A Linear-Time Complex-Valued Eigenvalue Solver for Full-Wave Analysis of Large-Scale On-Chip Interconnect Structures

Jongwon Lee, *Student Member, IEEE*, Venkataramanan Balakrishnan, *Senior Member, IEEE*,
Cheng-Kok Koh, *Senior Member, IEEE*, and Dan Jiao, *Senior Member, IEEE*

***Abstract*—This paper proposes a linear-time complex-valued eigenvalue solver for solving large-scale on-chip interconnect problems. The fast eigenvalue solution is achieved by eigenvalue clustering, fast system reduction with negligible computational cost, and fast linear-time solution of the reduced system. Numerical and experimental results are presented to demonstrate the accuracy and efficiency of the proposed method.**

***Index Terms*—Eigenvalue solver, finite-element methods, frequency domain, full-wave analysis, on-chip interconnects.**

I. INTRODUCTION

WITH continued breakthrough in processing technology, interconnect design has become one of the biggest challenges in the design of today's and next-generation integrated circuits. Over the past few decades, the modeling of on-chip interconnects has experienced a series of transitions: from lumped capacitance (*C*), lumped resistance and capacitance (*RC*), distributed *RC*, to distributed resistance, inductance, and capacitance (*RLC*) models. As the clock frequency of microprocessors entered the gigahertz regime, full-wave models have become increasingly important since it is necessary to analyze the chip response to harmonics that are up to five times the clock frequency. In particular, full-wave-based analysis can be used to characterize global electromagnetic coupling through the common substrate and power delivery network.

However, on-chip interconnect structures present many modeling challenges to electromagnetic analysis [1]. These challenges include large problem size, large number of nonuniform dielectric stacks with strong nonuniformity, large number of nonideal conductors, the presence of silicon substrate, highly skewed aspect ratios, etc. In recent years, both partial differential equation (PDE) based solutions and integral equation (IE) based solutions have been developed to address these challenges [2]–[17].

Among these techniques, the frequency-domain eigenvalue-based method in [16] is particularly geared toward full-wave modeling of large-scale on-chip interconnect structures. The original wave propagation problem involves a very large number of unknowns, N, in a 3-D computational domain. Formulated as a generalized eigenvalue problem, the technique in [16] *partially* addressed the complexity issue by seeking the solutions of only a few 2-D interconnect structures, each involving only $M(\ll N)$ unknowns in either the x–y- or y–z-plane (y is the stack growth direction). These solutions are then post-processed to obtain the solution of the original 3-D problem through an on-chip mode-matching technique. The procedure is rigorous and entails no approximation. Take the test-chip interconnect in [22] for example, M is 6678, while N is 10.1 million. In another example [22], M is 222K, while N is 336 million.

While the complexity is greatly reduced with the construction of M-parameter models in [16], the problem of finding the solution of the associated modeling problem in $O(M)$ complexity remains open. The computational bottleneck is the solution of a generalized eigenvalue problem. Efficient algorithms such as ARPACK [18] still require $O(M^2)$ storage and operations due to dense matrix–vector multiplications. The main contribution of this paper is an algorithm that provides a solution to the generalized eigenvalue problem with $O(M)$ complexity, thus paving the way for the full-wave simulation of very large scale integration (VLSI) circuits. The $O(M)$ complexity is achieved by the development of a direct matrix solver of linear complexity in the process of Arnoldi iteration.

In Section II, we give a brief overview of the frequency-domain eigenvalue-based method for full-wave modeling of on-chip interconnects. In Section III, we present the proposed linear-time eigenvalue solver. In Section IV, numerical and experimental results are given to demonstrate the accuracy and efficiency of the proposed solver. Section V relates to our conclusion.

II. REVIEW OF THE FREQUENCY-DOMAIN EIGENVALUE-BASED METHOD

Recognizing that although a 3-D on-chip interconnect structure may consist of a very large number of circuit elements, the number of modes that can be propagated in this structure is orders of magnitude smaller, a frequency-domain eigenvalue-based method was developed in [16] for full-wave modeling of large-scale 3-D on-chip interconnect structures. This method involves a number of important steps, as outlined below.

Manuscript received November 02, 2008; revised May 04, 2009. First published July 24, 2009; current version published August 12, 2009. This work was supported by the National Science Foundation (NSF) under Award 0747578 and Award 0702567.

The authors are with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907 USA (e-mail: djiao@purdue.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMTT.2009.2025457

