OPTIMAL CONVOLUTION SOR ACCELERATION OF WAVEFORM RELAXATION WITH APPLICATION TO SEMICONDUCTOR DEVICE SIMULATION

Mark Reichelt
Research Laboratory of Electronics, Massachusetts Institute of Technology
Cambridge, MA

SUMMARY

In this paper we describe a novel generalized SOR algorithm for accelerating the convergence of the dynamic iteration method known as waveform relaxation. A new convolution SOR algorithm is presented, along with a theorem for determining the optimal convolution SOR parameter. Both analytic and experimental results are given to demonstrate that the convergence of the convolution SOR algorithm is substantially faster than that of the more obvious frequency-independent waveform SOR algorithm. Finally, to demonstrate the general applicability of this new method, it is used to solve the differential-algebraic system generated by spatial discretization of the time-dependent semiconductor device equations.

INTRODUCTION

To achieve highest performance on a parallel computer, a numerical algorithm must avoid frequent parallel synchronization [1]. The waveform relaxation approach to solving time-dependent initial-value problems is just such a method, as the iterates are waveforms over an interval, rather than single timepoints [2, 3, 4]. Like any relaxation scheme, efficiency depends on rapid convergence, and there have been several investigations into how to accelerate WR [2, 5], including using multigrid [6] and conjugate direction techniques [7].

In this paper, we investigate using successive overrelaxation (SOR) to accelerate WR convergence. In particular, we show that the pessimistic results about waveform SOR derived in [2] can be substantially improved by replacing multiplication with a fixed SOR parameter by convolution with an SOR kernel. We derive the optimal SOR kernel using

* This work was supervised by Professors Jacob White and Jonathan Allen and supported by a grant from IBM, the Defense Advanced Research Projects Agency contract N00014-91-J-1698, and the National Science Foundation contract MIP-8858764 A02.

Fourier analysis techniques and then demonstrate the effectiveness of the approach for a model parabolic problem. Finally, we demonstrate the general applicability of the approach by using the method to solve the time-dependent drift-diffusion equations associated with modeling semiconductor devices.

We begin in Section 2 by reviewing waveform SOR, and in Section 3 we relate the algorithm to pointwise SOR to demonstrate the difficulty in accelerating WR with a fixed SOR parameter. In Section 4, we use Fourier analysis to derive the SOR kernel for the continuous WR algorithm, and give a proof of optimality. In Section 5 we briefly consider the effect of time-discretization, and in Section 6 we apply the method to device simulation. Finally, conclusions and acknowledgements are given in Section 7.

Waveform SOR

In this section, we consider applying waveform relaxation methods to the model linear initial-value problem

$$\left(\tfrac{d}{dt} + \boldsymbol{A}\right) \boldsymbol{x}(t) = \boldsymbol{b}(t) \quad \text{with} \quad \boldsymbol{x}(0) = \boldsymbol{x}_0, \tag{1}$$

where $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, $\boldsymbol{b}(t) \in \mathbb{R}^n$ is a given time-dependent right-hand side vector, $\boldsymbol{x}(t) \in \mathbb{R}^n$ is the unknown vector to be computed over simulation interval $t \in [0, T]$, and $\boldsymbol{x}_0 \in \mathbb{R}^n$ is an initial condition.

Given the relaxation splitting $\boldsymbol{A} = \boldsymbol{D} - \boldsymbol{L} - \boldsymbol{U}$, and subtracting successive waveform relaxation iterations, the waveform Gauss-Jacobi (WGJ) and waveform Gauss-Seidel (WGS) iteration equations, respectively, may be written as:

$$\left(\tfrac{d}{dt} + \boldsymbol{D}\right) \Delta\boldsymbol{x}^{k+1}(t) = (\boldsymbol{L} + \boldsymbol{U})\, \Delta\boldsymbol{x}^{k}(t) \tag{2}$$

$$\left(\tfrac{d}{dt} + \boldsymbol{D} - \boldsymbol{L}\right) \Delta\boldsymbol{x}^{k+1}(t) = \boldsymbol{U}\, \Delta\boldsymbol{x}^{k}(t), \tag{3}$$

where $\Delta\boldsymbol{x}^{k+1}(t) = \boldsymbol{x}^{k+1}(t) - \boldsymbol{x}^{k}(t)$ is used to eliminate the right hand side $\boldsymbol{b}(t)$.

The waveform SOR method for acceleration of WGS is a simple extension of algebraic SOR. To derive the waveform SOR iteration equation, compute a waveform $\hat{x}_i^{k+1}(t)$ on $t \in [0, T]$, as in WGS:

$$\left(\tfrac{d}{dt} + a_{ii}\right) \hat{x}_i^{k+1}(t) = b_i(t) - \sum_{j=1}^{i-1} a_{ij} x_i^{k+1}(t) - \sum_{j=i+1}^{n} a_{ij} x_i^{k}(t) \quad \text{with} \quad \hat{x}_i^{k+1}(0) = x_0, \tag{4}$$

and then update $x_i^k(t)$ in the iteration direction by multiplication with an overrelaxation parameter ω,

$$x_i^{k+1}(t) \leftarrow x_i^k(t) + \omega \cdot \left[\hat{x}_i^{k+1}(t) - x_i^k(t)\right]. \tag{5}$$

Combining equations (4) and (5) yields

$$\left(\tfrac{d}{dt} + a_{ii}\right) x_i^{k+1}(t) =$$

$$(6) \qquad (1-\omega)\left[\left(\tfrac{d}{dt} + a_{ii}\right) x_i^k(t)\right] + \omega \left[b_i(t) - \sum_{j=1}^{i-1} a_{ij} x_i^{k+1}(t) - \sum_{j=i+1}^{n} a_{ij} x_i^k(t) \right],$$

which, after subtracting successive waveform relaxation iterations, leads to

$$(7) \qquad (\tfrac{d}{dt} + D - \omega L)\, \Delta x^{k+1}(t) = [(1-\omega)(\tfrac{d}{dt} + D) + \omega U]\, \Delta x^k(t),$$

where $\Delta x^{k+1}(t) = x^{k+1}(t) - x^k(t)$.

