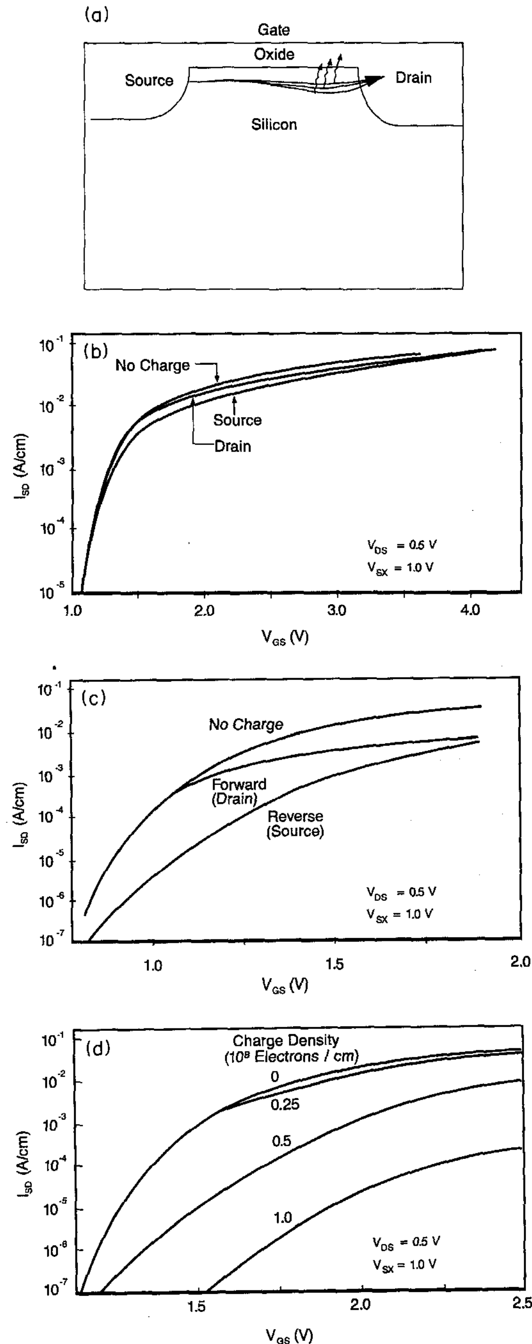


Fig. 5. Effects of charge trapped in the gate insulator. (a) Standard MOSFET. Hot electrons are injected into gate near the drain. (b) Experimental results. (c) Simulated results. (d) Results predicted by simulation for change in I_{SD} vs V_{GS}, due to change in fixed charge in oxide.

accumulating in the oxide, while the *drain* curve is the result obtained after the device was operated for a sufficiently long time for accumulation to occur. The *source* curve differs from the *drain* curve in that the bias between source and drain was reversed.

An accurate description of hot-electron behavior is now only becoming available with relatively recent advances made with Monte Carlo simulations (see Section 4.G). However, DD simulations carried out in the 70s were able to address certain aspects of hot-electron effects. For example, as described in Refs [180] and [181], the FIELDAY program was used to simulate the effect on device characteristics, as shown in Fig. 5(c), of introducing fixed charge into the oxide. As can be seen, good qualitative agreement was obtained between the experimental data of Fig. 5(b) and the simulation results of Fig. 5(c), thereby helping to affirm device designers' hypothesis of the source of the change in device characteristics. The utility with which a variety of process variations were able to be studied proved indispensable in the effort to optimize the performance of the structure. An example of the kind of rapid parametric investigations afforded by simulation is shown in Fig. 5(d), where the effects are illustrated of different magnitudes of fixed charge near the heavily doped source pocket.

4.C. Three-dimensional device modeling of a parasitic leakage mechanism in a 16-Mb DRAM cell

Here we describe an example where 3-D modeling capability was essential for accurately predicting device-related behavior. This example involves the simulation of leakage current from the storage node of a particular 16-Mb DRAM cell design that has been described elsewhere in the literature[182].

A top view of a portion of the memory array layout is drawn in Fig. 6. As can be seen, the overall cell structure is fairly complicated, and requires detailed modeling of the structure to predict actual device behavior. Between adjacent cells, oxide isolation regions serve to block possible conduction paths. The node of the storage capacitor for each cell is formed by first etching a deep trench into the silicon substrate. The trench is then lined with an insulator and filled with heavily doped polysilicon. The silicon substrate acts here as an effective capacitor plate that mirrors the charge on each polysilicon storage node, while a conducting bridge provides electrical continuity between the access transistor and the storage capacitor. The cell can be addressed by electrically selecting the bit line and word line that cross at the desired location in the memory array.

Even though the top portion of the perimeter of the storage trench is lined with a fairly thick layer of isolation oxide, which is referred to here as the collar oxide, there is some interaction between the storage node polysilicon and the adjacent transistor. Indeed, the state of the signal on the storage capacitor affects the characteristics of the transistor. One way of viewing the influence of the storage capacitor on the transistor's behavior is that the top portion of the storage node polysilicon acts as an effective gate on the parasitic FET that exists between the node and bit line diffusions.

To account for the effect of the presence of the trench capacitor on leakage current, the FIELDAY program was applied to the region enclosed by the

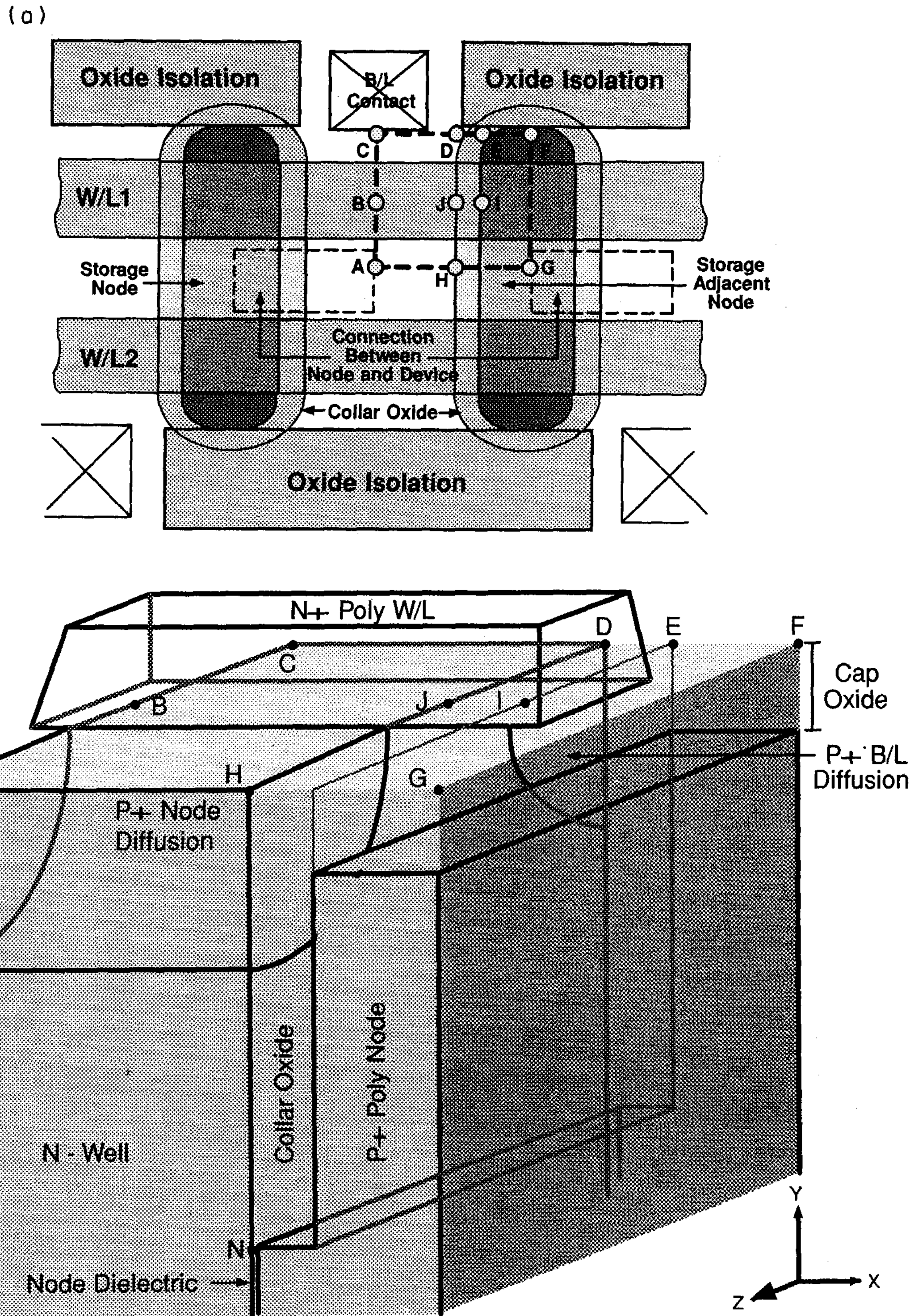


Fig. 6. (a) Top view showing modeled region. (b) Perspective view of modeled structure. In these figures, B/L and W/L refer to "bit line" and "word line", respectively.

dotted box in Fig. 6(a). A perspective drawing of this region is given in Fig. 6(b). The letters A, B, . . . , denote points in common between Figs 6(a) and (b). A finite element mesh was constructed so that a numerical solution could be obtained by the DD semiconductor set of equations. Part of this 3-D mesh is shown in Fig. 7(a), with the remainder of the mesh cut away for viewing purposes. Again, the letters A, B, . . . , in this figure correspond to the points shown in the previous figures.

Different mesh densities were used in order to obtain an estimate on numerical accuracy of

simulated results. The mesh sizes used here ranged from 6000 to 30,000 nodes. For a mesh with 20,000 nodes, the simulation time was about 10 CPU min per bias point on an IBM model 3090/600E computer running as a uniprocessor.

Figure 7(b) is a color plot of the doping concentration for the same mesh region shown in Fig. 7(a), while Fig. 7(c) illustrates the hole-minus-electron concentration under bias conditions that were found to yield the largest leakage current due to the presence of the adjacent storage node in a low-level state. The oxide region (yellow) closest to the observer has been cut away to reveal the underlying silicon. As can be seen, a conduction path forms along points H and J. This path continues over to point D, although this portion of the path is hidden from sight here because of the presence of the oxide region that has not been removed for viewing.

Our 3-D modeling of this 16-Mb cell clearly showed that the worst-case electrical biasing configuration for leakage occurred in the case of a stored high-level leaking down. With the node at a high level and the bit line diffusion at ground, the drain-to-source voltage, and hence the drain-induced barrier-lowering effect, is maximized. In this case, the parasitic leakage is gated by the polysilicon storage node connected to the adjacent device. Since the node of the device being studied is at a high level, the parasitic mechanism occurs on one sidewall only. On the other hand, the electrical biasing configuration for a low-stored level leaking up produces less leakage current despite parasitic conduction occurring on both sides of the device. In this situation the node itself and the node of the adjacent device are low, but the drain-to-source voltage is less than with the worst-case leakage configuration, since the storage nodes are written to approximately one threshold-voltage drop above zero.

Thus, 3-D modeling has shown that drain-induced barrier lowering has a significant effect on sidewall parasitic leakage. Moreover, the use of simulation has identified device parameters offering significant leverage in reducing leakage from this mechanism. Variations in temperature, trench sidewall fixed charge, doping profiles, trench isolation oxide thickness, trench polysilicon recess depth, channel length and n-well back bias have all been investigated. Figure 8 shows modeled device-leakage characteristics for three variations in physical parameters. This plot demonstrates that many options are available for meeting the leakage objective.

Good agreement with experimental data has also been obtained. Figure 9 compares the modeled and measured locus of points in the "adjacent trench voltage-channel length" plane that meet the leakage objective of 0.1 pA. As can be seen, agreement is good.

4.D. Three-dimensional device modeling of oxide-filled-trench isolation bounded MOSFETs at room and liquid nitrogen temperatures

This section describes a second simulation of a MOSFET that required a 3-D modeling capability. Moreover, the following simulation includes results that clearly illustrate the significantly different behavior of typical devices at room and low temperatures.