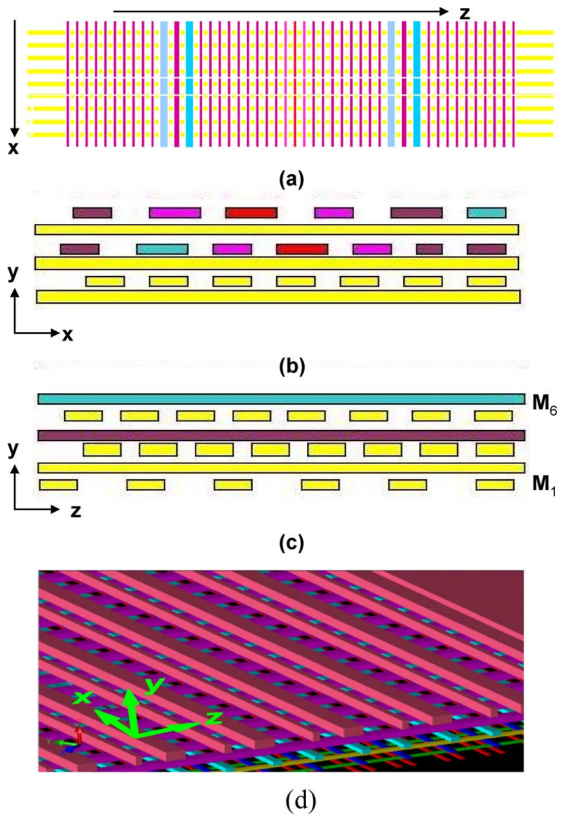


Fig. 1. 3-D interconnect structure. (a) Top view. (b) End view. (c) Side view (from [16]). (d) 3-D view.

A. Segmentation of the Interconnect Structure

A 3-D interconnect structure is sliced into segments. Each segment has a constant cross section. The segmentation direction, i.e., the longitudinal direction, is chosen from the x-, y-, and z-directions to minimize the number of unknowns in the transverse cross section. This is because the transverse cross section is numerically solved while the longitudinal direction is analytically processed in the frequency-domain eigenvalue-based method.

B. Identification of the Structure Seeds

After the segmentation, a set of structure seeds are identified. A structure seed is a unique cross section. Take a typical Manhattan-type bus structure made of six metal layers as an example, its top view, end view (cross-sectional view), and side view are shown in Fig. 1. Without loss of generality, assuming the bus is segmented along the z-direction. This results in a large number of segments in a large-scale on-chip structure. However, the number of unique cross sections is only a few. For a typical bus structure shown in Fig. 1, the number of structure seeds is only eight. Each structure seed can be represented by three digits, for example, 101. The first, second, and third digits correspond to orthogonal layers M5, M3, and M1, respectively. For each digit, value 0 denotes the absence of the lines in that layer, whereas value 1 denotes a presence. If the M5, M3, and M1 lines are aligned along the z-direction, the total number of structure seeds is only two, i.e., the number of unique x–y cross sections is only two. One refers to the presence of all of the M5, M3, and M1 lines. The other denotes their absence. The structure seeds are repeated in different lengths along the longitudinal direction, constructing the entire structure. The aforementioned scheme of segmentation and identification of structure seeds equally applies to interconnects with vias. In general, the number of seeds is orders of magnitude smaller than the number of segments.

C. Eigenvalue-Based Solution

In light of the fact that the electrical properties of interconnects are intrinsic in nature irrespective of the excitation, we construct an eigenvalue-based method for the analysis of an interconnect structure. Inside the interconnect structure, the electric field $\mathbf{E}$ satisfies the second-order vector wave equation

$$\nabla \times \left(\mu_r^{-1} \nabla \times \mathbf{E}\right) - k_0^2 \varepsilon_r \mathbf{E} + j\omega\mu_0\sigma\mathbf{E} = 0 \text{ in } \Omega \qquad (1)$$

subject to certain boundary conditions such as

$$\hat{n} \times \mathbf{E} = 0 \ \text{ on } \Gamma_1 \quad \hat{n} \times (\nabla \times \mathbf{E}) = 0 \ \text{ on } \Gamma_2. \qquad (2)$$

In (1), μ_r, ε_r, and σ denote the relative permeability, relative permittivity, and conductivity, respectively, Ω is the computational domain, which is the cross section of a structure seed including both dielectric and conducting regions, Γ_1 is the boundary where the Dirichlet boundary condition is applied, and Γ_2 is the boundary where the Neumann boundary condition is applied. A finite-element analysis of the boundary value problem defined in (1) and (2) results in the following generalized eigenvalue problem:

$$\begin{bmatrix} \mathbf{A}_{tt} & 0 \\ 0 & 0 \end{bmatrix} \begin{Bmatrix} e_t \\ e_z \end{Bmatrix} = -\frac{\gamma^2}{k_0^2} \begin{bmatrix} \mathbf{B}_{tt} & \mathbf{B}_{tz} \\ \mathbf{B}_{zt} & \mathbf{B}_{zz} \end{bmatrix} \begin{Bmatrix} e_t \\ e_z \end{Bmatrix} \qquad (3)$$

in which the eigenvalues correspond to the propagation constants γ, and the eigenvectors characterize the transverse electric field e_t and longitudinal electric field e_z. Matrices $\mathbf{A}$ and $\mathbf{B}$ are complex valued due to the penetration of fields into on-chip conductors. The entries of $\mathbf{A}$ and $\mathbf{B}$ are given by

