SEMICONDUCTOR PROCESS MODELING

The electronics industry's appetite for low-cost circuits with ever more functions and ever higher levels of integration puts a serious challenge to the semiconductor process design. Semiconductor manufacturing proceeds in sequences of processes that utilize many fabrication materials and technologies to design active devices and multilevel interconnect structures with desired electrical characteristics (1). The traditional design approach in semiconductor technology is to perform a set of experiments in order to determine the process parameters that are best suited for device and circuit specifications. However, as device dimensions continually shrink, experimental prototyping becomes extremely expensive and time-consuming because of the plant, equipment, personnel, materials, and supplies that are required. For this reason, it is increasingly recognized that the semiconductor industry has to ground the design of new technologies on predictive computational modeling.

Computational modeling is the act of producing an abstract description or representation of a problem or process in order to simplify the analysis of the problem or to enable the simulation of the process by using computer aids. With the help of computational modeling the evaluation and optimization of various design aspects are possible without resorting to costly and time-consuming trial fabrication and measurement steps. Moreover, it indirectly provides valuable insight into important physical quantities that cannot be measured directly. The benefits are shortened development cycles, reduced costs, and increased quality and reliability of the final industrial products. An important field of computational modeling related to semiconductor manufacturing belongs to process modeling.

The aim of process modeling is to predict geometries and material properties of the wafer structures and semiconductor devices as they result from the manufacturing process. It should be distinguished from the modeling activity of macroscopic processes within the fabrication equipment, which is referred to as equipment modeling (2). The equipment modeling principally serves to improve the equipment design, while process modeling considers microscopic processes or the wafer level of semiconductor manufacturing. The two traditional branches of process modeling are concerned with wafer topography and wafer bulk material layers. The objective of topography modeling is to predict the evolution and the final geometry of the wafer surface as it is affected by pattern definition and transfer processes. The bulk process modeling focuses on doping and material growth processes that substantially change the physical properties of the wafer material.

Process modeling plays an important role for the semiconductor technology design both in the development and in the characterization phase. In technology development it serves to refine a process recipe by evaluating its feasibility or by looking for improvements in the process flow. In technology characterization, the process modeling provides input data for device modeling. Device modeling accounts for carrier transport within the semiconductor device structure and for electrical characteristics of the device (3,4). Appropriate device models are further used in circuit modeling to predict the overall behavior of the electrical circuit. Integrating process, device, and circuit modeling enables one to predict the possible im-

pact of changes in the fabrication process on the circuit performance (5).

The first step in process modeling is to recognize a concept of mechanisms and relations that captures the essence of actual phenomena behind particular processes. This activity is often referred to as the formulation of physical models. Particular processes are typically characterized by a hierarchy of physical models. At the bottom of the hierarchy the models are derived from principles using mechanisms of atomic level or fundamental laws, while simple analytical models are on the top of the hierarchy. The models between allow a tradeoff of model generality for their simplicity. The physical models are commonly presented in mathematical form as systems of nonlinear partial differential equations (PDEs) or by algorithms. The analysis of physical models for process modeling is made conceptually more manageable by subdividing them into the models for photolithography, etching and deposition, ion implantation, bulk particle transport, and mechanical deformation.

In lithography processes, a wafer, covered by radiation-sensitive material (resist), is first exposed to light, X ray, electron-, or ion-beam radiation that alters the ability of the exposed material to resist an etching substance. A surface profile is then developed into the resist film with the etching rate determined from the radiation-produced latent image. Among different lithography technologies, photolithography holds the leading position in today's semiconductor industry. The models for light exposure and development processes in photolithography have to account for the light-intensity distribution in the photoresist film, for the chemical reaction that changes photoresist etching properties, and for the resulting photoresist profile after development.

The formation of multilayer wafer structures is principally based on the successive advancement of the wafer surface as a result of material removal in etching processes or material addition in deposition processes. The role of physical models is to relate the propagation velocity of the surface to material properties and processing conditions. The processing techniques used for etching and deposition range from isotropic chemical processes to directional physical processes, with mixed physicochemical techniques, such as reactive ion etching, in between. The propagation velocity of the wafer surface depends in a complicated way on the geometry of the wafer surface and on the source of particles in the processing equipment. The most important model parameters are the angle-dependent flux of source particles, the angle of particle incidence relative to the normal direction of the surface, and the visibility between the source and surface points.

In ion-implantation processes, the wafer is exposed to the beam of ions having energy enough to penetrate into the wafer material. As an energetic particle enters a solid target, it loses energy in a scattering process until it comes to rest. Of interest for process modeling are the distributions of stopped particles, the produced damage, and the energy, which is transferred to the target material. The damage occurs when ions collide with a lattice atom and when they cause it to leave its original site in the lattice. The consideration of the crystal structure within ion-implantation models is needed to account for the preferential penetration of ions along crystalline axes or planes, referred to as channeling. It results in deep tail regions beyond the profile within amorphous materials. The physical mechanisms of ion stopping are best employed in the atomic level models that trace individual ion trajectories and that can be implemented numerically by Monte Carlo methods. Instead of modeling individual ion trajectories, it is possible to formulate statistical distribution functions for implanted ions. It should be emphasized that the same modeling principles are applicable also to ion- and electron-beam lithography exposure processes.

One of the most important group of physical models is related to the transport of particles within the bulk region. Since semiconductor device characteristics depend drastically on the distribution of electrically active dopants, it is of prime importance to model dopant redistribution accurately in thermal processes such as diffusion, annealing, epitaxy, oxidation, or nitridation. However, particle transport is important for many other processes. The kinetics of oxidation, nitridation, and other native film-producing processes is essentially based on the transport of reactant particles through the growing films. The process of baking in photolithography is also based on particle transport. The principal physical mechanism for particle transport is diffusion. However, the governing equations for particle transport should also account for advection due to electric field, substrate material motion, and various chemical reactions among different particles. Bulk particle-transport models are commonly organized hierarchically with an increasing level of physical sophistication. They range from single-species diffusion equations to complex coupled systems of diffusion-drift-reaction PDEs for multiple species.

The models for mechanical deformation are necessary to follow the evolution of the stress field in different material layers during manufacturing. They are primarily formulated for the thermal oxidation process and other material growth processes based on the accelerated production of native films (nitridation or titanium silicidation, for example). However, the stresses that are induced by thermal cycling and subsequent material deposition are of the same importance. Generally, the cumulative mechanical stress represents an important factor that could affect the reliability of semiconductor devices and the interconnection system. Depending on the processing temperature the mechanical description of the material layers in semiconductor manufacturing varies from purely elastic solid to viscous fluids. The models and methods to determine stress distribution in semiconductor process modeling often originate in other engineering and science disciplines like metallurgy, geology, and mechanics of deformable bodies.

Besides physical models, it is of equal importance for semiconductor process modeling to formulate appropriate discrete models. The principal tasks to formulate the discrete model are the generation and the control of appropriate grid structures for arbitrarily shaped multilayer material domains and the derivation of the discrete analog of the governing mathematical description. Finally, the practical application of process modeling is enabled by simulation tools that integrate various physical and discrete models and that allow one to include appropriate numerical methods, user interfaces, and visualization techniques.

The numerical solution of the occurring PDEs or systems of PDEs requires a subdivision of the complete physical domain into small subdomains. These cells serve to formulate the discrete equivalent of the given problem. The algebraic problem then has to be solved by properly chosen numerical methods. These two phases, discretization and solution, are

strongly coupled to each other. Similar to other areas of computer simulation there are different approaches for the choice of cells, for the discrete approximation, and for the solution of the algebraic problem.

Within the discretization process the choice of structured and unstructured meshes can be considered. The finite-difference method (FD) replaces the derivatives within the differential operator by difference operators, derived on the given mesh with the aid of Taylor expansion. The finite-volume discretization (FV) is derived from the integral representation of the physical equations by applying the Gauss theorem on each grid cell (control volume) and using the Taylor expansion, again, for the derivatives that occur at cell boundaries. The variational formulation of the PDE with appropriately chosen test functions together with integration by parts and searching for the solution in a finite-dimensional function space leads to the finite element (FE) method.