Note that the iteration matrices implied by equations (2), (3) and (7) correspond exactly to the standard algebraic relaxation and SOR matrices with diagonal matrix D replaced by $(\frac{d}{dt} + D)$. Also note that waveform SOR as defined by (7) is *not* the same as the dynamic SOR iteration considered in [2], because, unlike WGJ or WGS, the waveform SOR iteration equations are not of the form

$$(8) \qquad \tfrac{d}{dt} \Delta x^{k+1} + M \Delta x^{k+1} = N \Delta x^k$$

where $M, N \in \mathbb{R}^{n \times n}$.

Relation to Pointwise SOR

Discretizing (1) in time using a multistep integration method yields

$$(9) \qquad \sum_{j=0}^{s} \alpha_j x[m-j] = h \sum_{j=0}^{s} \beta_j \left(b[m-j] - A x[m-j] \right),$$

where $\alpha_0 = 1$ and $x[m]$ denotes $x(t)$ at timepoint $t = mh$ with timestep h. Thus, the time-discretized model problem can be rewritten as a sequence of linear algebraic problems

$$\left[I + h\beta_0 A\right] x[m] =$$

$$(10) \qquad h\beta_0 b[m] - \sum_{j=1}^{s} \alpha_j x[m-j] + h \sum_{j=1}^{s} \beta_j \left(b[m-j] - A x[m-j] \right).$$

We now compare the convergence of the waveform SOR method to the convergence of pointwise SOR, in which algebraic SOR is used to solve the matrix problem at each timepoint.

The pointwise SOR iteration equations are derived by applying the relaxation splitting $A = D - L - U$ to equation (10) and taking the difference between the $(k+1)$st and kth

iterations. More precisely, the pointwise SOR iteration equation applied to solve (10) for $\Delta x^{k+1}[m] = x^{k+1}[m] - x^k[m]$ is

$$\begin{aligned} &\Big[(I + h\beta_0 D) - \omega h\beta_0 L\Big]\, \Delta x^{k+1}[m] = \\ &\qquad \Big[(1-\omega)(I + h\beta_0 D) + \omega h\beta_0 U\Big]\, \Delta x^k[m], \end{aligned} \tag{11}$$

where ω is the SOR parameter. It follows that the spectral radius of the iteration matrix generated by pointwise SOR at the mth timestep is

$$\rho\Big(\Big[(I + h\beta_0 D) - \omega h\beta_0 L\Big]^{-1}\Big[(1-\omega)(I + h\beta_0 D) + \omega h\beta_0 U\Big]\Big). \tag{12}$$

If waveform SOR is used to solve the model problem (1), and a multistep method is used to solve iteration equation (7), then $\Delta x^{k+1}[m]$, now denoting the discretized difference in waveform iterates, satisfies

$$\begin{aligned} &\sum_{j=0}^{s} \alpha_j \Big[\Delta x^{k+1}[m-j] - (1-\omega)\,\Delta x^k[m-j]\Big] = \\ &\qquad h\sum_{j=0}^{s} \beta_j \Big\{ -(D - \omega L)\,\Delta x^{k+1}[m-j] + [(1-\omega)\,D + \omega U]\,\Delta x^k[m-j]\Big\}. \end{aligned} \tag{13}$$

This can be rewritten as the discrete-time analogue of (7):

$$\begin{aligned} &\sum_{j=0}^{s} \Big[(\alpha_j I + h\beta_j D) - \omega h\beta_j L\Big]\Delta x^{k+1}[m-j] = \\ &\qquad \sum_{j=0}^{s} \Big[(1-\omega)(\alpha_j I + h\beta_j D) + \omega h\beta_j U\Big]\Delta x^k[m-j]. \end{aligned} \tag{14}$$

As the similarities of equations (11) and (14) suggest, if the time interval is finite, i.e. the number of timesteps is some finite L, then for a given timestep h and a given SOR parameter ω, the time-discretized waveform SOR method has the same asymptotic convergence rate as pointwise SOR.

Theorem 3.1. On a finite simulation interval, the iterations defined by (11) and (14) have the same asymptotic convergence rate.