As device sizes decrease, isolating devices from one another becomes more difficult. For dimensions below one micron, a significant limiting factor for developing smaller devices has been the so-called "bird's beak" that is due to the encroachment of the local oxidation of silicon (LOCOS)[183] isolation regions into active device regions. This effect causes thicker oxide in the active device region, thereby hampering device performance. In addition, dopant atoms migrate under the LOCOS into the device areas, further affecting the device behavior. An alternative approach involves etching a trench and filling

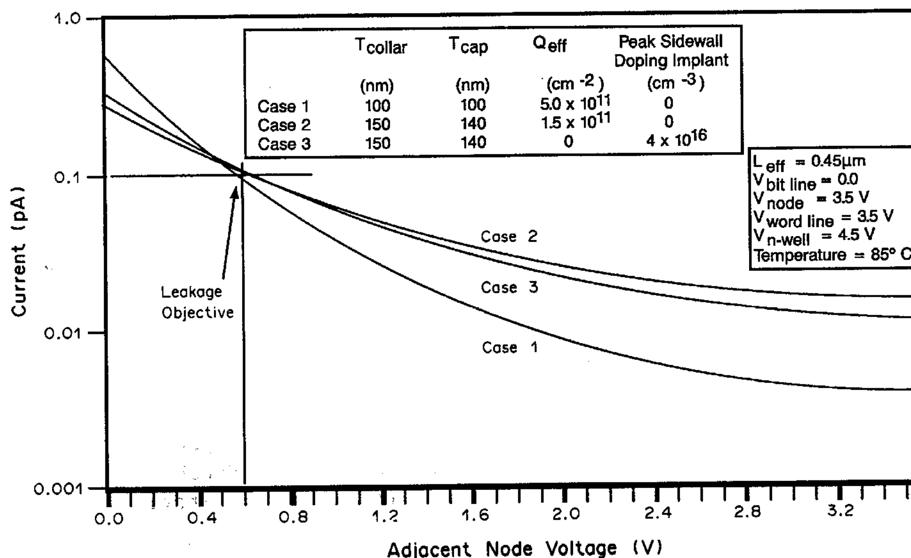


Fig. 8. Modeled leakage characteristics.

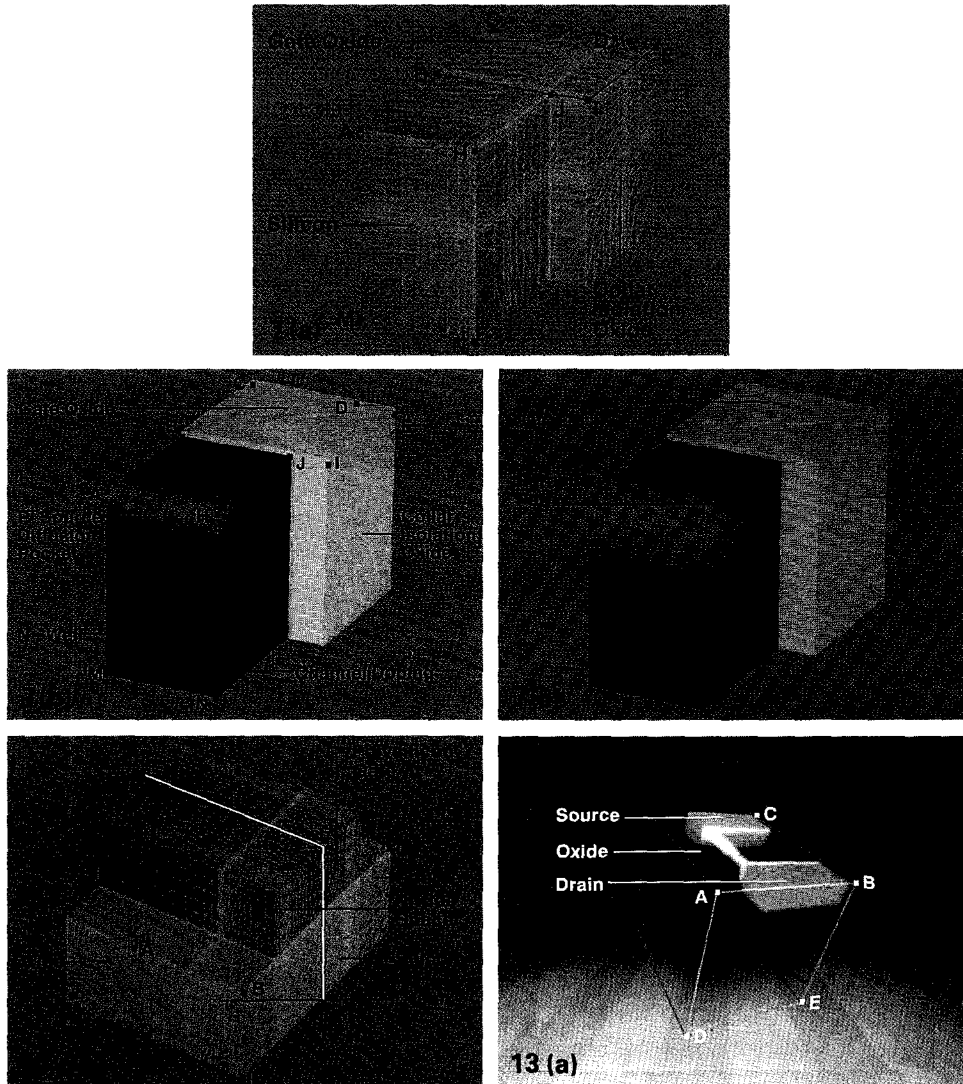


Fig. 7. (a) Finite-element mesh for model. Blue is silicon, yellow is oxide. Oxide has been removed from the portion closest to the observer for viewing underlying silicon. (b) Doping concentration. Green is p-type, red is n-type silicon. (c) Hole minus electron concentration showing formation of parasitic conduction path on trench sidewall.

Fig. 11. A typical 3-D mesh used for device modeling: gray—silicon; yellow—oxide; blue—polysilicon; red—nitride. The points A, B, . . . , E correspond to the same points in Fig. 13(a). The cross section indicated by the bright white line represents the left-hand side of the device sketched in Fig. 10(c).

Fig. 13. Electron density in a trench-isolated MOSFET, showing turn-on as the gate bias is increased from (a) to (c). Note that the device corner inverts at a lower gate bias than does the planar region. (See overleaf for Figs 13(b) and (c).)

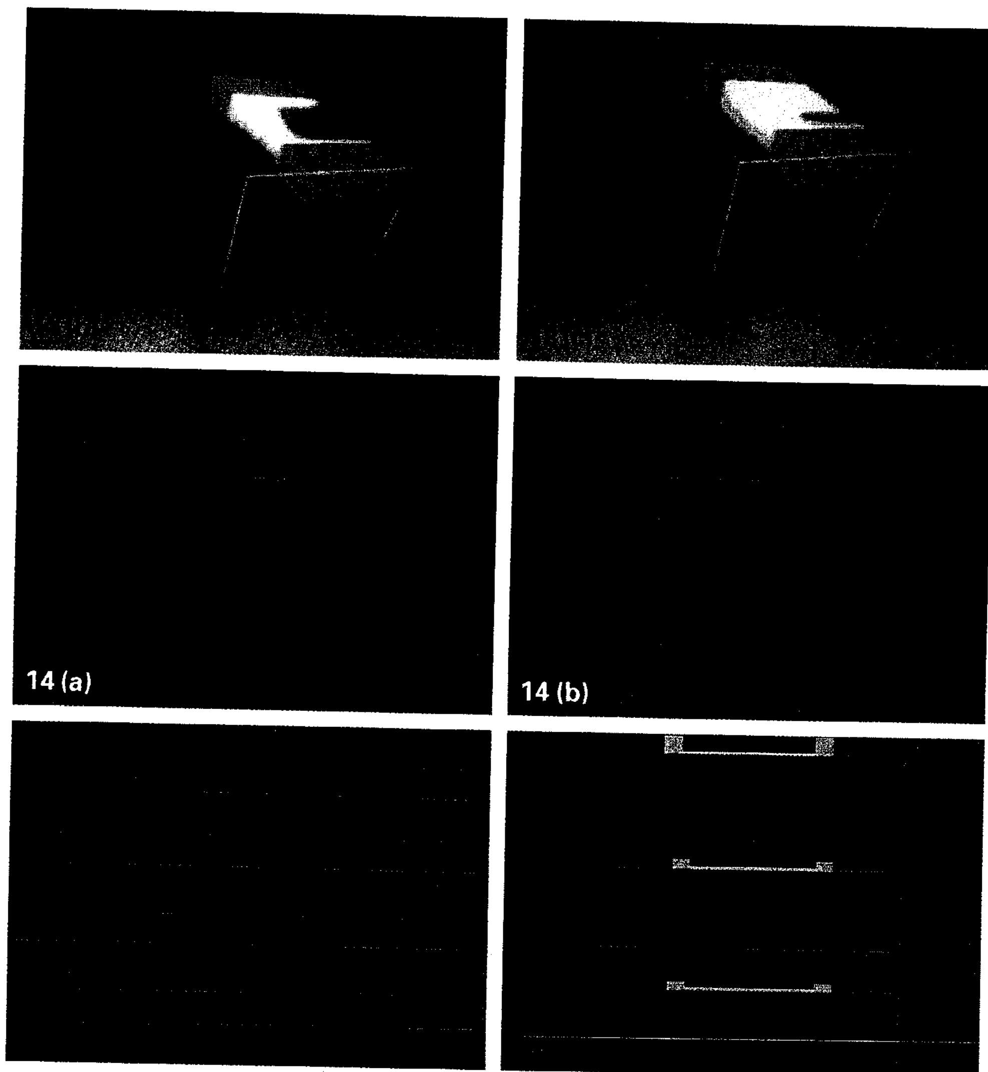


Fig. 14. Cross section of the electron density at mid-channel with the gate bias as the third axis. (a) $T = 300$ K. (b) $T = 77$ K.

Fig. 15. A sequence of contour plots of electron charge concentration that illustrate the evolution of the electron charge concentration after an α-particle strike. The right-hand side of each of the four plots represent the region where the *pn* junction lies. The α-particle enters from the right-hand side. The top cross section shows the initial situation before the α-particle hits. The next cross section down indicates the electron charge distribution 3 ps after the α-particle strike, while the next two cross sections show the situation at 90 and 180 ps, respectively.

Fig. 18. Simulation using FIELDAY to illustrate the effects of impact ionization and incomplete ionization for a 0.35-μm channel length *n*-MOSFET device with LDD structure. (Top) Doping concentration. Red is *n*-type, green is *p*-type silicon. (Middle) and (lower) Hole concentration at 77 K, for $V_{DS} = 2.0$ V, $V_{GS} = 4.0$ V, and $V_{SX} = 0.0$ V. (Middle) Simulation assuming complete ionization. (Lower) Simulation with incomplete ionization model.

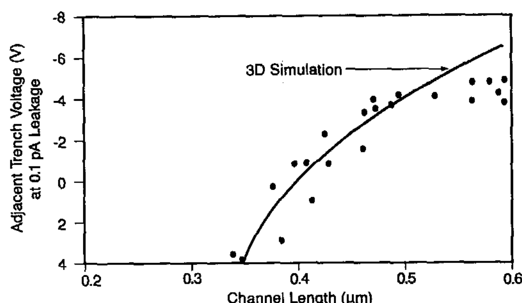


Fig. 9. Comparison of model data.

it with oxide to provide isolation[122]. This shallow trench isolation (STI) approach has been demonstrated to be a feasible technology for submicron isolation, as it eliminates the bird's beak. However, the simulation described here shows that device behavior changes due to this new approach to isolation.

For devices isolated from each other via a LOCOS structure, a polysilicon gate that passes across the device will curve upward and away from the device once it reaches the isolation oxide, as shown in Fig. 10(a). This figure represents the cross section of the region under the gate of a MOSFET, where the source and drain of the device are situated above and below Fig. 10(a). For this particular configuration of the polysilicon gate, the portion of the gate over the LOCOS region will have little effect upon the device behavior. Consequently, a simple 2-D simulation will give an adequate description of the behavior of the device, where the 2-D region is taken in the plane perpendicular to Fig. 10(a) and along the axis of the device, so that it includes the source and drain regions. (Of course, if a very narrow LOCOS structure is used, the encroachment of the bird's beak

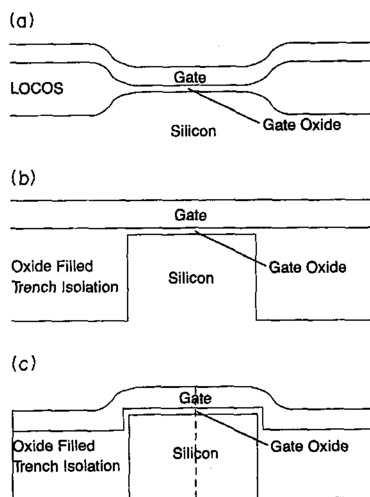


Fig. 10. Cross section through MOSFET structures at mid-channel, perpendicular to the channel axis: (a) LOCOS, isolated device; (b) trench-isolated device; (c) trench-isolated device with isolation oxide recess and gate wraparound. The left-hand side of the device in Fig. 10(c), which is enclosed by the dashed line, represents the cross section region indicated in Fig. 11 by the bright white line.

discussed above results in a non-uniform oxide thickness across the channel. This so-called "narrow-channel effect" would require 3-D modeling in order to adequately simulate the device behavior. In this case, the threshold voltage increases with decreasing channel width. However, such structures are undesirable and are generally not used, so we will not address this point any further here.)

