THE BELL SYSTEM TECHNICAL JOURNAL

DEVOTED TO THE SCIENTIFIC AND ENGINEERING ASPECTS OF ELECTRICAL COMMUNICATION

Volume 59 January 1980 Number 1

Two-Dimensional Concentration Dependent Diffusion

By D. D. WARNER and C. L. WILSON

(Manuscript received July 26, 1979)

The concentration dependent diffusion problem is studied in two dimensions for a variety of cases relevant to the fabrication of narrow-channel MOS transistors. A method for solving the partial differential equation is presented which allowed the solutions to be computed on an available minicomputer. This method enables analytic transformations to be exploited which improve the accuracy of the solutions and increase the speed of computation. The simulations demonstrate that there are three principal nonlinear effects: (i) a large translation of the diffusion front, (ii) a marked steepening of the front itself, and (iii) a very noticeable decrease in the ratio of the lateral to vertical diffusions. The ratio of the lateral to vertical diffusion as predicted by the model is compared with physical values experimentally determined using a scanning electron microscope.

I. INTRODUCTION

In 1965, Kennedy and O'Brien analytically investigated the impurity atom distribution near the edge of a diffusion mask.[1] The geometric effects are significant, and the results of their study are routinely applied in semiconductor device design.[2] The Kennedy and O'Brien model used the simple, linear, constant-coefficient, diffusion equation. In 1968, Hu and Schmidt introduced a concentration dependent diffusion model which included the effects of both the charged vacancy reaction and the impurity-induced electric fields. This model was investigated in one dimension, and it established that, in the regime in which semiconductors are actually fabricated, these nonlinear effects are quite significant for some impurities, such as arsenic.

In the decade since these studies, there have been significant improvements in fabrication technologies and lithographic techniques for both bipolar and MOS transistors. Furthermore, the importance of MOS transistors has also greatly increased. In this paper, we investigate the combination of nonlinear and geometric effects on the impurity atom distribution by studying a two-dimensional, concentration dependent diffusion model in a region containing the edge of the diffusion mask.

In Section II, the nonlinear diffusion coefficient, including the effects of "autodoping," electric fields, and multiply ionized impurities, is derived in a form suitable for the calculation of the desired impurity profiles. The modeling parameters, the geometries, and the boundary conditions for the partial differential equations are also described. Section III outlines the numerical method used to solve the model. Additional details of the method are presented in the appendix. Several combinations of initial and boundary conditions are analyzed in this paper. Section IV analyzes the case where the impurity concentration in the window is held at a constant value throughout the diffusion, and Section V analyzes the case where a fixed quantity of impurity has been implanted and then diffused. This section also presents the simulation which is to be compared with experimental results. The simulations demonstrate that there are three principal nonlinear effects: (*i*) a large translation of the diffusion front, (*ii*) a marked steepening of the front itself, and (*iii*) a very noticeable decrease in the ratio of the lateral to vertical diffusions.

To confirm the applicability of the physical model, a numerical solution of determined accuracy is compared with experimental measurements. Section VI contains a careful assessment of the accuracy of the numerical solution, while Section VII describes the experimental measurements of the lateral to vertical diffusion ratio. This experiment employed a scanning electron microscope (SEM), and the resolution obtained is superior to that previously obtained by optical measurements. These measurements are compared with the predictions calculated in Section V. The comparison corroborates the validity of the concentration dependent diffusion model for arsenic and demonstrates its relevance for contemporary processing technology. Section VIII briefly discusses the application of the concentration dependent diffusion model to general process simulation.

II. THE NONLINEAR, CONCENTRATION DEPENDENT DIFFUSION MODEL

The diffusion can be modeled, as in Hu and Schmidt,[3] with a concentration dependent diffusion coefficient which incorporates the charged vacancy reaction and the self-induced electric field. For a noninteracting impurity, the transport of the impurity atoms in an

isotropic medium is described by a continuity equation for the impurity current,

$$\frac{\partial N}{\partial t} = -\nabla \cdot (\mathbf{J}), \tag{1}$$

where N is the impurity concentration and $\mathbf{J}$ is the impurity current density. The current density is determined by diffusion and drift terms of the form

$$\mathbf{J} = -D_N \nabla N + Z\mu_N N\mathbf{E}, \tag{2}$$

where D_N is the diffusion coefficient, Z is the impurity charge, μ_N is the mobility, and $\mathbf{E}$ is the electric field.

The self-induced electric field term can be approximated from the concentration, N, by assuming local space charge neutrality,

$$n - p - ZN = 0. \tag{3}$$

Under these conditions, with the potential, ψ, referred to the center of the band gap, solving

$$n_i \exp\left(\frac{q\psi}{kT}\right) - n_i \exp\left(-\frac{q\psi}{kT}\right) - ZN = 2n_i \sinh\left(\frac{q\psi}{kT}\right) - ZN = 0$$

for ψ yields

$$\psi = \frac{kT}{q} \sinh^{-1}(ZN/2n_i),$$

where n_i is the intrinsic carrier concentration. Hence,

$$\mathbf{E} = -\nabla\psi = -\frac{kT}{q} \frac{Z\nabla N}{2n_i\sqrt{(ZN/2n_i)^2 + 1}}. \tag{4}$$

Using Einstein's relation and this expression for $\mathbf{E}$, (2) becomes

$$\mathbf{J} = -D_N\left(1 + \frac{Z^2 N}{2n_i\sqrt{(ZN/2n_i)^2 + 1}}\right)\nabla N. \tag{5}$$

Following Ref. 3, we combine the electric field term in eq. (2) with the diffusion coefficient to obtain

$$\frac{\partial \alpha}{\partial t} = \nabla \cdot (D_N h(\alpha) \nabla \alpha), \tag{6}$$

where

$$h(\alpha) = 1 + \frac{Z^2\alpha}{\sqrt{(Z\alpha)^2 + 1}} \tag{7}$$

and

$$\alpha = N/2n_i. \tag{8}$$

Finally, we incorporate the following formulation of D_N which was derived in Ref. 4 to account for the charged vacancy reaction,

$$D_N = D_0 \frac{(1 + \beta f)}{(1 + \beta)}. \tag{9}$$

Here D_0 is the phenomenological low concentration diffusion constant, f is the ratio of the electron or hole concentration to the intrinsic carrier concentration at the diffusion temperature ($f = n/n_i$ for donors and $f = p/n_i$ for acceptors), and β is a phenomenological coefficient discussed in detail in Ref. 4. By combining the equilibrium condition, $np = n_i^2$, with the assumption of local space charge neutrality, (3), we obtain the following approximation for f in terms of the concentration

$$f = |Z|\alpha + \sqrt{(|Z|\alpha)^2 + 1}. \tag{10}$$

Existing models of the type presented in this section have a long history of experimental validation. In his review paper,[5] Fair discusses in detail the cases of phosphorus, arsenic, and boron diffusion in silicon. In Ref. 3, Hu and Schmidt suggested that the correct values of β were $\beta = 100$ for donors and $\beta = 0.01$ for acceptors. They also presented experimental evidence which tended to support these values. Subsequent experimental work, which incorporates more accurate measurements at high concentrations, has continued to corroborate $\beta = 100$ for arsenic,[6,7] but has shown that for boron $\beta = 19$ is the correct value.[6,8,9] For impurities such as phosphorus, which have multiple vacancy reactions, this model is inappropriate. A model for phosphorus is discussed in detail in Ref. 10 and is beyond the scope of the present work.

The assumptions upon which the present theoretical model is based are given in Ref. 4. Briefly, these assumptions are: (*i*) the density of the charged vacancies is much lower than the density of the impurity atoms, (*ii*) the charged vacancy population is in a state of quasi-equilibrium determined by the local equilibrium between the impurity atoms and the charged vacancies, and (*iii*) the vacancy-impurity interactions are never more complex than pair interactions. These assumptions have been confirmed in the one-dimensional case by Fair and Tsai.[7]

In the present work, which is restricted to singly ionized impurities,

$$Z = -1 \quad \text{and} \quad \beta = 19 \quad \text{for acceptors,}$$

while

$$Z = 1 \quad \text{and} \quad \beta = 10^2 \quad \text{for donors.}$$

Thus we want to solve

$$\frac{\partial \alpha}{\partial t} = \nabla \cdot (D(\alpha) \nabla \alpha), \tag{11}$$

where the concentration dependent diffusion coefficient, $D(\alpha)$, is defined by eqs. (7) through (10), i.e.,

$$D(\alpha) = D_0 \frac{1 + \beta[\alpha + \sqrt{\alpha^2 + 1}]}{1 + \beta} \left[1 + \frac{\alpha}{\sqrt{\alpha^2 + 1}} \right]. \tag{12}$$

Figures 1 and 2 present plots of $D(\alpha)/D_0$ versus α for boron and arsenic respectively. Each figure displays the normalized diffusion coefficient for three ranges of α: 0.0 to 0.2, 0.0 to 2.0, and 0.0 to 20.

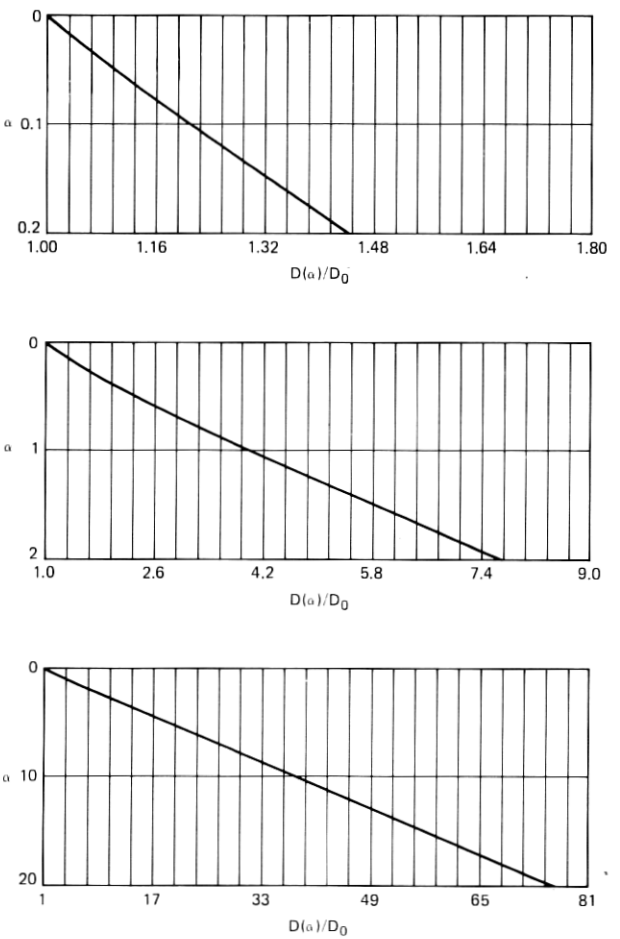

Fig. 1—The nonlinear diffusion coefficient for acceptor impurities.

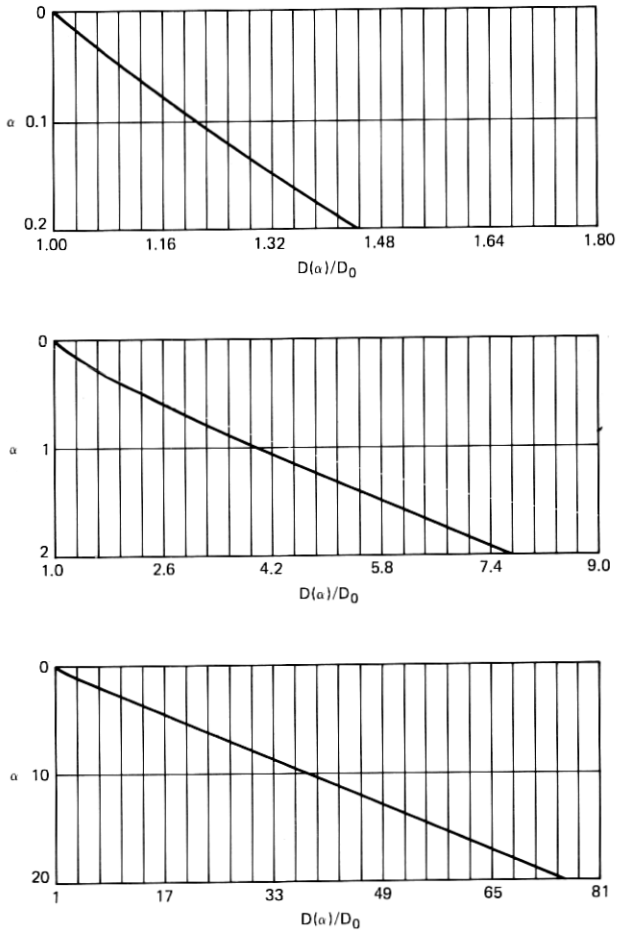

Fig. 2—The nonlinear diffusion coefficient for donor impurities.

Although the behavior of the diffusion coefficient is mildly nonlinear for small values of α, it quickly approaches a straight line for α larger than 2. The limiting behavior of $D(\alpha)/D_0$ for large α is $3.80\alpha + 0.100$ for boron and $3.96\alpha + 0.0198$ for arsenic. This nearly linear behavior is consistent with the experimental results summarized in Ref. 9. It should be noted, however, that the approximations for both $h(\alpha)$ and f assume uncompensated material, whereas the experiments in Ref. 9 are for heavily compensated material.

