JOURNAL OF COMPUTATIONAL PHYSICS 105, 324-332 (1993)

# On an Adaptive Time Stepping Strategy for Solving Nonlinear Diffusion Equations

K. CHEN, M. J. BAINES, AND P. K. SWEBY

*Institute of Computational Fluid Dynamics, Department of Mathematics, University of Reading, Whiteknights, P.O. Box 220, Reading RG6 2AX, Berkshire, England*

Received September 10, 1991; revised August 20, 1992

A new time step selection procedure is proposed for solving nonlinear diffusion equations. It has been implemented in the ASWR finite element code of Lorenz and Svoboda [10] for 2D semiconductor process modelling diffusion equations. The strategy is based on equidistributing the local truncation errors of the numerical scheme. The use of B-splines for interpolation (as well as for the trial space) results in a banded and diagonally dominant matrix. The approximate inverse of such a matrix can be provided to a high degree of accuracy by another banded matrix, which in turn can be used to work out the *approximate* finite difference scheme corresponding to the ASWR finite element method, and further to calculate estimates of the local truncation errors of the numerical scheme. Numerical experiments on six full simulation problems arising in semiconductor process modelling have been carried out. Results show that our proposed strategy is more efficient and better conserves the total mass. 

## 1. INTRODUCTION

In the numerical solution of time-dependent partial differential equations (PDEs), suitable selection of the time step size as well as the spatial mesh size (i.e., the mesh generation) is of crucial importance in ensuring good accuracy. We shall address the first selection problem in the two-dimensional case (2D). For the second selection problem, refer to [18] and the references therein. In the 1D case a practical time stepping strategy for a particular time stepping scheme has been discussed in [1], where we presented numerical results for the solution of 1D nonlinear semiconductor diffusion equations, using both finite difference methods and finite element methods for the spatial discretization.

Here we continue our study by extending previous 1D results to 2D nonlinear equations with particular reference to semiconductor process modelling simulation. However, as is well known, diffusion equations may arise from a number of application areas, where a Fick's type law applies. See [2]. Therefore our techniques are applicable to the numerical solution of a wide class of problems. The new time stepping strategy has been implemented in the ASWR finite element code of [10] and we report on the test results of performance of the modified ASWR code on some full 2D simulation problems. The code is capable of simulating, among other processes, 2D dopant diffusion of antimony, arsenic, boron, and phosphorus for either one dopant or multiple dopants.

We first state the nonlinear diffusion equations which will be considered here. For an $r$-dopant diffusion problem in a silicon medium $\Omega$, the concentrations of dopants in $\Omega$ at time $t$ may be described by

$$\frac{\partial C_k}{\partial t} = \mathrm{Div}[D_k \,\mathrm{grad}\, C_k + Z_k C_k \,\mathrm{grad}\, \Phi], \qquad k = 1, \ldots, r, \tag{1}$$

where $C_k = C_k(x, y, t)$ is the concentration for the $k$th dopant, $D_k$ is the diffusion coefficient, $Z_k = \pm 1$ depends on the dopant used ($-1$ for singly ionized acceptors, $+1$ for donors) and $\Phi = \Phi(\sum Z_k C_k)$ is the electrostatic potential. Denote by $C = \sum_{k=1}^{r} Z_k C_k$ the total concentration. Then the potential function is calculated by $\Phi = \log(n/n_i)$, where

$$n = \tfrac{1}{2}(C + \sqrt{C^2 + 4n_i^2})$$

is the electron concentration and $n_i$ is the intrinsic electron concentration at the process temperature. See [16]. As in [10], the transformation of $f_k = \log C_k$ will convert (1) to the system

$$\frac{\partial f_k}{\partial t} = \mathscr{L}_k f_k, \qquad k = 1, \ldots, r, \tag{2}$$

where

$$\mathscr{L}_k f_k = \mathrm{Div}[D_k \,\mathrm{grad}(f_k + Z_k \Phi)] + \mathrm{grad}\, f_k \cdot \mathrm{grad}(f_k + Z_k \Phi)$$

depends on $f_1, \ldots, f_r$. It is this system which we shall solve in what follows.

The time step selection is an important step in ensuring

0021-9991/93 $5.00

efficiency of numerical methods. There are not many strategies that are readily available. In the literature on solving nonlinear partial differential equations of parabolic type, the traditional method of lines (MOL) is usually used and the time step selection is determined by methods adopted for solving the system of ordinary differential equations. See [4, 7, 8]. This means that in general we have to solve a nonlinear system of algebraic equations for a typical implicit time stepping (refer to Section 3). Even so, it is often difficult to find a robust strategy for time stepping.

Consequently, for many practical codes it is common to use fixed time step sizes or heuristic time step selection procedures. For example, see [5, 12, 13].

Here, for a particular semi-implicit time stepping scheme based on operator splitting, we propose a strategy for automatic time step selection. The idea applies in principle to other schemes as well. It is based on equidistributing local truncation errors (LTEs) in both time and space discretizations. Therefore it is readily applicable to finite difference methods since such error estimates can usually be found. For finite element methods using a B-spline basis, we shall show that good approximations to LTEs are always possible so that our proposed strategy can be applied.

Other authors [6, 7] have proposed to estimate the temporal LTEs by finite differences at each time step, comparing two numerical solutions with different time steps (i.e., using Richardson's extrapolation), and then choosing time steps so that the temporal LTE at each time step is specified (fixed). These methods differ from our approach in two aspects. First, we only compute one numerical solution at each time step instead of two. Second, we equidistribute the temporal LTE with respect to the spatial LTE rather than fixing it throughout the time step history.