In the case of an STI-bounded device, the gate runs flat across the isolation, as shown in Fig. 10(b). Consequently, the part of the gate over the STI region creates more of a fringing effect on the corner and sidewall of the silicon mesa[184–186] than in the LOCOS case in Fig. 10(a). This edge effect becomes even stronger if the isolation region is partly eroded away, so that the gate wraps around the corner of the device mesa as in Fig. 10(c)[186, 187]. Now, the regions of the channel near the edges of the device mesa will behave much differently than the region at the center of the mesa. Consequently, 3-D modeling is required to accurately simulate the device characteristics.

The objective of the device simulation to be described here was to study the current turn-on behavior of trench-isolated devices in which the corner effects described above are clearly important. To carry out this task, a 3-D mesh was constructed for use in the FIELDAY device simulator[29], as shown in Fig. 11. This mesh contained approximately 30,000 nodes and 60,000 elements.

Device characteristics were studied at room temperature (300 K) and liquid nitrogen temperature (77 K). The low-temperature simulation was carried out to exercise the low-temperature modeling capabilities recently added to FIELDAY[38], and to examine the feasibility of employing isolation-trench technology in a low-temperature environment. Liquid nitrogen temperature (77 K) provides a convenient and commonly studied operating regime[188–190]. In submicron devices, reduction of leakage and suppression of short-channel effects are among the potential benefits. If a liquid nitrogen temperature technology is to be developed, accurate device modeling must be available to study the behavior of both conventional and potentially novel structures.

For each of these two cases, a low drain bias was chosen (0.2 V) while the gate was swept from 0.0 to 1.5 V in 0.05 V increments. The room-temperature simulation required 15 CPU h on an IBM model 3090/600E computer running as a uniprocessor, while the low temperature simulation required 68 CPU h.

In these simulations, a major result of interest was the electron density in the silicon. The presence of electrons indicates that a conducting channel has formed, and that a significant current will begin to flow in the device. Simulation results for this device at 300 and 77 K are shown in Figs 12(a) and (b), respectively. A 2-D result has been included to demonstrate the behavior of the planar portion of the

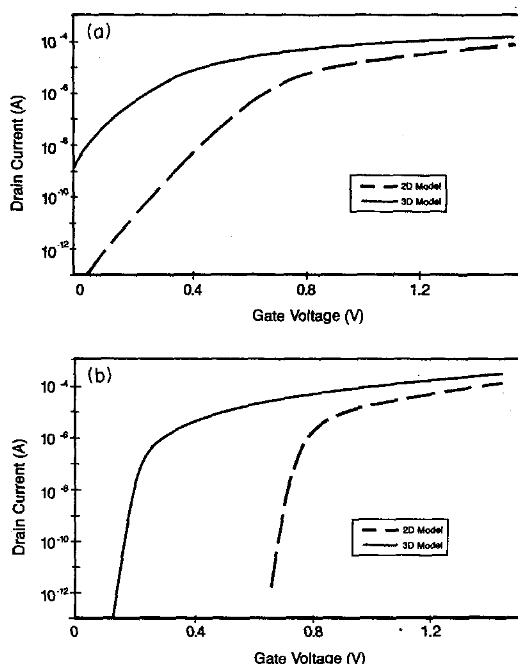


Fig. 12. FIELDAY simulation of a MOSFET, W/L = 0.5/0.7. (a) $T = 300$ K. (b) $T = 77$ K.

device. In the 300 K case, the lower turn-on present in the 3-D model clearly indicates the effect of the device corner. Due to the higher field at the corner, this region of the device has a lower threshold voltage than the center of the device. In the 77 K case, once again, due to the higher field, the device corner turns on well before the planar region. However, the corner conduction mechanism is clearly less sensitive to temperature than is the planar conduction mechanism. This interesting behavior can be further examined by visualization techniques. Further details on device aspects and experimental results may be found in Refs [185] and [187].

Color animation can produce images of electron density in the structure that convey an enormous amount of significant data in a single figure. The development of the inversion layer in a device may be found in Fig. 13. At low gate bias, only the device corner is inverted. As the gate bias increases, the extent of the inversion layer increases until, eventually, the planar portion of the device inverts. This behavior represents the physical basis of the double turn-on observed in these structures as described in Refs [185] and [187].

A cross section taken through the middle of the device perpendicular to the channel axis allows the development of the electron density to be examined as the gate bias is increased. An interesting approach is to take this x–y cross section and plot the gate bias on the third axis. Such a plot at 300 K is shown in Fig. 14(a). Here, the two conduction mechanisms are clearly displayed. At low gate bias, the device corner turns on. The extent of the electron density increases as the gate bias increases. At higher gate bias, the planar portion of the device turns on, as evidenced by the appearance of electrons across the entire device.

A similar plot at 77 K may be found in Fig. 14(b). Several important differences should be noted when compared with the 300 K case. When only the corner conduction mechanism is active, the electrons are more closely confined to the corner. As the gate bias increases, the extent of the electron density does not spread out laterally as much as in the 300 K case. Finally, the separation in the gate bias between the two mechanisms due to the lower temperature is clearly evident.

4.E. Single-event upset

Cosmic rays, as well as radioactive sources (e.g. the low-level radioactive source of lead that is used in the packaging of semiconductor devices), give rise to energetic ions that can travel through a semiconductor device. If one of these ions traverses a semiconductor *pn* junction, severe electric field distortions will occur that can enable a significant number of charge carriers to subsequently flow across the *pn* junction. If the device is used for storing charge, the stored-charge state can change, thereby resulting in a "soft error"[191–196].

Accurate characterization of this *single-event upset* phenomena has become increasingly important as device designers shrink cell dimensions. However, the inherent complexity of this phenomenon has impeded experimental progress toward fully understanding needed device design modifications to control soft errors. Numerical simulation, on the other hand, offers a unique opportunity to examine the physical behavior associated with *single-event upset* for a wide variety of processes and device operating conditions.

A fully 3-D silicon diode structure, for which the substrate resistivity was 1 Ω-cm and whose fabrication details may be found in Ref. [195], was modeled using the FIELDAY program. A $2 \times 2 \times 30\ \mu$m portion of the overall device, with a $1.4\ \mu$m junction depth, was discretized, reverse-biased to 25 V, and then "struck," normal to the surface of the n-contact, with a single 5-MeV α-particle.

Figure 15 illustrates, for a certain cross section taken through the center of the device, the time evolution of electron concentration after the impact of the α-particle. The top cross section in Fig. 15 represents a snapshot in time immediately preceding particle impact, while the second cross section from the top shows the situation 3 ps later. Here, the α-particle entered the semiconductor material on the right-hand side, which is where the *pn* junction lies. As can be faintly seen at the 3 ps cross section, a wake of excess charge has been formed due to the particle's penetration. The severe field distortion in this early stage of impact has started to collapse the depletion layer on the right-hand side of the figure. The charge configuration of the depletion layer continues to change with time, eventually resulting in a rapid transport of excess carriers to the surface contact:

This phenomenon has been termed "funneling" by Hsieh *et al.*[192], apparently inspired by the shape of the distorted field. The "funnel" is fully developed by 90 ps, as represented by the third cross section down from the top in Fig. 15. At this point in time, the contact current reaches its maximum value. The final cross section shown represents the situation at 180 ps. This contour plot marks the end of what is commonly referred to as (drift-dominated) *prompt* charge collection associated with the funneling phenomenon. As can be faintly seen on the right-hand side of this last plot, just to the left of the *pn* junction, the funnel has collapsed, so that the *pn* junction has essentially re-established its original charge characteristics. Diffusion then again becomes the dominant transport mechanism for carriers flowing across the *pn* junction.

Figure 16 shows simulation results for cumulative charge collection as a function of time. The physical origin of the characteristics of this curve can be clearly understood by examining Fig. 15. The initial delay of about 30 ps, during which there is no significant collected charge in Fig. 16, is indicative of the time delay required in the formation of the *funnel* structure at the *pn* junction. Transport of carriers to the contact is greatly enhanced once the funnel is fully established and then rapidly proceeds through the 180 ps point. The marked change in slope in Fig. 16 at 180 ps, ending the so-called prompt charge-collection phase, results from the funnel collapse illustrated in the bottom cross section in Fig. 15.

Wagner *et al.*[195,196] have recently published current and charge-collection data on structures similar to the problem modeled here. Figure 17 compares simulated vs measured current transients, and Table 1 shows a similar comparison for total collected charge, as well as some other current vs time characteristics. As can be seen, the agreement between simulation and measurement is quite good.

4.F. Impact ionization and low temperature simulation

The simulation example described here involves the use of FIELDAY on a 0.35-μm channel length MOSFET, with a lightly-doped source and drain structure. The simulation is carried out at $T = 77$ K.

Within the different regions of a typical device structure, carrier concentrations can vary over many

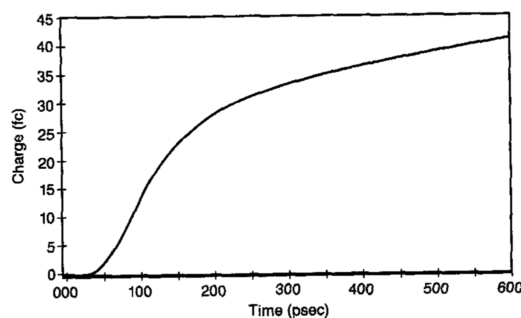


Fig. 16. Cumulative charge collection.

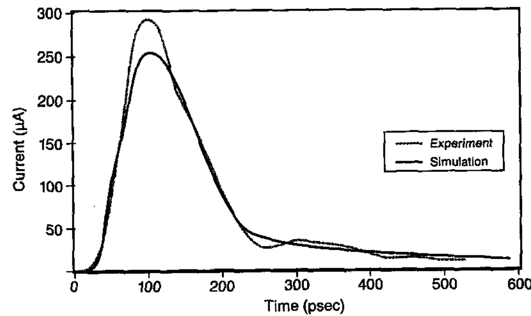


Fig. 17. Comparison of measured vs simulated current transients.

more orders of magnitude at 77 K than at room temperature. Consequently, a simulation at 77 K usually exhibits poorer convergence properties than the same device problem at, say, 300 K. Hence, 77 K simulations demand more attention on improved numerical methods, and more care in obtaining satisfactory numerical solutions (e.g. smaller voltage increments in tracing out I–V characteristics, finer meshes in critical device regions, etc.), than their room-temperature counterparts. The simulation example we describe here exhibits what we believe to be the implementation of a number of a fairly robust numerical methods for treating this device problem.

In addition, bias conditions were chosen sufficiently large that impact ionization could not be treated as a mere perturbation on the initial solution. Instead, this carrier generation mechanism had to be treated self-consistently with the remainder of the terms in the electron and hole current continuity equations. The problem converged quite well, which we attribute to an improved numerical method some of us have recently devised for implementing the impact-ionization model.

Figure 18(top) depicts the doping profile used for the device. Due to impact ionization, minority and majority carriers are generated near the drain end of the device. Figures 18(middle) and (lower) show the hole concentration under two different conditions; we will explain these two conditions shortly. As can be seen in these figures, the high-concentration region for the generated holes starts near the drain, which is on the right-hand side of the figure. In Fig. 18(middle), most of these minority carrier holes are then swept by the electric field to the substrate contact, which is represented here by the entire bottom surface of the device structure. The main bulk of the substrate is formed by these generated hole carriers. In Fig. 18(lower), not all of the holes reach the

Table 1. Comparison of measured vs simulated results

Physical quantity	Simulated	Measured
T_p (ps)	88	88
$Q_{collected}$ (fC)	41	39
I_{max} (μA)	265	293

substrate contact; some are swept by the electric field all the way back to the source, thereby resulting in additional impact ionization events in the high electric field region at the source end of the device.

The difference between Figs 18(middle) and (lower) results from using two different physical models for the fraction of ionized dopant atoms. In Fig. 18(middle), all the dopant atoms are assumed to be ionized, or, more precisely, $N_D^+ \simeq N_D$ and $N_A^- \simeq N_A$. The phenomenological incomplete ionization model described in Ref. [38] is used for Fig. 18(lower). As can be seen, a substantial difference exists between the two conditions, thereby indicating the significant physical influence on the behavior of devices at low temperature from the incompletely ionized regions of dopant atoms within the device. Indeed, further improvements are highly desirable in characterizing this phenomenon of incomplete ionization for the transition region between mid- to high-doping concentrations. However, we should note here that our simulation does not account for the effect of a strong electric field in ionizing the dopant atoms[197]. Taking this effect into account will undoubtedly influence our final result so as to make the difference between the two cases examined here less dramatic than what is illustrated.

4.G. Monte Carlo simulation

The last simulation that will be described here on device behavior was carried out using the program DAMOCLES. Section 4.A contained a brief description of this program, which effectively solves the Boltzmann transport equation (see Refs [43–46] for more details).

Here, we show specific plots that can be displayed using DAMOCLES. These plots aid in understanding the dynamical behavior of the carrier trajectories, and in characterizing the statistical distribution of various physical properties associated with the carriers. Our point here, as it has been with previous simulation examples, is to indicate pictorially the type of task that one can accomplish with an advanced simulation program of this caliber.