Proof. Let y^k denote the large vector consisting of the concatenation of vectors $\Delta x^k[m]$ at all L discrete timepoints, i.e. $y^k = \left[\Delta x^k[1]^T, \ldots, \Delta x^k[L]^T\right]^T$. Collecting together the equations (14) generated at each timepoint into one large matrix equation in terms of vectors y^{k+1} and y^k yields $M\Delta y^{k+1} = N\Delta y^k$ where $M, N \in \mathbb{R}^{Ln \times Ln}$ are block lower triangular banded matrices, with blocks of size $n \times n$, and with block bandwidth s. It is then easily seen that $M^{-1}N$ is block lower triangular, with diagonal blocks equal to

$$\Big[(I + h\beta_0 D) - \omega h\beta_0 L\Big]^{-1}\Big[(1-\omega)(I + h\beta_0 D) + \omega h\beta_0 U\Big]. \tag{15}$$

Therefore, $\rho(M^{-1}N)$ is given by (12), implying that the iterations defined by (11) and (14) have identical asymptotic convergence rates. □

Theorem 3.1 suggests that parameter ω for waveform SOR should be chosen to be precisely equal to the optimum parameter for the pointwise SOR method. However, this does not necessarily lead to fast convergence, as the following example illustrates.

Example 3.1. Let $t \in [0, 2048]$, $x(0) = 0$, and let matrix $A \in \mathbb{R}^{32\times 32}$ and time-dependent input vector $b(t) \in \mathbb{R}^{32}$ of the model problem (1) be given by

$$A = \begin{bmatrix} 2 & -1 & & \\ -1 & 2 & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{bmatrix}$$

(16)

$$b(t) = \begin{bmatrix} b_1(t) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{where} \quad b_1(t) = \begin{cases} 1 - \cos\left(\dfrac{2\pi t}{256}\right) & \text{if } t \le 256 \\ 0 & \text{otherwise.} \end{cases}$$

Consider the four problems generated by discretizing in time with the first-order backward difference formula, using 64, 128, 256, and 512 uniform timesteps of size $h = 32$, 16, 8 and 4 respectively.

Since the tridiagonal matrix A is symmetric and is consistently ordered [8, 9], the matrix $(I + h\beta_0 A)$ of the pointwise time-discretized model problem (10) is also consistently ordered, and the optimum pointwise SOR parameter ω_{opt} is given by

$$\omega_{opt} = \frac{2}{1 + \sqrt{1 - \mu_1^2}} \tag{17}$$

where $\mu_1 = \rho(H_{GJ})$ is the spectral radius of the pointwise Gauss-Jacobi iteration matrix $H_{GJ} = (I + h\beta_0 D)^{-1}(h\beta_0 L + h\beta_0 U)$. For the four problems with 64, 128, 256 and 512 timesteps, the optimum pointwise parameters ω_{opt} are 1.669, 1.586, 1.482 and 1.364 respectively.

Curves PT64, PT128, PT256 and PT512 of Figure 1 show the convergence of the waveform SOR method versus iteration for the four problems with their optimum pointwise SOR parameters ω_{opt}. Note that as the total number of timesteps is increased, the initial convergence rate is slower, approaching a limiting value of the convergence rate of the continuous Gauss-Seidel WR algorithm (shown as WR in Figure 1). In each case, the convergence rate of the waveform SOR eventually approaches the expected asymptotic value of $\omega_{opt} - 1$. Note that with a reasonable error accuracy tolerance such as 10^{-6} as a stopping point, the asymptotic convergence rate is *never* reached. For comparison, Figure 1

also shows the superposition of four convergence plots (CSOR) of the new convolution SOR method to be introduced in the following sections.

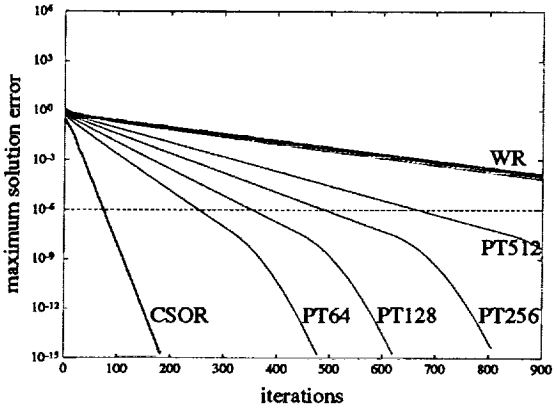

FIG. 1. *Convergence of waveform SOR using the pointwise optimal parameter (PT) compared to waveform relaxation (WR), and convolution SOR (CSOR), with 64, 128, 256 and 512 timesteps.*

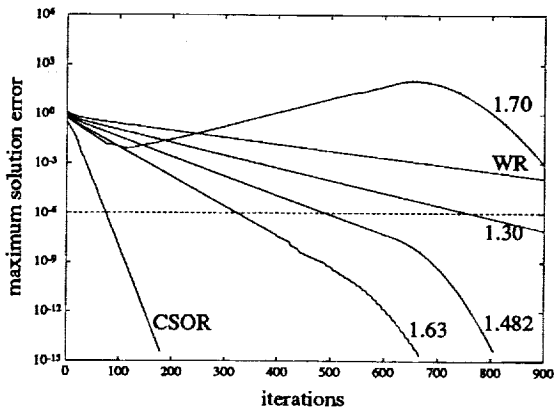

FIG. 2. *Effect on convergence of the 256-timestep waveform SOR of varying the SOR parameter from the pointwise optimum* $\omega_{opt} = 1.482$.

To illustrate the effect of choosing a different SOR parameter ω, Figure 2 shows the convergence versus iteration of the 256-timestep example for waveform SOR with values of the SOR parameter ω not equal to the pointwise optimum $\omega_{opt} = 1.482$. When $\omega = 1.30 < \omega_{opt}$, the convergence curve lies between the pointwise optimum curve and the WR convergence curve, i.e. both initial and asymptotic convergence rates are slower. By increasing the SOR parameter to $\omega = 1.63 > \omega_{opt}$, the initial convergence rate can be made faster at the expense of slowing down the asymptotic convergence rate. But as the $\omega = 1.70$ curve shows, once the SOR parameter is increased beyond some point, the waveform SOR method may appear to diverge before eventually converging. Also, the solution produced by the $\omega = 1.70$ example contains spurious oscillations, as shown in Figure 3. Note both the growth and translation of the oscillation with iteration.