$$\begin{aligned} \mathbf{A}_{tt,ij} &= \int\int_{\Omega} \left[-\frac{1}{k_0^2 \mu_r} \{\nabla_t \times \mathbf{N}_i\} \cdot \{\nabla_t \times \mathbf{N}_j\} \right. \\ &\qquad \left. + \bar{\varepsilon}_r \mathbf{N}_i \cdot \mathbf{N}_j \right] d\Omega \\ \mathbf{B}_{tt,ij} &= \int\int_{\Omega} \frac{1}{\mu_r} \mathbf{N}_i \cdot \mathbf{N}_j d\Omega \\ \mathbf{B}_{tz,ij} &= \int\int_{\Omega} \frac{1}{\mu_r} \mathbf{N}_i \cdot \nabla_t \xi_j d\Omega \\ \mathbf{B}_{zt,ij} &= \int\int_{\Omega} \frac{1}{\mu_r} \nabla_t \xi_i \cdot \mathbf{N}_j d\Omega \\ \mathbf{B}_{zz,ij} &= \int\int_{\Omega} \left[\frac{1}{\mu_r} \{\nabla_t \xi_i\} \cdot \{\nabla_t \xi_j\} - k_0^2 \bar{\varepsilon}_r \xi_i \xi_j \right] d\Omega \end{aligned} \qquad (4)$$

where $\bar{\varepsilon}_r$ denotes the complex permittivity that accounts for conductivity, $\mathbf{N}$ represents the edge basis function [19] used to expand the transverse field, ξ is the node basis function used to expand the longitudinal one, and Ω is the computational domain.

D. On-Chip Mode-Matching Technique

Once (3) is solved, the electric field in each segment can be obtained as

$$\mathbf{E}=\sum_{m=1}^{n}\left[\alpha_m \mathbf{e}_m(x,y)e^{-\gamma_m z}+\beta_m \mathbf{e}_m(x,y)e^{\gamma_m z}\right] \tag{5}$$

which is a superposition of all of the forward and backward propagation modes that can be supported by the structure. It should be noted that the $\mathbf{E}$ field in (5) has all three components $\mathbf{E}_x$, $\mathbf{E}_y$, and $\mathbf{E}_z$. The unknown coefficients α_m, and β_m in (5) are determined by imposing the following continuity condition at each junction that separates region 1 from region 2:

$$\begin{aligned}\sum_{m=1}^{K_1}\xi_{1m}\mathbf{e}_{1m,t}(x,y)&=\sum_{m=1}^{K_2}\xi_{2m}\mathbf{e}_{2m,t}(x,y)\\ \sum_{m=1}^{K_1}\eta_{1m}\mathbf{J}_{1m,z}(x,y)&=\sum_{m=1}^{K_2}\eta_{2m}\mathbf{J}_{2m,z}(x,y)\end{aligned} \tag{6}$$

where K_1 and K_2 are the number of modes in region 1 and region 2, respectively, and

$$\begin{aligned}\xi_{im}&=\alpha_{im}e^{-\gamma_{im}z}+\beta_{im}e^{\gamma_{im}z}\\ \eta_{im}&=\alpha_{im}e^{-\gamma_{im}z}-\beta_{im}e^{\gamma_{im}z}, \qquad i=1 \text{ or } 2\end{aligned} \tag{7}$$

and

$$\mathbf{J}_{im,z}=j\omega\varepsilon\mathbf{e}_{im,z}+\sigma\mathbf{e}_{im,z}, \qquad i=1 \text{ or } 2. \tag{8}$$

Testing (6) with appropriate functions results in (K_1+K_2) equations at the junction. Combining this set of equations at each junction with the loading conditions, the unknown coefficients α_m and β_m can be determined; hence, solving the field anywhere inside the interconnect structure.

III. PROPOSED LINEAR-TIME EIGENVALUE SOLVER

Equation (3) can be compactly written as

$$\mathbf{A}x=\lambda\mathbf{B}x. \tag{9}$$

Matrices $\mathbf{A}$ and $\mathbf{B}$ are sparse and of size $O(M)$. The number of eigenvalues that can make a difference in the final solution is generally much less than M, i.e., the number of modes that can propagate in an on-chip structure is generally much less than M. The number of propagation modes of a single stripline, for example, is only one when the electric size of the structure is small although the cross-section size can be very large. The cutoff frequencies of higher order modes are so far away from the cutoff frequency of the dominant mode that the higher order modes are attenuated quickly. When the size is increased or the frequency is increased, more modes can be propagated, but the number of propagation modes is still much less than M. As a result, the computing need here is to find K selected eigenpairs of the large sparse matrix system shown in (9), where K is the number of significant modes.

The Arnoldi iteration [18] is particularly suited for this computing task. Consider a standard eigenvalue problem

$$\mathbf{G}x=\lambda x. \tag{10}$$

A k-step Arnoldi process generates an orthonormal basis $\{v_j\}_{j=1}^{k}$ of the Krylov subspace $\kappa_k(v_1,\mathbf{G})$ spanned by $v_1,\mathbf{G}v_1,\ldots,\mathbf{G}^{k-1}v_1$, where v_1 is an initial unit norm vector. The projected matrix of $\mathbf{G}$ onto $\kappa_k(v_1,\mathbf{G})$ is represented by a $k\times k$ upper Hessenberg matrix $\mathbf{H}_k$, the Ritz pairs of which can be used to approximate the eigenpairs of $\mathbf{G}$. The algorithm of the k-step Arnoldi process is as follows.