Figure 19 shows one snapshot in time of a simulation of an n-channel MOSFET with an effective channel length of about 0.35 μm. The bias conditions chosen for this simulation were the following potential differences: 5.0 V between the drain and source, 2.5 V between the gate and source, and 0.0 V between the source and substrate contacts. Each dot in the figures represent a charged particle, which in turn is representative of an electron, or a set of electrons. This particular simulation tracked about 8000 charged particles simultaneously in space and time. The red and blue colors indicate the higher and lower energetic particles, respectively. At this particular moment in time, the ensemble of 8000 charged particles range in energy from about 0.0 to 2.676 eV. This energy range is indicated in the color energy scale on the RHS of Fig. 19(a).

The black areas in the source and drain were cut out of the simulation, since these regions contain an enormous number of electrons that are essentially in thermal equilibrium with the lattice at the temperature of 300 K assumed in the simulation. Considerable CPU time is saved by not tracking particles once they have entered these regions. Instead, boundary conditions appropriate for an ohmic contact region are assumed here (see Ref. [43]).

Moreover, in order to obtain accurate statistical information about the relatively few electrons that constitute the higher energy electrons within the device, a statistical enhancement factor of 10 was assumed for the channel region, and for a region around the heavily doped drain pocket. (See Refs [43] and [46] for more information on this method.)

For the particular bias conditions assumed here, the channel is pinched off at about 0.10 μm, or somewhat less, from the drain, as can be seen in Fig. 19(a). Thus, these bias conditions set the device somewhat past the onset of pinch-off conditions. Hence, few particles are present at the drain end of the channel. However, the particles that are in this region are quite energetic, as indicated by their red color. By averaging the results of the particle distribution over many time steps, good statistical results can be obtained.

Figure 19(b) shows the path taken by one of the particles. The scattering processes are labeled along the path. Figures 19(c) and (d) show blowups of the path of the particle near the source and drain ends of the device.

The result of averaging over many snapshots in time is shown in Figs 20(a) and (b). Here, contour plots are displayed for the average electron energy and for the electrostatic potential, respectively. As can be seen in Fig. 20(a), the higher energy electrons are found around the edge of the drain pocket. The longer the simulation is allowed to run, the greater the accuracy will be for characterizing the statistical distribution of physical properties for this ensemble of particles. Thus, Fig. 20(a) should evolve into a perfectly smooth structure after longer simulation times. However, the present simulation was already quite time-consuming: namely, about 100 CPU h on an IBM model 3090/600J computer with vector facilities, running in a uniprocessor mode. This lengthy simulation time indicates the weakness of such particle tracking programs; at present, however, in order to incorporate the detailed physics that exists within this program, such simulation times are inevitable.

By taking cross sections of one of the plots in Fig. 20, a line plot can be generated, as shown in Fig. 21(a). Both curves shown here represent the average energy of the electrons as a function of position along a line right below the oxide–silicon interface. The ragged curve describes the average energy of the ensemble of particles for one particular snapshot in time. Averaging over many time steps results in the smoother curve.

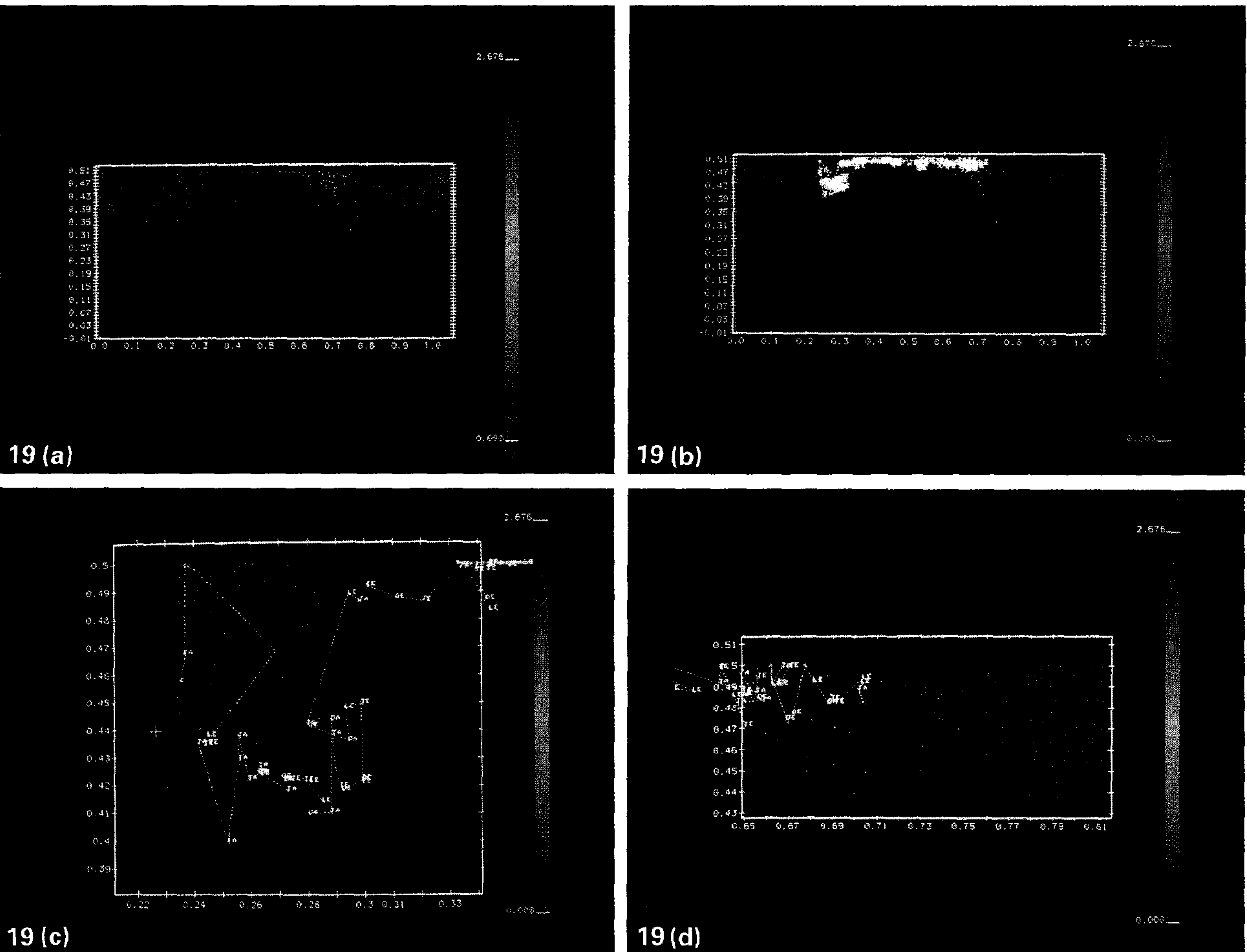


Fig. 19. A Monte Carlo simulation using DAMOCLES of a 0.35-μm channel length n-MOSFET. The oxide thickness is 12 nm. Gaussian profiles were assumed for the source and drain regions with a peak concentration of $2 \times 10^{20}\,\mathrm{cm}^{-3}$ n type. The substrate doping was $3 \times 10^{16}\,\mathrm{cm}^{-3}$ p type, and a *surface channel tailor implant* had a peak concentration of $1.83 \times 10^{17}\,\mathrm{cm}^{-3}$ p type. The bias conditions assumed here are $V_{\text{drain–source}} = 5.0$ V, $V_{\text{gate–source}} = 2.5$ V, and $V_{\text{source–sub}} = 0.0$ V. Approximately 8000 electrons are tracked. (a) One snapshot in time. (b) The past trajectory for one of the electrons in (a). (c) Blowup of (b) near the source end. Note that scattering events are labeled. (d) Blowup of (b) near the drain end.

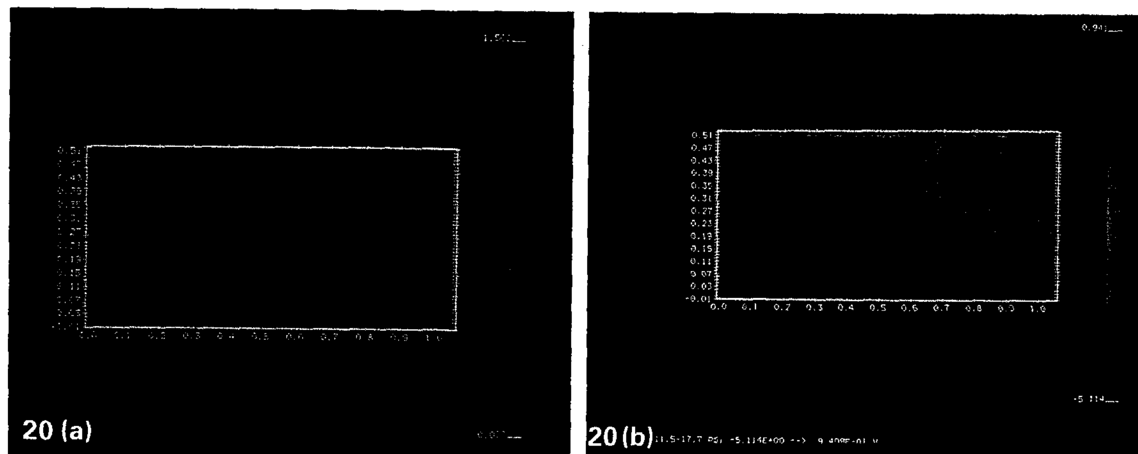


Fig. 20. Contour plots of calculated (a) average electron energy, and (b) the negative of the electrostatic potential. The average energy in (a) ranges from 0 eV (blue) to 1.5 eV (red), while the electrostatic potential in (b) ranges from +5.1 V (blue) to −0.94 V (red).

Figures 21(b) and (c) illustrate a typical example of detailed information that can be extracted from a Monte Carlo simulation. Figure 21(b) shows two electron-energy distributions near the drain of a 0.25 μm-long channel Si MOSFET at 300 K (see Ref. [198] for details on the device structure). A high drain bias is assumed in the simulation ($V_{DS} = 4.0$ V, $V_{GS} = 2.5$ V, and $V_{SX} = 0.0$ V) to investigate hot electron effects. Both distributions in Fig. 21(b) have reasonably accurate statistics down to the 10^{-8} range, which is the accuracy required to estimate gate currents. The difference in the two distributions lies in the absence of the long- and short-range inter-carrier Coulomb interaction in the simulation run represented by the dashed line. Sophisticated physical effects, like the long-range electron–plasmon interaction and the short-range screened electron–electron collisions, provide means to decrease the average electron energy, while actually increasing the number of carriers at high energy.

The potential drop from the source to the point near the drain where the electron-energy distribution was obtained is about 3.5 V. As can be seen in Fig. 21(b), if these interactions are not taken into account, few electrons are predicted to occur with energies greater than 3.5 V; however, with the interactions included in the simulation, a distinctly noticeable tail occurs beyond the 3.5 V point in the energy distribution. Since the high-energy tail of the electron distribution is the critical part to consider when investigating quantities such as gate current, these physical effects must be taken into account to obtain accurate results[199].

Figure 21(c) further emphasizes this point. Here, in proceeding from source to drain along the channel of the same device, we see that electrons "climb" the band-structure of Si. Indeed, a significant fraction of the electrons in the distribution leave the minimum along the Δ-symmetry line to enter the second conduction band (X_1). Some electrons in the ensemble even populate higher lying valleys (L_1, L_3, Γ_{15}), which must be correctly included in the model to analyze high-energy phenomena. This level of detail is necessary to understand hot-carrier-related reliability issues and stresses the need for more physically accurate modeling tools in the submicron regime.

5. DATA HANDLING, VISUALIZATION, AND THE REPRESENTATION OF COMPLEX GEOMETRICAL DEVICE STRUCTURES

State-of-the-art device and process simulation systems typically operate on large volumes of data. We expect that the simulation of the next generation of VLSI devices will require even more data due to more complex geometries and more stringent tolerances. Indeed, 2-D simulations will no longer be a reasonable approximation for typical situations; 3-D simulations will become essential. As a result, the VLSI modeling discipline is experiencing a data explosion. This explosion has necessitated the use of larger memory allocations and more efficient data handling schemes in the solution phase, as well as in the output data analysis phase.

In what follows, we cover three topics: (1) our approach to handling the data associated with process and device simulation, (2) scientific visualization of data, and (3) the representation of complex geometrical structures associated with semiconductor devices. These topics are quite different from each other, yet they are still related to the same problem of efficiently organizing the physical description of complex 2-D and 3-D structures.