The geometry of an idealized two-dimensional structure is presented in Fig. 3. It is a finite rectangle bounded by the lines $x = x_l$, $x = 0$, $y = y_l$ and $y = y_h$. The mask, which runs along $x = 0$ from y_l to 0, is assumed to be an impenetrable barrier for impurity atoms so that the flux across the interface is zero, i.e. $\partial\alpha/\partial x = 0$. The line $y = y_h$ is

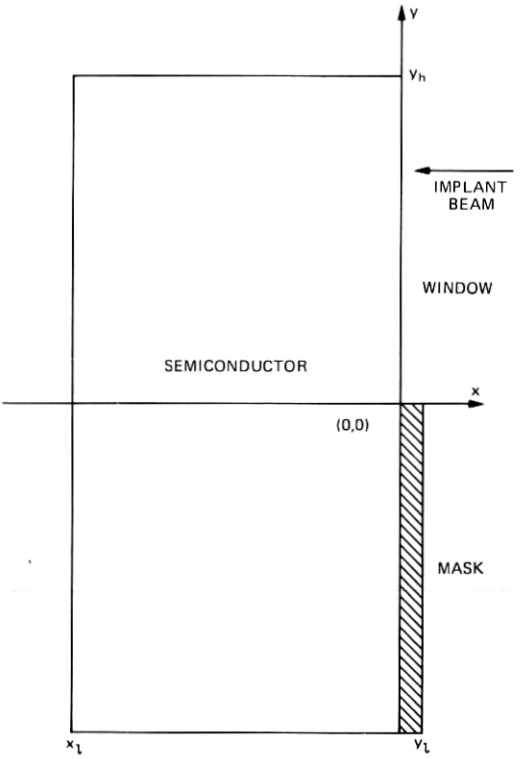

Fig. 3—The idealized structure.

assumed to bisect the window so that the appropriate boundary condition is also that the flux across the edge is zero. Along the edges $x = x_l$ and $y = y_l$ the concentration will be held at a nominal background concentration, typically 5 to 7 orders of magnitude less than the initial peak concentration. The initial conditions and the boundary conditions in the window, $x = 0$ and $0 \leqslant y \leqslant y_h$, will depend on the specific problem being investigated.

Two classes of diffusion problems are considered in this paper. The first class corresponds to the case where throughout the diffusion the concentration is held at a constant value, α_0, in the window, $x = 0$ and $0 \leqslant y \leqslant y_h$. The corresponding initial condition is that throughout the interior of the semiconductor the concentration is set at a background level of $10^{-5}\alpha_0$. This class of *constant-surface-concentration diffusion* problems is treated in Section IV. The second class corresponds to the case where an initial fixed quantity of impurity has been implanted, then diffused with a no-flux condition holding along the window. As an initial condition, we use an appropriately scaled Gaussian distribution corresponding to published results for actual implants. The

peak value of this distribution is denoted by α_0. This class of *implant diffusion* problems is treated in Section V.

In each numerical solution which is described, the constants α_0, β, and D_0, as well as other relevant parameters, are specified and the time evolution is carried forward to a specified time, t_{stop}. In discussing the solutions and their properties, it is frequently convenient to use the conventional dimensionless unit of *diffusion length*, where the dimensionless unit corresponds to physical length divided by $\sqrt{4D_0 t_{\text{stop}}}$. This dimensionless unit is commonly used in diffusion problems which admit self-similar solutions. However, we should underscore the obvious fact that the problems we have just described do not admit self-similar solutions. This follows both from the boundedness of the domain and from the nonlinear nature of the differential equation. The implication, of course, is that each physical problem requires a separate numerical simulation, and in general this is true. However, in the case of constant-surface-concentration diffusion for wide mask devices it may be possible to develop a family of self-similar solutions parameterized by the surface concentration α_0. This is discussed in Section VIII, which deals with the application of the nonlinear model to process simulation.

III. THE NUMERICAL METHOD

Equation (11) and its variations, which are derived later, are special instances of the following partial differential equation,

$$f(t, x, y, u, u_x, u_y, u_t, u_{xt}, u_{yt}) = D_x g_1(t, x, y, u, u_x, u_y, u_t, u_{xt}, u_{yt}) + D_y g_2(t, x, y, u, u_x, u_y, u_t, u_{xt}, u_{yt}), \quad (13a)$$

with initial conditions

$$u(0, x, y) = h(x, y) \quad (13b)$$

and boundary conditions

$$a(t, x, y)\mathbf{n}(x, y)\cdot(g_1, g_2) = b(t, x, y, u, u_t), \quad (13c)$$

where the unknown function, $u(t, x, y)$, is defined on the domain $t_{\text{start}} \leq t \leq t_{\text{stop}}$, $x_l \leq x \leq x_h$ and $y_l \leq y \leq y_h$, and where the vector, $\mathbf{n}(x, y)$, denotes the outward pointing normal along the boundary of the rectangle determined by x_l, x_h, y_l and y_h.

Given a well-posed problem of the form (13), the following numerical method is applicable. Represent the approximate solution, $\hat{u}(t, x, y)$, as a linear combination of tensor product B-splines with time-dependent coefficients

$$\hat{u}(t, x, y) = \sum_{j=0}^{m} \sum_{k=0}^{n} U_{jk}(t) B_j(x) B_k(y).$$

Using this representation, discretize (13) in space using Galerkin's method.[11] This creates a large first-order system of nonlinear ordinary differential equations. The time evolution of this system can then be determined by any technique suitable for the solution of "stiff," implicit systems of differential equations.[12,13] Our particular approach is to discretize in space with Galerkin's method applied to a second-order B-spline representation. An example of a second-order B-spline, usually called a chapeau function, is shown in Fig. 4. In this figure the mesh points which define the chapeau function occur at integer coordinates (0, 0), (1, 0), (0, 1), etc., while the quadrature points used in Galerkin's method occur at (½, ½), (−½, ½), etc. The time evolution is then developed using extrapolation of the strongly A-stable implicit Euler rule. Additional details are presented in the appendix.

By employing this general algorithm and by judiciously applying the chain rule for differentiation, we were able to develop our software so that the statement of the specific partial differential equation is separated from the general algorithm for solving it. There are two important consequences of this approach. First, it is extremely easy to investigate several models incorporating different physical effects. Second, it is easy to introduce analytic transformations of the original

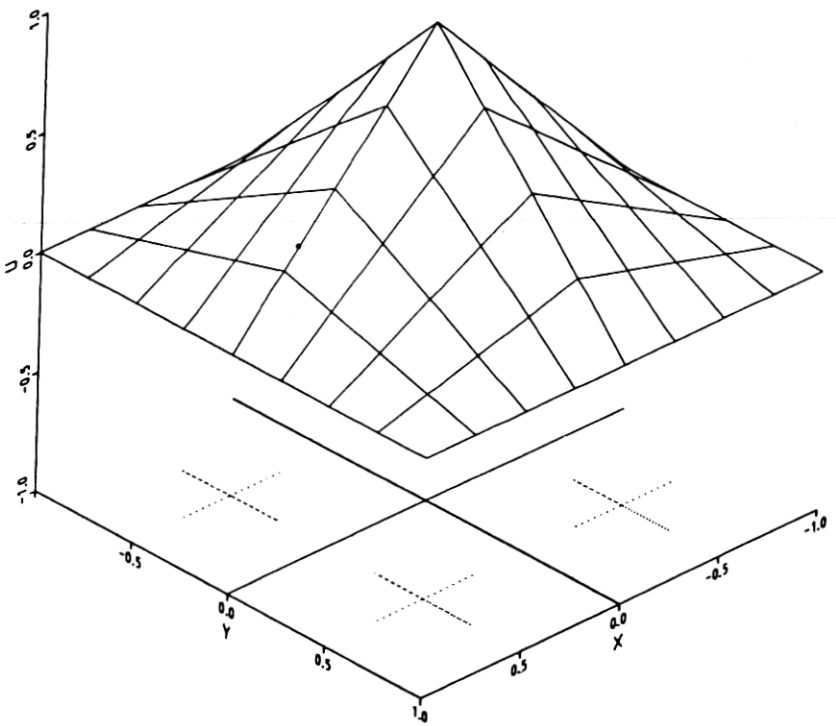

Fig. 4—A typical chapeau function.

problem which do not remove it from the canonical class defined by (13), but which may enhance the numerical accuracy and efficiency. Two such transformations are exploited in this paper. In Section 4.1 we introduce a logarithmic transformation which substantially enhances the accuracy of the solution. Then in Section 5.1 we introduce a Boltzmann-type, moving-coordinate system which enables us to obtain a solution in which the impurity dose is accurately conserved.

All the simulations presented in this paper have been calculated on a coarse mesh of 11 × 21 points and on a fine mesh of 21 × 41 points. The meshes are defined as the cross product of 11 (21) points, uniformly spaced between x_l and 0, with 21 (41) points, uniformly spaced between y_l and y_h. The difference between the solutions on the coarse and fine meshes provides a reasonable estimate of the error for the coarse mesh. In several cases, asymptotic relations may be applied to this estimate to yield a good estimate for the error on the fine mesh. In addition to comparing the solutions on two different meshes, we have also applied checks against the linear problems, which have known analytic solutions, as well as against the nonlinear one-dimensional problems which can be solved numerically to very high accuracy. These checks on the numerical error are discussed in detail in Section VI.

IV. CONSTANT-SURFACE-CONCENTRATION DIFFUSION

In this section we discuss the numerical solution of the *constant-surface-concentration diffusion* problem, i.e., the case where the impurity concentration in the window is held constant throughout the diffusion. Referring to Fig. 3, the initial conditions are that inside the semiconductor material (the rectangle determined by $x_l \leqslant x \leqslant 0$ and $y_l \leqslant y \leqslant y_h$), there is a background doping of $10^{-5}\alpha_0$, while along the window ($x = 0$, $0 \leqslant y \leqslant y_h$), the concentration is α_0. The boundary conditions for this problem are the following. In the window the concentration is maintained at α_0. Along the sides $x = x_l$ and $y = y_l$, the concentration is maintained at the background level, $10^{-5}\alpha_0$. The mask is assumed to be an ideal barrier, so that the boundary condition along the mask ($x = 0$, $y_l \leqslant y \leqslant 0$) is $\partial\alpha/\partial x = 0$. Finally, we assume the edge, $y = y_h$, bisects the window so that the boundary condition along here is $\partial\alpha/\partial y = 0$. For each diffusion problem presented, we have set the constant D_0 to 1 and we have calculated the time evolution forward to time $t_{\text{stop}} = 1$.

4.1 The logarithmic transformation

The first point to observe is that the solution, which should always be positive, ranges over five orders of magnitude. This means that relative error—not absolute error—is the appropriate criterion. How-

ever, the Galerkin discrètization in space tends to minimize absolute rather than relative error. This is also true of the vast majority of finite difference discretizations. One consequence is that an absolute error which is extremely small with respect to large values of the solution may be large enough to change the sign of small values. A relative error criterion is easily achieved if, instead of solving (11) directly, we derive and solve a partial differential equation for the logarithm of $\alpha(t, x, y)$ since, if $\alpha = \exp(u)$ and $\hat{\alpha} = \exp(u + \delta)$, then $\hat{\alpha} \approx \alpha(1 + \delta)$ for small δ. Let

$$\alpha(t, x, y) = e^{u(t,x,y)}$$

and let

$$\hat{D}(u) = D(e^u).$$

Then (11) can be rewritten as

$$u_t - \hat{D}(u)[u_x^2 + u_y^2] = \nabla \cdot (\hat{D}(u) \nabla u). \tag{14}$$

Although this equation contains the additional nonlinear terms, u_x^2 and u_y^2, it can be solved more efficiently than (11), and the answers obtained are significantly more accurate in the sense of relative error.

4.2 Linear, concentration independent diffusion

The first case we consider is the case studied by Kennedy and O'Brien, namely, constant-surface-concentration diffusion for the linear, concentration independent diffusion model on an idealized semi-infinite device.[11] Our interest in this case is twofold. First, this case offers a natural baseline against which to assess the significance of the nonlinear effects. Second, our numerical solution can be compared directly with the analytic solution provided by Kennedy and O'Brien. The accuracy of our discretization for this problem is then a rough estimate of the best accuracy achievable in the numerical solutions of the subsequent nonlinear problems.

For this problem, the function $D(\alpha)$ is simply the constant D_0, and eq. (11) simplifies to

$$\alpha_t = \nabla \cdot (D_0 \nabla \alpha). \tag{15}$$

Let $C(t, x, y)$ denote the solution developed by Kennedy and O'Brien in Ref. 1. This function satisfies equation (15) and the following constraints

$$C(t, x, y) = 0 \quad \text{for} \quad t = 0, \quad x < 0$$

$$C(t, x, y) = 1 \quad \text{for} \quad t \geqslant 0, \quad x = 0, 0 \leqslant y$$

$$C_x(t, x, y) = 0 \quad \text{for} \quad t \geqslant 0, \quad x = 0, y < 0.$$

In order to use the Kennedy and O'Brien solution to check our

numerical technique, we pose a problem on the finite rectangle whose solution is $C(t, x, y) + c_0$. To do this, simply require that $\alpha_x = 0$ along the mask and $\alpha(t, x, y) = C(t, x, y) + c_0$ along the remaining boundaries. Although this problem would form a mathematically precise test case, a somewhat more interesting test is to modify the boundary conditions to agree with those proposed for the nonlinear simulations, i.e., $\alpha(t, x, y) = c_0$ along $y = y_l$ and $x = x_l$, and $\alpha_y = 0$ along $y = y_h$. Inspection of $C(t, x, y)$ shows that this problem is only a slight perturbation of the first problem if $c_0 = 10^{-5}$ and the dimensions of the device, relative to t_{stop}, are greater than three diffusion lengths.

The solution to the problem with $D_0 = 1$, $t_{stop} = 1$, $x_l = y_l = -5$, and $y_h = 5$ is presented in Fig. 5. This numerical solution is not only qualitatively correct, but the error in the concentration is less than 7.5 percent, while the error in the log concentration is less than 2.8 percent. The precise definition of this error, which is relative to the maximum magnitude of the solution, is presented in Section VI. In that section it will also be pointed out that the larger errors are extremely localized and can be traced to two specific problems, the singularity at the edge of the mask and small displacements in the position of the diffusion front.