In Section 2 we discuss the finite element discretization of (2). In Section 3 we introduce a three-level predictor–corrector scheme for the temporal discretization and linearization of (2) and further investigate its (linear) stability property. In Section 4 we use the idea of approximate inversion to find an explicit and approximate finite difference form of the underlying finite element method, and then carry out the local truncation error (LTE) analysis. Using the LTE estimates, we describe in Section 5 two adaptive time stepping strategies, either of which can automatically select the time step. Here the second approach is designed specifically for conservation of the total mass, which is of physical importance. Numerical experiments on six full simulation problems are presented in Section 6.

## 2. FINITE ELEMENT SOLUTION

For simplicity, let us assume that the domain of interest is the unit square region $\Omega = [0, 1] \times [0, 1]$. The method can, however, be extended to the case of more general boundaries. Divide the domain $\Omega$ into $N \times N$ boxes $\Omega_{lm} = [x_{l-1}, x_l] \times [y_{m-1}, y_m]$ $(l, m = 1, ..., N)$. Then we choose a Petrov–Galerkin finite method which uses as

**test** space,

$$S_1 = \operatorname{span}\{\phi_{lm} \mid l, m = 1, ..., N\},$$

and **trial** space,

$$S_2 = \operatorname{span}\{D_{ik} \mid i, k = 0, ..., N+1\},$$

where

$$\phi_{lm} = \begin{cases} 1 & (x, y) \in \Omega_{lm} \\ 0 & \text{otherwise,} \end{cases}$$

and the $D_{ik}$'s are chosen from blending the 1D quadratic B-splines (as defined in [1]),

$$D_{ik}(x, y) = \bar{B}_{km} B_i(x) + \bar{B}_{il} B_k(y) - \bar{B}_{km} \bar{B}_{il},$$

where

$$\bar{B}_{km} = \frac{1}{(h_y)_m} \int_{y_{m-1}}^{y_m} B_k(y)\, dy,$$

$$\bar{B}_{il} = \frac{1}{(h_x)_l} \int_{x_{l-1}}^{x_l} B_i(x)\, dx,$$

and, for $\rho = i, k$ and correspondingly $z = x, y$,

$$B_\rho(z) = \begin{cases} \dfrac{(z - z_{\rho-2})^2}{(z_\rho - z_{\rho-2})(z_{\rho-1} - z_{\rho-2})} \\ \qquad \text{for} \quad z \in [z_{\rho-2}, z_{\rho-1}] \\ \dfrac{(z - z_{\rho-2})(z_\rho - z)}{(z_\rho - z_{\rho-2})(z_\rho - z_{\rho-1})} \\ \quad + \dfrac{(z_{\rho+1} - z)(z - z_{\rho-1})}{(z_{\rho+1} - z_{\rho-1})(z_\rho - z_{\rho-1})} \\ \qquad \text{for} \quad z \in (z_{\rho-1}, z_\rho] \\ \dfrac{(z_{\rho+1} - z)^2}{(z_{\rho+1} - z_\rho)(z_{\rho+1} - z_{\rho-1})} \\ \qquad \text{for} \quad z \in (z_\rho, z_{\rho+1}]. \end{cases}$$

The resulting finite element method (called the asymmetric separated weighted residual or ASWR in [10]) finds the solution

$$F_k = \sum_{i=0}^{N+1} \sum_{j=0}^{N+1} \alpha_{k,ij} D_{ij}(x, y), \qquad \text{for} \quad k = 1, ..., r, \quad (3)$$

from $S_2$ by solving

$$\left\langle \frac{\partial F_k}{\partial t} - \mathscr{L}_k F_k, \omega \right\rangle = 0, \qquad \forall \omega \in S_1,$$

i.e.,

$$\int_{\Omega_{lm}} \left( \frac{\partial F_k}{\partial t} - \mathscr{L}_k F_k \right) dx\, dy = 0, \qquad \text{for} \quad k = 1, ..., r. \quad (4)$$

In general, (4) leads to a system of nonlinear equations for the unknowns $\{\alpha_{k,ij}\}$ $(k = 1, ..., r; i, j = 0, 1, ..., N+1)$. However, as we shall see, only a linear algebraic system needs to be solved if we use the semi-implicit time stepping scheme described below.

## 3. A SEMI-IMPLICIT TIME STEPPING SCHEME

Let us consider a model equation of the one dopant case

$$\frac{\partial f}{\partial t} = D \cdot \mathrm{Div}(\mathrm{grad}\, f) + (V_1, V_2) \cdot \mathrm{grad}\, f. \quad (5)$$

Then the generalization of the predictor corrector scheme of [1, Section 4] to the 2D case takes the following form:

$$\begin{aligned} \frac{f^{j+\theta} - f^j}{\theta\, \Delta t} &= D \cdot \mathrm{Div}(\mathrm{grad}\, f^{j+\theta}) + (V_1, V_2) \cdot \mathrm{grad}\, f^j \\ \frac{f^{j+2} - f^j}{2\Delta t} &= D \cdot \mathrm{Div}(\mathrm{grad}\, f^{j+2}) + (V_1, V_2) \cdot \mathrm{grad}\, f^{j+\theta}. \end{aligned} \quad (6)$$

The above time stepping scheme (6), when applied to Eqs. (1), leads to differential equations involving essentially the Laplacian operator which are known to be amenable to the numerical conformal mapping methods for non-rectangular domains [15]. Also at each time level, only linear solvers are needed to solve the resulting linear systems.