The optimum pointwise SOR parameter ω_{opt} does not dramatically improve the convergence rate of waveform SOR because the matrix $M^{-1}N$ which describes the waveform SOR convergence is far from normal. This suggests that although the spectral radius of the iteration matrix determines the *asymptotic* convergence rate of waveform SOR, it does not determine the practically observable convergence rate. The convergence rate could be characterized, for example, by computing the pseudo-eigenvalues [10] of the waveform SOR iteration matrix. In the following section, we take an alternate approach.

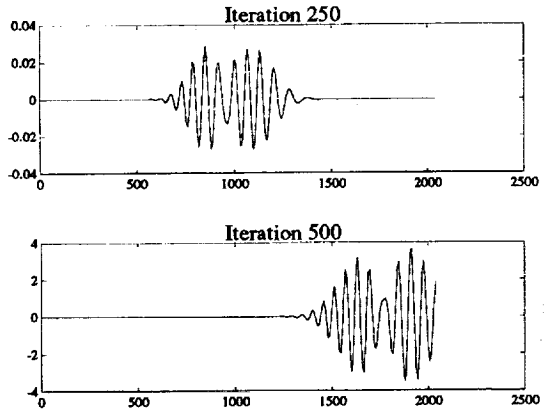

FIG. 3. *Delta waveform* $\Delta x_{16}^{k+1}(t) = x_{16}^{k+1}(t) - x_{16}^{k}(t)$ *versus time after iterations 250 and 500, for the 256-timestep waveform SOR method using* $\omega = 1.70$, *showing the growth and translation of an oscillating region.*

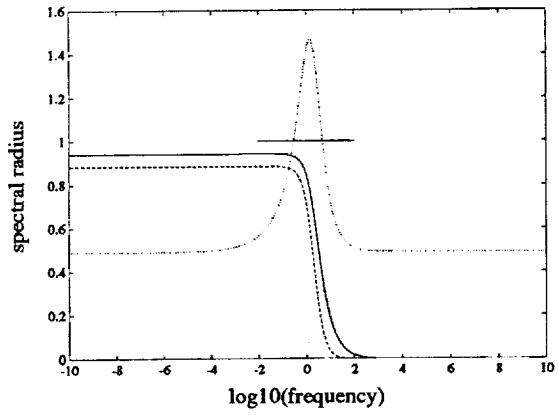

FIG. 4. *The spectral radii as functions of frequency* Ω *of the Gauss-Jacobi WR (solid), Gauss-Seidel WR (dashed) and waveform SOR (dotted) iteration matrices for an* 8×8 *version of the continuous-time problem of Example 3.1.*

FOURIER ANALYSIS

In [2], the spectral radius of dynamic iteration operators which map $\boldsymbol{x}^k$ to $\boldsymbol{x}^{k+1}$, such as those given by equations (2), (3), and (8), was related to their Fourier transform. In this section, we make a more detailed use of Fourier analysis to derive a frequency-dependent SOR parameter for the waveform SOR operator of equation (7).

The Fourier transform of $\boldsymbol{x}^k(t)$ is given by

$$\boldsymbol{x}^k(i\Omega) = \int_{-\infty}^{\infty} \boldsymbol{x}^k(t) e^{-i\Omega} \, dt = \mathcal{F}\{\boldsymbol{x}^k(t)\}, \tag{18}$$

where Ω is frequency. Standard Fourier identities can be used to show that $\Delta \boldsymbol{x}^{k+1}(i\Omega) = \boldsymbol{H}(i\Omega) \; \Delta \boldsymbol{x}^k(i\Omega)$, where for WGJ (2), WGS (3) and waveform SOR (7), the iteration operator $\boldsymbol{H}(i\Omega)$ is given by

$$\boldsymbol{H}_{GJ}(i\Omega) = (i\Omega \boldsymbol{I} + \boldsymbol{D})^{-1}(\boldsymbol{L} + \boldsymbol{U}) \tag{19}$$

$$\boldsymbol{H}_{GS}(i\Omega) = (i\Omega \boldsymbol{I} + \boldsymbol{D} - \boldsymbol{L})^{-1}\boldsymbol{U} \tag{20}$$

$$\boldsymbol{H}_{SOR}(i\Omega) = (i\Omega \boldsymbol{I} + \boldsymbol{D} - \omega \boldsymbol{L})^{-1}[(1-\omega)(i\Omega \boldsymbol{I} + \boldsymbol{D}) + \omega \boldsymbol{U}] \tag{21}$$

respectively. The obvious interpretation of equations (19)–(21) is that the spectral radius $\rho(\boldsymbol{H}(i\Omega))$ yields the asymptotic convergence rate for errors in the frequency component Ω.