In principle, each of the discretization techniques just mentioned can be combined with any type of grid. The final selection of the grid and the discretization method should depend on the geometry of the domain, the PDE (including boundary conditions) to be solved, and the coordinate system that is used for the description of the continuous problem. In practice, finite elements (and finite volumes) are used together with unstructured meshes, whereas finite differences and finite volumes are traditionally combined with Cartesian and logically rectangular boundary-fitted grids. Such boundary-fitted grids require the transformation of the underlying problem to new curvilinear coordinates. The grid within the corresponding computational domain is well structured, and the boundary conditions can be discretized easily although a transformed problem has to be considered.

In two-dimensional process simulation all approaches mentioned previously have been applied. The construction of boundary-fitted grids in process simulation exploits algebraic transformation rules, transformations which are based on elliptic systems of PDEs and variational approaches. Within the complete sequence of simulation steps the grid has to guarantee efficient algorithms and accurate solutions. Any grid-generation technique has to take care of problems arising from strongly varying quantities, multilayer devices, geometrical singularities, and time-dependent structures. These typical problems for process simulation and the desired efficiency automatically lead to the requirement of grid adaptation. As the mesh size cannot be determined in advance the solution process on a given relatively coarse mesh has to provide the information about where to refine the mesh. This type of adaptation strongly depends on error estimators that have to be developed both for structured and unstructured meshes. As the discretization error or approximations to it, which are commonly used for numerically sound grid-adaptation criteria, depend both on the local mesh size and on the local order of discretization, there are two possible ways of improving the accuracy: first, the order of approximation could be increased, and second, the local mesh size could be decreased. For practical reasons of programming the latter approach is chosen in most cases.

Within practically used design environments the steps of grid generation, grid adaptation, and solution of the resulting algebraic systems of equations have to be performed automatically and without an interaction from outside. This is mandatory for technology computer-aided design (TCAD) where complete processing sequences are intended to be simulated.

PHYSICAL MODELS

Photolithography

The propagation velocity of the resist profile during photolithography development is related to the latent image, which is produced in the resist material during the light-exposure processing phase. The latent image is described as a distribution of relative photoactive compound (PAC) concentration, $M(\boldsymbol{r}, t)$, representing the fraction of PAC that remains in the resist at position $\boldsymbol{r}$ after exposure time t. A physical model has to take into account (1) the intensity of the electromagnetic (EM) field that develops in the resist and in underlying material layers due to reflection, refraction, and absorption phenomena and (2) photochemical kinetics activated by the absorption of EM energy (bleaching), which produces the degradation and nonuniform distribution of the PAC. Both effects should be considered simultaneously because PAC degradation reduces the absorption coefficient of the resist and consequently modifies the optical properties of the resist.

Dill's so-called ABC model for the absorption coefficient α within the resist and the bleaching kinetics is defined by (6)

$$\alpha = AM(\boldsymbol{r}, t) + B \tag{1}$$

$$\frac{\partial M(\boldsymbol{r}, t)}{\partial t} = -I(\boldsymbol{r}, t)M(\boldsymbol{r}, t)C \tag{2}$$

where A, B, and C are model parameters. $I(\boldsymbol{r}, t)$ is the EM field intensity, which is related to the electric field $\boldsymbol{E}$ and the magnetic field $\boldsymbol{H}$ by $I \propto |\boldsymbol{E} \times \boldsymbol{H}|$. Since the bleaching rate is sufficiently small compared to the speed of light, the EM field is modeled as quasistatic and time harmonic obeying the Maxwell equations in the form

$$\nabla \times \boldsymbol{H}(\boldsymbol{r}) = [-j\omega\epsilon(\boldsymbol{r}, \alpha) + \sigma(\boldsymbol{r})]\boldsymbol{E}(\boldsymbol{r}) \tag{3}$$

$$\nabla \times \boldsymbol{E}(\boldsymbol{r}) = j\omega\mu_0\boldsymbol{H}(\boldsymbol{r}) \tag{4}$$

where ϵ, μ_0, and σ are the permittivity, permeability, and conductivity of the material layers, respectively. Space-dependent permittivity and conductivity take into account the presence of different material layers below the resist. The variation of optical properties of the resist with PAC degradation is taken into account through a permittivity that depends on the absorption coefficient. The boundary condition on the interface between resist and air is defined from the areal image. The EM intensity distribution at the plane of perfect focus is produced by the mask and the imaging system of the photolithography equipment. The conditions on other boundaries are formulated from their reflection and absorption properties. In some special cases the EM field intensity can be calculated using analytical solutions of the Maxwell equations (7).

Baking the resist after exposure is common practice and smooths the PAC concentration, which initially shows wavelike oscillations. This processing step is typically modeled using analytical solutions to the PAC diffusion equation, which is a function of baking temperature and duration. Finally, the propagation velocity of the resist profile should be related to

the postbake PAC concentration. An often-used phenomenological relation is a three-parameter model:

$$Q(\boldsymbol{r}) = \exp[E_1 + E_2 M(\boldsymbol{r}) + E_3 M(\boldsymbol{r})^2] \tag{5}$$

where E_1, E_2, and E_3 are experimental constants of the resist that depend on the developer and on processing conditions. For additional information on Eq. (5) and other phenomenological models for the propagation velocity of the resist see Ref. 8.

Etching and Deposition

The goal of modeling in etching and deposition processes is to determine the velocity by which additional layers of materials are stripped away from or added onto the wafer surface. The schematic representation of the wafer surface and the geometric parameters that are relevant for etching and deposition processes are shown in Fig. 1.

The wafer surface and the equipment source surface above the wafer are defined by position vectors $\boldsymbol{x}$ and $\boldsymbol{x}'$, respectively. ψ is the angle variation in the source ray from the vertical axis and θ is the angle between the surface normal and the source ray. A general expression for the etching or deposition rate at the wafer surface is (9)

$$Q(\boldsymbol{x}) = \int_{-\pi/2}^{\pi/2} f(\psi) g(\theta) V(\boldsymbol{x}, \boldsymbol{x}) d\psi \tag{6}$$

where $f(\psi)$ is the angular flux distribution function of the incident particles while $g(\theta)$ is the surface reaction velocity function. $V(\boldsymbol{x}, \boldsymbol{x}')$ is a visibility function that indicates whether the point $\boldsymbol{x}$ on the surface can be seen ($V = 1$) or not ($V = 0$) from the source point $\boldsymbol{x}'$.

The particular form of the angular functions $f(\psi)$ and $g(\theta)$ depends on various etching and deposition mechanisms and on the equipment. For example, in unidirectional reaction we have $f(\psi) = \delta(\psi)$, where $\delta(\cdot)$ is the Dirac delta function. The corresponding reaction velocity is proportional to the sputtering yield function (10). In the case of an incident neutral-particle component, which is not affected by an electric field, the distribution of incident particles is nonuniform and defined by the hypercosine function $f(\psi) \propto \cos^n \psi$, where the parameter n describes the distribution profile. For the corresponding reaction velocity function, $g(\theta) \propto \cos \theta$ holds if the reaction proceeds into the incident direction.

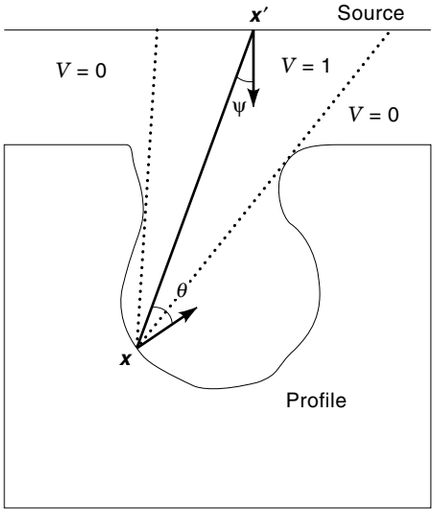

Figure 1. General etching and deposition process geometry and model parameters.

For isotropic chemical reactions the angular distribution of incident particles and the reaction velocity are regarded as uniform and Eq. (6) results in a constant surface propagation velocity Q_0. However, the reaction rate is additionally modified as (11)

$$Q(K) = \frac{2Q_0}{\sqrt{1 + 2KQ_0} + 1} \tag{7}$$

in order to preserve an accurate volume expansion or depletion in consideration of the wafer curvature K.