To carry out a stability analysis of (6), we approximate the space derivatives by finite differences leading to

$$\begin{aligned} \frac{f^{j+\theta}_{mn} - f^j_{mn}}{\theta\, \Delta t} &= D \left( \frac{\delta^2}{\Delta x^2} + \frac{\delta_y^2}{\Delta y^2} \right) f^{j+\theta}_{mn} \\ &\quad + \left( V_1 \frac{\delta_x}{2\Delta x} + V_2 \frac{\delta_y}{2\Delta y} \right) f^j_{mn} \\ \frac{f^{j+2} - f^j_{mn}}{2\Delta t} &= D \left( \frac{\delta_x^2}{\Delta x^2} + \frac{\delta_y^2}{\Delta y^2} \right) f^{j+2}_{mn} \\ &\quad + \left( V_1 \frac{\delta_x}{2\Delta x} + V_2 \frac{\delta_y}{2\Delta y} \right) f^{j+\theta}_{mn}, \end{aligned} \quad (7)$$

where the usual finite difference operators are defined by

$$\begin{aligned} \delta_x^2 f^\tau_{mn} &= f^\tau_{m+1,n} + f^\tau_{m-1,n} - 2f^\tau_{m,n}, \\ \delta_y^2 f^\tau_{mn} &= f^\tau_{m,n+1} + f^\tau_{m,n-1} - 2f^\tau_{m,n}, \\ \delta_x f^\tau_{mn} &= f^\tau_{m+1,n} - f^\tau_{m-1,n}, \\ \delta_y f^\tau_{mn} &= f^\tau_{m,n+1} - f^\tau_{m,n-1}. \end{aligned}$$

We can analyse the stability of scheme (7). Here by "stability" we mean that, given fixed $\Delta x$, $\Delta y$, and $\Delta t$, numerical errors of the time scheme will not be amplified. The proof of stability essentially consists of showing that the amplification factor is less than one (see [14] and note the presence of the intermediate time level $j + \theta$). We have

THEOREM 1 (2D stability). *The predictor–corrector scheme* (7) *is stable for any* $\theta > 0$, *provided that*

$$\Delta t \leqslant \min\left( \frac{D}{V_1^2}, \frac{D}{V_2^2} \right).$$

*Proof.* The complete proof is technical but not difficult. We shall give an outline below. Using the von Neumann's analysis (i.e., writing errors in terms of Fourier expansions on time levels $j$, $j - \theta$, and $j + 2$), we obtain the following amplification factor (by comparing levels $j$ and $j + 2$ for the Fourier exponent $(\beta, \gamma)$)

$$\mu = \frac{\mathrm{RHS}}{\mathrm{LHS}}, \quad (8)$$

where

$$\begin{aligned} \mathrm{LHS} &= 1 + 8\left( \sigma_x \sin^2 \frac{\xi}{2} + \sigma_y \sin^2 \frac{\eta}{2} \right) \\ \mathrm{RHS} &= 1 + \frac{4(\omega_x \sin \xi + \omega_y \sin \eta)\mathbf{i}}{1 + 4\theta(\sigma_x \sin^2(\xi/2) + \sigma_y \sin^2(\eta/2))} \\ &\quad - \frac{8\theta(\omega_x \sin \xi + \omega_y \sin \eta)^2}{1 + 4\theta(\sigma_x \sin^2(\xi/2) + \sigma_y \sin^2(\eta/2))} \end{aligned}$$

with $\mathbf{i} = \sqrt{-1}$, and

$$\begin{aligned} \xi &= \beta\, \Delta x, & \eta &= \gamma\, \Delta y, \\ \sigma_x &= D\, \Delta t/\Delta x^2, & \sigma_y &= D\, \Delta t/\Delta y^2, \\ \omega_x &= V_1\, \Delta t/(2\Delta x), & \omega_y &= V_2\, \Delta t/(2\Delta y). \end{aligned}$$

We shall try to show that $|\mu|^2 < 1$. Let us denote by $A = \sigma_x \sin^2(\xi/2) + \sigma_y \sin^2(\eta/2)$ and $B = (\omega_x \sin \xi + \omega_y \sin \eta)^2$. Then it can be shown that

$$4\theta^2(B - A) < \frac{A}{B} + (\theta - 1)$$

if $\Delta t \leqslant \min(D/V_1^2, D/V_2^2)$. On the other hand, $|\mu|^2 < 1$ if

$$4\theta^2(B-A) < \frac{A}{B} + (\theta - 1) + 4(2\theta + 1)\frac{A^2}{B} + 16\theta(\theta + 2)\frac{A^3}{B} + 64\theta^2\frac{A^4}{B}.$$

Therefore we have $|\mu| < 1$ under the sufficient condition stated and $\theta > 0$. The proof is complete. ▮

## 4. LOCAL TRUNCATION ERROR ANALYSIS

We shall now apply the finite element method of Section 2 to the model equation (5) combined with the time stepping scheme (6) and further analyze the local truncation error. Our direct analysis will also reveal the approximate finite difference structure for the ASWR finite element method. Estimates of truncation errors can be used in designing a practical time step selection strategy (see Section 5).

Integrating the corrector equation of (6) over the box $\Omega_{lm}$, we obtain the ASWR equation

$$\frac{1}{2\Delta t}\int_{\Omega_{lm}} (f^{j+2} - f^j)\,dx\,dy = D\int_{\Omega_{lm}} \mathrm{Div}(\mathrm{grad}\, f^{j+2})\,dx\,dy + \int_{\Omega_{lm}} (V_1, V_2)\,\mathrm{grad}\, f^{j+\theta}\,dx\,dy, \tag{9}$$

where $f^\tau = \sum \alpha_{i,k}^\tau D_{ik}(x, y)$. Taking into consideration the piecewise behaviour of $D_{ik}$, we obtain

$$\sum_{i=l-1}^{l+1}\sum_{k=m-1}^{m+1} \frac{\alpha_{i,k}^{j+2} - \alpha_{i,k}^j}{2\Delta t}\int_{\Omega_{lm}} D_{ik}(x, y)\,dx\,dy = D\sum_{i=l-1}^{l+1}\sum_{k=m-1}^{m+1} \alpha_{i,k}^{j+2}\int_{\Omega_{lm}} \mathrm{Div}(\mathrm{grad}\, D_{ik})\,dx\,dy + \sum_{i=l-1}^{l+1}\sum_{k=m-1}^{m+1} \alpha_{i,k}^{j+\theta}\int_{\Omega_{lm}} (V_1, V_2) \times \mathrm{grad}\, D_{ik}\,dx\,dy. \tag{10}$$

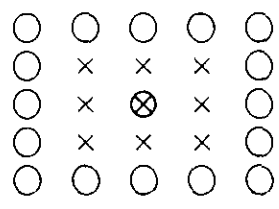

FIG. 1. Nine-point stencil for variable $\alpha$.