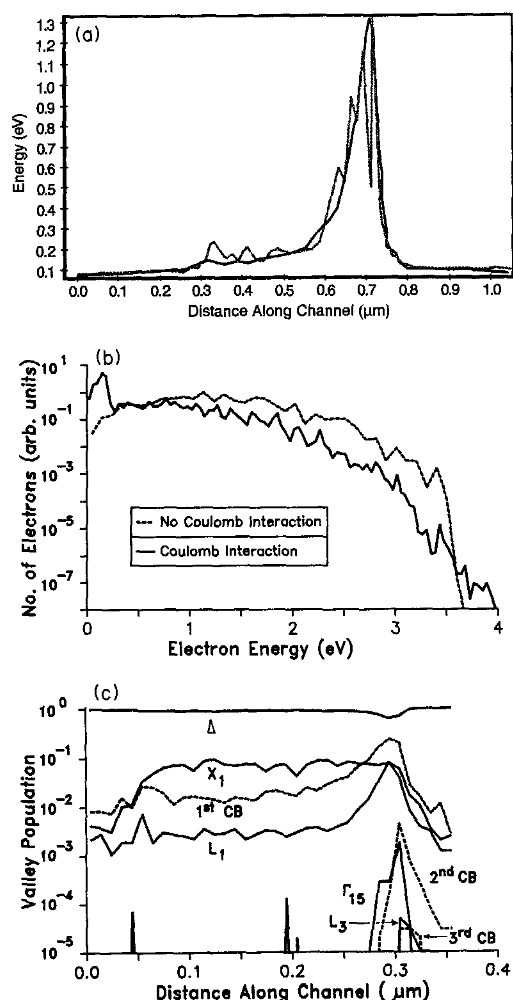


Fig. 21. (a) Average energy of electrons taken along a cut right under the gate oxide. The ragged curve represents one snapshot in time. The smooth curve was obtained after averaging over many time steps. (b) Hot-electron energy distribution in a $0.25\,\mu$m-long channel device. (c) Population of the various valleys of Si along the channel of the same $0.25\,\mu$m device. For a full understanding of this figure, a band structure diagram should be consulted, such as on p. 75 in Ref. [127]. The top curve labeled by Δ refers to the minimum in the first conduction band that is slightly displaced from the symmetry point X, along the symmetry line Δ. The curve labeled by X_1 refers to the electrons in the minimum in the second conduction band at the symmetry point X_1. The dotted lines labeled by 1st CB, 2nd CB, and 3rd CB refer to electrons outside the valley regions, and lying in the remainder of the 1st, 2nd, and 3rd conduction bands (CB), respectively. The remaining curves refer to electrons populating (a) the singly degenerate minimum at L_1 (about 1 eV above the bottom of the 1st CB), (b) the triply degenerate minima at Γ_{15} (somewhat higher than 2 eV above the bottom of the 1st CB), and the doubly degenerate minima at L_3 (somewhat higher than 3 eV above the bottom of the 1st CB).

5.A. The data base

The data break down into roughly three categories: (1) the device geometry represented mathematically by bounded regions with assigned material and/or electrical properties, (2) the physical quantities of interest such as dopant concentrations, the associated carrier concentrations and currents, electrical potential, etc. which are interpolated onto this computational device geometry, and (3) the input physical parameters such as impurity diffusion coefficients, the parameters in carrier-mobility models, temperature, and required boundary conditions.

The mathematical representation of the geometry can be achieved in several ways. In the scheme generally used for semiconductor device and process simulation, each of the material regions of the device consists of a discretization of the device into nodes and elements, or "mesh", that relate the governing equations to the computational device geometry. (For FEDSS and FIELDAY, we use the finite element or "control volume" method to effect the mapping.) Roughly speaking, the finer the mesh, the more accurately the numerical algorithms will converge to the exact solution of the equations describing the phenomena. For a Monte Carlo particle-tracking scheme, only the boundaries of regions in a device need to be defined, although the fields still need to be obtained via a method such as the spatial discretization scheme described above. Also for a Monte Carlo program, a statistically significant sample of the individual "diffusing" species must be tracked, which typically requires a large amount of storage. In the case of process simulation these species are dopant atoms, while for device simulation, these species are electron and hole charge carriers.

An important point to note is that the bookkeeping problem presented by process and device simulation literally multiplies in going to higher spatial dimensions (even ignoring the time domain for now). As a crude figure of merit for the finite element method, whereas five years ago a device simulation in 2-D would strain computational limits at 5000 nodes, we now encounter problems requiring as many as 200,000 to 300,000 nodes for practical 3-D design problems.

Thus the volume of data for a single analysis can be large; for multiple analyses, the problems of data volume can be overwhelming. The practical problem of manipulating these large amounts of data became more evident over time for us, particularly as we were challenged with more aggressive device-design problems. An efficient means needed to be found for organizing data so that most, if not all, of the data pertinent to a given simulation run might be reviewed, reused, or "restarted" at a later time. This capability is of particular importance when new aspects of a device or cell design might become apparent, or when an engineering change is anticipated.

In order to keep analysts from being overwhelmed by data, we have devised a comprehensive, flexible data base. This system provides for data storage and retrieval, and for pre- and post-processing of the data

in an integrated, interactive environment. The data is organized into five categories or "files":

- A general "variable" file for parameters related to physical constants and coefficients required by the models,
- A "doping" file for impurity atom distribution (either a result of process simulation or created from superposition of analytic functions),
- A "geometry" file that contains the nodes and elements of the spatial discretization,
- A "results" file that contains the discretized (nodal) values of such quantities as electronic potential, electron and hole concentrations, dopant atom concentrations, etc.,
- A "dictionary" file that provides a single indicator to any given simulation "run" and a cross-reference to data for that run contained in the other four data files mentioned above.

This data base is defined before running a simulation (or series of simulations). Similarly, the mesh representing the device structure at the initial stage of simulation can be created either interactively at a work station using a mouse/joystick, or in batch mode via a file specifying bounding points and lines for the structure. In many cases the input structure for a device simulation is passed through the data base directly to FIELDAY from FEDSS. As a simulation progresses, data is stored and cataloged according to "run" and "record" numbers that keep track of the chronology and storage location of the simulation, respectively. We also have a data-base monitoring facility that permits the user to view the indexing and time/date stamps of the data base at any point. In fact, one can monitor the data base (or even view results) as a simulation progresses.

Several years of experience have shown our data base to provide an effective means of managing simulation data. However, we note here that other groups have since proposed alternative methods for handling simulation data. Most notable has been Duvall's proposal for the Profile Interchange Format (PIF)[200], a hierarchical data format for storing simulation data. This format is more general than our own, but also more complex. However, a significant point to the PIF proposal is to establish a standard protocol for exchanging data among simulation tools developed at different sites. Consequently, important characterisitics of PIF include its flexibility and generality, which are likely to aid its acceptance when PIF parsers become more widely available.

To view 1-D or 2-D results from our data base, we have developed an interactive postprocessor that displays at the user's option: linear, log-linear, contour, log-contour, surface (in perspective), or vector plots of the full or partial ("windowed") results. For 3-D data, a slice in a plane of the solution domain may be chosen and results viewed on this cut plane. However, these techniques do not fully utilize the value of the data, nor the power of today's graphics display techniques. New methods of data analysis have therefore been invoked to supplement the more traditional ones. Indeed, such a diverse set of new graphical techniques has become available, and continues to be developed at such a rapid pace, that a new field has been spawned known as scientific visualization. We next briefly describe some of the main features of this area of development, along with some of our own work in this field.

5.B. Recent directions in scientific visualization for semiconductor technology simulation

5.B.1. General discussion. Scientific visualization software[201] for semiconductor technology simulation must meet several requirements. First, it must provide a simple human interface with a seamless connection to the simulation data base. Second, it must be highly interactive. The type of visualization required for analysis varies with the type of structure being studied and the sorts of phenomena under analysis. The user must therefore be able to quickly manipulate the visualization package in order to study the phenomena of interest. This requirement also impacts the computer hardware being used for visualization. The graphics must be fast enough to allow the user to try many different visualization methods in a reasonable amount of time. Third, the visualization software must provide a set of flexible graphical and post-processing operations that can be used for a wide variety of studies. Finally, the visualization software should be built on standard hardware and software platforms to the greatest extent possible, so as to minimize cost and improve accessibility.

There is probably no single best hardware platform for scientific visualization. The use of a local area network (LAN) with work stations and mainframes that run Unix (many versions of Unix exist, such as AIX) and X windows is therefore an attractive option, since it permits a variety of machines to be chosen. Most graphics can be done locally on a workstation or over the network using X. A few high-end graphics supercomputers can be made available for more demanding problems.

The use of Unix, X windows, and PHIGS graphics provides an excellent software platform for scientific visualization. Unix and X are fast becoming *de facto* standards for engineering and scientific computing, and provide a number of toolkits for developing good, human-interactive interfaces with network communication. In some cases nonstandard graphics libraries are required to make full use of specialized graphics hardware.

A number of scientific visualization techniques are being used at IBM for semiconductor technology simulation. These include:

- Animation: A set of 2-D or 3-D images may be displayed in rapid sequence so data pattern changes become more evident. The sequence may involve the time evolution of another

variable, such as gate bias or oxidation temperature. Animation of image rotations and of changes in perspective are highly effective in aiding detection and interpretation of patterns or pattern formation in 3-D data.

- Multiwindow displays: Simultaneous display of different steps in a sequence can help highlight changes in data. Multiwindow displays can also be used to compare different sets of data, whether they are different variables from a given simulation or the same variables from more than one simulation.
- Stereo pairs: 3-D data are often hard to interpret when printed on paper. Stereo image pairs aid perception of 3-D data in a volume when viewed with the proper equipment.
- Opacity/translucency: Judicious use of opacity and translucency can permit the viewer to "see through" a volume and perceive internal structures or obscured data.
- Postprocessing: The ability to view general mathematical functions of the data generated by simulation can be used to pinpoint features that would otherwise be invisible in the raw data. In device simulation, for example, it is possible to use a classification function to display depletion regions as distinct from neutral and inverted regions. This feature provides more contrast than the simple use of contour plots.
- Coordinate selection: Often some features of a 3-D simulation become more evident when displayed via an alternative set of coordinates. A 2-D plane through a device structure, for example, may be displayed with the coordinate in the third dimension chosen to represent potential. The resulting 3-D image can then be used to show in a vivid manner how potential is distributed throughout the cross section of the device. (See for example, Fig. 14 in Section 4.D.) Likewise, if temperature or other quantities are displayed, interesting perspectives are obtained on the device characteristics.

The set of techniques outlined above should not be considered complete since new visualization methods are continually being developed. However, these techniques do serve as a representative set of methods that we have found useful in our own work.

5.B.2. Sample results. Figures 22(a)–(c) and 13(a)–(c) illustrate how visualization techniques may be used in analyzing semiconductor technology simulation data. Figures 22(a) and 13 show how 3-D perspective with color can be used to highlight internal device physics. Figure 22(a) shows the potentials throughout a trench-isolated device when the gate is at a low voltage and the drain is at a high voltage. Electron concentrations in the same device are displayed in Fig. 13. Here, concentrations at three gate voltages illustrate how the channel forms. These images represent only three steps in a sequence that are scanned quickly by the graphics work station. The animation that results from scanning aids perception of the qualitative behavior of channel formation. Figure 22(b) illustrates the use of translucency and opacity to select features of interest. An intermediate range of potentials is shown as opaque red; higher and lower potentials are shown with translucent colors. In this image the three-dimensional shape of the intermediate potential shell can easily be perceived. Figure 22(c) shows how different types of data may be combined to pick out features of interest. In this image, high concentrations of electrons in the channel are represented as an opaque red. The translucent blue defines the extent of the device, which has been cut away at the drain end.

5.C. Representing complex geometrical device structures

We now turn to the problem of displaying and interactively modifying the complicated structural forms associated with device geometries. For example, the problem described in Section 4.C (see Fig. 6) illustrates a simplification of a cell structure; clearly, the peripheral structure surrounding an actual semiconductor device can become extremely complicated. To perform simulations for variations in such structures, adequate representations must be achieved in a flexible manner.

Our present scheme for defining device geometry in two dimensions consists of interactively creating stick-figures that define the borders of different material regions for the structure of interest. These figures are then triangulated to form the mesh consisting of the nodes and elements. To accomplish this task, we employ a suite of semi- and fully-automatic mesh generators. This scheme can be used for some limited 3-D structures by extruding the 2-D mesh into the third dimension, and then doing some crude volume modifications. However, for 3-D objects of more arbitrary shape, this method is unworkable; instead, a solid model of the object is required. Solid models provide a precise, mathematical description of an object that can be utilized by an automatic mesh generator by other geometry manipulation tools.

Creation of a solid model can be accomplished using constructive solid geometry. Many commercially available programs exist with this capability. However, these programs are oriented towards creation of relatively large, macroscopic engineering structures, and so are difficult to use for creating small integrated circuit device structures. Fortunately, we have access to a solid modeler that makes the creation of IC-specific solid models relatively painless. This program, OYSTER[40–42], uses a process-like description along with the mask-level information to create solid models of a device structure. For example, the command "grow oxide 500 Å"

causes the OYSTER program to pick the proper mask level and to extrude the profiles for that level by 500 Å over whatever geometry already exists. Other process steps such as deposit, etch, wash, apply, develop, and lift-off evoke geometric actions by OYSTER to create a solid model of the structure. To create solid models that are even more precise (i.e. "less square"), edge profiles of shapes can be defined and used by the program at the option of the user. Thus, with the use of OYSTER, a fairly realistic and complete geometry of a structure can be created.