Ion Implantation

When penetrating through the wafer surface into the solid material, energetic ions lose energy and change their direction by the elastic interaction with the nuclei of the target's lattice atoms. They move on straight paths and lose energy by inelastic interactions with lattice electrons. The elastic nuclear scattering is modeled quite effectively using the two-body binary collision theory from classical mechanics. Let the ion with mass m_1 and kinetic energy E approach an initially stationary target atom with mass m_2. The interaction of the ion and the target atom is defined only by the screening Coulomb potential $V(r)$, being a function of the distance r between them. The "universal" potential proposed in Ref. 12 gives excellent results for a wide variety of atom combinations. The scattering problem is defined by the impact parameter p. It represents the closest distance at which the ion and atom would approach each other if there were no interacting forces. The energy transferred to recoil the target atom in nuclear collision is

$$\Delta E_n = \frac{4Em_1m_2}{m_1 + m_2} \sin^2 \frac{\theta}{2} \tag{8}$$

where θ is the scattering angle in the center-of-mass coordinate system. It is an integral function of E, p, and $V(r)$. A real ion scattering angle γ by which the incoming particle is deflected in the laboratory coordinate system is defined by

$$\cos \gamma = \frac{1 - 0.5(1 + m_2/m_1)\Delta E_n/E}{\sqrt{1 - \Delta E_n/E}} \tag{9}$$

This angle is obtained from the conservation laws of energy and momentum together with results of spherical trigonometry that take into account the three-dimensional nature of atomic collisions. The physics of electronic stopping is quite complex. A widely accepted model assumes that the loss of electronic energy is proportional to the velocity of the ion in analogy to frictional drag forces. This model can be expressed as

$$\Delta E_e = k_0 L \sqrt{E} \tag{10}$$

where L is the distance of travel between nuclear collisions and k_0 is a model parameter that takes into account different pairs of ions and target atoms.

The trajectory of an ion through the target material can be calculated from the formulas for the loss of nuclear and electronic energy and from those for the nuclear scattering angle as discussed above. Meaningful statistical information on the final distribution of implanted ions is obtained by accumulating the results of a large number of hypothetical ion trajectories based on Monte Carlo methods (13,14). In practical implementations of Monte Carlo methods, the scattering angle is commonly given in the form of a look-up table for $\sin^2(\theta/2)$. In amorphous targets L is taken to be the average distance between target atoms and p, normalized by L, is considered to be a random variable. In the case of crystalline targets, both p and L are directly obtained from the lattice geometry. Moreover, in crystalline targets the electron density is much greater close to the lattice sites than in the region between atomic sites. An easy way to model this effect is to introduce an empirically justified dependence of the parameter k_0 on the impact parameter p. The particle model based on Monte Carlo methods can be used also to evaluate the distribution of point defects produced by ion implantation and amorphization of the crystal lattice due to the accumulation of point defects.

An alternative approach to the modeling of ion implantation is to treat energetic ions as a fluid with definite density. The model is described with appropriate statistical distribution functions. They are governed by transport equations based on the total nuclear and electronic differential scattering cross section obtained from the particle ion-implantation model. Let $F(\boldsymbol{P}, \boldsymbol{r})$ be the probability that an ion with momentum $\boldsymbol{P}$ is located at position $\boldsymbol{r}$. The Boltzmann transport equation (BTE), with a scattering term formulated with differential cross section for electronic and nuclear scattering, describes how the distribution function F changes through the target materials. For more details on the application of BTE in ion-implantation modeling see Ref. 15.

An appropriate transport equation can be formulated also for a distribution function $f(d, E, \gamma)$, giving the probability that a particle with energy E stops at a distance d and angle γ compared to its current position and direction. It is a common approach to solve such a transport equation for the moments of the distribution function along certain directions. The moments are typically projected in vertical (initial ion direction) and corresponding lateral direction and given as projection range, standard deviation, skewness, and kurtosis corresponding roughly to depth, width, asymmetry, and flatness of the distribution function, respectively. The results of such calculations are available for various ion-target combinations in Ref. 16. In practice, the moments are used to calibrate parameters of various probability distribution functions. Quite popular are Gaussian, joint half Gaussian, and the family of Pearson distributions. Several analytical models are also proposed to model ion distributions in multilayer structures. Of special interest for multidimensional ion-implantation modeling is the evaluation of the distribution function $f(\boldsymbol{r}, \boldsymbol{x}) = f(|\boldsymbol{r} - \boldsymbol{x}|, E, \gamma)$ for a single ion entering the wafer surface at a point $\boldsymbol{x}$. It is referred to as a point response function. A multidimensional distribution of stopped ions is then obtained by a convolution of the point response function over the wafer surface S:

$$C(\boldsymbol{r}) = N_d \int_S f(\boldsymbol{r}, \boldsymbol{x}) dS \tag{11}$$

where N_d is the implantation dose. In the two-dimensional case, point response functions are commonly modeled as products of appropriate vertical and lateral distribution functions.

Bulk Particle Transport

The most important particle types for bulk processing are dopants, point defects, and chemical reactants used in native film growth processes. Figure 2 schematically shows particle-transport mechanisms for the thermal oxidation process. This thermal oxidation process is based on the oxidant transport through existing oxide layers (17).

The mutual interdependence of thermal oxidation or nitridation and dopant transport is of importance for an accurate process modeling. Namely, during thermal oxidation, the dopants are inherently subject to both diffusion and advective transport in oxide layers due to the oxide flow. The fundamental mechanism for the dopant diffusion transport is the interaction with point-defect particles (vacancies and interstitials) (18). It has become clear that some "abnormal" behavior of dopant diffusion is caused by the nonequilibrium point defects. The nonequilibrium point-defect concentrations are either induced by the diffusion process itself or injected into the substrate by external treatment as oxidation or nitridation, or as a result of ion implantation. The main difficulty for a proper modeling is the formulation of dopant–point-defect pairs for which the quantities have different transport properties.

The transport of the jth particle ($1 < j < N$) is generally governed by the continuity equation

$$\frac{\partial C_j}{\partial t} + \nabla \cdot \boldsymbol{F}_j = R_j \tag{12}$$

where C_j, $\boldsymbol{F}_j$, and R_j are the concentration, flux, and reaction term of the jth particle, respectively, and t is time. The reaction term R_j is used to model various chemical interactions

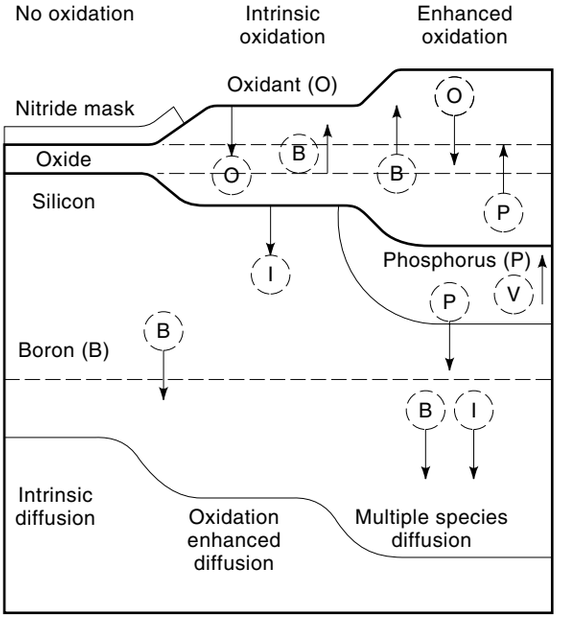

Figure 2. Particle transport in a hypothetical thermal oxidation process. The dashed lines denote initial oxide depth and boron junction depth.

among particles such as dopant clustering, formation, and dissociation of vacancy-interstitial and dopant–point-defect pairs.