Now direct calculations of the above coefficients simplify (10) to

$$\begin{aligned}
&\frac{\delta_t}{72\Delta t}\left[\alpha_{l-1,m-1}^{j+2} + \alpha_{l+1,m-1}^{j+2} + \alpha_{l-1,m+1}^{j+2} + \alpha_{l+1,m+1}^{j+2} + 4(\alpha_{l,m-1}^{j+2} + \alpha_{l,m+1}^{j+2} + \alpha_{l-1,m}^{j+2} + \alpha_{l+1,m}^{j+2}) + 16\alpha_{l,m}^{j+2}\right] \\
&= \frac{D}{3\Delta x^2}\left[\alpha_{l-1,m-1}^{j+2} + \alpha_{l+1,m-1}^{j+2} + \alpha_{l-1,m+1}^{j+2} + \alpha_{l+1,m+1}^{j+2} + \alpha_{l,m-1}^{j+2} + \alpha_{l,m+1}^{j+2} + \alpha_{l-1,m}^{j+2} + \alpha_{l+1,m}^{j+2} - 8\alpha_{l,m}^{j+2}\right] \\
&\quad + \frac{V_1}{2\Delta x}\left[(\alpha_{l+1,m+1}^{j+\theta} - \alpha_{l-1,m+1}^{j+\theta}) + (\alpha_{l+1,m-1}^{j+\theta} - \alpha_{l-1,m-1}^{j+\theta}) + 4(\alpha_{l+1,m}^{j+\theta} - \alpha_{l-1,m}^{j+\theta})\right] \\
&\quad + \frac{V_2}{2\Delta y}\left[(\alpha_{l+1,m+1}^{j+\theta} - \alpha_{l+1,m-1}^{j+\theta}) + (\alpha_{l-1,m+1}^{j+\theta} - \alpha_{l-1,m-1}^{j+\theta}) + 4(\alpha_{l,m+1}^{j+\theta} - \alpha_{l,m-1}^{j+\theta})\right],
\end{aligned} \tag{11}$$

where $\delta_t \beta^{j+2} = \beta^{j+2} - \beta^j$. The computational stencil of nine points is shown in Fig. 1. To find the local truncation error for (11), we may use the Taylor theorem. Unfortunately the unknowns ($\alpha$'s) are not the solution functions ($f$). They are only related to the solution by the following interpolation relationship

$$\sum_{i=l-1}^{l+1}\sum_{k=m-1}^{m+1} \alpha_{i,k}^\tau D_{ik}(\bar{x}_l, \bar{y}_m) = f(t_\tau, \bar{x}_l, \bar{y}_m) \tag{12}$$

where $(\bar{x}_l, \bar{y}_m)$ is taken to be the centre point of box $\Omega_{lm}$ for $l,\ m = 1, \ldots, N$. For the Dirichlet boundary condition $f(t, x, y)|_{\partial\Omega} = 0$, the above system may be written as

$$F\alpha^\tau \equiv \frac{1}{72}\begin{pmatrix} U_0 & U_1 & & & \\ U_1 & U_2 & U_1 & & \\ & \ddots & \ddots & \ddots & \\ & & U_1 & U_2 & U_1 \\ & & & U_1 & U_0 \end{pmatrix}\begin{pmatrix} \alpha_{1,1}^\tau \\ \alpha_{1,2}^\tau \\ \vdots \\ \alpha_{N,N}^\tau \end{pmatrix} = \begin{pmatrix} f_{1,1}^\tau \\ f_{1,2}^\tau \\ \vdots \\ f_{N,N}^\tau \end{pmatrix}, \tag{13}$$

where

$$U_1 = \begin{pmatrix} 6 & 1 & & & \\ 1 & 7 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 7 & 1 \\ & & & 1 & 6 \end{pmatrix},$$

$$U_2=\begin{pmatrix} 23 & 7 & & & \\ 7 & 40 & 7 & & \\ & \ddots & \ddots & \ddots & \\ & & 7 & 40 & 7 \\ & & & 7 & 23 \end{pmatrix},$$

$$U_0=\begin{pmatrix} 27 & 6 & & & \\ 6 & 33 & 6 & & \\ & \ddots & \ddots & \ddots & \\ & & 6 & 33 & 6 \\ & & & 6 & 27 \end{pmatrix}.$$

The matrix $F$ can be seen to be diagonally diagonally dominant and in block tridiagonal form. The elements of the inverse of such a matrix are known to decay exponentially to zero in magnitude away from the main diagonal. Refer to [3, 11].

By calculating the inverse $F^{-1}$ numerically, we can express the unknowns $\alpha^{\tau}_{i,k}$'s in terms of the solution (RHS vector) $f^{\tau}_{i,k}$'s. Here we propose to use a band approximation of $F^{-1}$ to solve Eq. (13) by neglecting those subdiagonal entries of $F^{-1}$ away from the main diagonal which are less than some small positive number $\varepsilon$ in magnitude. This enables us to express accurately the $\alpha^{\tau}_{i,k}$'s in terms of local $f^{\tau}_{i,k}$'s. For (13), such an approximation is given by (taking $\varepsilon=10^{-2}$)