4.3 Nonlinear, concentration dependent diffusion

We now consider the nonlinear problem with the charged vacancy reaction and impurity-induced electric fields. Two cases of arsenic diffusion will be treated in this section. In the first case, a low-concentration case, the concentration in the window is held at $\alpha_0 = 1$, while in the second, high-concentration case, the concentration in the window is held at $\alpha_0 = 20$. Because of the similarity of the diffusion coefficients, results for boron will be qualitatively the same. For both cases, D_0 and t_{stop} are set to 1. For the low-concentration case, the dimensions of the device are $x_l = y_l = -5$ and $y_h = 5$, while for the high-concentration case the dimensions are $x_l = y_l = -8.5$ and $y_h = 8.5$.

Figures 6 and 7 present the results for $\alpha_0 = 1$ and $\alpha_0 = 20$, respectively. As might be expected, for the low-concentration case the diffusion is enhanced, but the differences from the linear case are modest. However, for the high-concentration case the diffusion enhancement is very pronounced. The diffusion front has become extremely abrupt and has been translated from the region of 1 to 2 diffusion lengths from the mask edge in the linear case to a distance of 6 to 7 diffusion lengths in the high concentration case.

The following rationale provides some qualitative insight into these nonlinear effects. When the concentration in the window is held constant, the diffusion coefficient near the window also remains constant and the solution in the immediate vicinity of the window behaves like the solution to a linear problem with an effective diffusion coeffi-

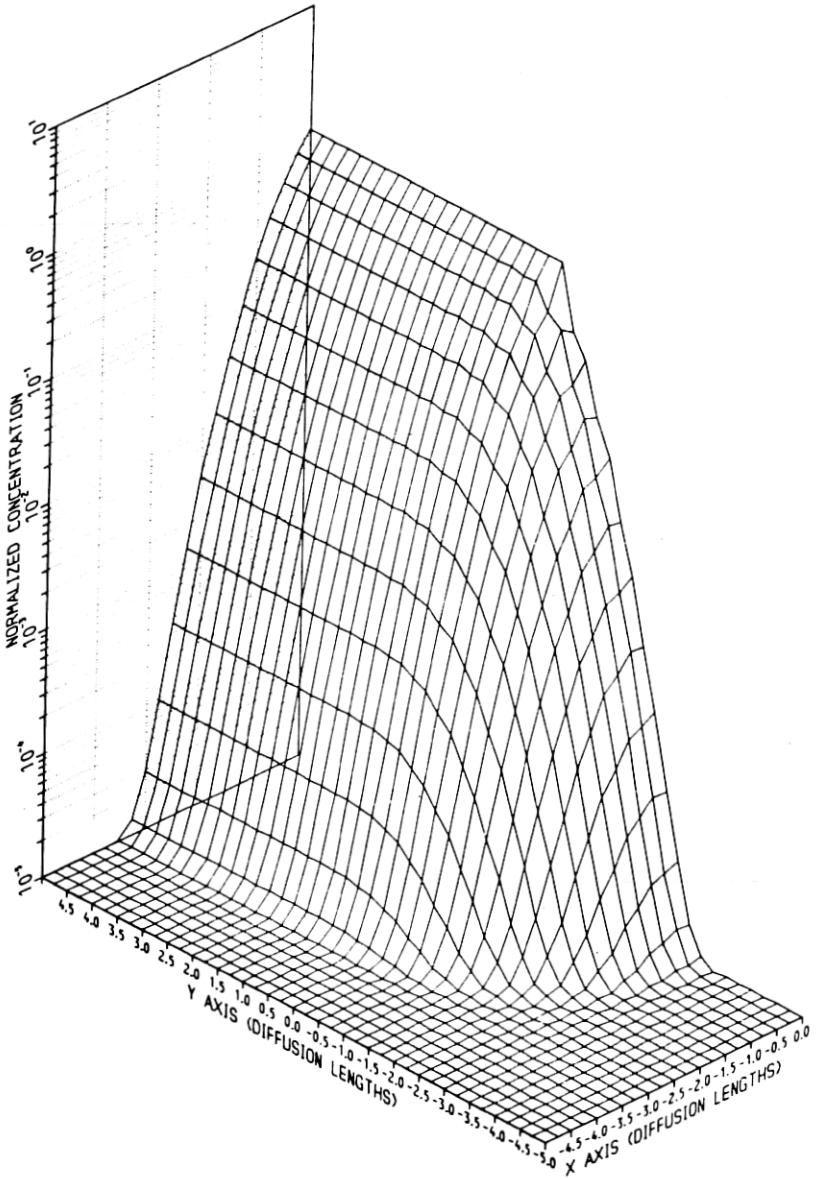

Fig. 5—Constant surface concentration diffusion, linear model.

cient of $3.96 D_0 \alpha$. However, this coefficient is falling linearly with the solution. In the critical region where $\alpha = 0.1$ the solution becomes similar to that of a linear problem with a coefficient of D_0. This causes the gradient of the solution to rapidly increase and generates an abrupt junction with a sharp corner.

In summary, the combination of nonlinear and two-dimensional

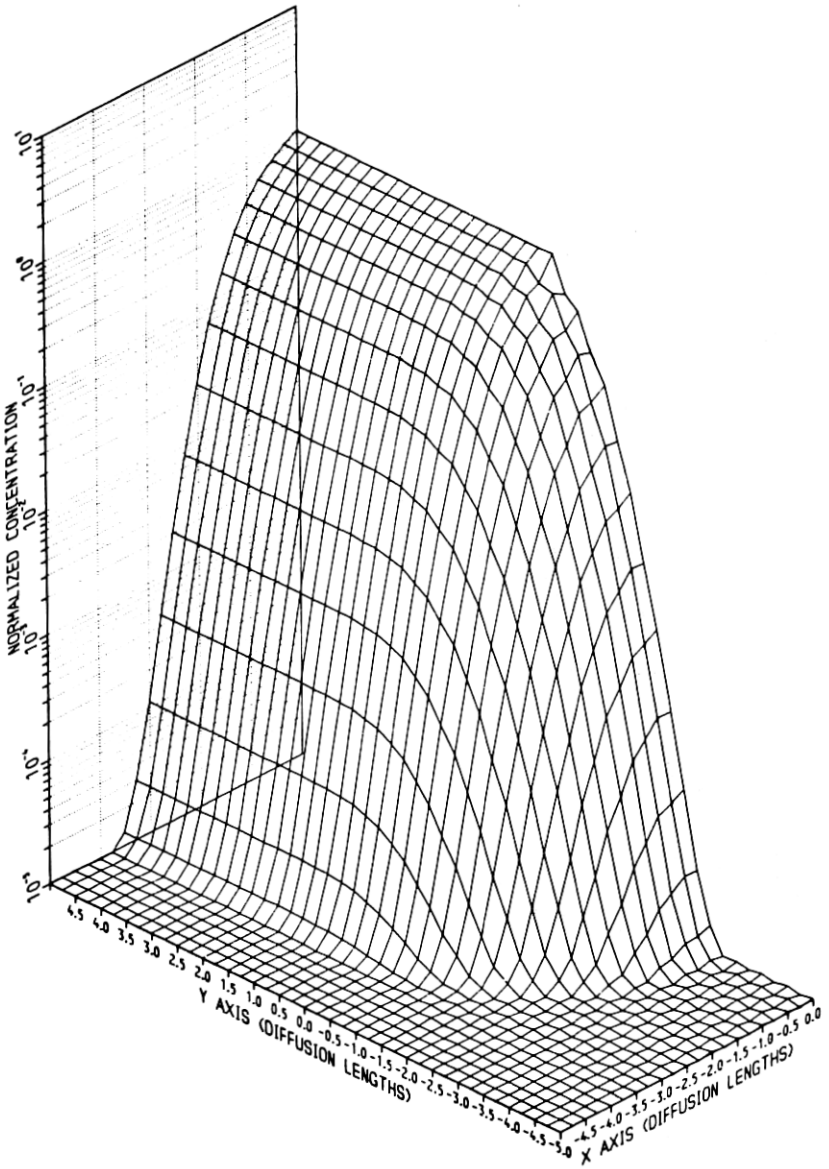

Fig. 6—Constant surface concentration diffusion, nonlinear model with donor impurities and $\alpha_0 = 1$.

effects for the constant-surface-concentration diffusion implies that, as the surface concentration increases, the diffusion front is translated further into the device, the shape of the front becomes more abrupt, and the curvature in the corner region is increased. Moreover, these effects are modest for low impurity concentrations but extremely pronounced for high impurity concentrations.

V. DIFFUSION OF AN IMPLANTED IMPURITY

In this section, we consider the case where an initial impurity dose has been implanted in the device just inside the window ($x = 0$, $0 \leq y \leq y_h$); see Fig. 3. The impurity is then diffused into the semiconductor. Along the mask ($x = 0$, $y_l \leq y \leq 0$) and the edges $x = x_l$, $y = y_l$, and

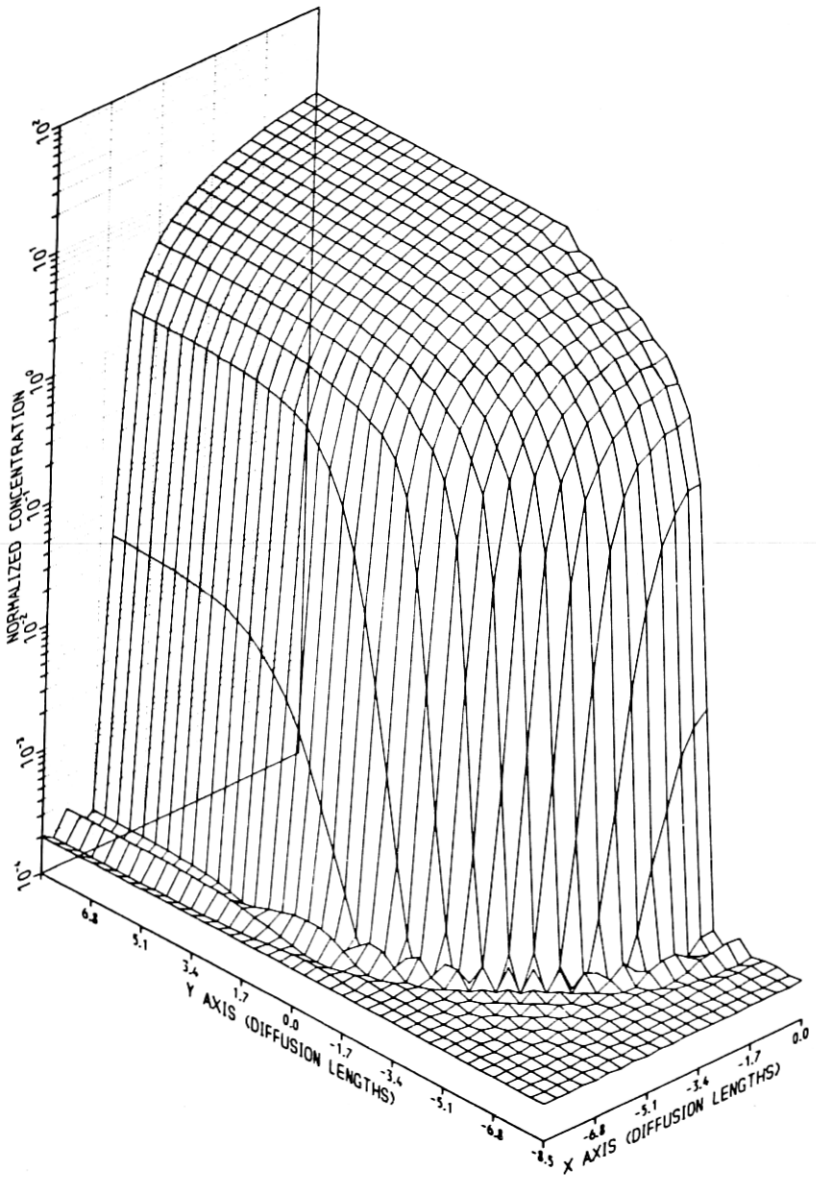

Fig. 7—Constant surface concentration diffusion, nonlinear model with donor impurities and $\alpha_0 = 20$.

$y = y_h$, the boundary conditions are the same as for the constant-surface-concentration problem discussed in Section IV. However, the boundary condition in the window is now a no-flux condition, $\partial\alpha/\partial x = 0$. In addition, we now have a nontrivial condition, which is the description of the impurity implant.

In this section, we are concerned primarily with two simulations corresponding to a boron implant and an arsenic implant. The initial conditions for each implant will correspond to an appropriately scaled Gaussian whose mean and standard deviation are functions of the beam energy and the dose. We also assume that under the mask the implant decays quickly as a Gaussian in y. The case corresponding to a boron implant is somewhat hypothetical, but the case corresponding to an arsenic implant is carefully modeled after the key processing steps used to fabricate the junction studied in Section VII.

5.1 The Boltzmann coordinate transformation

For this problem, it is useful to introduce yet another analytic transformation. A carefully selected moving coordinate system enables us to model the solution accurately throughout the transient evolution. This, in turn, improves our numerical conservation of the implanted dose. The natural transformation is

$$\zeta = \frac{x}{\sqrt{4D_0 t}}, \quad \xi = \frac{y}{\sqrt{4D_0 t}}, \quad \tau = t,$$

with the associated differential operators

$$\frac{\partial}{\partial x} = \frac{1}{\sqrt{4D_0 t}} \frac{\partial}{\partial \zeta}, \quad \frac{\partial}{\partial y} = \frac{1}{\sqrt{4D_0 t}} \frac{\partial}{\partial \xi},$$

and

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial \tau} - \frac{\zeta}{2\tau} \frac{\partial}{\partial \zeta} - \frac{\xi}{2\tau} \frac{\partial}{\partial \xi}.$$

This transforms eq. (14) into

$$\tau v_\tau - \frac{1}{2}[\zeta v_\zeta + \xi v_\xi] - \frac{D(v)}{4D_0}[v_\zeta^2 + v_\xi^2] = \nabla \cdot \left[\frac{D(v)}{4D_0} \nabla v\right]. \tag{16}$$

5.2 Linear, concentration independent diffusion

Again, the first case we consider is the case studied by Kennedy and O'Brien, namely, instantaneous source diffusion for the linear, concentration independent diffusion model on an idealized semi-infinite device.[1] Let $C(t, x, y)$ denote the solution developed by Kennedy and O'Brien. This function satisfies (15) on the left half-plane with the boundary condition $C_x(t, x, y) = 0$ for $t \geq 0$, $x = 0$, and the initial

condition that it should yield an instantaneous unit dose per unit length for $t = 0, y \geq 0$.