Figure 4 is a plot of the spectral radii of $\boldsymbol{H}_{GJ}(i\Omega)$, $\boldsymbol{H}_{GS}(i\Omega)$ and $\boldsymbol{H}_{SOR}(i\Omega)$ for an 8×8 version of the continuous-time problem given in Example 3.1, using $\omega = 1.49$ for $\boldsymbol{H}_{SOR}(i\Omega)$. From the plot it is clear that very high frequency components of the error are damped much more quickly than low frequency components. However, the spectral radius $\rho(\boldsymbol{H}_{SOR}(i\Omega))$ is greater than one over a range of frequencies, and therefore the waveform SOR iteration magnifies errors in this frequency range. This effect was predicted in [2] and is easily seen in Figure 3.

This situation can be remedied by using a generalized SOR algorithm, in which equation (5) is replaced by an overrelaxation convolution with a time-dependent SOR parameter $\omega(t)$,

$$x_i^{k+1}(t) \leftarrow x_i^k(t) + \int_{-\infty}^{\infty} \omega(t-\tau) \cdot \left[\hat{x}_i^{k+1}(\tau) - x_i^k(\tau)\right] d\tau. \tag{22}$$

The Fourier transform of the SOR operator is then given by

$$\boldsymbol{H}_C(i\Omega) = \left[i\Omega \boldsymbol{I} + \boldsymbol{D} - \omega(i\Omega)\boldsymbol{L}\right]^{-1}\left[\left(1-\omega(i\Omega)\right)\left(i\Omega \boldsymbol{I} + \boldsymbol{D}\right) + \omega(i\Omega)\boldsymbol{U}\right], \tag{23}$$

where $\omega(i\Omega)$ is the Fourier transform of the time-dependent $\omega(t)$. We refer to the SOR algorithm represented by iteration matrix (23) as the *convolution SOR* algorithm (CSOR). The theorem below, which is the main result of this paper, gives a formula for determining the optimal frequency-dependent SOR parameter $\omega(i\Omega)$.

Theorem 4.2. If the spectrum of $\boldsymbol{H}_{GJ}(i\Omega)$ lies on the line segment $[-\mu_1(i\Omega), \mu_1(i\Omega)]$ with $|\mu_1| < 1$, then the spectral radius of $\boldsymbol{H}_C(i\Omega)$ is minimized at frequency Ω by a unique optimum $\omega(i\Omega) = \omega_{opt}(i\Omega) \in \mathbb{C}$ given by

$$\omega_{opt}(i\Omega) = \frac{2}{1 + \sqrt{1 - \mu_1(i\Omega)^2}} \tag{24}$$

where $\sqrt{\cdot}$ denotes the root with the positive real part.

Proof. For brevity, the argument $(i\Omega)$ will be omitted in the following, and $\boldsymbol{H}_C(\omega)$ will denote the convolution SOR operator (at frequency Ω) computed using SOR parameter ω.

Let $\mu_i = r_i \mu_1$ denote each eigenvalue of $\boldsymbol{H}_{GJ}$, where $r_i \in [-1, 1]$. Classical SOR theory [8, 9] guarantees that for each $\mu_i = r_i\mu_1$, there is an eigenvalue λ_i of $\boldsymbol{H}_C(\omega)$ which satisfies

$$\lambda_i - \omega r_i \mu_1 \sqrt{\lambda_i} + (\omega - 1) = 0, \tag{25}$$

and therefore, from the quadratic formula,

$$\sqrt{\lambda_i} = \frac{r_i \mu_1 \omega}{2} + \sqrt{\left(\frac{r_i \mu_1 \omega}{2}\right)^2 - \omega + 1}. \tag{26}$$

Let ω to be the conjectured optimal ω_{opt}. Combining equation (24) with (26) yields

$$\sqrt{|\lambda_i|} = \left|\tfrac{1}{2}\mu_1\omega_{opt}\left[r_i + \sqrt{r_i^2 - 1}\right]\right| = \left|\tfrac{1}{2}\mu_1\omega_{opt}\right|, \tag{27}$$

where the rightmost equality follows from the fact that $\left|r_i + \sqrt{r_i^2 - 1}\right| = 1$ for $r_i \in [-1, 1]$. And as (27) holds for all i,

$$\rho(H_C(\omega_{opt})) = |\lambda_i| = \left|\left(\tfrac{1}{2}\mu_1\omega_{opt}\right)^2\right| = |\omega_{opt} - 1|\,. \tag{28}$$

Equation (28) implies that $\rho(H_C(\omega))$ cannot be decreased below $\rho(H_C(\omega_{opt}))$ by using an ω such that $|\omega - 1| > |\omega_{opt} - 1|$. This follows from the fact that, in general [8, 9],

$$\rho(H_C(\omega)) \geq |\omega - 1| \tag{29}$$

for any ω.

To show that $\rho(H_C(\omega))$ also cannot be decreased by choosing a value of ω such that $|\omega - 1| < |\omega_{opt} - 1|$, consider the eigenvalue λ_j corresponding to μ_1:

$$\sqrt{\lambda_j} = f_+(\omega) = \frac{\mu_1\omega}{2} + \sqrt{\frac{\mu_1^2\omega^2}{4} - \omega + 1} \tag{30}$$

and note that $f_+ : \mathbb{C} \to \mathbb{C}$, given by equation (30), is a single-valued, continuous function that is analytic except at

$$\omega_1, \omega_2 = \frac{2}{1 \pm \sqrt{1 - \mu_1^2}}. \tag{31}$$

Since $|\mu_1| < 1$, points ω_1 and ω_2 lie in the interior and exterior, respectively, of the circle $|\omega - 1| = 1$ in the complex ω-plane. Note that ω_1 equals the conjectured ω_{opt} from equation (24).