$$\begin{aligned}\alpha^{\tau}_{l,m}=10^{-4}[&-133(f^{\tau}_{l,m-3}+f^{\tau}_{l,m+3}+f^{\tau}_{l-3,m}+f^{\tau}_{l+3,m})\\ &-185(f^{\tau}_{l-1,m-2}+f^{\tau}_{l-1,m+2}+f^{\tau}_{l+1,m-2}\\ &+f^{\tau}_{l+1,m+2}+f^{\tau}_{l-2,m-1}+f^{\tau}_{l-2,m+1}\\ &+f^{\tau}_{l+2,m-1}+f^{\tau}_{l+2,m+1})\\ &+831(f^{\tau}_{l,m-2}+f^{\tau}_{l,m+2}+f^{\tau}_{l-2,m}+f^{\tau}_{l+2,m})\\ &+852(f^{\tau}_{l-1,m-1}+f^{\tau}_{l-1,m+1}\\ &+f^{\tau}_{l+1,m-1}+f^{\tau}_{l+1,m+1})\\ &-3830(f^{\tau}_{l,m-1}+f^{\tau}_{l,m+1}+f^{\tau}_{l-1,m}+f^{\tau}_{l+1,m})\\ &+20600f^{\tau}_{l,m}. \end{aligned} \tag{14}$$

The computational stencil of $25f$ points for each $\alpha^{\tau}_{l,m}$ is shown in Fig. 2. The availability of $\alpha^{\tau}_{i,k}$'s (for $\tau=j$, $j+\theta, j+2$) from (14) allows us to work out the local truncation error of (11) in a straightforward manner. This has been found to be

$$\begin{aligned}\tau(t_j,\bar{x}_l,\bar{y}_m)=&R_0\,\Delta t+R_{11}\,\Delta x^2+R_{12}\,\Delta x\,\Delta y\\ &+R_{22}\,\Delta y^2+\cdots,\end{aligned} \tag{15}$$

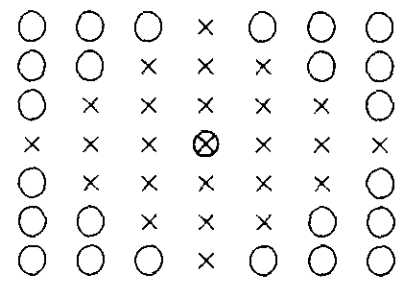

FIG. 2. Twenty-five-point stencil for variable $f$ (with each $\alpha$).

where

$$R_0=E_1-\theta E_2;$$

$$E_1=\left(V_1\frac{\partial}{\partial x}+V_2\frac{\partial}{\partial y}\right)^2 f-D^2\nabla^2(\nabla^2 f);$$

$$E_2=\left(V_1\frac{\partial}{\partial x}+V_2\frac{\partial}{\partial y}\right)^2 f-D^2\left(V_1\frac{\partial}{\partial x}+V_2\frac{\partial}{\partial y}\right)\nabla^2 f;$$

$$R_{11}=\frac{D}{6}\frac{\partial^2}{\partial x^2}\nabla^2 f-\frac{D}{12}\frac{\partial^4 f}{\partial x^4};$$

$$R_{12}=-\frac{D}{3}\frac{\partial^4 f}{\partial x^2\,\partial y^2};$$

$$R_{22}=\frac{D}{6}\frac{\partial^2}{\partial y^2}\nabla^2 f-\frac{D}{12}\frac{\partial^4 f}{\partial y^4}.$$

In a typical semiconductor diffusion model, the potential term $\Phi$ is also present and the equation corresponding to (5) is of the following form:

$$\begin{aligned}\frac{\partial f}{\partial t}=&D\cdot\mathrm{Div}(\mathrm{grad}(f+Z\Phi))\\ &+[(\bar{V}_1,\bar{V}_2)+ZD\ \mathrm{grad}\ \Phi]\ \mathrm{grad}\, f.\end{aligned} \tag{16}$$

The above equation may be identified with (5) with the perturbation of $f\to f+Z\Phi$ and $V_1=\bar{V}_1+ZD(\partial\Phi/\partial x)$ and $V_2=\bar{V}_2+ZD(\partial\Phi/\partial y)$. Therefore the time step selection strategies of the next section can be used for solving Eq. (16). In order to apply our model analysis to the full nonlinear equation

$$\begin{aligned}\frac{\partial f}{\partial t}=&\mathrm{Div}[D\ \mathrm{grad}(f+Z\Phi)]\\ &+D\ \mathrm{grad}\, f\cdot\mathrm{grad}(f+Z\Phi),\end{aligned} \tag{17}$$

local linearizations have to be introduced. For example, the diffusion coefficient $D$ may be viewed as locally constant

and $(V_1, V_2) = D \operatorname{grad}(f + Z\Phi)$ so that results of the linear analysis are applicable. We next consider the selection of the time step $\Delta t$ based on these truncation error estimates.

## 5. TIME STEP SELECTION

The automatic selection of the time step $\Delta t$ follows the principle that the temporal error ($\Delta t$ term) should be of comparable magnitude to the spatial discretization error ($\Delta x$ and $\Delta y$ terms). In this way, the overall error is determined only by the spatial discretization accuracy.

Apart from using the errors in certain equation variables, we shall also consider the relative mass balance error. This quantity is defined at time $t$ by

$$Q_t = \frac{(M_t - M_0)}{M_0}, \tag{18}$$

where $M_0$ and $M_t$ denote the total mass present in the silicon at times 0 and $t$, respectively. Here the total mass at time $t = t_j$ based on the numerical solution $f^j$, is defined by

$$M_t = \int_\Omega C^j(x, y)\, dx\, dy \equiv \int_\Omega \exp[f^j(x, y)]\, dx\, dy. \tag{19}$$

Ideally we wish to have *mass conservation* of the total mass $M_t$, which is of physical importance. By this we mean that $M_t = M_0$ at any time $t$. If the underlying PDE is in conservation form such as (1) in $C$, then most numerical methods (including the Euler scheme) for solving such an equation will conserve the total mass. But our transformed PDE (2) in $f$ is not in conservation form. Conservation of mass in $f$ as well as in $C = \exp(f)$ cannot be maintained, in general. We hope to conserve the mass as much as possible numerically by suitably choosing the time step $\Delta t$, given a priori spatial discretizations.