Algorithm: The *k*-step Arnoldi process

```
1. $v_1 = v_1/\|v_1\|$
2. for $j = 1, 2, \ldots, k$ do
    2.1. $w = \mathbf{G}v_j$;
    2.2. for $i = 1, 2, \ldots, j$ do                (11)
        $h_{ij} = v_i^* w$;
        $w = w - h_{ij}v_i$.
    2.3. $h_{j+1,j} = \|w\|$, $v_{j+1} = w/h_{j+1,j}$.
```

The complexity of this algorithm is $O(Mk^2)$ if $\mathbf{G}$ is sparse. However, in our problem, $\mathbf{G}$ is dense because it is equal to $\mathbf{B}^{-1}\mathbf{A}$, and $\mathbf{B}^{-1}$ is dense, as can be seen from (3). Therefore, the complexity of a straightforward implementation of the Arnoldi process is $O(Mk^2+M^3+M^2k)$, where the $O(M^3)$ complexity accounts for the generation of $\mathbf{G}$, and the $O(M^2k)$ complexity accounts for the k dense matrix–vector multiplication operations. As a result, the cost of step 2.1 in (11) can dominate the total computational expense. The key contribution in this paper is the reduction of this computation to $O(M)$. We will first reduce the system matrix from 2-D to 1-D, then solve the reduced system and recover other unknowns. Three efficient algorithms will be developed to accomplish these two tasks with $O(M)$ complexity.

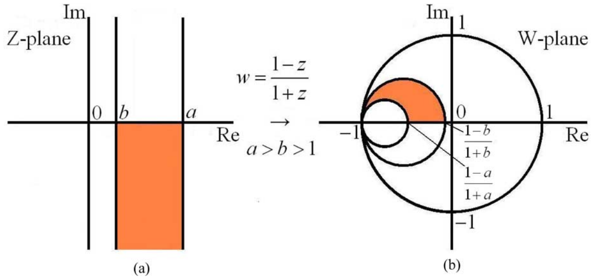
A. Eigenvalue Clustering

If the conductors are perfect, i.e., fields do not penetrate into conductors, the real part of the eigenvalues of (3) is bounded between the minimum and maximum relative permittivity. Since the conductors are lossy in a real on-chip environment, the real part is, in fact, bounded in a larger region, but still is bounded as shown in Fig. 2(a). However, the imaginary part can be widely scattered in complex plane due to the conductor loss induced attenuation that is modulated by complicated coupling from surrounding wires. This hinders the fast convergence of an Arnoldi process. To overcome this problem, we transform (9) to

$$\mathbf{A}'x=\lambda'\mathbf{B}'x \tag{12}$$

where

$$\mathbf{A}'=\mathbf{B}-\mathbf{A} \quad \mathbf{B}'=\mathbf{B}+\mathbf{A} \quad \lambda'=\frac{1-\lambda}{1+\lambda}. \tag{13}$$



(a) (b)

Fig. 2. Mapping of the eigenvalues from z plane to w.

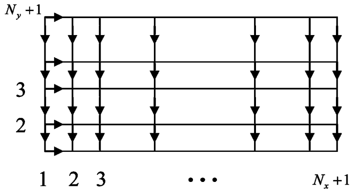


Fig. 3. Mesh and ordering.

By doing so, we cluster the eigenvalues that are originally scattered in the shaded region shown in Fig. 2(a) to a localized region shown in Fig. 2(b). Moreover, by a shift-invert operation, we further transform (12) to

$$\left(\mathbf{A}' - \tau\mathbf{B}'\right)^{-1}\mathbf{B}'x = \frac{1}{\lambda' - \tau}x \tag{14}$$

in which τ is an initial guess of λ', which is determined from the propagation constants of typical on-chip interconnect structures. Thus, the magnitude of the eigenvalues that are of physical interest becomes the maximum one, and hence, the convergence of the Arnoldi process is further expedited.

B. Reduction From 2-D to 1-D Lines in Complexity Much Less Than $O(M)$

In (9), both $\mathbf{A}$ and $\mathbf{B}$ are sparse matrices. However, their sparse patterns are not amenable for direct use in our computational technique. As established in (3), all the edge unknowns are ordered first and node unknowns are ordered next. The resultant matrix $(\mathbf{A}' - \tau\mathbf{B}')$ in (14) involves four block sub-matrices, each of which has its own sparse pattern that cannot be exploited readily. Therefore, we will transform $(\mathbf{A}' - \tau\mathbf{B}')$ to a banded matrix by permuting the ordering of the underlying variables. Further regularity of structure can be realized by discretizing the computational domain into rectangular elements. Since typical on-chip interconnects have Manhattan geometry, discretization with rectangular elements is indeed natural.

In Fig. 3, we plot a mesh to explain the unknown ordering scheme. There is an unknown associated with each edge, which is denoted as an edge unknown. There is also an unknown associated with each node, which is denoted as a node unknown. We discretize the computational domain into N_x segments along x and N_y segments along y. We denote the y-direction edge unknowns as e_y, x-direction edge unknowns as e_x, and z-direction node unknowns as e_z. We then first order e_y of line 1 (y orientated), e_z on line 1; and along x we proceed to e_x between line 1 and line 2; e_y of line 2, e_z on line 2, and so on. By doing so, we generate a banded matrix formed by submatrices in all segments. The sub-matrix in each x-segment (the region formed between two vertical lines) can be represented as

$$\begin{array}{c|c|c|c|}
 & e_{y,i} \quad e_{z,i} & e_{x,i} & e_{,yi+1} \quad e_{z,i+1} \\
\hline
\begin{matrix} e_{y,i} \\ e_{z,i} \end{matrix} & \mathbf{Q_i} & \mathbf{D_i} & \mathbf{E_i} \\
\hline
e_{x,i} & \mathbf{D_i^T} & \mathbf{T_i} & \mathbf{F_i} \\
\hline
\begin{matrix} e_{y,i+1} \\ e_{z,i+1} \end{matrix} & \mathbf{E_i^T} & \mathbf{F_i^T} & \mathbf{C_i} \\
\hline
\end{array} \quad . \tag{15}$$