Let D denote the interior of the curve given by the perimeter of the circle $|\omega - 1| = 1$, except with a cut along the line defined by the circle's center and ω_1. The cut follows the line from the perimeter down to ω_1, and then back up the other side to the perimeter, as shown in Figure 5. The function f_+ is nonzero everywhere within D, since equation (25) implies that a zero can occur only at $\omega = 1$, and $f_+(1) = \mu_1$. Therefore, the minimum modulus theorem [11] implies that $|f_+(\omega)|$ attains its minimum value somewhere on the boundary of D. Finally, the lower bound in (29) implies that $\omega_1 = \omega_{opt}$ in (24) is the only point on D which can achieve as low a $\rho(H_C(\omega))$ as given in (28), completing the proof. □

Note that when the eigenvalues μ lie on a real line segment, this is yet another alternative proof of a classic SOR Theorem [8, 9, 12]. Also note that, in general, the optimal overrelaxation parameter $\omega(i\Omega)$ is complex.

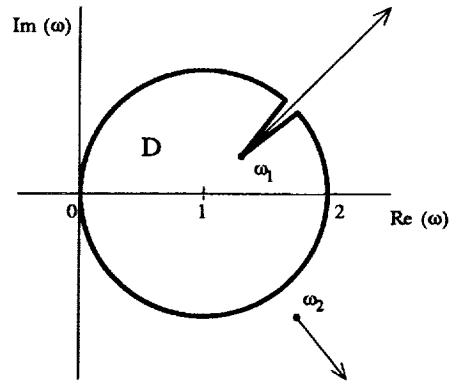

FIG. 5. *The region D and branch cuts in the complex ω-plane.*

The conditions of optimal SOR parameter Theorem 4.2 are satisfied by a large class of matrices.

Corollary 4.3. If $\boldsymbol{A}$ in (1) is consistently ordered, symmetric, and has constant diagonal $\boldsymbol{D} = d\boldsymbol{I}$, then the optimal SOR parameter is given by (24),

$$\omega_{opt}(i\Omega) = \frac{2}{1 + \sqrt{1 - \left(\dfrac{d\mu_1}{d + i\Omega}\right)^2}}, \tag{32}$$

where μ_1 denotes the spectral radius of $\boldsymbol{D}^{-1}(\boldsymbol{L} + \boldsymbol{U})$.

Proof. To show that Theorem 4.2 applies, note that (19) implies that for a constant diagonal $\boldsymbol{A}$, the $\boldsymbol{H}_{GJ}(i\Omega)$ eigenvalues $\mu(i\Omega)$ are given by $\mu(i\Omega) = d\mu_0/(d + i\Omega)$, where μ_0 are the eigenvalues of $\boldsymbol{D}^{-1}(\boldsymbol{L} + \boldsymbol{U})$. Since μ_0 lie on the real axis, the $\mu(i\Omega)$ lie on a line rotated in the complex plane. □

Corollary 4.4. If $\boldsymbol{A}$ in (1) is consistently ordered, symmetric, and has constant diagonal $\boldsymbol{D} = d\boldsymbol{I}$, then the optimal time-dependent SOR convolution waveform $\omega(t)$ is real.

Proof. Equation (32) implies that $\omega_{opt}(i\Omega)$ is a conjugate-symmetric function of Ω. □

DISCRETE-TIME MODIFICATION

For the sake of brevity, we consider only the first-order backward difference formula, in which case equation (14) becomes

$$\begin{aligned} &\Delta x^{k+1}[m] + h(D - \omega L)\Delta x^{k+1}[m] - \Delta x^{k+1}[m-1] = \\ &(1-\omega)\,\Delta x^k[m] + \left[(1-\omega)\,hD + h\omega U\right]\Delta x^k[m] - (1-\omega)\,\Delta x^k[m-1], \end{aligned} \tag{33}$$

where h is the uniform timestep. The z-transform of $x[m]$, defined by

$$x(z) = \sum_{m=-\infty}^{\infty} x[m]z^{-m} = \mathcal{Z}\{x[m]\}, \tag{34}$$

may be used to show that $\Delta x^{k+1}(z) = \boldsymbol{H}_C(z)\ \Delta x^k(z)$, where the z-dependent convolution SOR operator is

$$\boldsymbol{H}_C(z) = \left(\frac{1-z^{-1}}{h}\boldsymbol{I} + \boldsymbol{D} - \omega(z)\boldsymbol{L}\right)^{-1}\left[\left(1-\omega(z)\right)\left(\frac{1-z^{-1}}{h}\boldsymbol{I} + \boldsymbol{D}\right) + \omega(z)\boldsymbol{U}\right].$$

Since $\omega(z)$ depends on z, overrelaxation becomes a convolution sum

$$x_i^{k+1}[m] \leftarrow x_i^k[m] + \sum_{j=-\infty}^{\infty} \omega[m-j]\cdot\left(\hat{x}_i^{k+1}[j] - x_i^k[j]\right), \tag{35}$$

where $\omega(z) = \mathcal{Z}\{\omega[m]\}$. To determine the optimal $\omega(z)$, we have the following theorem, whose proof is analogous to that of Theorem 4.2.