### 5.1. *Error Control in f*

This approach, as introduced in [1, Section 4.4], is to use the transformed equation (5) in the variable $f$. For Eq. (5), the time step selection based on (15) should satisfy

$$\Delta t \leqslant \mathrm{TOL}/\max |R_0|, \tag{20}$$

where TOL is chosen to be proportional to the estimate of $|R_{11}\, \Delta x^2 + R_{12}\, \Delta x\, \Delta y + R_{22}\, \Delta y^2|$. Of course, the above choice should also be subject to the stability condition of Theorem 1 being satisfied.

### 5.2. *Mass Error Control in C*

The idea here is first to relate the local truncation error of Section 4 to the dopant mass error, and then to identify contributions from temporal and spatial discretizations in such a mass error in order to select an appropriate time step.

Assume that stability is satisfied throughout our calculations. Then the following relation holds for the global error

$$f - f^{j+2} = K\tau + O(\Delta t^2), \tag{21}$$

where $f = f(x, y, t_{j+2})$ and $f^{j+2}$ represent respectively the exact and numerical solution at time level $j+2$, $K$ is a stability constant, $\tau$ is the local truncation error (see (15)), and $O(\Delta t^2)$ denotes the negligible high order error terms. Note from Section that $C = \exp(f)$ and $C^{j+2} = \exp(f^{j+2})$. Taking the exponential of both sides of (21) gives rise to

$$C = C^{j+2} \exp[K\tau + O(\Delta t^2)]. \tag{22}$$

Now integrate both sides of (22) over the entire domain to obtain

$$M_0 = M_t + \int_\Omega \exp[f^{j+2}] \times \{\exp[K\tau + O(\Delta t^2)] - 1\}\, dx\, dy, \tag{23}$$

where we have used the definitions of (18) and (19).

We shall try to conserve the total mass, i.e., to minimise $M_0 - M_t$ while allowing the largest possible time step $\Delta t$ for efficiency. Define the mass error (ME) by

$$\mathrm{ME} = M_0 - M_t \equiv \int_\Omega \exp[f^{j+2}] \times \{\exp[K\tau + O(\Delta t^2)] - 1\}\, dx\, dy \tag{24}$$

whose leading term on expansion is

$$\mathrm{ME}_1 = K \int_\Omega \exp[f^j](R_0\, \Delta t + \mathrm{SE})\, dx\, dy, \tag{25}$$

where substitution of (15) has been performed and the spatial error is denoted by $\mathrm{SE} = R_{11}\, \Delta x^2 + R_{12}\, \Delta x\, \Delta y + R_{22}\, \Delta y^2$. More specifically, Eq. (25) can be rewritten as

$$\mathrm{ME}_1 = K(\mathrm{ME}_t + \mathrm{ME}_s), \tag{26}$$

with $\mathrm{ME}_t = \Delta t \int_\Omega \exp[f^j]\, R_0\, dx\, dy$ and $\mathrm{ME}_s = \int_\Omega \exp[f^j]\, \mathrm{SE}\, dy\, dy$ representing mass error contributions from $t$emporal and $s$patial discretizations, respectively.

The strategy based on $C$ mass error control is therefore to equidistribute the total mass error ME in time and space by forcing

$$\mathrm{ME}_t = \mathrm{ME}_s,$$

i.e., by selecting the time step

$$\Delta t = \int_\Omega \exp[f^j]\ \mathrm{SE}\ dx\ dy \Big/ \int_\Omega \exp[f^j]\ R_0\ dx\ dy. \quad (27)$$

### 5.3. *Choice of the Predictor Step*

The choice of the predictor step may be arbitrary as far as the stability of the predictor corrector scheme (6) is concerned (see Theorem 1). The idea here is to minimise in some sense a measure of the solution error. We have considered two approaches.

The first approach is based on minimization of the local truncation error. As $\theta$ is a constant appearing in the pointwise local truncation error estimates, we cast the problem of choosing $\theta$ as a minimisation problem in the least squares sense. This minimization is only for $|R_0|$, since the minimization of $\tau$ of (15) would require a knowledge of $\Delta t$ that we do not have at this stage. Let us rewrite $R_0$ as $a_{jm} - \theta b_{jm}$ at point $(\bar{x}_j, \bar{y}_m)$, where $a_{jm} = E_1$ and $b_{jm} = E_2$ (see (15)). Construct a quadratic function of $\theta$ as

$$\begin{aligned} P(\theta) &= \sum_{m=1}^{N} \sum_{j=1}^{N} (a_{jm} - \theta b_{jm})^2 \\ &= \theta^2 \left( \sum_{m=1}^{N} \sum_{j=1}^{N} b_{jm}^2 \right) - 2\theta \left( \sum_{m=1}^{N} \sum_{j=1}^{N} a_{jm} b_{jm} \right) \\ &\quad + \sum_{m=1}^{N} \sum_{j=1}^{N} a_{jm}^2. \end{aligned} \quad (28)$$

The solution of minimizing $P(\theta)$ is generally given by

$$\theta^* = \sum_{m=1}^{N} \sum_{j=1}^{N} a_{jm} b_{jm} \Big/ \sum_{m=1}^{N} \sum_{j=1}^{N} b_{jm}^2. \quad (29)$$

However, we are only interested in positive $\theta^*$ which is only possible if $\sum_{m=1}^{N} \sum_{j=1}^{N} a_{jm} b_{jm} > 0$. Otherwise we have to look for the smallest positive $\theta$ which minimizes $|R_0|$ pointwise or simply fix $\theta$. If the choice of fixed $\theta = 1$ is used, then the time stepping scheme resembles the more conventional three-level scheme (refer to [6, 14]).