Each sub-matrix overlaps with its neighbors through matrix $\mathbf{Q}$ and $\mathbf{C}$. Although the overall matrix formed by sub-matrices of the form in (15) is a banded matrix, computation that involves a banded matrix remains expensive when the size is large. To overcome this problem, we first eliminate all the edge unknowns between lines, i.e., horizontal (x-orientated) edge unknowns. Eliminating these unknowns is equivalent to the following block matrix operation:

$$\begin{aligned}
\mathbf{A}_{i,r} &= \mathbf{Q}_i - \mathbf{D}_i\mathbf{T}_i^{-1}\mathbf{D}_i^T \\
\mathbf{B}_{i,r} &= \mathbf{E}_i - \mathbf{D}_i\mathbf{T}_i^{-1}\mathbf{F}_i \\
\mathbf{C}_{i,r} &= \mathbf{B}_{i,r}^T \\
\mathbf{D}_{i,r} &= \mathbf{C}_i - \mathbf{F}_i^T\mathbf{T}_i^{-1}\mathbf{F}_i^T.
\end{aligned} \tag{16}$$

The right-hand side of (14) needs to also be updated in the reduction process as follows:

$$\begin{aligned}
b_{i1,r} &= b_{i1} - \mathbf{D}_i\mathbf{T}_i^{-1}b_{i2} \\
b_{i3,r} &= b_{i3} - \mathbf{F}_i^T\mathbf{T}_i^{-1}b_{i2}.
\end{aligned} \tag{17}$$

It is apparent from (16) that in order to eliminate all the horizontal unknowns, one has to fill in matrices $\mathbf{Q}$, $\mathbf{T}$, $\mathbf{C}$, $\mathbf{D}$, $\mathbf{E}$, and $\mathbf{F}$ for each segment. In addition, one has to evaluate $\mathbf{DT}^{-1}\mathbf{D}^T$, $\mathbf{DT}^{-1}\mathbf{F}$, and $\mathbf{F}^T\mathbf{T}^{-1}\mathbf{F}$ for each segment. The required computational cost can be very high when the number of segments is large. It turns out that such computational cost is negligible because the matrices involved exhibit the following properties.

a) Matrix $\mathbf{D}$ is the same for all the segments.
b) Matrices $\mathbf{F}$ and $\mathbf{D}$ are correlated: $\mathbf{F} = -\mathbf{D}^T$
c) Matrix $\mathbf{Q}$ is equal to matrix $\mathbf{C}$ in each segment.
d) Matrix $\mathbf{T}$ is linearly proportional to the segment length.
e) Matrix $\mathbf{T}$ only needs to be formed and inverted for each unique structure seed.

Although these matrix properties are similar to those in [17], the underlying reasons for them are quite different since the matrices in [17] involved 3-D structures.

As an immediate result of the aforementioned factors, the computational cost of eliminating all the horizontal unknowns is reduced to that of solving $\mathbf{DT}^{-1}\mathbf{D}^T$ for each structure seed. The dimension of matrix $\mathbf{T}$ is $N_y + 1$. When $\mathbf{T}$ is large, the factorization could cost $O\left(N_y^3\right)$ in both time complexity and space complexity. The operation of $\mathbf{T}^{-1}\mathbf{D}^T$ also costs $O\left(N_y^3\right)$, which is expensive. Here we will reduce the complexity to $O\left(N_y^2\right)$.

A careful examination of $\mathbf{T}$ reveals that it is a tridiagonal matrix. As can be seen from Fig. 3, each horizontal edge unknown only has crosstalk with its upper and lower neighbors among all the horizontal unknowns. The inverse of a tridiagonal matrix belongs to the class of hierarchically semiseparable matrices. For a symmetric tridiagonal matrix of order n, there exist two sequences $\{u_i\}$, $\{v_i\}$, $i = 1, 2, \ldots, n$ such that [20]

$$\mathbf{T}^{-1} = \begin{pmatrix} u_1v_1 & u_1v_2 & u_1v_3 & \cdots & u_1v_n \\ u_1v_2 & u_2v_2 & u_2v_3 & \cdots & u_2v_n \\ u_1v_3 & u_2v_3 & u_3v_3 & \cdots & u_3v_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ u_1v_n & u_2v_n & u_3v_n & \cdots & u_nv_n \end{pmatrix}. \tag{18}$$

Denoting $\mathbf{T}$ as $tri\,(a_{1:n}, -b_{1:n-1})$, the sequences $\{u_i\}$, $\{v_i\}$, $i = 1, 2, \ldots, n$ can be generated in $O(n)$ operations as follows:

$$\begin{aligned} d_n &= a_n \quad d_i = a_i - \frac{b_{i+1}^2}{d_{i+1}}, \qquad i = n-1, \ldots, 1 \\ v_1 &= \frac{1}{d_1} \quad v_i = v_{i-1}\frac{b_i}{d_i}, \qquad i = 2, \ldots, n \\ \chi_1 &= a_1 \quad \chi_i = a_i - \frac{b_i^2}{\chi_{i-1}}, \qquad i = 2, \ldots, n \\ u_n &= \frac{1}{\chi_n v_n} \quad u_{n-i} = \frac{b_{n-i+1}}{\chi_{n-i}} u_{n-i+1,}, \qquad i = 1, \ldots, n-1. \end{aligned} \tag{19}$$