The flux of the jth particle is modeled as

$$\boldsymbol{F}_j = -\sum_{i=1}^{N}(D_{ij}\nabla C_i + \mu_{ij}C_i Z_i \nabla\psi) + \mathbf{v}C_j + \boldsymbol{d}_j \tag{13}$$

where D_{ij} and μ_{ij} are diffusivity and mobility matrices, ψ is the built-in electric potential, and Z_j is the charge state of the jth particle. The first term in Eq. (13) accounts for the diffusion transport mechanism while the second term incorporates the drift transport due to the electric field. The diagonal diffusion and mobility terms account for the self-induced transport of the particle and obey the Einstein relationship $D_{jj}/\mu_{jj} = V_T$ with the thermal voltage V_T. The off-diagonal terms take into account corresponding fluxes that are driven by other particles' diffusion and drift. Dopant fluxes, for instance, may be driven by gradients in the point-defect concentration. The third term accounts for convective flux due to the motion of substrate material with velocity $\boldsymbol{v}$. This velocity is important for modeling the particle transport through growing material films, for example, the dopant diffusion in the oxide region during thermal oxidation. The last flux term $\boldsymbol{d}_j$ allows one to incorporate other driving forces that are not directly related to the particle concentration. An example of such driving forces is the gradient of the mechanical potential due to stress generation during processing. Alternatively, the influence of the mechanical stress on the particle transport could be incorporated by stress-dependent transport coefficients. It is common practice to avoid solving the Poisson equation for the electric potential ψ. Instead, it is evaluated from the local charge-neutrality condition. In this case, the effect of the electric field can be entirely included in the diffusion term and affects only the values of D_{ij}. The coefficients of the equations may be functions of processing temperature, particle concentrations, time, and spatial coordinates.

Mechanical Deformation

Stress modeling has been introduced for thermal oxidation processes that consider oxide layers as incompressible fluids at higher processing temperatures (19) and as elastic material at low processing temperatures (20). Elastic materials are generally considered to be those that obey Hooke's law: the shearing stress is proportional to the shearing deformation. In viscous fluids, the shearing stress is linearly related to the rate of shearing deformation. Intermediate states are considered as emerging from elastic or viscous limits towards viscoelastic deformation models.

The slow (creeping) motion of material layers during mechanical transformation is governed by momentum conservation equation

$$\nabla \cdot \hat{\sigma} = \boldsymbol{f} \tag{14}$$

where $\hat{\sigma}$ is the total stress tensor and $\boldsymbol{f}$ is the interior force per unit volume (density of forces). It is convenient to split the total stress tensor into two decoupled components:

$$\hat{\sigma} = -p\hat{I} + \hat{s} \tag{15}$$

where p is the mean pressure, a scalar quantity. $\hat{I}$ is the identity tensor and $\hat{s}$ is the symmetric deviatoric stress tensor. The first term in Eq. (15) represents the dilatation and the isotropic part of the total stress. The change in pressure is proportional to the relative change in the material density. From the mass-continuity equation follows

$$\delta p = -k\nabla \cdot \boldsymbol{u} \tag{16}$$

where k is the modulus of compressibility and $\boldsymbol{u}$ is the displacement vector. For incompressible material ($k \to \infty$) the bounded change of pressure values in Eq. (16) implies the additional condition

$$\nabla \cdot \boldsymbol{u} = 0 \tag{17}$$

This additional equation is required for a complete description of the pressure distribution.

In order to obtain a consistent description of the problem, an additional constitutive relationship for the deviatoric stress tensor is required. It depends on the specific material and the processing conditions considered. In the Maxwell linear viscoelastic model

$$\frac{1}{G}\frac{ds}{dt} + \frac{1}{\mu}s = \nabla\boldsymbol{v} + (\nabla\boldsymbol{v})^t - \frac{2}{3}\hat{I}\nabla \cdot \boldsymbol{v} \tag{18}$$

where G is the modulus of rigidity or shear modulus, μ is the dynamic viscosity coefficient, and $\boldsymbol{v} = d\boldsymbol{u}/dt$ is the velocity of deformation. This is a first-order description of the material properties in the wide range of processing temperature. With $\mu \gg G$ we have the Cauchy-Navier equation for linear elasticity while for $G \gg \mu$ the Stokes fluid equation applies.

DISCRETE MODELS

Front-Propagation Techniques

One of the central concerns in computer implementation of topography models is an accurate and stable technique for tracing the propagation of the topography surface fronts. The initial surface geometry and the surface propagation velocity in its normal direction are sufficient for a complete description of the evolution and of the final state of the wafer surface. A variety of techniques are available for surface propagation in topography modeling. These techniques can be classified by three general categories: string methods, cell-based methods, and level-set methods.

In the string methods, the propagation front is given in a discrete parametrized version. In two dimensions it is approximated by a finite number of points (nodes) joined by straight line segments, while in three dimensions, usually, a nodal triangularization is developed. The surface grid nodes are propagated based on the normal velocity of the front. The surface curvature is evaluated from the discrete surface representation.

In the cell-based methods, a large domain surrounding the propagating front is divided into an array of rectangular cells. Each cell is characterized by a volume fraction number between 0 and 1 that represents the fraction of the material that is contained in each cell. The discrete position of propagating surfaces can be reconstructed, at any time, from these

volume fractions. An example of such a surface reconstruction for a three-dimensional isolation structure is shown in Fig. 3.

Level-set methods (21) implicitly describe the propagation of the surface by the zero level set of the function $\phi(\boldsymbol{r}, t)$:

$$\phi(\boldsymbol{r}, t) = 0 \tag{19}$$

The function $\phi(\boldsymbol{r}, t)$ is obtained as unique solution of the Hamilton-Jacobi type initial-value problem:

$$\frac{\partial \phi(\boldsymbol{r}, t)}{\partial t} + Q(\boldsymbol{r})|\nabla \phi(\boldsymbol{r}, t)| = 0 \tag{20}$$

$$\phi(\boldsymbol{r}, t = 0) = \pm d \tag{21}$$

where $Q(\boldsymbol{r})$ is the propagation velocity of the surface in its normal direction and d is the distance from point $\boldsymbol{r}$ to the surface at $t = 0$ while the plus (minus) sign in Eq. (21) indicates different sides of the surface. The surface velocity Q in Eq. (20) should be defined in the whole space. It is straightforward for photolithography development but requires an appropriate extension of the speed function in etching and deposition processes in which the surface speed function is known only on the surface.

Grid-Generation Techniques

For the numerical solution of partial differential equations that govern process simulation, it is important to satisfy accurately the boundary and the interface conditions. The first and very often used technique in process simulation is to use a coordinate transformation that maps the physical domain (x, y) onto a stationary, usually rectangular, computational domain (ξ, η) (22). However, the usage of closed-form transformation functions has a limited versatility in handling generalized geometries or three-dimensional domains.

The most frequently used transformation techniques are based on the mapping function satisfying a system of Poisson equations

$$\xi_{xx} + \xi_{yy} = P(\xi, \eta) \qquad \text{and} \qquad \eta_{xx} + \eta_{yy} = Q(\xi, \eta)$$

where P and Q are source terms that allow flexible control over the resulting (ξ, η)-coordinate system. The numerical properties of the coordinate system can be adjusted to special needs if higher-order elliptic systems (biharmonic equations) are used as mapping functions. The mapping equations are transformed by interchanging the roles of dependent and independent variables and are solved in a simplified computational space. This implies the transformation of the underlying equations and boundary conditions.

Compared to the numerical grid generators based on elliptic PDEs, the variational techniques offer an even improved control of desired grid characteristics like smoothness, orthogonality, and cell area. The principle of this method is the minimization of a linear combination of integrals that serve as a measure for the different grid characteristics.

In the multilayer process simulation in which the problem area is composed of regions with different physical properties and internal moving boundaries it is fair to perform grid generation in a multiblock or multizone manner. The different physical layers (or zones) can be associated with separate computational domains (blocks). A discussion of multizone grid generation combined with the variational method is contained in Ref. 23.

The properties of the boundary-fitted grids in the interior of the domain can be improved significantly with the multiblock approach of block-structured grids. Such grids are composed of an arbitrary set of subgrids. Each of the subgrid is logically rectangular. To set up a block-structured grid on a given general domain, the domain is subdivided into blocks and a rectangular boundary-fitted grid is generated for each block. The advantages of block-structured grids are manifold. In addition to the great geometrical flexibility their easy use in practice is obvious: being logically rectangular, the description as a Fortran array is possible without using pointers or indirect addressing. Furthermore, the block structure allows very general approaches to refine grids and to use them in the context of multilevel adaptive techniques and on parallel computers. Even grid generation can be done in parallel. The benefit of decomposing the computational domain into single blocks and surrounding them by overlap areas is obvious: for parallel applications in any dimension only lower-dimensional data have to be communicated in order to refresh the values within the overlap area. A reduction of communication can be achieved if a properly chosen order of updating the different blocks is chosen.