For our numerical problem, we consider eq. (15) with the initial condition

$$\alpha(t_{\text{start}}, x, y) = C(t_{\text{start}}, x, y) + c_0$$

and the boundary conditions

$$\alpha(t, x, y) = c_0 \quad \text{for} \quad x = x_l \quad \text{or} \quad y = y_l$$

$$\alpha_y(t, x, y) = 0 \quad \text{for} \quad y = y_h$$

$$\alpha_x(t, x, y) = 0 \quad \text{for} \quad x = 0,$$

where $t_{\text{start}} = 10^{-2}$ and $c_0 = 10^{-5}$. Inspection of $C(t, x, y)$ shows that if y_h, relative to t_{stop}, is larger than three diffusion lengths, than $\alpha(t, x, y)$ should be nearly identical to $C(t, x, y) + c_0$.

The solution of eq. (16) for $D_0 = 1$, $t_{\text{stop}} = 1$, $x_l = y_l = -5$, and $y_h = 5$ is presented in Fig. 8. The key property of the Boltzman coordinate transformation is that the solution to (16) with the corresponding semi-infinite boundary conditions does not change its shape as time evolves. The amplitude of the solution simply decays as $1/\sqrt{4D_0 t}$. Thus the solution to our finite problem should simply decay in the transformed coordinate system with no change of shape. This is indeed the case. The numerical accuracy proves to be limited only by the accuracy requested in the time evolution. For the example shown, the worst error is less than 0.4 percent.

5.3 Nonlinear, concentration dependent diffusion

We now consider two nonlinear problems, modeling diffusions of boron and arsenic implants. Referring to Fig. 3, the boundary conditions are that the concentration is held at the background doping along $x = x_l$ and $y = y_l$. Along $x = 0$ and $y = y_h$, a no-flux condition must hold. If this condition is valid at $t = t_{\text{stop}}$, then for the transformed system we obtain

$$\nu(\tau, \zeta, \xi) = c_0 \quad \text{for} \quad \zeta = \zeta_l \quad \text{or} \quad \xi = \xi_l$$

$$\nu_\zeta(\tau, \zeta, \xi) = 0 \quad \text{for} \quad \zeta = 0$$

$$\nu_\xi(\tau, \zeta, \xi) = 0 \quad \text{for} \quad \xi = \xi_h,$$

where $\tau \geq \tau_{\text{start}}$.

For the boron simulation we consider a hypothetical device with $x_l = y_l = -2.0$ μm and $y_h = 2.0$ μm. The diffusion temperature is taken to be 1400°K. The implant is taken to be the familiar Gaussian with a spread of 0.04 μm and a peak concentration of $\alpha_0 = 18.45$, where the concentration is normalized to the intrinsic carrier concentration. The

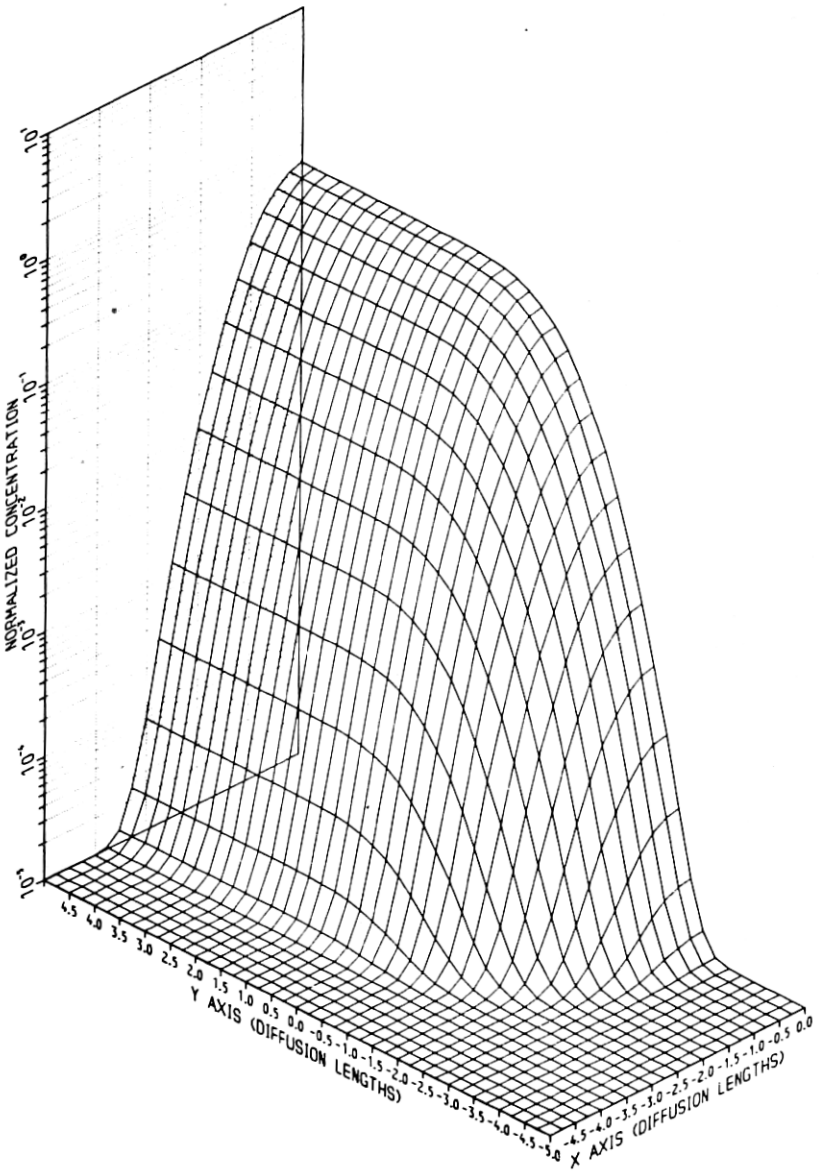

Fig. 8—Implant diffusion, linear model.

background doping is set seven orders of magnitude below this peak concentration. The length of the diffusion is 900 seconds, and the value of t_{start} is set at 65.0 seconds. The choice of t_{start} is dictated by the resulting value of the low-concentration diffusion constant, which in

this case is $D_0 = 3.37 \times 10^{-13}$ cm^2/s, and the requirement that the boundary conditions will be valid on the window defined by ζ_l ξ_l, and ξ_h. The result of the boron simulation is shown in Fig. 9.

For the arsenic simulation, the dimensions of the device are taken to be $x_l = y_l = -0.9$ μm and $y_h = 0.9$ μm. The diffusion temperature is

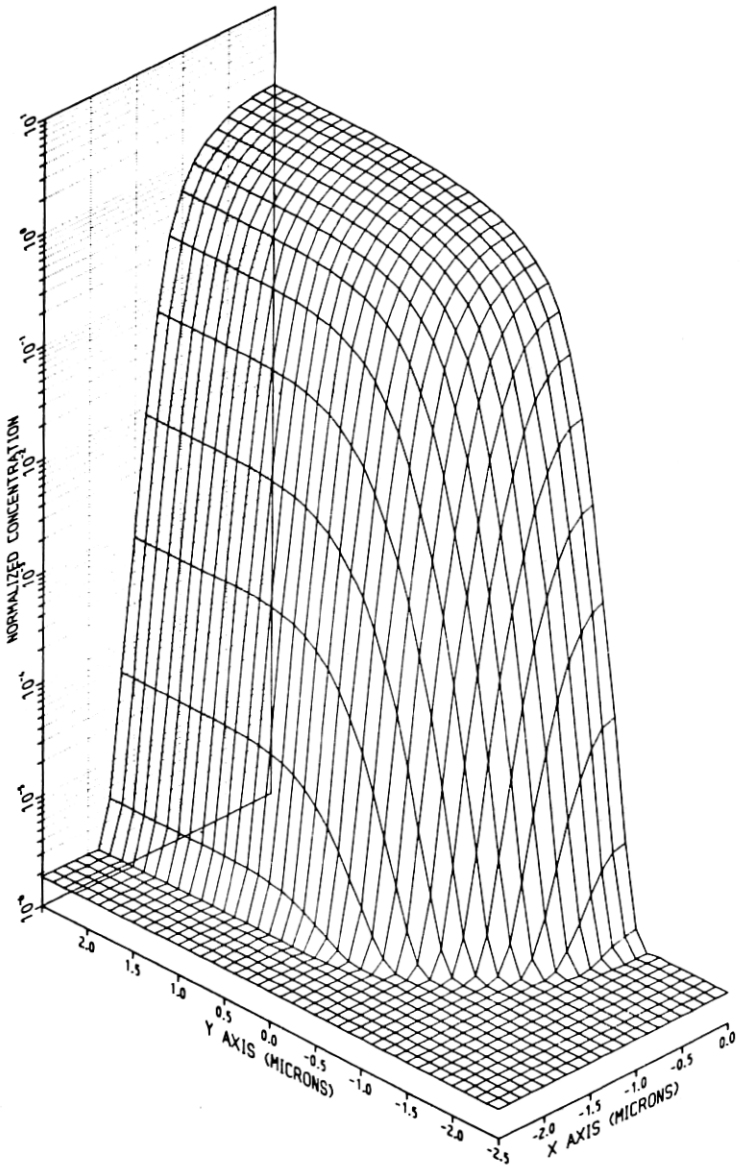

Fig. 9—Implant diffusion, nonlinear model for boron.

1373°K. The implant is assumed to have a spread of 0.017 μm and a peak concentration of 114. Again, the background doping is set seven orders of magnitude below the peak concentration and the length of the diffusion is 900 seconds. In this case, $D_0 = 2.03 \times 10^{-14}$ cm^2/s, and t_{start} is set to 67.2 seconds. The result of the arsenic simulation is shown in Fig. 10.

As the peak concentration drops, the nonlinear model becomes increasingly linear. However, the solutions shown in Figs. 9 and 10 clearly demonstrate that for contemporary fabrication the combination of peak concentration and diffusion time cause the solutions to exhibit several nonlinear effects. In particular, the shape of the diffusion front, the curvature of the corner, and the ratio of lateral-to-vertical diffusion distances are at variance with the linear predictions even for boron.

VI. NUMERICAL ACCURACY

A numerical simulation should be subjected to the same careful scrutiny as a physical experiment to establish error bounds and insure that the results are free from any systematic bias. This is particularly important in the current study where we are employing a new algorithm on a relatively small computer to investigate the two-dimensional effects predicted by a nonlinear model. In this section, we discuss the checks we have applied to this study and present various estimates for the accuracy of the numerical results.

First, we establish some meaningful but concise error descriptions. Let $\alpha(t, x, y)$ denote the concentration and $u(t, x, y)$ the log concentration, $\exp(u(t, x, y)) = \alpha(t, x, y)$. Let $\|\alpha\|$ and $\|u\|$ denote the maximum magnitude of each function over the rectangle $x_l \leq x \leq 0$ and $y_l \leq y \leq y_h$. If $\hat{\alpha}$ and $\hat{u}$ are computed solutions, define $\epsilon(t, x, y) = |\hat{\alpha}(t, x, y) - \alpha(t, x, y)|$ and $\delta(t, x, y) = |\hat{u}(t, x, y) - u(t, x, y)|$. Then two concise error measurements are $\|\epsilon\|/\|\alpha\|$ and $\|\delta\|/\|u\|$, where $\|\epsilon\|$ and $\|\delta\|$ denote the maximum magnitude of these functions over the rectangle. It is important to realize that these conventional criteria, although mathematically precise, represent the worst-case error and may be unduly pessimistic or even misleading.