The second approach is to track the mass balance history and dynamically adjust $\theta$ from step to step in order to achieve mass conservation; see [10]. To illustrate, let us define at time $t = t_j$ the *approximate* total mass in the silicon medium $\Omega$ by

$$m_j = \sum_{l=1}^{N} \sum_{m=1}^{N} \exp[\bar{x}_l, \bar{y}_m, f(t_j)]\ \Delta x_l\ \Delta y_m \quad (30)$$

and the relative $j$th step mass error by

$$B_j = (m_j - m_{j-1})/m_{j-1}, \quad (31)$$

where $m_j$ is the result of (19) applied with the mid-point quadrature rule. Using $\theta_0 = 1$, it has been proposed in [10] to use

$$\theta_j = \theta_{j-1}(1 + 100B_j) \quad (32)$$

if $B_j > 0$ and $B_j > B_{j-1}$ and

$$\theta_j = \theta_{j-1}[1 - 100(B_{j-1} - B_j)] \quad (33)$$

if $B_j > 0$ and $B_{j-1} - B_j < B_j/3$. Similar choices are made for the case of $B_j < 0$; otherwise set $\theta_j = \theta_{j-1}$.

In the next section, we experiment on our time step selection strategy using both the minimization and mass balance choices for the predictor step.

## 6. NUMERICAL EXPERIMENTS

We have taken four typical semiconductor processing structures as shown in Figs. 3–6, where the physical sizes are measured in micrometers. The initial profiles for our test problems are obtained from the ion implantation menu of COMPOSITE [9], which are Pearson IV distributions. Detailed data specifications are shown in Table I, where Tests 5 and 6 use two dopants.

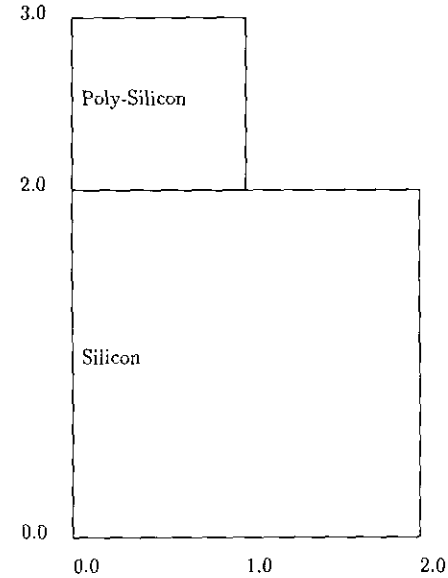


**FIG. 3.** Substrate for Test Examples 1 and 2.

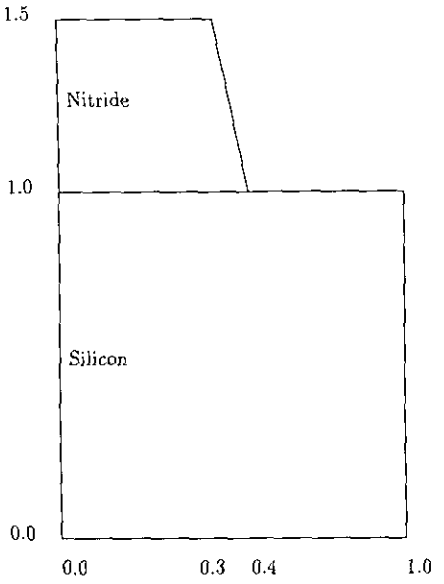


FIG. 4. Substrate for Test Example 3.

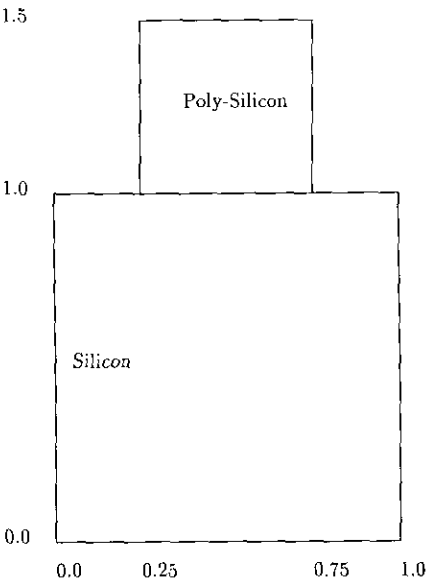


FIG. 6. Substrate for Test Examples 5 and 6.

We have run all six test examples with both the existing ASWR method of [10] and our modified version of the ASWR method. With our modified version, we have used the following strategies:

- For boron/phosphorus/antimony implants, the selection method of Section 5.1 is used;
- For arsenic implant, the selection method of Section 5.2 is used;
- The predictor step uses the second approach of Section 5.3.

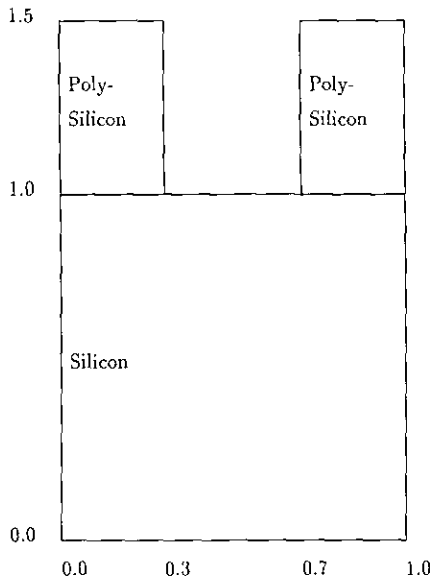


FIG. 5. Substrate for Test Example 4.