Theorem 5.5. If the spectrum of $\boldsymbol{H}_{GJ}(z)$ lies on the line segment $[-\mu_1(z), \mu_1(z)]$ with $|\mu_1| < 1$, then the spectral radius of $\boldsymbol{H}_C(z)$ is minimized at z by the unique optimum $\omega(z) = \omega_{opt}(z) \in \mathbb{C}$ given by

$$\omega_{opt}(z) = \frac{2}{1+\sqrt{1-\mu_1(z)^2}} \tag{36}$$

where $\sqrt{\cdot}$ denotes the root with the positive real part.

In Example 3.1, matrix $\boldsymbol{A}$ has constant diagonal $\boldsymbol{D} = d\boldsymbol{I}$, so that

$$\omega_{opt}(z) = \frac{2}{1+\sqrt{1-\left(\dfrac{d\mu_1}{d+\dfrac{1-z^{-1}}{h}}\right)^2}}, \tag{37}$$

where μ_1 denotes the spectral radius of $\boldsymbol{D}^{-1}(\boldsymbol{L}+\boldsymbol{U})$. Thus, to compute the optimal convolution SOR sequence $\omega[m]$ for the four CSOR plots of Figure 1, equation (37) was used to compute $\omega(z)$, and then the inverse z-transform of $\omega(z)$ was computed analytically by series expansion.

A device is assumed to be governed by the Poisson equation, and the electron and hole continuity equations:

$$\nabla^2 u + c_1 (p - n + N) = 0 \tag{38}$$

$$\nabla^2 n - \nabla n \nabla u - n \nabla^2 u = c_2 \frac{\partial n}{\partial t} \tag{39}$$

$$\nabla^2 p + \nabla p \nabla u + p \nabla^2 u = c_3 \frac{\partial p}{\partial t} \tag{40}$$

where u is the normalized electrostatic potential, n and p are the electron and hole concentrations, N is a background concentration, and c_1, c_2, c_3 are physical constants [13].

Given a rectangular mesh that covers a two-dimensional slice of a MOSFET, a common approach to spatially discretizing the device equations is to use a finite-difference formula to discretize the Poisson equation, and an exponentially-fit finite-difference formula to discretize the continuity equations [13]. On an N-node rectangular mesh, the spatial discretization yields a differential-algebraic system of $3N$ equations in $3N$ unknowns.

The convolution SOR method was implemented in the WR-based device transient simulation program WORDS [14]. WORDS uses red/black block Gauss-Seidel WR, where the blocks correspond to vertical mesh lines. The equations governing nodes in the same block are solved simultaneously using the first order backward-difference formula. The implicit algebraic systems generated by the backward difference formula are solved with Newton's method, and the linear equation systems generated by Newton's method are solved with sparse Gaussian elimination.

The three MOS devices of Figure 6 were used to construct six simulation examples, each device being subjected to either a drain voltage pulse with the gate held high (the **D** examples), or a gate voltage pulse with the drain held high (the **G** examples). All examples ranged from low to high drain current, and in the **G** examples, the gate displacement current was substantial because the applied voltage pulses changed at a rate of .2 $\sim$ 2 volts per picosecond.

device	description	mesh	unknowns
kar	abrupt junction	19 × 31	1379
ldd	lightly-doped drain	15 × 20	656
soi	silicon-on-insulator	18 × 24	856

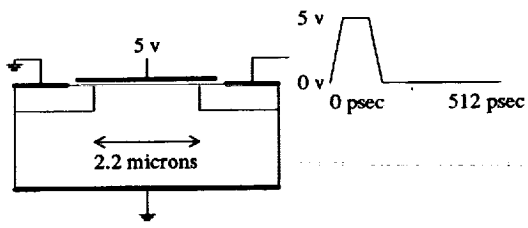

FIG. 6. *Description of devices and illustration of the drain-driven* **karD** *example.*

Figure 7 shows the convergence of the six examples as a function of iteration for WR, ordinary waveform SOR (using the pointwise optimum parameter), and the convolution SOR algorithm. The convolution SOR sequence $\omega[m]$ was calculated by linearizing (38)–(40) about the initial condition, estimating the spectral radius of the iteration matrix as a function of z, applying Theorem 5.5 and inverse transforming. Both overrelaxation methods were applied only to the potential variable u. All simulations began with 64 initial WR iterations, and used 256 equally-spaced timesteps. In Figure 7, convergence was measured using the terminal current error.

Despite the nonlinearity of the semiconductor equations, the convolution SOR algorithm converged substantially faster than either WR or ordinary waveform SOR, demonstrating the robustness of the approach.