A competitive approach is to exploit the intrinsic geometric flexibility of unstructured grids. They can be used for almost any shape of the domain, but they require special techniques based on different Delauny criteria to produce high-quality meshes. Unstructured grids allow a flexible description of arbitrarily shaped domains. In contrast to structured approaches that need in the case of complex computational domains a rather complicated blocking of the domain into blocks, the use of unstructured grids avoids this difficulty. Furthermore, a large amount of highly desired automation has been achieved for this type of grid generation. The price for this is some memory overhead compared to structured grids and a data structure that, in general, is difficult to opti-

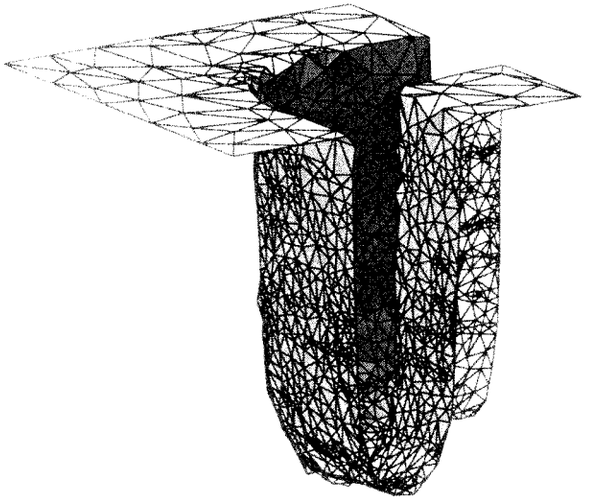

Figure 3. Smoothed and reduced surface triangulation of a 3D isolation structure (trench) containing 3759 triangles. From Ref. 27, ©1996 IEEE. (Source: Institute for Microelectronics, TU Vienna).

mize with respect to certain classes of supercomputer architectures. This has to be considered especially for time-dependent geometries in which the data structure of unstructured meshes may produce difficulties with respect to an efficient implementation. In order to make the handling of the data structure more efficient for two-dimensional domains quadtree-based grid structures have been proposed. Due to predefined templates for geometrical situations the initial mesh of cells that are almost everywhere rectangular is triangulated. One of the main advantages of this grid-generation technique is the ability to conform the grid to a new geometry efficiently.

In contrast to device simulation (24), three-dimensional (3D) process simulation has not yet reached a comparable maturity but several research groups strongly influence this new area (25–27). It can be observed that many of the principles for 2D grid generation carry over to 3D applications. For instance, the Voronoi diagram of polygons in 2D has to be replaced by the Voronoi diagram of polyhedra. Similarly, the Delauny network of triangles becomes a Delauny network of tetrahedra, and the often-used circle criterion generalizes to the sphere criterion. In 3D the numerical complexity of grid-generation algorithms is increasingly important. If a grid-generation method is of order $O(N^2)$, with N being the number of grid points, this may be acceptable in two-dimensional applications, but in 3D the grid-generation scheme should have a better order. Several approaches with a complexity of $O(N \log N)$ have been reported (28,29). The problem of 3D grid generation is the increasing complexity of the data structure. When the grid generation in two dimensions starts from a quadtree-based data structure the same principle in 3D leads to octrees, which require more internal managing of data information. This reflects the increasing number of geometrical possibilities to compose a given 3D body from similar subbodies (octahedrons, tetrahedrons, cubes, prisms, bricks). The complexity further increases if the created octree mesh is triangulated. There is no unique splitting of an octant into tetrahedrons that satisfy the Delauny sphere condition. It can be decomposed into either five or six tetrahedra. If tetrahedra are subdivided by connecting the midpoints of the edges, this leads to four smaller tetrahedra at the corners and to an interior octahedron, which itself can be decomposed into octahedra or into octahedra and tetrahedra. The latter decomposition is an example for a so-called mixed-element decomposition method (30). Looking to the neighborhood relations of the elements shows that the management of the data structure is significantly complicated compared to the 2D case.

In contrast to such formal structured approaches, also in 3D fully unstructured meshes have to be considered. An $O(N \log N)$ algorithm that is based on the advancing-front tetrahedralization technique leads to automatically generated meshes of high quality.

Grid-Adaptation Techniques

One objective of grid adaptation is to produce a defined level of accuracy in a solution with a minimum number of discretization cells. Grid-adaptation techniques exploit the idea of equidistribution, which seeks to distribute some measure of the discrete solution error or at least of the local discretization error (LDE) equally over the grid structure. The LDE τ_h usually is of order $O(h^p)$ with h the local mesh size and p the local order of approximation. In more detail

$$\tau_h(h, p, x, y, z) = O(h(x, y, z)^{p(x,y,z)})$$

As the adaptation of the local discretization order is technically complicated (it requires a locally adapted discretization scheme), the most promising approach is to refine (or coarsen) locally an existing grid by adapting the local mesh size. For process simulation, the grid-adaptation procedure is significantly complicated due to the time-dependent solutions and geometries. A fairly challenging question in any grid-adaptation method is the formulation of reliable discrete error estimators and local refinement criteria. This holds independent of the chosen mesh.

Although the LDE is often considered as the most natural monitor function for the formulation of refinement criteria, its use is not always effective (31). Additional problem-dependent criteria are introduced. Quantities such as the magnitude of local gradients used as refinement indicators may work in some cases. In general, such a criterion is not applicable. A number of error equidistribution strategies for the purpose of grid adaptation of tensor-product grids in the solution of parabolic PDEs are reviewed and developed in Ref. 32.

One class of estimators is motivated by the relation between the error $e_h = (u)_h - w_h$ [on a given mesh with discretization parameter h, $(u)_h$ is the solution of the continuous problem $Lu = f$ restricted to the mesh, and w_h a discrete approximation to it] and the discrete residual $r_h = f_h - L_h w_h$. The so-called residual equation $L_h e_h = r_h$ for linear operators can be derived. Assuming an existing inverse operator L_h^{-1} and taking norms yield the inequality

$$\|e_h\| \leq \|L_h^{-1}\| \cdot \|r_h\|.$$

Together with a stable discretization this implies that a refinement that reduces the residual will also reduce the error. This principle can be used on any subdomain, especially on each grid cell to construct an error indicator.

For FE grid adaptation an approximation of the global discretization error can be obtained by a method that is comparable to defect correction techniques (33) in computational fluid dynamics: the discrete problem is solved twice, once with a first-order approximation using linear elements and a second time using second-order elements. The difference between the two discrete solutions is used as an approximation to the global discretization error. The disadvantage of such a method is the second solution of the additional discrete problem. The numerical work for doing so may rule out the possibility of accelerating the solution process with the help of grid adaptation. An alternative approach (34) is based on solving the additional higher-order problem only locally in each grid element. This is a consequent exploitation of the fact just mentioned that the discretization error also depends on the local order of approximation.

Another idea that is of the same quality originates in multilevel adaptive techniques (35). Exploiting the natural grid hierarchy it can be shown with the help of asymptotic expansions that the difference between the coarse-grid operator applied to the restricted fine-grid approximation and the fine-grid operator applied to the fine-grid solution, evaluated on

the coarse grid,

$$L_H \hat{I}_h^H u_h - I_h^H L_h u_h$$

defines the relative local discretization error. This quantity is cheaply computed on the coarse H grid and approximates the discretization error with respect to the h grid up to higher-order terms. This H grid quantity indicates where to refine the h grid by a $h/2$ refinement.

Discretization Schemes

As long as structured grids are used both finite-difference and finite-volume discretizations have been applied. Unstructured meshes have been the natural basis for finite-element and finite-volume approaches. The discretization parameter mesh size has to be adjusted in such a way that required accuracy conditions are satisfied. Of course, an obvious rule for an efficient simulation is not to use more points than needed for the desired accuracy. According to this, discretization schemes that are adapted to particular features of the solution represent a powerful technique to reduce the number of grid points.

The standard discretization methods for the diffusion process simulation assume a linear variation of concentrations between adjacent grid nodes. This assumption may be quite inaccurate on coarse grids. Sophisticated FD schemes that exploit the exponential flux behavior (36,37) have lead to a significant reduction of grid points. A remarkable analogy to the Scharfetter-Gummel scheme (3) in device simulation can be observed.