We have applied three types of checks to the simulations. First, we have computed the numerical solution to linear problems for which we already know an excellent analytic approximation. This serves first as a consistency check and second as a gauge of the best achievable accuracy for the nonlinear results. Second, we have computed each solution on a coarse mesh of 11 × 21 points and on a fine mesh of 21 × 41 points. Four-thirds of the difference between these two solutions yields a good estimate on the error in the coarse mesh solution. With some exceptions which are discussed later, the error in the fine mesh solution should be one-third the difference. These estimates of the

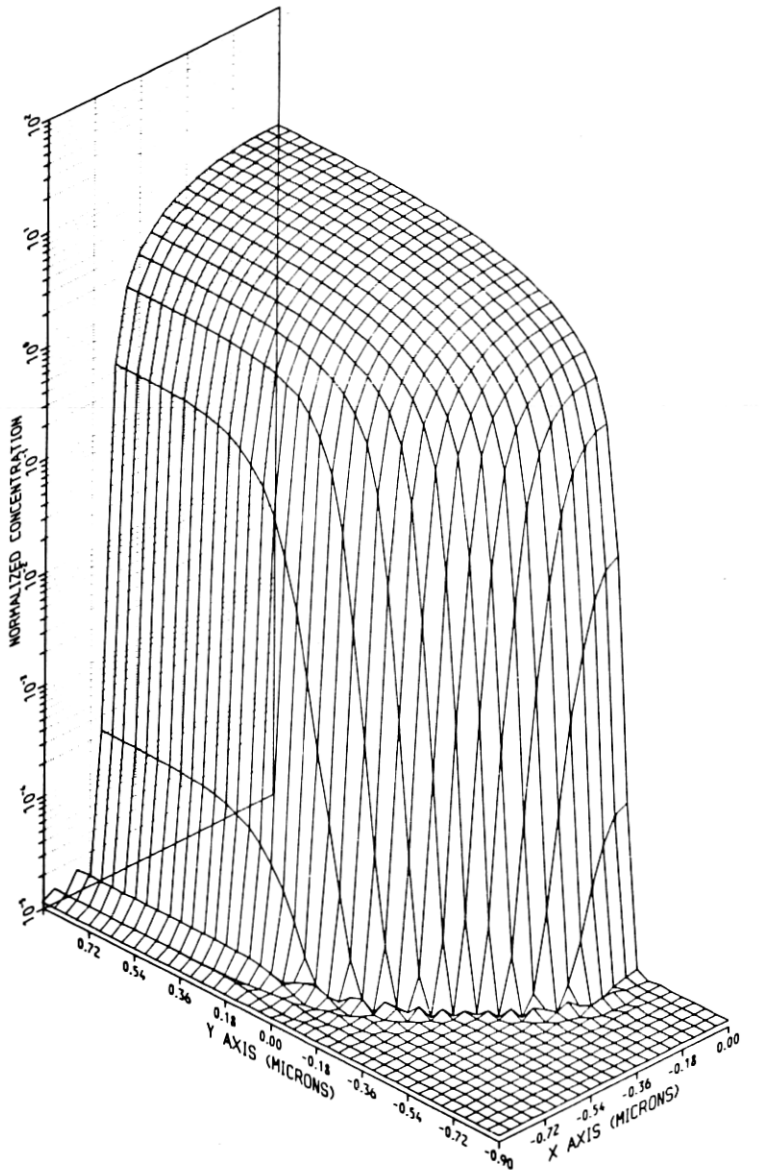

Fig. 10—Implant diffusion, nonlinear model for arsenic.

error follow from the fact that we are applying a second-order discretization in space. Finally, the solutions along the line $y = y_h$ should correspond to the solution of a one-dimensional problem, provided, of course, that the dimensions of the device are sufficiently large. Our third check is to compute the solutions to these one-dimensional

problems to very high accuracy. We use a very well-tested one-dimensional PDE package, POST,[14] and meshes of 11, 21, 41, and 81 points. These highly accurate solutions are then compared with the two-dimensional solutions to provide a check along this boundary. This last approach yields a consistency check both on the numerical software and on the dimensions of the device in that the one-dimensional numerical solution for 21 mesh points should be identical to the corresponding two-dimensional solution. In every case, the one- and two-dimensional solutions for the same mesh do, in fact, agree to the precision requested in the time evolution.

6.1 Errors in the implant diffusions

Since the error analysis is simpler in the case of the implant simulations, we address them first. Figure 11 graphs the four one-dimensional solutions for boron. A number of interesting facts can be

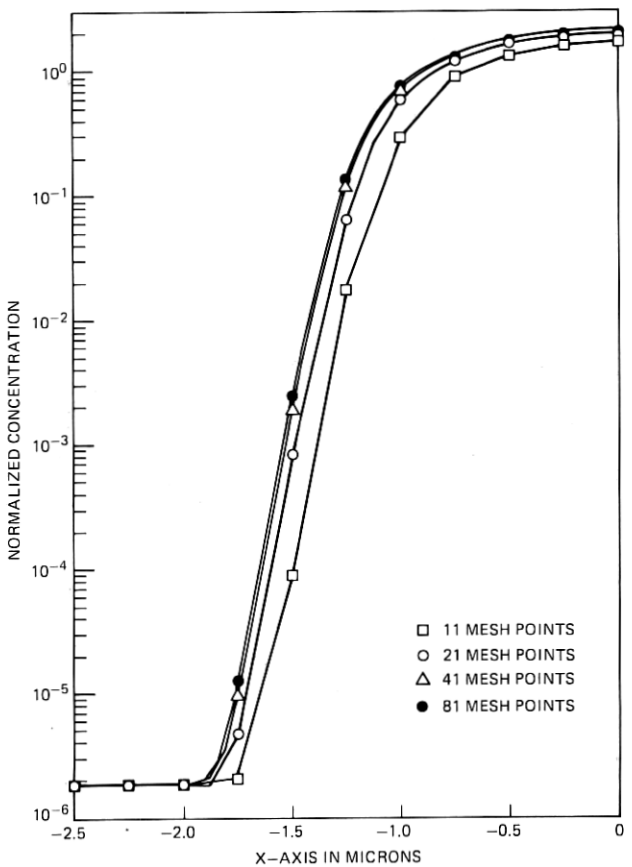

Fig. 11—One-dimensional solutions, nonlinear model for boron.

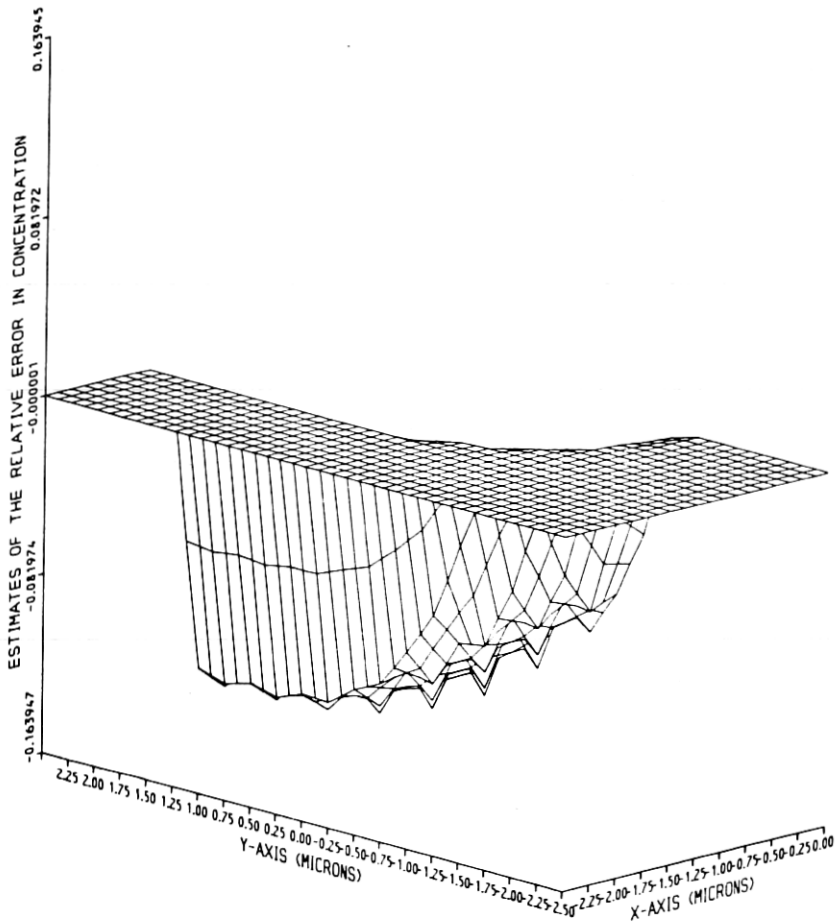

Fig. 12—Relative difference in the concentration, nonlinear model for boron.

observed. First, the solutions are rapidly converging as the mesh is refined. In fact, closer inspection shows that the asymptotic h^2 rate of convergence is being achieved, that is, the error is being decreased by a factor of one-fourth each time the mesh spacing is halved. Second, the 21-point solution, while still discernibly in error, is quite good. The maximum error in the concentration is only 9.6 percent, while the maximum error in the log concentration is only 8.4 percent. Third, we see that for boron the chief contribution to the error is an underestimate of the amplitude. This is probably due to the fact that the coarser meshes do not conserve the dose well. We also note that the maximum error in the log concentration occurs in the steepest part of the front where a small change in the x coordinate introduces a very large

change in the value of the concentration. In fact, it is probably far more relevant to remark that the maximum error in the position of the front is less than 0.06 μm, a relative error of about 4.0 percent if the front is assumed to be located 1.5 μm inside the window.

Figures 12 and 13 are plots of the relative differences between the coarse (11 × 21) and fine (21 × 41) mesh, two-dimensional solutions. Both figures show that the maximum error is quite localized and that it should be well estimated by the one-dimensional errors. Consequently, we can safely estimate the maximum errors in the fine mesh solution by applying the h^2 relation, i.e. $\delta_{\text{fine}} \approx (u_{\text{coarse}} - u_{\text{fine}})/3$. The error estimates for boron are summarized in Table I.

Because of the increased nonlinearity of the donor model as well as

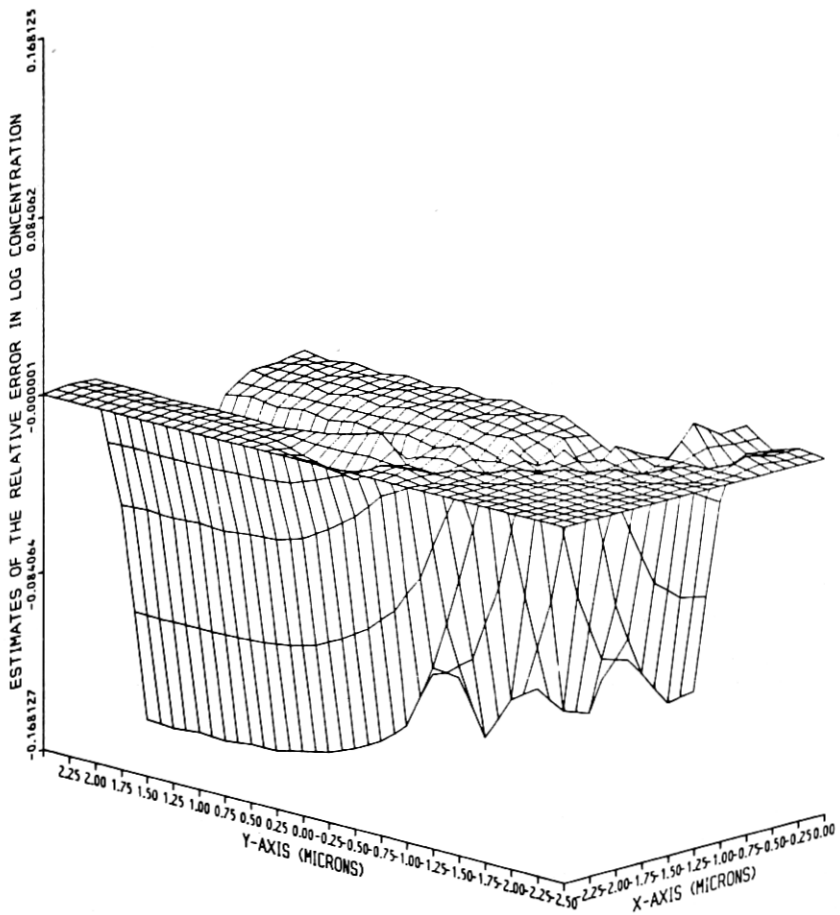

Fig. 13—Relative difference in the log concentration, nonlinear model for boron.

Table I—Error estimates for the implant simulations

Error Estimate	Boron Implant	Arsenic Implant
One-Dimensional		
$\|\delta\|/\|\mu\|$	8.4%	2.4%
$\|\epsilon\|/\|\alpha\|$	9.6%	2.2%
Two-Dimensional		
$\|\delta\|/\|\mu\|$	5.6%	32.6%
$\|\epsilon\|/\|\alpha\|$	5.5%	33.0%
Frontal Displacement		
Absolute error	0.06 μm	0.018 μm
Relative error	4.0%	2.7%

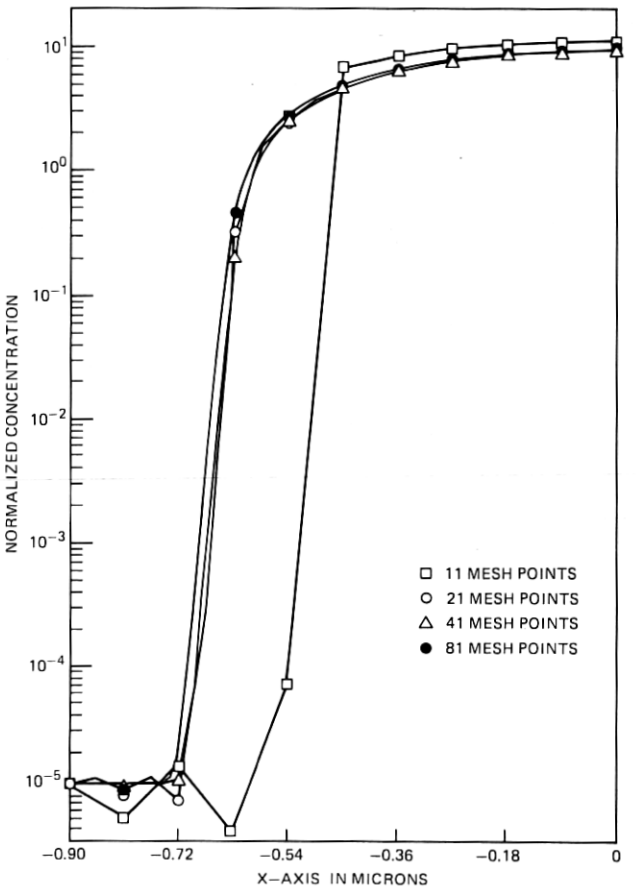

Fig. 14—One-dimensional solutions, nonlinear model for arsenic.

the much higher initial peak concentration, the case of the arsenic implant is considerably more difficult numerically than boron. Figure 14 shows the four one-dimensional solutions for this case. Two facts leap out. First, the 11-point solution is extremely inaccurate, while the

21-point solution is in fact more accurate than the corresponding solution for boron. The maximum estimated errors in the concentration and log concentration are 2.2 percent and 2.4 percent, respectively. Second, the front is somewhat steeper which, in turn, amplifies the sensitivity of the error to small changes in the position of the front. The front is displaced by about 0.018 μm, a relative error of about 2.7 percent if the front is assumed to be located 0.67 μm inside the window. In spite of the relatively good accuracy, the errors indicate that the asymptotic h^2 rate of convergence has not yet been achieved.