Although the inverse of a tridiagonal matrix is dense, with n^2 entries, it can be compactly represented by $2n-1$ parameters in $\{u_i\}$ and $\{v_i\}$ $(u_1 = 1)$. In addition, matrix $\mathbf{D}$ is sparse. Hence, the cost of $\mathbf{D}\mathbf{T}^{-1}\mathbf{D}^T$ scales as $O\left(N_y^2\right)$.

Performance Analysis: The time complexity and space complexity of the aforementioned scheme are both $O\left(N_y^2\right)$. Since the direction y is the stack growth direction, N_y dictates the discretization of the stack. This number is generally much less than M. For example, the interconnect systems of 90-nm technology node involve only 8–9 metal layers. Hence, compared to $O(M)$, the cost of $O\left(N_y^2\right)$ is negligible. Furthermore, the cost does not grow with the problem size within each generation as the stack is fixed for each generation.

C. Solving Reduced System Matrix in $O(M)$ Complexity

The reduced system matrix forms a block tridiagonal matrix of order $(2N_y + 1)(N_x + 1)$, which can be denoted by $\mathbf{S} = tri\,(\mathbf{X}_{1:N_x+1}, \mathbf{Y}_{1:N_x})$. Here each $\mathbf{X}_i, \mathbf{Y}_i \in \mathbf{C}^{(2N_y+1)\times(2N_y+1)}$. Thus, $\mathbf{S} \in \mathbf{C}^{(2N_y+1)(N_x+1)\times(2N_y+1)(N_x+1)}$ with $N_x + 1$ diagonal blocks of size $2N_y + 1$ each. Since the right-hand side of (14) changes at each iteration step of an Arnoldi process, we are specifically interested in its direct solution. Similar to tridiagonal matrices, elegant theoretical results that describe the inverses of block tridiagonal matrices exist. For a symmetric $nm \times nm$ block tridiagonal matrix $\mathbf{S}$, there exist two sequences of $m \times m$ matrices $\{\mathbf{U}_i\}, \{\mathbf{V}_i\}$ such that for $j \geq i$, $\left(\mathbf{S}^{-1}\right)_{ij} = \mathbf{U}_i\mathbf{V}_j^T$. Thus,

$$\mathbf{S}^{-1} = \begin{pmatrix} U_1V_1^T & U_1V_2^T & \cdots & U_1V_n^T \\ V_2U_1^T & U_2V_2^T & \cdots & U_2V_n^T \\ \vdots & \vdots & \ddots & \vdots \\ V_nU_1^T & V_nU_2^T & \cdots & U_nV_n^T \end{pmatrix}. \tag{20}$$

While theoretically elegant, the computation of parameters $\{\mathbf{U}_i\}$ and $\{\mathbf{V}_i\}$ is beset by numerical problems for even modest-sized problems. The root cause of such an instability is that $\{\mathbf{U}_i\}$ and $\{\mathbf{V}_i\}$ scale exponentially with increasing problem size. Here, we will adopt a variant of the ratio-based approach, which is numerically stable [21]

$$\begin{aligned} \mathbf{U}_i &= \mathbf{R}_i\mathbf{U}_{i+1} \\ \mathbf{V}_{i+1} &= \mathbf{V}_i\mathbf{S}_i, \qquad i = 1, 2, \ldots, N_x \\ \mathbf{R}_1 &= \mathbf{X}_1^{-1}\mathbf{Y}_1 \\ \mathbf{R}_i &= \left(\mathbf{X}_i - \mathbf{Y}_{i-1}^T\mathbf{R}_{i-1}\right)^{-1}\mathbf{Y}_i, \qquad i = 2, \ldots, N_x \\ \mathbf{S}_{N_x} &= \mathbf{Y}_{N_x}\mathbf{X}_{N_x+1}^{-1} \\ \mathbf{S}_i &= \mathbf{Y}_i\left(\mathbf{X}_{i+1} - \mathbf{S}_{i+1}\mathbf{Y}_{i+1}^T\right)^{-1}, \qquad i = N_x - 1, \ldots, 1. \end{aligned} \tag{21}$$

As can be seen from (21), the computational cost of obtaining $\{\mathbf{U}_i\}$ and $\{\mathbf{V}_i\}$ is $O\left(N_xN_y^3\right)$