Advanced discretization schemes also can be formulated directly in the physical domain near nonplanar moving boundaries and interfaces as an alternative to boundary-fitted grids. The main idea of these immersed interface discretization methods is to resolve the problem of nonplanar and moving boundaries and interfaces in the physical domain without spoiling the grid regularity. The discrete equations from FD and FV methods or the elements in FE methods near the interfaces involve grid points from either side of the interface. Because various quantities (concentrations or some of their derivatives) may be discontinuous across the interface, the standard discretization schemes would lead to poor results. The global advantage of using fixed and regular grid structures is offset locally in the development of accurate discretization schemes near interfaces: depending on the desired accuracy the immersed boundary discretization techniques use five (38) or six neighboring points (35), respectively, instead of four of them for the formulation of the discrete equation.

Solving Techniques

After linearization of the mostly nonlinear problems using Newton methods or variants of it (39), linear systems of algebraic equations have to be solved. Only in special situations can the direct solver be applied. Solving the linear algebraic systems iteratively is the standard way. For this purpose, classical iterative methods like the Jacobi, Gauss-Seidel, and weighted-variant techniques [successive overrelaxation (SOR), derived from the Gauss-Seidel method] are used. Because of their limited convergence, which often behaves as $O(1-h)$ or $O(1-h^2)$, better convergent methods are requested.

As the convergence depends on the spectrum of the operator matrix, so-called preconditioners are introduced to transform the spectrum away from one. A properly chosen preconditioner improves the convergence of the iterative method in such a way that the extra work pays off. In combination with preconditioners, conjugate-gradient (CG) methods and the generalized minimal residual (GMRES) or biconjugate-gradient methods (BiCG, BiCGStab) are used to produce robust solvers (40) within the whole range of parameters of the application.

Although the approaches mentioned previously have reached a high level of sophistication that guarantees good convergence and robustness such as those in the black-box solver, their numerical complexity is not optimal. Especially with respect to 3D simulations the search for optimal, general, and robust methods that also incorporate a natural parallelism goes on.

The multigrid method, an algorithm that possesses these properties, has shown its potential for process simulation (35,41). The application of the geometric multigrid requires a hierarchy of grids. The applicability of this type of multigrid becomes more and more limited the more unstructured features the grids possess. The idea of the hierarchical basis multigrid (42) together with algorithms used to refine and coarsen unstructured meshes (43,44) overcomes these problems. A second multigrid approach that is completely independent of the underlying mesh is the so-called algebraic multigrid (AMG) (45,46). The idea of the AMG is to construct a sequence of smaller and smaller algebraic subproblems from the original one. The coarsening criterion only depends on the algebraic coupling of different equations of the system. AMG can be parallelized.

PROCESS MODELING AND SIMULATION TOOLS

Practical computer implementation of semiconductor process modeling requires appropriate software tools. These tools provide an environment to analyze the validity of physical and discrete models or to simulate particular fabrication steps or process flow sequences. Historically, there was a trend towards developing comprehensive, stand-alone process simulators. They can be principally classified by the space dimension (1D, 2D, or 3D) that is used to describe the geometry and the relevant physics. Moreover, the topography and bulk process modeling is commonly implemented in separate programs. At the forefront of topography process simulation are programs such as SAMPLE and PROLITH, compared in Ref. 8, and SPEEDIE (11). Perhaps the most widely used process simulation programs for bulk process modeling belongs to the SUPREM family developed at Stanford University. The most popular are 1D version SUPREM III (47) and 2D version SUPREM IV (48). However, there are some programs such as COMPOSITE (10), which combines both topography and bulk process modeling. In the model development and verification phase, it is important to have convenient input for an incorporation of new models. To this end, one can use robust general-purpose PDE solvers such as the program environment L_iSS (41) or process simulation programs like PROMIS (49) with well-defined interfaces for a fast model evaluation. These latter programs have been used to produce the simulation re-

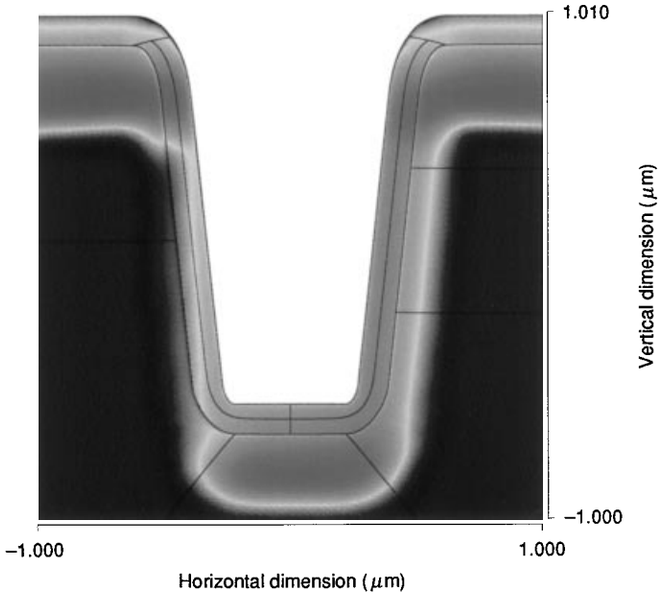

Figure 4. Boron profile obtained by Monte Carlo simulation of ion implantation using PROMIS and transferred to a block-structured grid. The tilted ion-implantation process is performed in the silicon trench structure, which is covered by a pad oxide layer.

sults for coupled diffusion and thermal oxidation processes (50) that are shown in Figs. 4 and 5.

In order to obtain acceptable yield and reliability from fabrication processes and hence their economic viability, it is imperative that the statistical variation in geometric and material properties of the fabricated wafers and devices has to be kept to a minimum (51). The process disturbances could be caused by fluctuations in the environment that surrounds the wafer within the fabrication equipment, by variations of the chemical compounds and material used in the manufacturing process, and by defects present in the semiconductor substrate or lithography-related disturbances. Statistical process modeling and simulation tools provide a cost-effective way for identifying how the actual yield of a process depends on the fluctuation in process parameters. While deterministic process simulation tools consider process parameters as physical constants at a given processing temperature, in statistical process simulators they are treated as random variables. The random variables are defined by the mean values that are used in deterministic simulators and variances that measure the spread around mean parameter values caused by inherent process instabilities. For the sake of statistical process simulation a set of hierarchical random-number generators are used to produce input process parameters. One method of obtaining the desired statistical distributions that are required for the yield prediction is to run deterministic process simulators like SUPREM III many times for different input parameters. However, this procedure is time-consuming even for 1D process simulation. In order to overcome this deficiency, a statistical process simulator FABRICS (52), developed at Carnegie-Mellon University, mostly uses analytical models or discrete models based on efficient numerical approaches.

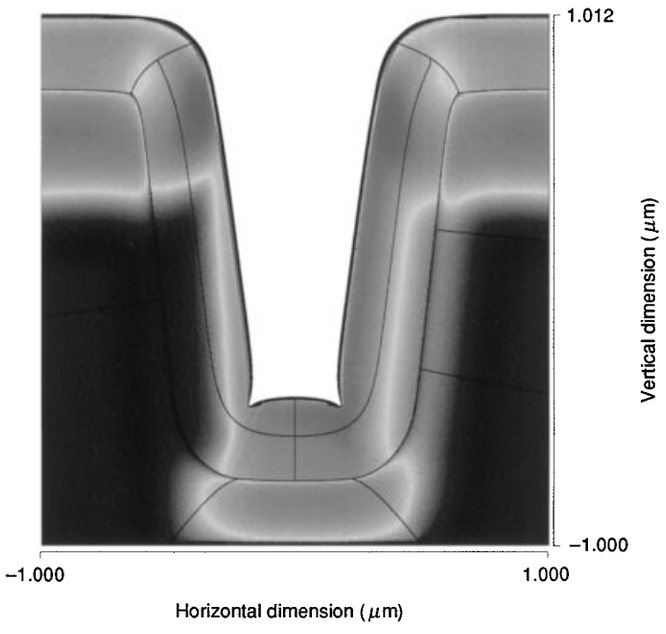

Figure 5. Boron profile and the oxide shape at the end of the thermal oxidation process of the trench structure obtained by L_iSS.