Figures 15 and 16 are plots of the relative differences between the coarse (11 × 21) and fine (21 × 41) mesh, two-dimensional solutions. Both plots indicate that the dominant error is caused by the displace-

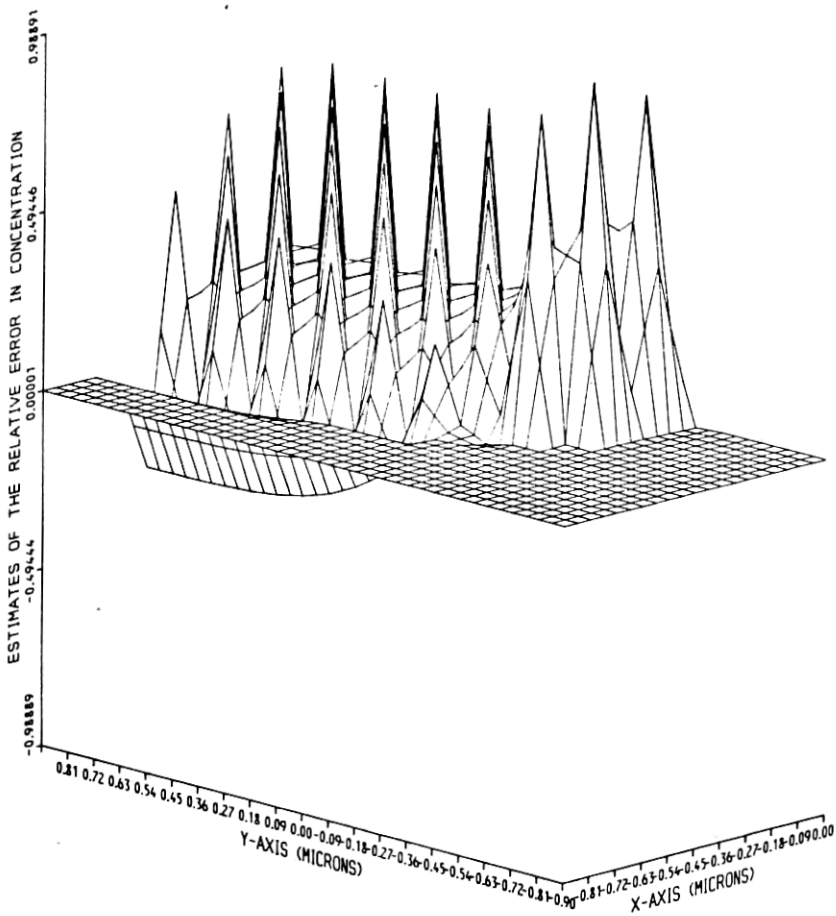

Fig. 15—Relative difference in the concentration, nonlinear model for boron.

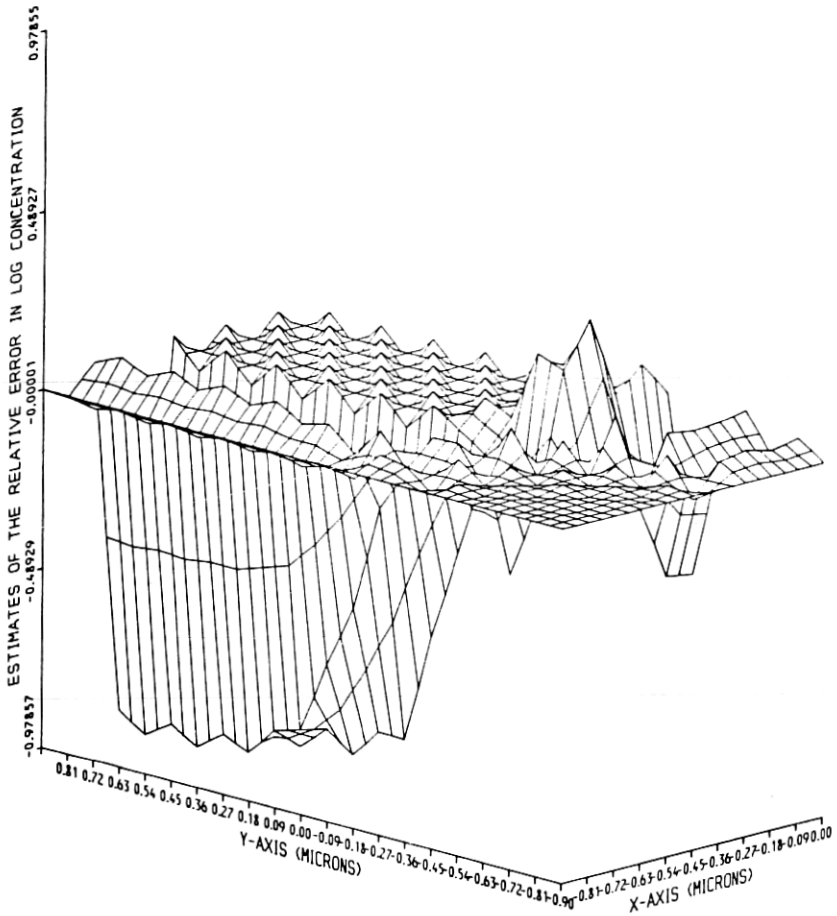

Fig. 16—Relative difference in the log concentration, nonlinear model for arsenic.

ment of the front. Figure 16 also indicates that the error in the log concentration should be well estimated by the one-dimensional analysis.

Table I summarizes the error analysis for the implant study. The one-dimensional and frontal displacement error estimates are taken directly from Figs. 11 and 14. The two-dimensional error estimates are simply the maximum differences shown in Figs. 12, 13, 15, and 16, multiplied by the asymptotic correction factor, ⅓. In the case of boron, the one-dimensional analysis indicates that, although the errors are beginning to display the proper asymptotic behavior, the two-dimensional estimate is overoptimistic. In the case of arsenic on the other hand, the one-dimensional analysis shows that, although the errors are not yet behaving asymptotically, the two-dimensional estimates are

grossly pessimistic. Clearly, the poor placement of the front in the 11 × 21 point mesh completely dominates the error. We conclude that the 21 × 41 point mesh is quite adequate for engineering work. We also note that merely reporting the maximum error is too simplistic.

6.2 Errors in the constant-surface-concentration diffusions

In the case of constant-surface-concentration diffusion, the concentration in the window is pinned throughout the diffusion. This means that the dominant error in the one-dimensional region is simply the frontal displacement. However, this also means that the true solution is singular at the edge of the mask ($x = 0$, $y = 0$). Such boundary singularities invalidate the h^2 asymptotic property of the numerical method we are employing (as well as virtually any other numerical method that does not deal with the singularity directly). Fortunately,

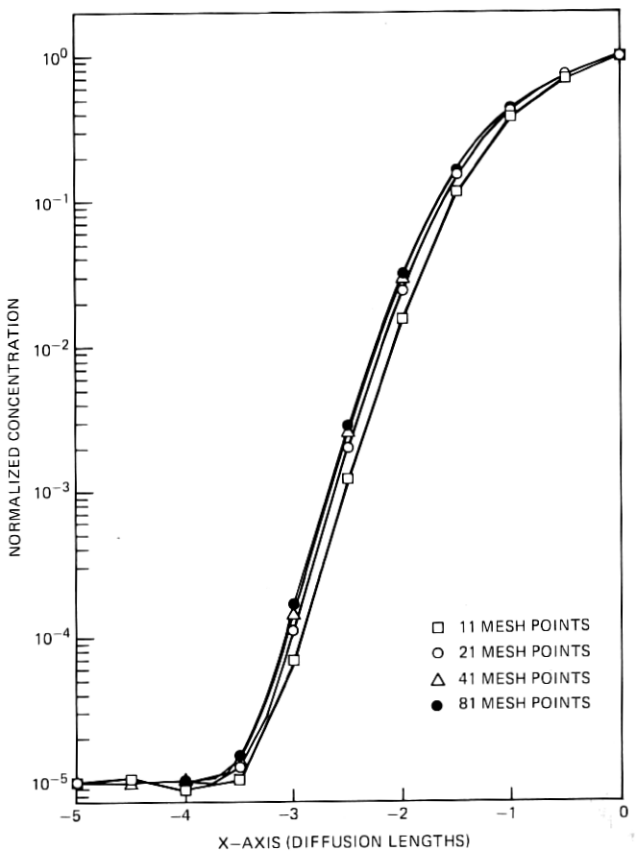

Fig. 17—One-dimensional solutions, constant surface concentration diffusion, $\alpha_0 = 1$.

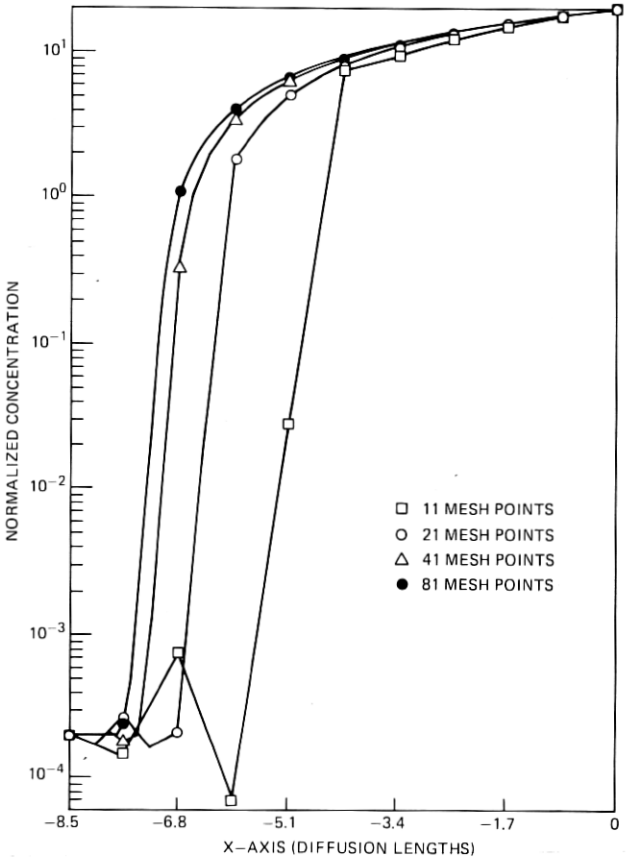

Fig. 18—One-dimensional solutions, constant surface concentration diffusion, $\alpha_0 = 20$.

the effect of such a singularity is generally restricted to the immediate vicinity of the singularity.

Figures 17 and 18 present the four one-dimensional solutions for the low and high concentration cases, respectively. The frontal displacement problem is particularly acute in the high concentration case. The error in the position of the front is about 11.6 percent. While this may well be acceptable as far as locating the front, the fact that the front is so steep means that the value of $\delta / \| u \|$ in this region is essentially meaningless.

Another way of describing the error in a problem of this nature is to apply a type of backward error analysis. It may be more informative to compare the solution with a more accurate solution evolved forward to a different time. For each solution of the four constant-surface-concentration cases, we have picked a mesh point in the diffusion front

and have then required that the 11-, 41-, and 81-point solutions take on the same value at this position as the 21-point solution. Figure 19 presents the results for this one-dimensional study in the high concentration case. From this point of view, the error in the stopping time is 18.1 percent, while the maximum error in the concentration is only 1.0 percent, and the maximum error in the log concentration is 2.3 percent.

We now consider the effect of the mask edge singularity. That the mask singularity has an effect is clearly evidenced from the small irregularities near the mask edge in Figures 5 to 7. Figure 20 presents the relative error in the concentration between the numerical solution to the linear problem and the analytic solution. Two things are apparent. First, the major contribution to the error is the mask singularity, and second, its effects are quite localized. This second observation is important if we are to place any confidence in the

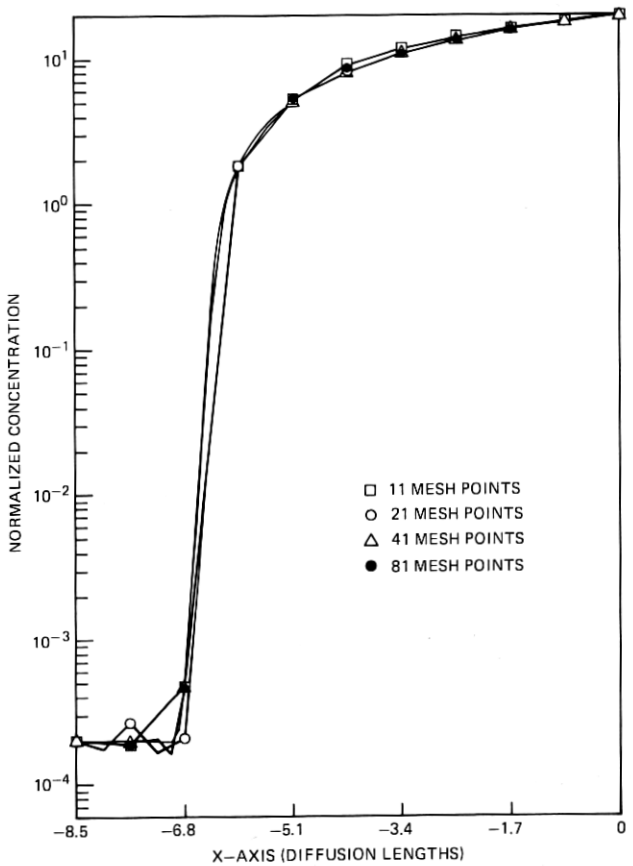

Fig. 19—One-dimensional solutions, t_{stop} varying, constant surface concentration diffusion $\alpha_0 = 20$.