Once a solid model has been created, it can be saved and used for visualization, or by other software that utilizes the geometric information in the solid model. For example, we use the solid models of the resulting device structure as input to our automatic mesh generator so that an analyst does not need to devote inordinate energy to this difficult chore. Both of these capabilities are under development and have provided exciting results to date. An example of an OYSTER model of part of a simple MOSFET is shown in Fig. 23(a). (Note that a complete discussion of this device and the process steps used to create it are given in Ref. [40].) The top layers of the device structure have been removed to aid viewing. Here, the substrate is shown in pink, gate oxide in yellow, isolation oxide in red, and polysilicon in green. Also seen in blue are portions of the source and drain metal contacts. Note that with OYSTER, individual structures can be displayed, as well as any desired collection of materials and structures. These individual structures can be viewed from different vantage points and shown as separate entities from the original structure (see Fig. 23b). In addition, cutting planes can be passed through the structure and saved for later use.

We have used OYSTER-created solid models as input to our 3-D automatic mesh generator. Currently, the mesh generator creates the minimal mesh necessary to capture the geometry of interest. For the MOSFET example, we show some of the meshes created with the software. In Fig. 23(b), we see the tetrahedral mesh for the polysilicon finger in green. Also shown is the isolation oxide in light blue, and three metal components in dark blue (the three metal components appear to be nearly touching one another in Fig. 21b). These metal pieces are used to contact the terminals of the device; they are not shown in Fig. 21(a) in order to provide an unobstructed view of the rest of the device. (See Ref. [40] for more details on the device structure.)

As is evident from the figures in Fig. 21(b), the mesh generator does a fine job of capturing the geometry. These meshes are adequate for calculations of fields and capacitances, but are not sufficiently fine for device simulation. We are currently developing automated methods to refine these meshes to provide more accurate meshes for device simulation.

6. CONCLUDING REMARKS

A number of pictorial examples were presented here that illustrate how semiconductor process and device computer simulation aids in the design and analysis of the semiconductor devices used in state-of-the-art integrated circuits. A brief background on past and present simulation development activities was also given, along with a number of relevant references that the reader may wish to refer to for more detailed information.

If used properly, simulation can substantially help in cutting development costs. Indeed, some researchers are of the opinion that cost savings can be as large as 40% of the development cost[202]. After all, parametric variations can rapidly be made by simulation that would be enormously expensive if attempted by experimentation alone. For example, the BiCMOS example in Section 3.E contained over 100 distinct process steps, with each of these steps containing numerous possible variations in process controls. Optimizing these parameters by experimentation alone would be quite impractical.

However, we note here that the cost savings depends considerably on the skill of the person using a simulation program, since care and imagination are still required in: extracting relevant regions in which to simulate complicated device structures (see, for example, the 3-D simulation in Section 4.C), using pertinent material properties (parameters and models), varying appropriate process and device operating conditions to obtain optimum device performance, and ensuring that accurate numerical answers are produced. In addition, a basic understanding of the physics incorporated in the simulation program is helpful to the device designer in understanding the limitations of the computer program.

Finally, we note here the fascinating relationship that exists between the simulation of semiconductor devices and improvements in computing power. The two sides influence and complement each other. Simulating the fabrication and behavior of semiconductor devices aids in developing smaller and faster devices, which improves available computational resources. Greater computing power means more accurate simulations can be constructed, which are presently needed as device sizes shrink to submicron regimes, thereby requiring new physical effects to be taken into account.

Acknowledgements—All of us sincerely thank T. Way for his aid in obtaining the pictures shown here. His photography skill was invaluable. Also, we thank D. Hall for her summer work linking FEDSS and SAMPLE. J. Mandelman thanks S. Geissler for data on the 3-D parasitic device modeling. Additionally, and in particular, D. Foty and T. Linton sincerely appreciate the aid from E. Farrell and P. Appino in obtaining the scientific visualization pictures. Also, D. Foty thanks T. Furukawa and D. M. Kenney for discussions of device physics and O. Bula appreciates the expertise of J. Bruce for his aid in the use of SAMPLE. Finally, we want to acknowledge and thank the number of people who have worked directly with us in the past on improving the

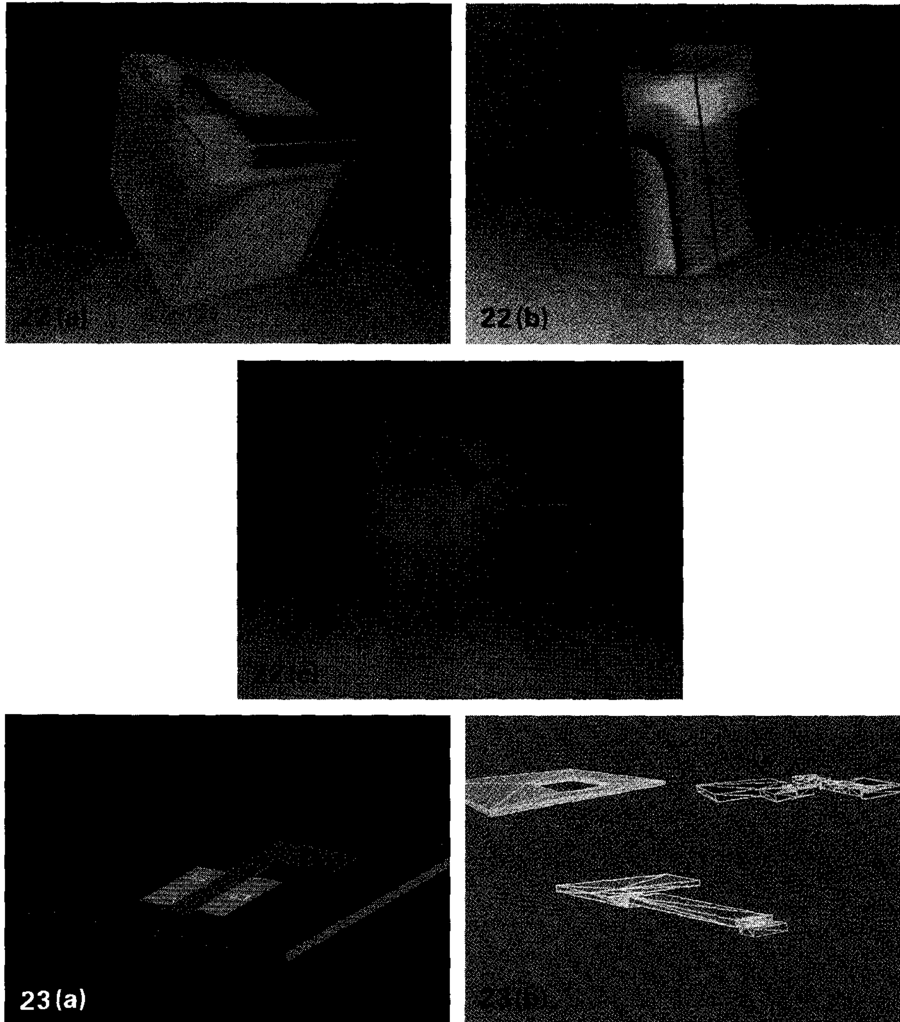


Fig. 22. Scientific visualization examples. (a) Electrostatic potentials in the trench-isolated MOSFET of Section 4.D. Low potential is shown as blue and high potential as red. (b) Electrostatic potentials in the modeled region in Section 4.C. The opaque red shell is a selected range of intermediate potentials. (c) Electron concentration for the modeled region of Section 4.D. The opaque red region shows the location of high electron concentrations in the channel.

Fig. 23. OYSTER representation of a MOSFET. (a) The green is a polysilicon line. The yellow represents the gate oxide. The two blue regions are metal contacts to the source and drain regions. (b) Pieces extracted from the FET structure and meshed.

capabilities of some of the simulation programs described here. In particular, we thank L. Borucki for his significant work on FEDSS.

REFERENCES

1. B. R. Penumalli, in *Process and Device Modeling* (Edited by W. L. Engl), p. 1. North-Holland, Amsterdam (1986).
2. J. D. Plummer, in *Semiconductor Processing* (Edited by D. C. Gupta), p. 21. American Society for Testing and Materials, Baltimore (1984).
3. S. Selberherr and E. Langer, *Microelectronics J.* **20,** 113 (1989).
4. W. L. Engl, O. Manck and A. W. Wieder, in *Process and Device Modeling for Integrated Circuit Design*, p. 3. Noordhoff, Leyden (1977).
5. W. L. Engl, H. K. Dirks and B. Meinerzhagen, *Proc. IEEE* **71,** 10 (1983).
6. S. J. Polak, C. Den Heijer and W. H. A. Schilders, *Int. J. Numerical Methods Engng* **24,** 763 (1987).
7. A. R. Newton, *Proc. IEEE* **69,** 1189 (1981).
8. K. Kani, in *Numerical Analysis of Semiconductor Devices, Proceedings of the NASECODE I Conference* (Edited by B. T. Browne and J. J. Miller), p. 104. Boole Press, Dublin (1979).
9. S. Selberherr, in *Process and Device Modeling* (Edited by W. L. Engl), p. 265. North-Holland, Amsterdam (1986).
10. S. Selberherr and C. A. Ringhofer, *IEEE Trans. Comput. Aided Design.* **CAD-3,** 52 (1984).
11. R. E. Bank, W. M. Coughran Jr, W. Fichtner, D. J. Rose and R. K. Smith, in *Process and Device Modeling* (Edited by W. L. Engl), p. 229. North-Holland, Amsterdam (1986).
12. P. J. Mole, *Inst. Phys. Conf.*, Ser. No. 82, p. 149 (1987).
13. G. Baccarani, M. Rudan, R. Guerrieri and P. Ciampolini, in *Process and Device Modeling* (Edited by W. L. Engl), p. 107. North-Holland, Amsterdam (1986).
14. S. Selberherr, *Analysis and Simulation of Semiconductor Devices*. Springer, Wien (1984).
15. C. M. Snowden, *Introduction to Semiconductor Device Modeling*. World Scientific, Singapore (1986).
16. C. M. Snowden (Editor), *Semiconductor Device Modeling*. Springer, London (1989).
17. M. Kurata, *Numerical Analysis for Semiconductor Devices*. Lexington Press, Lexington, MA (1982).
18. P. A. Markowich, *The Stationary Semiconductor Device Equations*. Springer, Wien (1986).
19. M. S. Mock, *Analysis of Mathematical Models of Semiconductor Devices*. Boole Press, Dublin (1983).
20. G. Baccarani and M. Rudan (Editors), *Simulation of Semiconductor Devices and Processes*, Vol. 3. Technoprint, Bologna (1988).
21. W. L. Engl (Editor), *Process and Device Modeling*. North-Holland, Amsterdam (1986).
22. K. A. Salsburg and H. H. Hansen, *IEEE Trans. Electron Devices* **ED-30,** 1004 (1983).
23. L. Borucki, H. H. Hansen and K. Varahramyam, *IBM J. Res. Develop.* **29,** 263 (1985).
24. L. Borucki and J. Slinkman, in *NASECODE IV: Proceedings of the Fourth International Conference on the Numerical Analysis of Semiconductor Devices and Integrated Circuits* (Edited by J. J. H. Miller), p. 226. Boole Press, Dublin (1985).
25. J. P. Biersack and L. G. Haggmark, *Nucl. Instrum. Meth.* **174,** 257 (1980).
26. J. F. Ziegler, J. P. Biersack and U. Littmark, *The Stopping and Range of Ions in Solids*, Vol. 1. Pergamon Press, New York (1984).
27. P. E. Cottrell and E. M. Buturla, in *Numerical Analysis of Semiconductor Devices, Proceedings of the NASECODE I Conference* (Edited by B. T. Browne and J. J. H. Miller), p. 31. Boole Press, Dublin (1979).
28. E. M. Buturla and P. E. Cottrell, *Solid-St. Electron.* **23,** 331 (1980).
29. E. M. Buturla, P. E. Cottrell, B. M. Grossman and K. A. Salsburg, *IBM J. Res. Develop.* **25,** 218 (1981).
30. E. M. Buturla, P. E. Cottrell, B. M. Grossman, C. T. McMullen and K. A. Salsburg, in *Numerical Analysis of Semiconductor Devices and Integrated Circuits, Proceedings of the NASECODE II Conference* (Edited by B. T. Browne and J. J. H. Miller), p. 160. Boole Press, Dublin (1981).
31. S. E. Laux, in *NASECODE III: Proceedings of the Third International Conference on the Numerical Analysis of Semiconductor Devices and Integrated Circuits* (Edited by J. J. H. Miller), p. 161. Boole Press, Dublin (1983).
32. B. M. Grossman and M. J. Hargrove, *IEEE Trans. Electron Devices* **30,** 1092 (1983).
33. B. M. Grossman, E. M. Buturla and P. E. Cottrell, in *Computing Methods in Applied Science and Engineering*, Vol. VI (Edited by R. Glowinski and J. L. Lions), p. 697. North-Holland, Amsterdam (1984).
34. S. E. Laux and R. G. Byrnes, *IBM J. Res. Develop.* **29,** 289 (1985).
35. S. E. Laux, *IEEE Trans. Electron Devices* **32,** 2028 (1985).
36. S. E. Laux and B. M. Grossman, *IEEE Trans. Electron Devices* **32,** 2076 (1985).
37. S. E. Laux, E. M. Buturla, W. H. Chang, P. L. Corson, P. E. Cottrell, E. J. Farrell and J. A. Mandelman, Modeling using the "FIELDAY Program", in *Large Scale Computational Device Modeling* (Edited by K. Hess), p. 43. University of Illinois Press (1986).
38. D. C. Cole and J. B. Johnson, *IEEE Proc. of the Workshop on Low Temperature Semiconductor Electonics*, p. 73 (1989).
39. E. M. Buturla, J. Johnson, S. Furkay and P. Cottrell, in *NASECODE VI: Proceedings of the Sixth International Conference on the Numerical Analysis of Semiconductor Devices and Integrated Circuits* (Edited by J. J. H. Miller), p. 291. Boole Press, Dublin (1989).
40. G. M. Koppelman and M. A. Wesley, *IBM J. Res. Develop.* **27,** 149 (1983).
41. R. N. Wolfe, M. A. Wesley, J. C. Kyle Jr, F. Gracer and W. J. Fitzgerald, *IBM J. Res. Develop.* **31,** 277 (1987).
42. R. C. Evans, G. Koppelman and V. T. Rajan, *IBM J. Res. Develop.* **31,** 343 (1987).
43. M. V. Fischetti and S. E. Laux, *Phys. Rev. B* **38,** 9721 (1988).
44. S. E. Laux and M. V. Fischetti, *IEEE Electron Device Lett.* **EDL-9,** 467 (1988).
45. M. V. Fischetti and S. E. Laux, in *Simulation of Semiconductor Devices and Processes*, Vol. 3 (Edited by G. Baccarani and M. Rudan), p. 349. Technoprint, Bologna (1988).
46. S. E. Laux, M. V. Fischetti and D. J. Frank, *IBM J. Res. Develop.* In press.
47. W. G. Oldham, S. N. Nandgaonkar, A. R. Neureuther and M. M. O'Toole, *IEEE Trans. Electron Devices* **ED-26,** 717 (1979).
48. W. G. Oldham, A. R. Neureuther, C. Sung, J. L. Reynolds and S. N. Nandgaonkar, *IEEE Trans. Electron Devices* **ED-27,** 1455 (1980).
49. R. A. Ferguson and A. R. Neureuther, *Proc. KTI Microelectronics Seminar*, p. 41 (1987).
50. D. E. Lyons, S. F. Meier, L. Winemberg, G. Addiego, A. R. Neureuther and W. G. Oldham, *Proc. KTI Microelectronics Seminar*, p. 261 (1988).