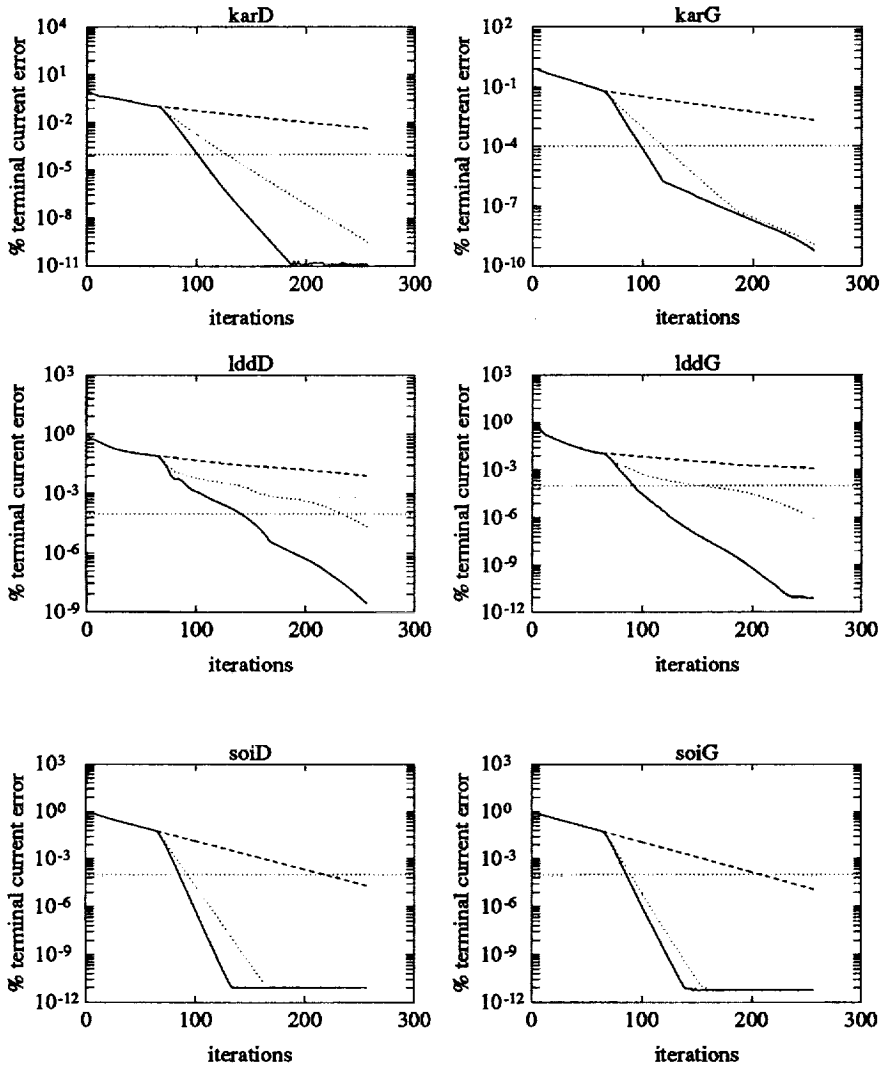

FIG. 7. *Terminal current error of the six examples as a function of iteration for WR (dashed), ordinary waveform SOR (dotted), and convolution SOR (solid).*

Conclusion

In this paper, a new waveform overrelaxation algorithm was presented and applied to solving the differential-algebraic system generated by spatial discretization of the time-dependent semiconductor device equations. In the experiments included, the convolution SOR algorithm converged robustly, and substantially faster than ordinary WR.

The author would like to acknowledge extensive conversations with his advisor, Professor Jacob White, and also thank Professors Alar Toomre, Donald Rose, Paul Lanzcron, Andrew Lumsdaine and Olavi Nevanlinna for many valuable suggestions.

REFERENCES

[1] A. Agarwal, "Limits on interconnection network performance," *IEEE Trans. Parallel Distrib. Sys.*, pp. 398–412, October 1991.

[2] U. Miekkala and O. Nevanlinna, "Convergence of dynamic iteration methods for initial value problems," *SIAM J. Sci. Statist. Comput.*, vol. 8, pp. 459–482, July 1987.

[3] J. K. White and A. Sangiovanni-Vincentelli, *Relaxation Techniques for the Simulation of VLSI Circuits.* The Kluwer International Series in Engineering and Computer Science, Boston: Kluwer Academic Publishers, 1987.

[4] E. Lelarasmee, A. E. Ruehli, and A. L. Sangiovanni-Vincentelli, "The waveform relaxation method for time domain analysis of large scale integrated circuits," *IEEE Trans. CAD*, vol. 1, pp. 131–145, July 1982.

[5] R. D. Skeel, "Waveform iteration and the shifted Picard splitting," *SIAM J. Sci. Statist. Comput.*, vol. 10, no. 4, pp. 756–776, 1989.

[6] C. Lubich and A. Osterman, "Multi-grid dynamic iteration for parabolic equations," *BIT*, vol. 27, pp. 216–234, 1987.

[7] A. Lumsdaine, *Theoretical and Practical Aspects of Parallel Numerical Algorithms for Initial Value Problems, with Applications.* PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, 1992.

[8] D. M. Young, *Iterative Solution of Large Linear Systems.* Orlando, FL: Academic Press, 1971.

[9] R. S. Varga, *Matrix Iterative Analysis.* Automatic Computation Series, Englewood Cliffs, NJ: Prentice-Hall, 1962.

[10] L. N. Trefethen, "Pseudospectra of matrices," in *Numerical Analysis 1991* (D. F. Griffiths and G. A. Watson, eds.), Harlow, Essex, England: Longman Scientific & Technical, 1992.

[11] E. B. Saff and A. D. Snider, *Fundamentals of Complex Analysis for Mathematics, Science, and Engineering.* New Jersey: Prentice-Hall, Inc., 1976.

[12] R. C. Y. Chin and T. A. Manteuffel, "An analysis of block successive overrelaxation for a class of matrices with complex spectra," *SIAM J. Numer. Anal.*, vol. 25, pp. 564–585, June 1988.

[13] S. Selberherr, *Analysis and Simulation of Semiconductor Devices.* New York: Springer-Verlag, 1984.

[14] M. Reichelt, J. White, and J. Allen, "Waveform relaxation for transient two-dimensional simulation of MOS devices," in *Proc. International Conference on Computer-Aided Design*, (Santa Clara, CA), pp. 412–415, November 1989.