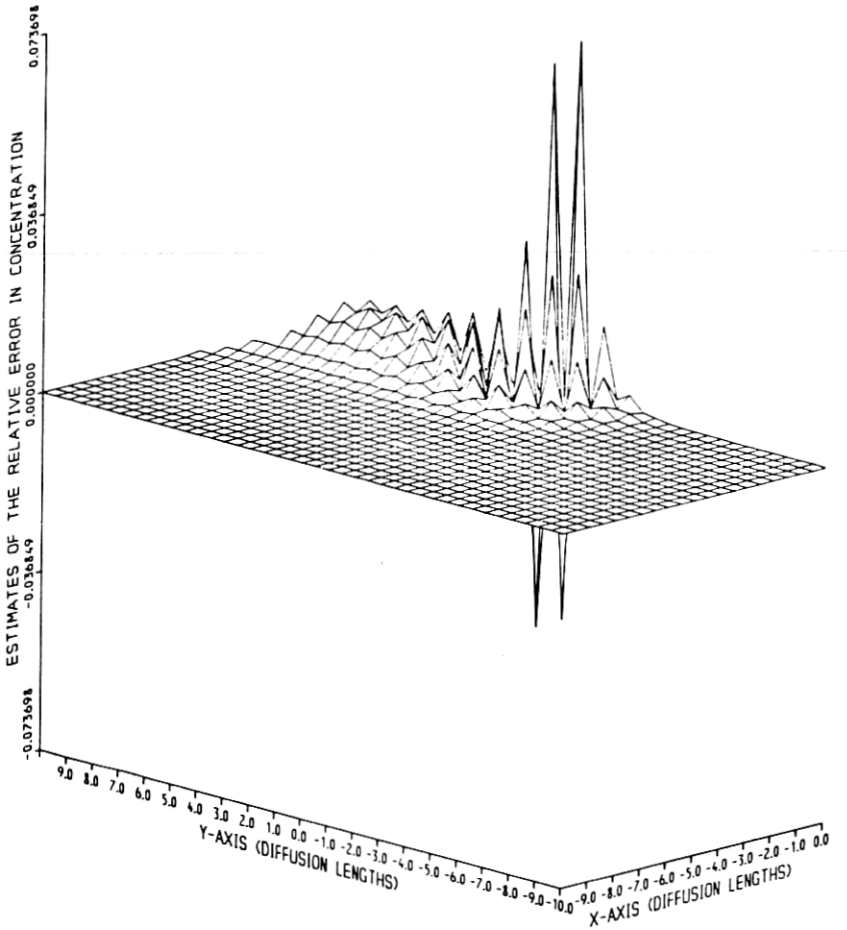

Fig. 20—Relative error in the concentration, constant surface concentration diffusion, linear model.

prediction of the position, y_f ($y_l \leq y_f \leq 0$), of the "lateral out-diffusion front." Figure 21 presents the relative error in the log concentration for the linear problem. Figure 21 indicates that, for the nonlinear model, the dominant error in the log concentration will be due to the frontal displacement. Comparisons of the coarse and fine mesh solutions yield plots that are qualitatively the same as the last two figures. The error in the concentration is dominated by the mask edge singularity, and the error in the log concentration is dominated by the frontal displacement.

The error estimates for the constant-surface-concentration diffusions are summarized in Table II. The absolute error in the location of the diffusion front is given in diffusion lengths. The relative time error

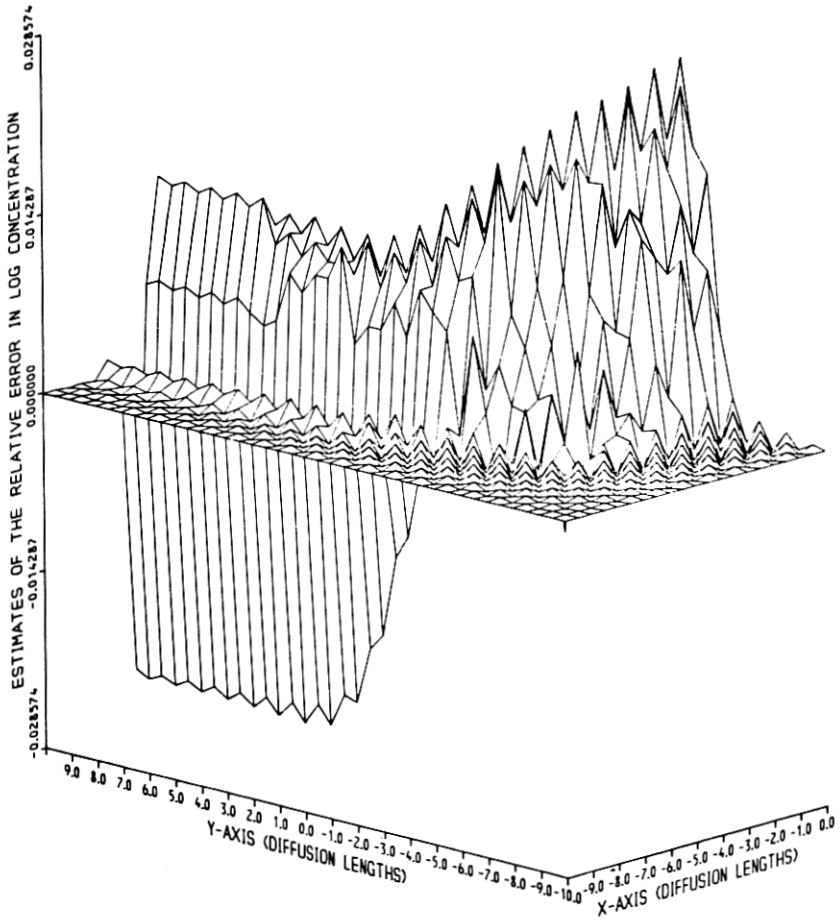

Fig. 21—Relative error in the log concentration, constant surface concentration diffusion, linear model.

Table II—Error estimates for the constant-surface-concentration simulations

Error Estimate	$\alpha_0 = 1$	$\alpha_0 = 20$
One-Dimensional		
$\lVert\delta\rVert/\lVert\mu\rVert$	1.2%	2.3%
$\lVert\epsilon\rVert/\lVert\alpha\rVert$	0.3%	1.0%
$\Delta t/t_{stop}$	5.9%	22.2%
Two-Dimensional		
$\lVert\delta\rVert/\lVert\mu\rVert$	3.6%	34.1%
$\lVert\epsilon\rVert/\lVert\alpha\rVert$	12.8%	9.6%
$\Delta t/t_{stop}$	9.7%	37.2%
Frontal Displacement		
Absolute error	0.08	0.85
Relative error	3.6%	11.6%

estimates for the two-dimensional cases are determined by evolving the coarse mesh solution forward until it interpolates the fine mesh solution at a prescribed point in the diffusion front. The one-dimensional error estimates are simply the maximum differences between the 21-point and 81-point solutions, and the two-dimensional estimates are simply the maximum differences between the 11×21 point and the 21×41 point solutions. Obviously, the two-dimensional estimates are unduly pessimistic. However, no asymptotic relations appear to be applicable because of the difference in stopping times as well as the presence of the mask edge singularity. Again it seems reasonable to conclude that the accuracy obtained by the 21×41 point mesh is adequate for engineering work.

The analysis presented in this section would appear to conclusively establish four facts: (*i*) the algorithm outlined in Section III is effective, (*ii*) the various diffusion simulations are free from systematic errors, (*iii*) for acceptor and donor impurities in low concentrations, the simulators can obtain answers accurate to a few percent, and (*iv*) for high concentrations of impurities, the simulators can obtain answers which are qualitatively correct and probably suitable for engineering design. At the same time, the fact that for high concentrations the errors were on the order of 10 to 30 percent clearly indicates that more accurate results in this regime will require more sophisticated techniques or larger computers.

We have indicated two analytic transformations of the partial differential equation which have made it more tractable numerically. Several other ideas have also suggested themselves and should be explored some time in the future. First, boundary singularities, such as the one caused by the mask edge, are well-known problems, and for simple linear problems there are two standard remedies, either explicitly model the analytic singularity or locally refine the mesh in the vicinity of the singularity. For nonlinear problems, the same ideas should apply and we have briefly explored both possibilities, but we have not as yet achieved any notable success. Second, the underlying mesh used in the calculations presented in this paper consists of points uniformly distributed along the x and y axes. Clearly, a graded mesh could further enhance the accuracy of the solution. However, while a graded mesh will not adversely affect the asymptotic behavior of the errors, care must be exercised in grading the mesh to insure that the conditioning of the Galerkin matrix is not unduly degraded. Third, in view of the generality of the underlying differential equation, a natural idea would be to explore curvilinear coordinate systems. The level curves of the solutions immediately suggest a system based on the conformal map, $w = (z + z^{-1})^{1/2}$, which gives rise to an elliptic-hyperbolic coordinate system.

VII. PROFILE MEASUREMENTS

The calculations of concentration dependent diffusion profiles discussed here depend, for their validity, on the applicability of the underlying physical theory developed by Hu.[7] Experimental measurements of the vertical diffusion have verified the one-dimensional version of the model.[3] Verification of the two-dimensional model requires that the lateral diffusion effects also be examined. This section presents the results of such an experiment.

Measurement of lateral diffusion effects by angle lapping is impossible since the vertical extent of the doping profiles precludes magnification by angle lapping. The resolution of optical microscopes is insufficient to provide accurate measurements of the small lateral distances involved, ~0.5 μm. The resolution problem is solved by making measurements of required distances using a scanning electron microscope (SEM). The remaining difficulty arises from exposing the junctions to be measured and preparing samples in a way which will allow the junctions to be detected.

Two methods of junction exposure were attempted. Conventional angle lapping of the junction has been previously used to expose junctions.[15] This technique results in smooth, easily cleaned surfaces, but control of the specific location of the surface cut is difficult. A second method of laser scribing to ± 25 μm on either side of the junction and cleaving provides accurate ± 25-μm surface cut location, but leaves a surface which can contain numerous surface cracks which trap charge and obscure the surface during SEM studies. Both methods allow operation of the device with substantially the same electrical characteristics as before the sample preparation.

Two kinds of imaging techniques were used to make the junction position visible. MacDonald and Everhart[15] have shown that the width of p-n junction depletion layers can be measured using the SEM. Furthermore, Child, Ranasinghe and White[16] have shown that the depletion layer in MOS transistors can be directly measured using the SEM in the beam-induced conductivity (BIC) mode.[16] The MacDonald and Everhart technique was used on angle-lapped samples, but the resolution obtained was not adequate for the present study. The BIC mode of exposure was used on laser-scribed samples and provides acceptable images; however, scattering of the incident electron beam and collection of BIC electrons away from the depletion layer by diffusion make interpretation of the images unduly complex for junction profile measurements.[17]

The most successful method of profile measurement involves differential etching of doped material after cleaving or angle lapping. The most successful etch used was nitric acid containing 1 to 2 percent hydrofluoric acid. Etching for 2 to 3 seconds under strong illumination

is adequate. A 5 percent hydrogen-peroxide solution with 1 to 2 percent hydrofluoric acid was also used successfully with 3-minute etch time. The nitric-based solution has the added advantage of acting as an optical stain which makes sample preparation easier. After etching, samples are not operated electrically, but are coated with 100 to 200 Å of gold to improve electron emission.

The electron microscope used is equipped with image-processing equipment which allows contrast ratios in the final image to be improved by contrast adjustments of pseudo-color image processing. A typical image is shown in Fig. 22. The thin rectangle at the top is the polysilicon gate electrode with a width of 3.43 μm $\pm$ 0.14 μm. Lateral diffusion of the arsenic source and drain implants under the gate results in the source and drain junctions etched out below the gate. Vertical junction depth is 0.57 μ $\pm$ 0.025 μm; the ratio of lateral-to-vertical junction depth is 0.66 $\pm$ 0.09. These means and standard deviations were obtained by making several SEM photographs and calculating the means and standard deviations from this distribution of measurements.

For those lengths which are resolvable by optical means, the length ratios given above are within one standard deviation of the SEM data. Optical measurements of our diffusion are not possible, however, and

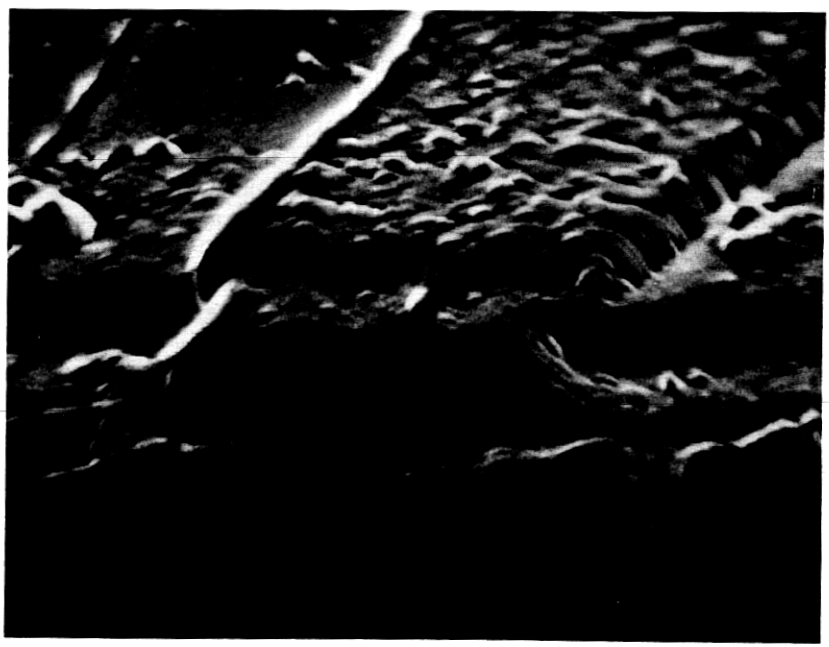

Fig. 22—SEM photograph of an exposed junction.

in general measurements on the SEM have at least five times the resolving power of the optical measurements, greater depth of field, and freedom from optical interfacing effects in lapped oxides.