51. R. A. Ferguson, C. A. Spence, Y. Shacham-Diamand and A. R. Neureuther, *SPIE Proc.* **1086,** 262 (1989).
52. C. A. Spence, R. A. Ferguson, A. R. Neureuther and W. G. Oldham, *Solid-St. Electron.* **33,** 625 (1990).
53. M. R. Pinto, C. S. Rafferty, H. R. Yeager and R. W. Dutton, PISCES IIB—Poisson and continuity equation solver, Tech. Rep., Stanford Electronics Lab., Stanford University, Palo Alto, CA (1986).
54. Z. Yu and R. W. Dutton, SEDAN III—A generalized electronic material device analysis program, Tech. Rep., Stanford Electronics Lab., Stanford University, Palo Alto, CA (1985).
55. G. Baccarani, R. Guerrieri, P. Ciampolini and Massimo Rudan, in *NASECODE IV: Proc. Fourth Int. Conf. Numerical Analysis of Semiconductor Devices and Integrated Circuits*, p. 3. (Edited by J. J. Miller), Boole Press, Dublin (1985).
56. R. Guerrieri, P. Ciampolini, A. Gnudi, M. Rudan and G. Baccarini, *IEEE Trans. Electron Devices.* **ED-33,** 1201 (1986).
57. A. Forghieri, R. Guerrieri, P. Ciampolini, A. Gnudi, M. Rudan and G. Baccarini, *IEEE Trans. Comput.-Aided Design* **CAD-7,** 231 (1988).
58. S. Selberherr, W. Fichtner and H. W. Pötzl, in *Numerical Analysis of Semiconductor Devices, Proc. NASECODE I Conf.*, p. 275. (Edited by B. T. Browne and J. J. H. Miller). Boole Press, Dublin (1979).
59. S. Selberherr, A. Schütz and H. W. Pötzl, *IEEE Trans. Electron Devices* **ED-27,** 1540 (1980).
60. M. Thurner and S. Selberherr, in *NASECODE V: Proceedings of the Fourth International Conference on the Numerical Analysis of Semiconductor Devices and Integrated Circuits*, p. 327 (Edited by J. J. H. Miller). Boole Press, Dublin (1987).
61. W. Hänsch and S. Selberherr, *IEEE Trans. Electron Devices* **ED-34,** 1074 (1987).
62. D. A. Antoniadis, S. E. Hansen, R. W. Dutton and A. G. Gonzales, Tech. Rep. 5019-1, Stanford Electronics Lab., Stanford University, Palo Alto, CA (1977).
63. D. A. Antoniadis and R. W. Dutton, *IEEE Trans. Electron Devices* **ED-26,** 490 (1979).
64. C. P. Ho, J. D. Plummer, S. E. Hansen and R. W. Dutton, *IEEE Trans. Electron Devices* **ED-30,** 1438 (1983).
65. M. E. Law, C. S. Rafferty and R. W. Dutton, *SUPREM IV User's Manual.* Stanford University, Stanford (1986).
66. M. E. Law and R. W. Dutton, *IEEE Trans. Comput.-Aided Design* **7,** 181 (1988).
67. M. E. Law, Ph.D. Thesis, Stanford University (1988).
68. R. B. Fair and R. Subrahmanyan, *Proc. SPIE* **530,** 88 (1985).
69. R. B. Fair, *IEEE Trans. Electron Devices* **ED-35,** 285 (1988).
70. D. Collard and K. Taniguchi, *IEEE Trans. Electron Devices* **ED-33,** 1454 (1986).
71. S. Odanaka, H. Umimoto, M. Wakabayashi and H. Esaki, *IEEE Trans. Comput.-Aided Design* **7,** 675 (1988).
72. J. Lorenz, J. Pelka, H. Ryssel, A. Sachs, A. Seidl and M. Svoboda, *IEEE Trans. Electron Devices* **ED-32,** 1977 (1985).
73. More information on some of these programs, as well as descriptions of other programs not mentioned here, can be found in *Software Tools for Process, Device and Circuit Modelling* (Edited by W. Crans). Boole Press, Dublin (1989).
74. In particular, see R. Hamann, M. Schlüter and C. Chiang, *Phys. Rev. Lett.* **43,** 1494 (1979).
75. P. Hohenberg and W. Kohn, *Phys. Rev. B* **136,** 864 (1964); W. Kohn and L. J. Sham, *Phys. Rev. A* **140,** 1133 (1965). More recent articles can be found in *The Inhomogeneous Electron Gas* (Edited by N. H. March and S. Lundqvist). Plenum, New York (1984).
76. R. Carr, P. J. Kelly, A. Oshiyama and S. T. Pantelides, *Phys. Rev. Lett.* **52,** 1814 (1984); **54,** 360 (1985).
77. Y. Bar-Yam and J. D. Joannopoulos, *Phys. Rev. B* **30,** 2216 (1984); *J. Electron. Mater.* **14a,** 261 (1985).
78. C. S. Nichols, C. G. Van de Walle and S. T. Pantelides, *Phys. Rev. Lett.* **62,** 1049 (1989).
79. See, for example, *Deep Centers in Semiconductors* (Edited by S. T. Pantelides), Chap. 1. Gordon and Breach, New York (1986). Also see, S. T. Pantelides, *Mat. Res. Soc. Symp. Proc.*, Vol. **71** (Edited by M. Wittmer *et al.*), p. 147 (1986).
80. K. C. Pandey, *Phys. Rev. Lett.* **57,** 2287 (1986).
81. An excellent reference for this method of calculation is, R. W. Hockney and J. W. Eastwood, *Computer Simulation Using Particles.* McGraw-Hill, New York (1981). Other good sources also exist. See, for example, F. F. Abraham, *J. Vac. Sci. Technol. B* **2,** 534 (1984).
82. C. H. Bennett, in *Diffusion in Solids* (Edited by A. S. Nowick and J. J. Burton), p. 73. Academic Press, New York (1975); *Thin Solid Films* **25,** 65 (1975); *J. Comput. Phys.* **22,** 245 (1976); in *Algorithms for Chemical Computations* (Edited by R. E. Christoffersen), p. 63. American Chem. Soc., Washington D.C. (1977); *Colloq. Metall.* **19,** 656 (1976).
83. A. DaFano and G. Jacucci, *Phys. Rev. Lett.* **39,** 950 (1977); *J. Nucl. Mater.* **69/70,** 549 (1978).
84. G. Jacucci, in *Diffusion in Crystalline Solids* (Edited by G. E. March and A. S. Nowick), Chap. 8. Academic, New York (1984).
85. I. Batra, F. F. Abrahams and S. Ciraci, *Phys. Rev. B* **35,** 9552 (1987).
86. J. H. Weiner and R. E. Forman, *Phys. Rev. B* **10,** 315 (1974).
87. G. E. Murch, in *Diffusion in Crystalline Solids* (Edited by G. E. Murch and A. S. Nowick), Chap. 7. Academic, New York (1984).
88. K. W. Kehr and K. Binder, in *Simulation of Diffusion in Lattice Gases and Related Phenomena* (Edited by K. Binder). Springer, Berlin (1985).
89. R. A. Tahir-Kheli, *Phys. Rev. B* **35,** 5503 (1987).
90. K. W. Kehr and K. Binder, in *Applications of the Monte Carlo Method in Statistical Physics, Vol. 36, Topics in Current Physics* (Edited by K. Binder), p. 181. Springer, New York (1987).
91. A. Akiyama, T. Hosoi, I. Ishihara, S. Matsumoto and T. Niimi, *IEEE Trans. Comput.-Aided Design* **CAD-6,** 185 (1987).
92. F. F. Morehead and R. F. Lever, *Appl. Phys. Lett.* **48,** 151 (1986).
93. B. J. Mulvaney and W. B. Richardson, *Appl. Phys. Lett.* **51,** 1439 (1987).
94. M. Orlowski, *Appl. Phys. Lett.* **53,** 1323 (1988).
95. R. T. Bate, *Scientific Am.* **258,** 96 (1988).
96. K. Varahramyan, IBM Burlington Tech. Rep. (1990).
97. C. C. Williams, J. Slinkman, W. P. Hough and H. K. Wickramasinghe, *Appl. Phys. Lett.* **55,** 1662 (1989).
98. S. H. Goodwin-Johansson, R. Subrahmanyan, C. E. Floyd and H. Z. Massoud, *IEEE Trans. Comput.-Aided Design* **8,** 232 (1989).
99. B. E. Deal and A. S. Grove, *J. appl. Phys.* **36,** 3770 (1965).
100. J. Lindhard, M. Scharff and M. Schiott, *Mat. Fys. Medd. Dan. Vidensk. Selsk.* **33,** 1 (1963).
101. K. Lehovec and A. Slobodskoy, *Solid St. Electron.* **3,** 45 (1961).
102. M. M. Atalla and E. Tannenbaum, *Bell Syst. Tech. J.* **39,** 933 (1960).
103. A. S. Grove, O. Leistiko and C. T. Sah, *J. appl. Phys.* **35,** 2695 (1964).
104. S. M. Hu and S. Schmidt, *J. appl. Phys.* **39,** 4272 (1968).