Since $\{\mathbf{U}_i\}$ and $\{\mathbf{V}_i\}$ constitute a compact representation of the inverse of a block tridiagonal matrix, the matrix vector multiplication can be performed in an efficient way. For example, $\mathbf{S}^{-1}b$ can be conducted in the following manner:

$$\begin{array}{lll} \mathbf{U}_1(\mathbf{V}_1^Tb_1 + \underline{\mathbf{V}_2^Tb_2 + \mathbf{V}_3^Tb_3 + \mathbf{V}_4^Tb_4 + \cdots \mathbf{V}_{N_x}^Tb_{N_x})} & & \\ \cdots \quad \mathbf{U}_2(\mathbf{V}_2^Tb_2 + \underline{\mathbf{V}_3^Tb_3 + \mathbf{V}_4^Tb_4 + \cdots \mathbf{V}_{N_x}^Tb_{N_x})} & & \\ \cdots \quad \mathbf{U}_3(\mathbf{V}_3^Tb_3 + \underline{\mathbf{V}_4^Tb_4 + \cdots \mathbf{V}_{N_x}^Tb_{N_x})} & & \\ \cdots & & \\ \cdots \quad \cdots & & \mathbf{U}_{N_x}\mathbf{V}_{N_x}^Tb_{N_x}. \end{array} \tag{22}$$

In (22), for clarity, only the upper triangular part is shown. The underlined terms are those that can be incrementally computed from the previous step if one starts from the last row. As a result, for each row, there is only one matrix vector multiplication that needs to be calculated. The same is true for the lower triangular part. Thus, the cost for computing $\mathbf{S}^{-1}b$ is $O\left(N_xN_y^2\right)$.

Performance Analysis: With the aforementioned scheme, the time complexity of step 2.1 in (11) is $O\left(N_xN_y^3 + kN_xN_y^2\right)$. Although the inverse of $\mathbf{S}$ is a dense matrix, with $\{\mathbf{U}_i\}$ and $\{\mathbf{V}_i\}$, which are $2N_x + 1$ matrices of size N_y^2, $\mathbf{S}^{-1}$ can be stored in $O\left(N_xN_y^2\right)$ memory, while a traditional technique would require $O\left(N_x^2N_y^2\right)$ memory. In terms of M, the computational complexity is $O\left(MN_y^2 + kMN_y\right)$. As N_y is much less than M and the number of dominant eigenvalues k is small, $O\left(MN_y^2 + kMN_y\right) \approx O(M)$. Similarly, the storage complexity is $O\left(MN_y\right) \approx O(M)$.

IV. NUMERICAL RESULTS

To evaluate the performance of the proposed linear-time eigenvalue solver, a number of on-chip interconnect structures were simulated, of which two were obtained from the test chip reported in [22] and two were artificially created.

First, an interconnect structure as shown in Fig. 4 was simulated. The dimensions of this structure were set according to typical on-chip geometrical dimensions. Four dielectric layers were involved. The thicknesses of the dielectric layers, from bottom to top, were 0.6, 0.5, 0.5, and 0.2 μm, respectively. The relative permittivities were 2.5, 4, 3, and 4, respectively. On the

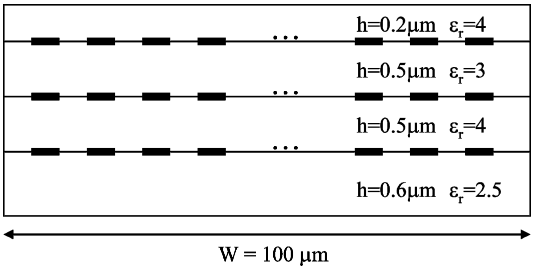


Fig. 4. Interconnect example of 300 wires.

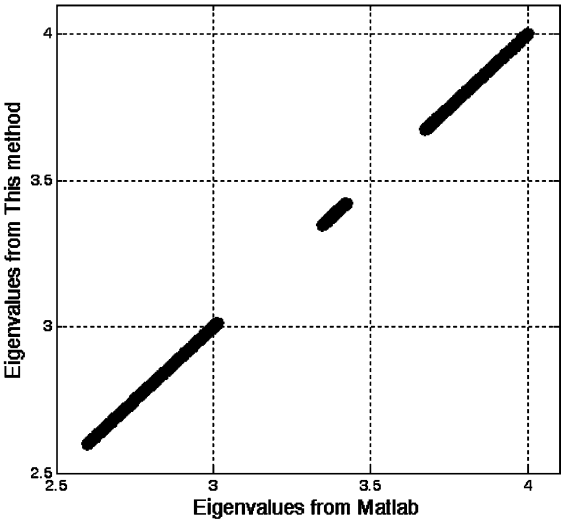


Fig. 5. Eigenvalues simulated by the proposed method in comparison with those generated by MATLAB.

top of the first, second, and third dielectric layers, 100 interconnect wires were placed, and hence, in total 300 interconnect wires were involved in this test structure. Each wire was 0.4975 μm wide, and 0.4975 μm apart from each other horizontally. The frequency of interest was 1 GHz. The proposed linear-time eigenvalue solver extracted 300 eigenvalues accurately, as can be seen from Fig. 5. As expected, for this example involving perfect conductors, all the eigenvalues are distributed between the minimum and maximum relative permittivity. In this simulation, the number of Arnoldi iterations was chosen as 320. The value of τ was chosen as 3.5. The overall CPU time of the proposed solver was shown to be 1.5 times faster than that of MATLAB, or more specifically ARPACK [18], a state-of-the-art large-scale sparse eigenvalue solver, for eigenvalue computation. For a fair comparison, we provided MATLAB with the same τ and required it to compute only 300 eigenvalues.