With the moving of individual process simulation tools from the predominantly academic research and development groups to production-oriented environments, it is recognized that the integration and maintenance of such codes becomes difficult and requires a significant level of user experience. This fact has motivated the rapid development of *frameworks.* A framework is a software environment supporting the use of multiple simulators while working independently of any particular simulator. Frameworks provide convenient data transfer between different simulators, a uniform user interface, comprehensive optimization capabilities, visualization, and well-defined procedures for adding new tools. An important motivation for frameworks is standardization. Most current frameworks exploit tool-independent data representation and hardware-independent software standards. Discussions among developers from industry and academia have led to the definition and acceptance of the standard process interchange format (PIF). A detailed description of the most important and frequently used frameworks is given in Ref. 53.

STATE OF TECHNOLOGY AND FUTURE TRENDS

Over a period of more then two decades the field of process modeling has become an essential enabling technology in semiconductor industry. Although impressive progress in the development of process modeling has been achieved, there is much more potential to be exploited. The principal deficiency is the lack of predictive capabilities. Historically, process modeling has lagged behind the needs of leading process development by one process generation. The improved models, required for a new technology, usually are not available before the technology itself can be processed and is more or less under control. The demands facing process modeling are the increasing complexity of processes, the variety of materials, and the multitude of techniques and concepts. Many physical and geometrical effects considered to be negligible on a larger

scale become first-order effects on a smaller one. It is evident that the progress of process modeling has to accelerate in the future; also the application of process simulation should be more effective than at present.

Process modeling has to provide general concepts, guidance, and insights at a very early stage of process or technology development for the engineers. The models based on first-order approaches coupled with computer modeling capability in the critical stage between invention and application are crucial for the semiconductor industry. Particularly, as the most important needs for future process modeling, the Semiconductor Industry Association Roadmap priorities are (1) automatic grid-generation and -adaptation algorithms, (2) defect-mediated dopant profile evolution, (3) combined equipment and feature-scale topography models, (4) 2D and 3D doping profile measurement tools, (5) etch model predictability, and (6) silicidation models. Great effort is directed today towards the development of 3D process simulation tools.

Defect-based dopant models for implantation, diffusion, and activation must start with underlying first-principle calculation and characterization methods. Modeling of photolithography exposure, mechanical deformation, and bulk-particle-transport processes presumably will have to deal with physical models based on PDEs and corresponding solution methods. But models of atomic level and hence Monte Carlo simulation algorithms will become increasingly important. Ion-implantation modeling is an area for which it is likely in the near future to convert entirely to Monte Carlo–based calculations. Monte Carlo methods are inherently three-dimensional. They work effectively for arbitrary multilayer target structures and can also provide reliable information on produced point-defect distributions.

For modern semiconductor technologies, interconnections have become at least as important as the active semiconductor devices for the determination of the overall chip performance. Interconnect technology includes dielectric and metal-film formation as well as the etch process. The accurate evaluation of the process variations and their effects on the performance and on the reliability of interconnects essentially depends on the integration of equipment and feature-scale topography modeling of deposition, lithography, and etching. This also includes a critical need for an improved physical modeling of topography processes. Photolithography continues to be the mainstream processing technology for pattern definition and transfer. Special attention has to be given to the modeling of electromagnetic problems in photolithography. This is particularly computationally intensive because typical feature sizes of interest are on the order of the one wavelength. The formulation of predictive models for deposition and etching is also essential for the interconnect modeling. These models are expected to have more variations than lithography or bulk processes and thus need improved statistical analysis methods and tools.

The lack of accurate experimental verification is an important obstacle for process model development and model calibration that should be overcome in future. For example, the measurement techniques used for the investigation of doping profiles are not very accurate and most of them are inherently 1D. The problem is even more emphasized with damage distributions that are induced by implantation and their evolution during subsequent annealing processes. This phenomenon cannot be measured directly and is only verified indirectly by its effect on dopant distributions. It is evident that a better understanding of the physics of the bulk particle transport increasingly demands further improvements in metrology. This limitation in measurement technology severely hampers the development of accurate multidimensional process modeling tools.

The trend towards 3D with more complex models, leading to larger systems of coupled PDEs, to more complex topologies, and to multilayer structures, is obvious (25). This requires computing power as provided in an ideal way by scalable parallel architectures. Therefore, parallelization is an innovative technique that can be used for new algorithmic developments. The first steps will be made on shared-memory machines by a straightforward loop parallelization of initially sequential programs. A typical approach to parallelize grid-oriented PDE applications for large-scale parallelization is grid partitioning (54). This technique is essentially independent of the particular partial differential equation or system to be solved. The overlap area for each subgrid is updated within the solution process. This reduces the amount of communication. For an efficient parallelization, however, load balancing and locality have to be taken into account. A satisfactory load balancing presupposes that all processors are responsible for approximately the same number of discrete equations and variables. This requirement can be better satisfied the more regular the data structures are. Additionally, for low communication cost the algorithm should offer a large amount of locality.

Many improvements both on the physical and on the discrete approximation level can be expected in the near future. The combination of these improvements requires flexible and reliable software. The next-generation process simulation tools have to be designed to be modular in such a way that innovative models or algorithms can easily be added. Recent advances in object-oriented software engineering seem to be the natural development framework for process modeling. The object-oriented programming approach significantly simplifies the tool development by providing a simple and unified access mechanism to objects that represent wafer and device structure without going into details of the data structures used. This approach also provides the possibility for code structuring that may allow an active participation of a large community in the development of widely used software packages. To realize the idea of virtual factories by new TCAD tools it is necessary to look back to developments of the past. The complexity of model development, automatic grid generation, adaptive meshing, regridding of time-dependent domains, search for optimal solvers, parallel programming, pre- and postprocessing of single simulation steps, and approximately complete simulation of processing steps poses new challenges to the developers of software tools. Especially the software development for process simulation on parallel machines has to exploit the experience of other disciplines in which parallelization has a long tradition. Apart from the need of portability with respect to parallel programming, a definite must is to separate modeling, discrete description, and solving from one other. Such a concept of keeping the formulation of the application or discretization away from the particular solver has been used for a parallel programming environment (41). This idea, which definitely represents the approach of the future, is used to develop an object-oriented PDE solver for TCAD applications (55). There the lessons to

be learned on the way to the next-generation simulation software are described.