105. J. F. Gibbons, W. S. Johnson and S. W. Mylroie, *Projected Range Statistics*. Wiley, New York (1975).
106. D. K. Brice, *Ion Implantation Range and Energy Deposition Distributions*. Plenum, New York (1975).
107. R. B. Fair and G. R. Weber, *J. appl. Phys.* **44,** 273 (1973).
108. R. B. Fair and P. N. Pappas, *J. Electrochem. Soc.* **122,** 1241 (1975).
109. R. B. Fair and J. C. C. Tsai, *J. Electrochem. Soc.* **124,** 1107 (1977).
110. D. P. Kennedy and R. R. O'Brien, *IBM J. Res. Develop.* **9,** 179 (1965).
111. D. P. Kennedy and P. C. Murley, *IBM J. Res. Develop.* **10,** 6 (1966).
112. B. R. Penumalli, *IEEE Trans. Electron Devices* **ED-30,** 986 (1983).
113. D. Chin, M. R. Kump, H. G. Lee and R. W. Dutton, *IEEE Trans. Electron Devices* **ED-29,** 336 (1982).
114. M. T. Robinson and I. M. Torrens, *Phys. Rev. B* **9,** 5008 (1974).
115. M. T. Robinson, *User's Guide to MARLOWE (Version 12)*. Oak Ridge National Laboratory (1984).
116. P. M. Fahey, P. B. Griffin and J. D. Plummer, *Rev. Mod. Phys.* **61,** 289 (1989).
117. FEDSS Reference Guide, Version 1.0, IBM Corp., GC40-1653-00, pp. 16–17 (1989).
118. S. M. Hu, *J. appl. Phys.* **55,** 4095 (1984).
119. C. S. Rafferty, L. Borucki and R. W. Dutton, *Appl. Phys. Lett.* **54,** 1516 (1989).
120. D. Chin, S. Y. Oh, S. M. Hu and R. W. Dutton, *IEEE Trans. Electron Devices* **ED-30,** 744 (1983).
121. N. C. Lu, P. E. Cottrell, W. J. Craig, S. Dash, D. L. Critchlow, R. L. Mohler, B. J. Machesney, T. H. Ning, W. P. Noble, R. M. Parent, R. E. Scheuerlein, E. J. Sprogis and L. M. Terman, *IEEE J. Solid St. Circuits* **SC-21,** 627 (1986).
122. B. Davari, C. Koburger, T. Furukowa, Y. Taur, W. Noble, J. Warmock and J. Mauer, *IEDM Tech. Digest*, p. 92 (1988).
123. W. Shockley, *Bell System Tech. J.* **28,** 435 (1949).
124. W. Shockley, *Electrons and Holes in Semiconductors*, Chap. 12. D. Van Nostrand, New York. (1950).
125. W. Van Roosbroeck, *Bell System Tech. J.* **29,** 560 (1950).
126. N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, Chap. 29. Holt, Rinehart and Winston, Philadelphia (1976).
127. R. F. Pierret, *Advanced Semicondutor Fundamentals, Vol. VI, Modular Series on Solid State Devices*, Chap. 6. Addison-Wesley, Reading, MA (1987).
128. A. H. Marshak and C. M. Van Vliet, *Proc. IEEE* **72,** 148 (1984).
129. J. Bardeen and W. Shockley, *Phys. Rev.* **80,** 72 (1950).
130. H. Kroemer, *RCA Rev.* **18,** 332 (1957).
131. O. W. Purbo, D. T. Cassidy and S. H. Chisholm, *J. appl. Phys.* **66,** 5078 (1989).
132. D. L. Scharfetter and H. K. Gummel, *IEEE Trans. Electron Devices* **ED-16,** 64 (1969).
133. See, for example, J. P. Kreskovsky, *IEEE Trans. Electron Devices* **ED-34,** 1128 (1987).
134. H. K. Gummel, *IEEE Trans. Electron Devices* **ED-11,** 455 (1964).
135. H. W. Loeb, R. Andrew and W. Love, *Electron. Lett.* **4,** 352 (1968).
136. J. E. Schroeder and R. S. Muller, *IEEE Trans. Electron Devices* **ED-15,** 954 (1968).
137. J. W. Slotboom, *Electron. Lett.* **5,** 677 (1969).
138. D. P. Kennedy and R. R. O'Brien, *IBM J. Res. Develop.* **13,** 662 (1969).
139. D. Vandorpe and N. H. Xuong, *Electron. Lett.* **7,** 47 (1971).
140. D. P. Kennedy and P. C. Murley, *IBM J. Res. Develop.* **17,** 2 (1973).
141. M. S. Mock, *Solid-St. Electron.* **16,** 601 (1973).
142. A. Yoshii, H. Kitazawa, M. Tomizawa, S. Horiguchi and T. Sudo, *IEEE Trans. Electron Devices* **ED-29,** 184 (1982).
143. N. Shigyo, M. Konaka and R. L. M. Dang, *Electron. Lett.* **ED-18,** 274 (1982).
144. J. R. Barker and D. K. Ferry, *Solid-St. Electron.* **23,** 519 (1980).
145. Numerous articles exist on mobility models. See the references cited in Refs [14] and [34], as well as, for example, Refs [146–152] below.
146. S. A. Schwarz and S. E. Russek, *IEEE Trans. Electron Devices* **ED-30,** 1629, 1634 (1983).
147. N. D. Arora and G. S. Gildenblat, *IEEE Trans. Electron Devices* **ED-34,** 310 (1987).
148. T. Nishida and C. T. Sah, *IEEE Trans. Electron Devices* **ED-34,** 310 (1987).
149. A. Rothwarf, *IEEE Trans. Electron Device Lett.* **EDL-8,** 499 (1987).
150. C. Lombardi, S. Manzini, A. Saporito and M. Vanzi, *IEEE Trans. Comput.-Aided Design* **CAD-7,** 1164 (1988).
151. M. S. Lin, *IEEE Trans. Electron Devices* **ED-32,** 700 (1985); **35,** 2406 (1988).
152. H. Shin, A. F. Tasch Jr, C. M. Maziar and S. K. Banerjee, *IEEE Trans. Electron Devices* **ED-36,** 1117 (1989).
153. R. Stratton, *Phys. Rev.* **126,** 126 (1962).
154. K. Bløtekjær, *IEEE Trans. Electron Devices* **ED-17,** 38 (1970).
155. F. Reif, *Fundamentals of Statistical and Thermal Physics*, Section 14.5. McGraw-Hill, New York (1965).
156. K. Huang, *Statistical Mechanics*, Section 5.2. Wiley, New York (1963).
157. R. K. Cook and J. Frey, *COMPEL* **1,** 65 (1982).
158. M. Rudan and F. Odeh, *COMPEL* **5,** 149 (1986).
159. M. Rudan, F. Odeh and J. White, *COMPEL* **6,** 151 (1987).
160. E. M. Azoff, *Solid-St. Electron.* **30,** 913 (1987).
161. C. J. Wilson, *IEEE Trans. Electron Devices* **ED-35,** 180 (1988).
162. B. Meinerzhagen and W. L. Engl, *IEEE Trans. Electron Devices* **ED-35,** 689 (1988).
163. Y. Feng and A. Hintz, *IEEE Trans. Electron Devices* **ED-35,** 1419 (1988).
164. E. M. Azoff, *IEEE Trans. Electron Devices* **ED-36,** 609 (1989).
165. C. L. Gardner, J. W. Jerome and D. J. Rose, *IEEE Trans. Comput.-Aided Design* **CAD-5,** 501 (1989).
166. P. Price, *IBM J. Res. Develop.* **14,** 12 (1970).
167. P. Price, *Solid-St. Electron.* **21,** 9 (1978).
168. D. K. Ferry, in *Handbook of Semiconductors*, Vol. 1 (Edited by W. Paul), p. 563. North-Holland, Amsterdam (1982).
169. P. J. Price, *Semiconductors and Semimetals*, Vol. 14, p. 249. Academic, New York (1979).
170. C. Jacoboni and L. Reggiani, *Rev. Mod. Phys.* **55,** 645 (1983).
171. C. Jacoboni and P. Lugli, *The Monte Carlo Method for Semiconductor Device Simulation*. Springer, New York (1989).
172. S. Bandyopadhyay, M. E. Klausmeier-Brown, C. M. Maziar, S. Datta and M. S. Lundstrom, *IEEE Trans. Electron Devices* **ED-34,** 392 (1987).
173. J. M. Higman, K. Hess, C. G. Hwang and R. W. Dutton, *IEEE Trans. Electron Devices* **ED-36,** 930 (1989).
174. The literature on this subject is considerable. However, for a strong flavor of this field of research see: T. Ando, A. B. Fowler and F. Stern, *Rev. Mod. Phys.* **54,** 437 (1982); S. E. Laux and F. Stern, *Appl. Phys. Lett.* **49,** 91 (1986).

175. S. Datta, *Quantum Phenomena, Vol. VIII, Modular Series on Solid State Devices*. Addison-Wesley, Reading, MA (1989).
176. L. Reggiani, *Physica B, C* **134B,** 123 (1985).
177. H. L. Grubin, D. K. Ferry and C. Jacoboni (Editors), *Proceedings of the International Workshop on Quantum Transport Theory with Applications to Nanometer Electronics, Italy*. Plenum, New York (1988).
178. R. Brunetti, C. Jaboboni and F. Rossa, *Phys. Rev. B* **39,** 10781 (1989).
179. J. R. Barker, in *Semiconductor Device Modeling* (Edited by C. M. Snowden), p. 207. Springer, London (1989).
180. P. E. Cottrell and E. M. Buturla, *IEDM Tech. Digest*, p. 51 (1975).
181. P. E. Cottrell, R. R. Troutman and T. H. Ning, *IEEE Trans. Electron Devices* **ED-26,** 520 (1979).
182. D. Kenney, E. Adler, B. Davari, J. DeBrosse, W. Frey, T. Furukawa, P. Geiss, D. Harmon, D. Horak, M. Kerbaugh, C. Koburger, J. Lasky, J. Rembetski, W. Schwittek and E. Sprogis, *Proc. 1988 VLSI Symp. on VLSI Technology*, San Diego, California, p. 25 (1988).
183. J. A. Appels, E. Kooi, M. M. Paffen, J. J. H. Schatorje and W. H. C. G. Verkuylen, *Philips Res. Reports* **25,** 118 (1970).
184. N. Shigyo, T. Wada, S. Fukuda, K. Hieda, T. Hamamoto, H. Watanabe, K. Sunouchi and W. Tango, *IEDM Tech. Digest*, p. 636 (1987).
185. T. Furukawa and J. Mandelman, *Electrochemical Society Fall Meeting Proc.*, p. 329 (1988).
186. K. Hieda, F. Horiguchi, H. Watanabe, K. Sunouchi and T. Hamamoto, *IEEE Trans. Electron Devices* **ED-36,** 1615 (1989).
187. D. Foty, J. Mandelman and T. Furukawa, *Electrochemical Society Fall Meeting Proc.*, p. 295 (1989).
188. F. Gaensslen, V. Rideout, E. Walker and J. Walker, *IEEE Trans. Electron Devices* **ED-24,** 218 (1977).
189. IEEE Transactions on Electron Devices, Special Issues on Low Temperature Semiconductor Electronics (1987), (1989).
190. *Proceedings of the Symposium on Low Temperature Electronics and High Temperature Superconductors* (Edited by S. Raider). The Electrochemical Society, Pennington, NJ (1988).
191. T. C. May and M. H. Woods, *IEEE Trans. Electron Devices* **ED-26,** 2 (1979).
192. C. M. Hsieh, P. C. Murley and R. R. O'Brien, *IEEE Electron Device Lett.* **EDL-2,** 103 (1981).
193. G. A. Sai-Halasz, M. R. Wordeman and R. H. Dennard, *IEEE Trans. Electron Devices* **ED-29,** 725 (1982).
194. G. C. Messenger, *IEEE Trans. Nuclear Sci.* **NS-29,** 2024 (1982).
195. R. S. Wagner, J. M. Bradley, N. Bordes, C. J. Maggiore, D. N. Sinha and R. B. Hammond, *IEEE Trans. Nuclear Sci.* **NS-34,** 1240 (1987).
196. J. M. Wagner, N. Bordes, J. M. Bradley, C. J. Maggiore, A. R. Knudson and A. B. Campbell, *IEEE Trans. Nuclear Sci.* **NS-35,** 1578 (1988).
197. L. Reggiani and P. Lugli, *Inst. Phys. Conf.*, Ser. No. 95, p. 545 (1989).
198. G. A. Sai-Halasz, M. R. Wordeman, D. P. Kern, E. Ganin, S. Rishton, D. S. Zicherman, H. Schmid, M. R. Polcari, H. Y. Ng, P. J. Restle, T. H. P. Chang and R. H. Dennard, *IEEE Electron Device Lett.* **EDL-8,** 463 (1987).
199. For further discussion, see M. V. Fischetti, S. E. Laux and D. J. DiMaria, *Appl. Surf. Sci.* **39,** 578 (1989).
200. S. G. Duvall, *IEEE Trans. Comput.-Aided Design* **CAD-7,** 741 (1988).
201. For relevant recent references to this field that we are calling scientific visualization, see the conference proceedings for the SPIE/SPSE Symposium on Electonic Imaging (1990). For earlier work, see, for example, E. J. Farrell, S. E. Laux, P. L. Corson and E. M. Buturla, *IBM J. Res. Develop.* **29,** 302 (1985).
202. See Ref. [14], p. 2.