The diffusion of the arsenic implant for the samples studied was numerically simulated. As described in Section 5.3, the initial implant was modeled by a Gaussian with a peak concentration of 114.39 (relative to the intrinsic carrier concentration) and a spread of 0.017 μm. The implant was then diffused for a period of 900 seconds at a temperature of 1373.16°K. In the samples studied, the ratio of surface concentration to bulk concentration was 10^4. Applying inverse linear interpolation to the results of the simulation (see Figs. 10 and 14), we calculate that the lateral junction depth is 0.486 μm $\pm$ 0.021 μm, the vertical junction depth is 0.687 μm $\pm$ 0.018 μm, and the ratio of lateral to vertical junction depths is 0.707 $\pm$ 0.051. The lateral junction depth is computed from the two-dimensional solution, while the vertical junction depth is computed from the more accurate one-dimensional solutions. Thus, the nonlinear model predicts greater penetration than observed experimentally. This is consistent both with the physical assumptions presented in Section II and with the fact that the simulation neglects oxide growth during the drive-in cycle. However, the shape of the diffusion front, as evidenced by the ratio of the lateral-to-vertical junction depths, is in excellent agreement with the predictions of the nonlinear model.

VIII. PROCESS SIMULATION

The results presented in this paper demonstrate that accurate simulations of realistic diffusions involving a single impurity can be obtained with limited computing resources. Clearly, numerical simulations should lead to deeper understanding in more complex processing situations such as multiple diffusions of interacting species and diffusions involving narrow masks and/or windows. Such simulations may, in fact, prove essential in developing technologies.

With respect to the limited computing resources, all the simulations presented in this paper were performed on two minicomputers, the Harris Slash 7 and the Interdata 8/32. Both provide 256K words of directly addressable primary memory. In the case of the Harris Slash 7, this was accomplished through a virtual memory system. In our implementation, which relies on direct banded-matrix techniques, primary memory has proven to be the crucial resource. Typical run times for the fine mesh solutions have been on the order of one to two hours. However in the case of high concentrations of donor impurities the run times have been on the order of 10 to 16 hours. Several months after the completion of this study, some of the simulations were rerun on a Cray-1 computer, the largest and fastest commercially available

scientific computer. The most difficult of the simulations required only 7.25 minutes. Thus, in a production environment simple economics will favor large computers.

Finally we note that if equilibrium solutions exist for (16), then the constant-surface-concentration diffusion problem for the ideal semi-infinite device has a family of self-similar solutions, parameterized by the surface concentration, α_0. For a suitable selection of α_0's, such equilibrium solutions could be calculated numerically. The resulting collection of solutions could then supplement the linear model presented in Ref. 2.

IX. CONCLUSIONS

In this paper we investigated the combination of nonlinear and geometric effects on the impurity atom distribution by studying a two-dimensional, concentration dependent diffusion model in a region containing the edge of the diffusion mask. This study involved four major efforts. First, we developed software for solving the nonlinear partial differential equation describing the diffusion model, and we used this software to perform several numerical simulations on two different minicomputers, the Harris Slash 7 and the Interdata 8/32. Second, we applied a number of checks on the accuracy of the results from the simulations and concluded that the results are sufficiently accurate for device design using current fabrication technologies. In particular, we find that large minicomputers are adequate for process simulation of this limited scope. Third, we analyzed the simulation results in the light of the behavior of the nonlinear diffusion coefficient, $D(\alpha)$. From its behavior, one would anticipate that the nonlinear effects would become increasingly pronounced as the ratio of the impurity concentration to the intrinsic carrier concentration became larger, and that they should be more pronounced for donor impurities. The simulations confirmed this and demonstrated that there are three principal nonlinear effects: (*i*) a large translation of the diffusion front, (*ii*) a marked steepening of the front itself, and (*iii*) a very noticeable decrease in the ratio of the lateral-to-vertical diffusions. One inescapable consequence is that, as the mask width approaches the junction depth, channel widths will be determined by these nonlinear, two-dimensional effects. Fourth, one of the authors, C. L. Wilson, obtained experimental measurements of actual lateral-to-vertical diffusion ratios for arsenic diffused in silicon. These measurements confirm the validity of the concentration dependent diffusion model.

X. ACKNOWLEDGMENTS

We take this opportunity to thank the following people who were of immense help at various stages in this project: H. K. Gummel who

provided assistance with the Harris Slash 7 computer on which the diffusion simulations were run; D. M. Ritchie, S. I. Feldman, and P. J. Weinberger who assisted with the Interdata 8/32 computer on which we performed the error studies including the stopping-time analysis; J. H. Bruning and D. D. Cook whose assistance in collecting the SEM data was absolutely essential; R. L. Johnson who helped immensely with the sample preparation; N. L. Schryer and L. C. Kaufman who contributed both excellent advice and important software; and finally J. G. Ruch who supplied many helpful suggestions and ideas. We are indeed grateful to each of these individuals for contributing their time, energy, and expertise.

APPENDIX

In this appendix we present the details of our semi-discretization algorithm for solving problems such as the nonlinear diffusion equation. We start with a partial differential equation in general divergence form.

$$f(t, x, y, u, u_x, u_y, u_t, u_{xt}, u_{yt}) = D_x g_1(t, x, y, u, u_x, u_y, u_t, u_{xt}, u_{yt}) + D_y g_2(t, x, y, u, u_x, u_y, u_t, u_{xt}, u_{yt}) \quad (17a)$$

with initial conditions

$$u(0, x, y) = h(x, y) \quad (17b)$$

and boundary conditions

$$a(t, x, y)\mathbf{n}(x, y) \cdot (g_1, g_2) = b(t, x, y, u, u_t), \quad (17c)$$

where the unknown function, $u(t, x, y)$, is defined on the domain $0 \leq t \leq t_{\text{stop}}$, $x_l \leq x \leq x_h$, and $y_l \leq y \leq y_h$, and where the vector, $\mathbf{n}(x, y)$, denotes the outward pointing normal along the boundary of the rectangle determined by x_l, x_h, y_l, and y_h.

Given meshes

$$x_L = x_0 < x_1 < \cdots < x_{m-1} < x_m = x_R$$

and

$$y_L = y_0 < y_1 < \cdots < y_{n-1} < y_n = y_R,$$

we will approximate $u(t, x, y)$, the solution of (13), by a linear combination of second-order B-splines

$$\hat{u}(t, x, y) = \sum_{j=0}^{m} \sum_{k=0}^{n} U_{jk}(t) B_j(x) B_k(y). \quad (18)$$

A second-order B-spline, $B_j(x)$, also called a *chapeau* function, is a piecewise linear function which is 1 at x_j and 0 at all other mesh points. Outside the interval $[x_L, x_R]$, the spline is defined to be 0. Inside the

interval, it is defined by requiring it to be continuous. The set of product functions $\{B_j(x)B_k(y):0 \leq j \leq m,\ 0 \leq k \leq n\}$ is commonly referred to as the tensor product basis.

The basic idea is to choose the coefficients, $U_{jk}(t)$, so that the remainder function,

$$\begin{aligned} r(t, x, y, \hat{u}) = {} & f(t, x, y, \hat{u}, \hat{u}_x, \hat{u}_y, \hat{u}_t, \hat{u}_{xt}, \hat{u}_{yt}) \\ & - D_x g_1(t, x, y, \hat{u}, \hat{u}_x, \hat{u}_y, \hat{u}_t, \hat{u}_{xt}, \hat{u}_{yt}) \\ & - D_y g_2(t, x, y, \hat{u}, \hat{u}_x, \hat{u}_y, \hat{u}_t, \hat{u}_{xt}, \hat{u}_{yt}), \end{aligned}$$

is orthogonal to the basis elements $\{B_j(x)B_k(y)\}$. Thus we obtain the following system of equations

$$\int\int_R [f - D_x g_1 - D_y g_2] B_{jk}\, dx\, dy = 0, \tag{19}$$

where $j = 0, \cdots, m$; $k = 0, \cdots, n$ and B_{jk} denotes the basis function, $B_j(x)B_k(y)$. This is a system of $(m + 1) \times (n + 1)$ nonlinear equations in the $(m + 1) \times (n + 1)$ unknowns, $U_{jk}(t)$. We can rewrite (19) as

$$\begin{aligned} \int\int_R & [fB_{jk} + (D_x B_{jk})g_1 + (D_y B_{jk})g_2]\, dx\, dy \\ & = \int\int_R D_x(B_{jk}g_1) + D_y(B_{jk}g_2)\, dx\, dy. \end{aligned}$$

Applying Green's theorem to the right-hand side, we obtain

$$\begin{aligned} \int\int_R & [fB_{jk} + (D_x B_{jk})g_1 + (D_y B_{jk})g_2]\, dx\, dy \\ & = \int_\Gamma \mathbf{n}\cdot(B_{jk}g_1, B_{jk}g_2)\, ds. \end{aligned} \tag{20}$$

By the definition of the B_{jk}, the right-hand side of (19) will vanish whenever $0\ j < m$ or $0 < k < n$. However, for those splines which do not vanish on Γ, this integral will enable us to incorporate the boundary conditions (17c) into the Galerkin formulation.

The next step is to replace the integrals in (19) with simple quadrature formulas. For a detailed discussion of the technical details, see Ref. 11. This reduces (20) to an $(m + 1) \times (n + 1)$ implicit system of ordinary differential equations which may be written in the form

$$\mathbf{F}(t, \mathbf{U}, \mathbf{U}') = 0,$$

with

$$\mathbf{U}(0) = \mathbf{U}_0. \tag{21}$$

Systems of ordinary differential equations derived from such semi-discretization processes are usually *stiff*.[12] The solution of stiff equations typically requires the use of an implicit method. Any of a number of standard techniques[13] can be applied. We start with the strongly A-stable implicit Euler rule. Then, using Schryer's step-size and order monitor,[18] we extrapolate trial answers obtained by this method to obtain a powerful variable-order, variable-step time evolution.

This reduces the problem to the repeated solution of a large system of nonlinear equations. These are solved using a variant of Newton's method in which the Jacobian matrix is only infrequently recomputed. The resulting linear systems are solved with direct methods. This approach insures that the nonlinear iterations converge quickly at the expense of large storage requirements for the factored Jacobian matrix. The use of direct techniques also allows us to ignore certain mesh restrictions which might be necessary to insure the convergence of a fully iterative technique such as successive overrelaxation. As noted in Section IV, this additional robustness is particularly important during the initial transient analysis on the nonlinear diffusion problem.

REFERENCES

1. D. P. Kennedy and R. R. O'Brien, "Analysis of the Impurity Atom Distribution Near the Diffusion Mask for a Planar p-n Junction," IBM J., *9* (1965), pp. 179–186.
2. H. F. Wolfe, *Silicon Semiconductor Data*, Oxford: Pergamon Press, 1967.
3. S. M. Hu and S. Schmidt, "Interactions in Sequential Diffusion Processes in Semiconductors," J. Appl. Phys., *39* (1968), pp. 4272–4283.
4. S. M. Hu, "General Theory of Impurity Diffusion in Semiconductors via the Vacancy Mechanism," Phys. Rev., *180* (1969), pp. 773–784.
5. R. B. Fair, "Recent Advances in Implantation and Diffusion Modeling for the Design and Process Control of Bipolar ICs," Proc. Third Int. Symp. Silicon Material Science and Technology, 1977, pp. 968–987.
6. B. L. Crowder, J. F. Ziegler, F. F. Morehead, and G. W. Cole, in *Ion Implantation in Semiconductors and Other Materials*, B. L. Crowder, ed., New York: Plenum Press, 1973, p. 267.
7. R. B. Fair and J. C. C. Tsai, "Profile Parameters of Implanted-diffused Arsenic Layers In Silicon." J. Electrochem. Soc., *123* (1976), pp. 583–586.
8. R. B. Fair, "Boron Diffusion in Silicon–Concentration and Orientation Dependence, Background Effects, and Profile Estimation," J. Electrochem. Soc., *122* (1975), pp. 800–805.
9. R. B. Fair and P. N. Pappas, "Diffusion of Ion-Implanted B in High Concentration P- and As-Doped Silicon," J. Electrochem. Soc., *122* (1975), pp. 1241–1244.
10. R. B. Fair and J. C. C. Tsai, "Quantitative Models for Diffusion of Phosphorus in Silicon and Emitter Dip Effect," J. Electrochem. Soc., *124* (1977), pp. 1107–1118.
11. G. Strang and G. J. Fix. *An Analysis of the Finite Element Method*, Englewood Cliffs, N. J.: Prentice-Hall, 1973.
12. C. W. Gear, *Numerical Initial Value Problems in Ordinary Differential Equations*, Englewood Cliffs, N. J.: Prentice-Hall, 1971.
13. D. D. Warner, "The Numerical Solution of the Equations of Chemical Kinetics," J. Phys. Chem., December, 1977.
14. N. L. Schryer, "Numerical Solution of Time-Varying Partial Differential Equations in One Space Variable," Bell Laboratories Computing Science Technical Report No. 53, 1976.
15. N. C. MacDonald and T. E. Everhart, "Direct Measurement of the Depletion Layer Width Variation vs Applied Bias for a p-n Junction," Appl. Phys. Lett., *7* (1965), pp. 267–269.
16. M. R. Child, D. W. Ranasinghe and D. White, "Applications of the Scanning

Electron Microscopy in the Development of Microtechnology," Proceedings of Advanced Technology in Failure Analysis, 1977.
17. G. V. Spivak, G. V. Saparin, and L. F. Kamolova, "The Physical Fundamentals of the Resolution Enhancement in the SEM for CL and EBIC modes," Scanning Electron Microscopy, *1* (1977), pp. 191–199.
18. N. L. Schryer, "An Extrapolation Step-Size and Order Monitor for Use in Solving Differential Equations," Proc. ACM National Meeting, San Diego, 1974.