We remark that the above combination of ideas offers a robust method from experimental observations, although it gives by no means the best results in some cases. The numerical results are summarised in Table II, where information on the number of time steps taken, the CPU user time on a Sun-sparc 1 and the mass balance error $Q_r$ is given. The mesh data given are for the finite difference (FD) mesh, and the number of boxes for the finite element (FE) discretization is about half of the number of FDM mesh lines in both directions. For example, a $65 \times 65$ FD mesh corresponds to $[(65-1)/2] \times [(65-1)/2]$ FE boxes.

Results have clearly demonstrated that our modified time stepping strategy generally shows better performance compared with the existing ASWR strategy, which uses heuristic time stepping ideas (see [10]). Our method of estimating the local truncation errors for the ASWR Petrov–Galerkin finite element method with B-splines appears to be new and simple.

**TABLE I**

*Ion Implantation and Diffusion Data*

| Test | Dopants | Energy (Kev) | Dose ($cm^{-2}$) | Temperature/ | Time (Min) |
|---|---|---|---|---|---|
| 1 | Boron | 30 | 1.0E12 | 1000 | 40 |
| 2 | Phosphorus | 50 | 1.0E15 | 1000 | 50 |
| 3 | Antimony | 25 | 1.0E14 | 1100 | 2 |
| 4 | Arsenic | 40 | 1.0E14 | 1100 | 2 |
| 5 | Boron | 30 | 1.0E13 | 1100 | 1 |
| | Phosphorus | 25 | 1.0E12 | | |
| 6 | Boron | 20 | 1.0E12 | 1000 | 5 |
| | Phosphorus | 50 | 1.0E15 | | |

**TABLE II**

ASWR Test Results of Six Examples

| Method | Test | Mesh | Time steps | CPU | Error $Q_i$ |
|---|---|---|---|---|---|
| Existing | 1 | $85 \times 85$ | 30 | 535 | 1.5E − 3 |
| ASWR | 2 | $65 \times 65$ | 123 | 1290 | −7.6E − 4 |
| | 3 | $65 \times 65$ | 86 | 971 | −1.3E − 3 |
| | 4 | $129 \times 129$ | 69 | 3000 | −3.3E − 4 |
| | 5 | $85 \times 85$ | 62 | 2000 | −1.0E − 3 |
| | | | | | −5.0E − 2 |
| | 6 | $65 \times 65$ | 50 | 992 | −1.1E − 2 |
| | | | | | 1.9E − 3 |
| Modified | 1 | $85 \times 85$ | 13 | 277 | 2.8E − 3 |
| ASWR | 2 | $65 \times 65$ | 43 | 574 | 2.4E − 5 |
| | 3 | $65 \times 65$ | 28 | 428 | −1.9E − 4 |
| | 4 | $129 \times 129$ | 59 | 2890 | −1.5E − 4 |
| | 5 | $85 \times 85$ | 22 | 1150 | 2.0E − 3 |
| | | | | | 1.5E − 3 |
| | 6 | $65 \times 65$ | 16 | 443 | −4.7E − 3 |
| | | | | | −5.7E − 4 |

## ACKNOWLEDGMENTS

The authors thank the anonymous referees for some helpful comments. Financial support from the Department of Trade and Industry (DTI) through the Science and Engineering Research Council (SERC) is gratefully acknowledged.

## REFERENCES

1. K. Chen, M. J. Baines, and P. K. Sweby, Numerical Analysis Report 10/91, Dept. of Mathematics, University of Reading, Reading, England, 1991; submitted.
2. J. Crank, *The Mathematics of Diffusion*, 2nd ed. Oxford Univ. Press, London, 1975.
3. S. Demko, W. F. Moss, and P. W. Smith, *Math. Comput.* **43** (188), 491 (1984).
4. K. Eriksson and C. Johnson, *SIAM J. Numer. Anal.* **24** (1), 12 (1987).
5. J. P. Kreskovsky and H. L. Grubin, *J. Comput. Phys.* **68**, 420 (1987).
6. M. R. Kump and R. W. Dutton, *IEEE Trans. Computer Aided Design* **CAD-7** (2), 191 (1988).
7. M. E. Law and R. W. Dutton, *IEEE Trans. Computer Aided Design* **CAD-7** (2), 181 (1988).
8. J. D. Lambert, *Computational Methods in Ordinary Differential Equations* (Wiley, London, 1973).
9. J. Lorenz, J. Pelka, H. Ryssel, A. Sach, A. Seidel, and M. Svoboda, *IEEE Trans. Electron Devices* **ED-32** (10), 1977 (1985).
10. J. Lorenz and M. Svoboda, in *Simulation of Semiconductor Devices and Processes*, Vol. 3, edited by G. Daccarani *et al.* (Tecnoprint, Italy, 1988), p. 243.
11. G. Meurant, *SIAM J. Appl. Math.*, to appear.
12. R. R. O'Brien, C. M. Hsieh, J. S. Moore, R. F. Lever, R. F. Murley, K. W. Brannon, G. R. Srinivasan, and R. W. Knepper, *IBM J. Res. Dev.* **29** (3) (1985).
13. S. J. Polak, W. H. A. Schilders, and P. Markowich, *Int. J. Numer. Methods Eng.* **24**, 763 (1987).
14. R. D. Richtmyer and K. W. Morton, *Difference Methods for Initial Value Problems*, 2nd ed., (Wiley, New York, 1967).
15. A. Seidl and M. Svoboda, *IEEE Tran. Electron Devices* **ED-32** (10), 1960 (1985).
16. S. Selberherr, *Analysis and Simulation of Semiconductor Devices* (Springer-Verlag, New York/Berlin, 1984).
17. M. Svoboda, Ph. D. thesis, Faculty of Mathematics, Ludwig Maximillians University, Munich, Germany, 1988 (unpublished).
18. J. F. Thompson, Z. U. A. Warsi, and C. W. Mastin, *Numerical Grid Generation: Foundations and Applications* (North-Holland, Amsterdam, 1985).