BIBLIOGRAPHY

1. S. M. Sze (ed.), *VLSI Technology,* New York: McGraw-Hill, 1983.
2. M. Meyyappan (ed.), *Computational Modeling in Semiconductor Processing,* Boston: Artech House, 1995.
3. S. Selberherr, *Analysis and Simulation of Semiconductor Devices,* Wien: Springer-Verlag, 1984.
4. R. W. Dutton and Z. Yu, *Technology CAD—Computer Simulation of IC Processes and Devices,* Boston: Kluwer Academic Publishers, 1993.
5. G. F. Carey et al., *Circuit, Device and Process Simulation, Mathematical and Numerical Aspects,* Chichester: Wiley, 1996.
6. F. H. Dill, Optical lithography. *IEEE Trans. Electron Devices,* **ED-22**: 440–444, 1975.
7. C. A. Mack, Analytical expression for the standing wave intensity in photoresist, *Appl. Opt.,* **25**: 1958–1961, 1986.
8. O. D. Crisalle et al., A comparison of the optical protection lithography simulators in SAMPLE and PROLITH, *IEEE Trans. Semicond. Manuf.,* **5**: 14–26, 1992.
9. S. Tazawa, S. Matsuo, and K. Saito, A general characterization and simulation method for deposition and etching technology, *IEEE Trans. Semicond. Manuf.,* **5**: 27–33, 1992.
10. J. Pelka, K. P. Müller, and H. Mader, Simulation of dry etch processes by COMPOSITE, *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.,* **7**: 154–159, 1988.
11. J. McVittie et al., *SPEEDIE: User's manual,* Stanford, CA: Stanford University, 1995.
12. J. F. Ziegler, J. Biersack, and U. Littmark, *The Stopping and Ranges of Ions in Solids,* New York: Pergamon, 1985, Vol. 1.
13. G. Hobler and S. Selberherr, Monte Carlo simulation of ion implantation into two- and three-dimensional structures, *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.,* **8**: 450–459, 1989.
14. W. Bohmayr et al., Trajectory split method for Monte Carlo simulation of ion implantation, *IEEE Trans. Semicond. Manuf.,* **8**: 402–407, 1995.
15. M. D. Giles, Ion implantation calculations in two dimensions using the Boltzmann transport equation, *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.,* **CAD-5**: 679–684, 1986.
16. J. F. Gibbons, W. S. Johnson, and S. W. Mylroie, *Projected Range Statistics,* Stroudsburg, PA: Dowden, Hutchinson, and Ross, 1975.
17. B. E. Deal and A. S. Grove, General relationship for the thermal oxidation of silicon, *J. Appl. Phys.,* **36**: 3770–3778, 1965.
18. R. B. Fair, Physics and chemistry of impurity diffusion and oxidation of silicon. In D. Kahng (ed.), *Silicon Integrated Circuits, Part B.* New York: Academic Press, 1981.
19. D. Chin et al., Two-dimensional oxidation, *IEEE Trans. Electron Devices,* **ED-30**: 744–749, 1983.
20. A. Poncet, Finite-element simulation of local oxidation of silicon, *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.,* **CAD-4**: 41–53, 1985.
21. J. A. Sethian, *Level Set Methods: Evolving Interfaces in Geometry, Fluid Mechanics, Computer Vision, and Material Science,* Cambridge, UK: Cambridge University Press, 1996.
22. K. Wimmer et al., Transformation methods for nonplanar process simulation, in W. Fichtner and D. Aemmer (eds.), *Simulation of Semiconductor Devices and Processes,* Konstanz, Germany: Hartung-Gorre Verlag, 1991, Vol. 4, pp. 131–137.
23. M. K. Moallemi and H. Zhang, A general numerical procedure for multilayer multistep IC process simulation, *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.,* **13**: 1379–1390, 1994.
24. G. Garreton et al., Unified grid generation and adaptation for device simulation, in H. Ryssel and P. Pichler (eds.), *Simulation of Semiconductor Devices and Processes,* Wien: Springer-Verlag, 1995, Vol. 6.
25. J. Lorenz (ed.), *3-Dimensional Process Simulation,* Wien: Springer-Verlag, 1995.
26. T. Chen, D. W. Yergeau, and R. W. Dutton, Efficient 3D mesh adaptation in diffusion simulation, in *Proc. Int. Conf. Simul. Semicond. Process Devices,* Tokyo, 1996, IEEE Cat. No 96TH8095, 1996.
27. P. Fleischmann et al., Grid generation for three-dimensional process and device simulation, in *Proc. Int. Conf. Simul. Semicond. Process. Devices,* Tokyo, 1996, IEEE Cat. No 96TH8095, 1996.
28. S. Halama, The viennese integrated system for technology CAD applications architecture and critical software components, PhD. thesis, Technical University, Vienna, Austria, 1994.
29. P. Fleischmann and S. Selberherr, A new approach to fully unstructured three-dimensional Delauny mesh generation with improved element quality, in *Proc. Int. Conf. Simul. Semicond. Process. Devices,* Tokyo, 1996, IEEE Cat. No 96TH8095, 1996.
30. E. Leitner and S. Selberherr, Three-dimensional grid adaptation using a mixed-element decomposition method, in H. Ryssel and P. Pichler (eds.), *Simulation of Semiconductor Devices and Processes,* Wien: Springer-Verlag, 1995, Vol. 6.
31. P. Pichler et al., Simulation of critical IC fabrication steps, *IEEE Trans. Electron Devices,* **32**: 1940–1953, 1985.
32. K. Chen, Error equidistribution and mesh adaptation, *Soc. Ind. Appl. Math. J. Sci. Comput.,* **15**: 798–818, 1994.
33. R. Ismail and G. Amaratunga, Adaptive meshing schemes for simulating doping diffusion, *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.,* **9**: 276–289,1990.
34. R. E. Bank and A. Weiser, Some a posteriori error estimators for elliptic partial differential equations, *Math. Comput.,* **44**: 283–301, 1985.
35. W. Joppich and S. Mijalković, *Multigrid Methods for Process Simulation,* Wien: Springer-Verlag, 1993.
36. C. C. Lin, M. E. Law, and R. E. Lowther, Automatic grid refinement and higher order flux discretization for diffusion modeling, *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.,* **12**: 1209–1216, 1993.
37. S. Mijalković, Exponentially fitted discretization schemes for process simulation on coarse grids, *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.,* **15**: 484–492, 1996.
38. R. J. Leveque and Z. Li, The immersed interface method for elliptic equations with discontinuous coefficients and singular sources, *Soc. Ind. Appl. Math. J. Numer. Anal.,* **31**: 1019–1044, 1994.
39. R. E. Bank and D. J. Rose, Global approximate Newton methods, *Numer. Math.,* **37**: 279–295, 1981.
40. R. Barrett et al., *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods,* Philadelphia: SIAM, 1994.
41. M. G. Hackenberg et al., Simulation of thermal oxidation and diffusion processes by the parallel PDE solver L_iSS, in *Proc. Int. Conf. Simul. Semicond. Process. Devices,* Tokyo, 1996, IEEE Cat. No 96TH8095, 1996, also GMD-Arbeitspapier 1039, Sankt Augustin, December 1996.
42. R. E. Bank, T. F. Dupont, and H. Yserentant, The hierarchical basis multigrid method, *Numer. Math.,* **52**: 427–458, 1988.
43. R. E. Bank and J. Xu, An algorithm for coarsening unstructured meshes, *Numer. Math.,* **73**: 1–36, 1996.

44. R. E. Bank, *PLTMG: A Software Package for Solving Elliptic Partial Differential Equations. User's Guide 7.0,* Philadelphia: SIAM, 1994.
45. A. Brandt, Algebraic multigrid theory: The symmetric case, in *Proc. Int. Multigrid Conf.,* Copper Mountain, CO, 1983.
46. J. W. Ruge and K. Stüben, Algebraic multigrid (AMG). In S. F. McCormick (ed.), *Multigrid Methods,* Vol. 5 of Frontiers in Applied Mathematics, Philadelphia: SIAM, 1986.
47. C. P. Ho et al., VLSI process modeling—SUPREM III, *IEEE Trans. Electron Devices,* **30**: 1438–1453, 1983.
48. M. E. Law and R. W. Dutton, Verification of analytic point defect models using SUPREM-IV, *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.,* **7**: 181–190, 1988.
49. G. Hobler, P. Pichler, and K. Wimmer, *PROMIS 1.6: User's Guide,* Tech. Rep., Vienna, Austria: Technical University, 1991.
50. M. G. Hackenberg et al., Coupled simulation of oxidation and diffusion in VLSI wafer fabrication. In A. Sydow (ed.), *Proceedings of the 15th World Congress on Scientific Computing, Modelling and Applied Mathematics—IMACS,* Berlin: Wissenschaft und Technik Verlag, 1997, Vol. 3, pp. 587–592.
51. S. W. Director, W. Maly, and A. J. Strojwas, *VLSI Design for Manufacturing: Yield Enhancement,* Boston: Kluwer Academic Publishers, 1990.
52. S. R. Nassif, A. J. Strojwas, and S. W. Director, FABRICS II: A statistically based IC fabrication process simulator, *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.,* **CAD-3**: 40–46, 1984.
53. F. Fasching, S. Halama, and S. Selberherr (eds.), *Technology CAD Systems,* Wien: Springer-Verlag, 1993.
54. O. A. McBryan et al., Multigrid methods on parallel computers—a survey of recent developments. *Impact Comput. Sci. Eng.,* **3**: pp. 1–75, 1991.
55. D. W. Yergeau, R. W. Dutton, and R. J. G. Goossens, A general OO-PDE solver for TCAD applications. Paper presented at 2nd Annu. Object-Oriented Numer. Conf., Sunriver OR, 1994.

WOLFGANG JOPPICH
German National Research Center
for Information Technology

SLOBODAN MIJALKOVIĆ
University of Niš

SEMICONDUCTOR PROCESS MODELING.

See NEURAL NETS FOR SEMICONDUCTOR MANUFACTURING.