With the accuracy and efficiency of the proposed eigenvalue solver validated, we simulated a test-chip interconnect example, which was of 300-μm width [22]. It involved a 10-μm-wide strip in the M2 layer, one ground plane in the M1 layer, and one ground plane in the M3 layer. This strip was 50 μm to the M2 returns at the left- and right-hand sides. The strip was 2000-μm long. The reference ground is located at the bottom of M1. The S-parameters were extracted by the proposed linear-time eigenvalue solver at the near and far ends of the M2 center wire and compared with measured data. As can be seen clearly from Fig. 6, there is an excellent agreement.

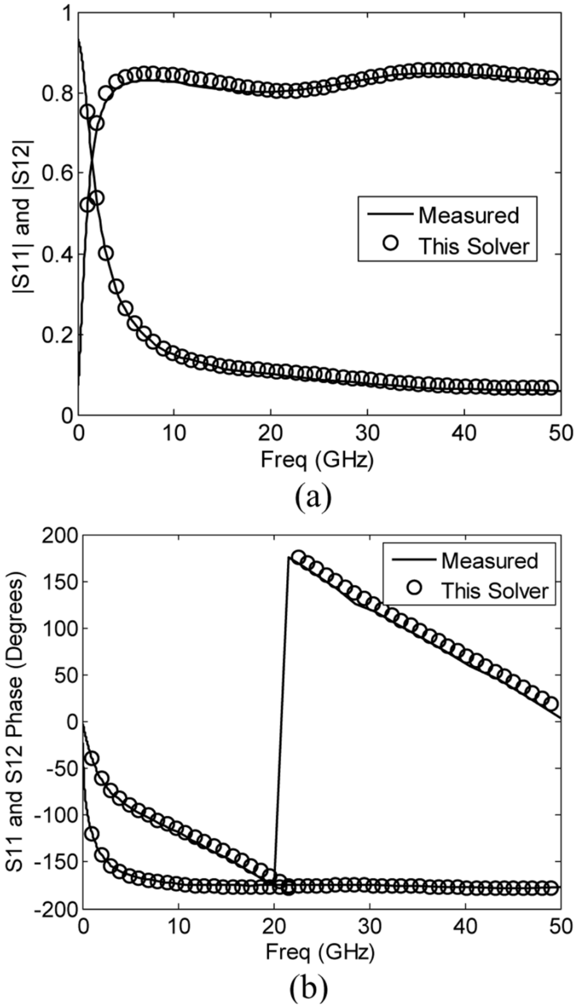


Fig. 6. Simulation of a test-chip interconnect. (a) S-parameter magnitude. (b) S-parameter phase.

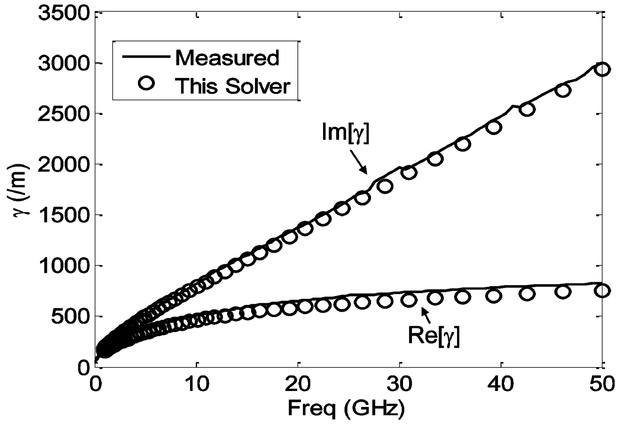


Fig. 7. Complex propagation constant simulated by the proposed method in comparison with measured data.

We also compare the complex eigenvalues extracted by the proposed solver at different frequency points with measured propagation constants, as shown in Fig. 7, which again reveals an excellent agreement. In Fig. 8, we plot the total CPU time of the proposed linear-time eigenvalue solver at one frequency point in comparison with that of a conventional Arnoldi-based

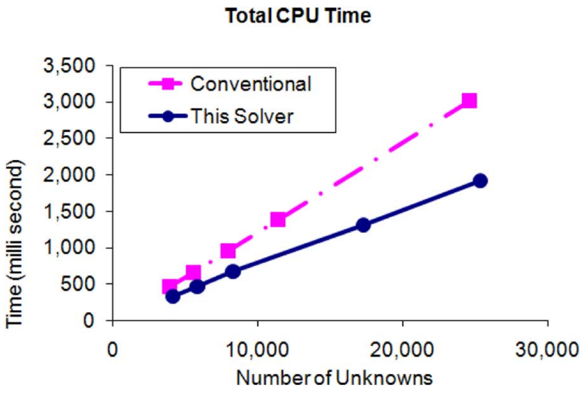


Fig. 8. Total CPU time comparison for a test-chip interconnect structure.

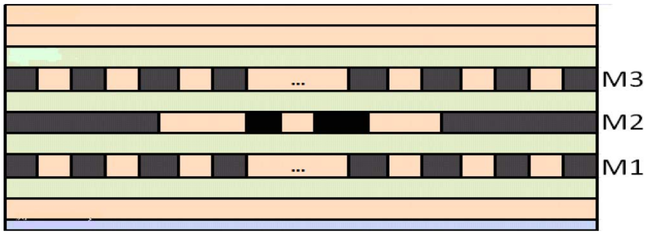
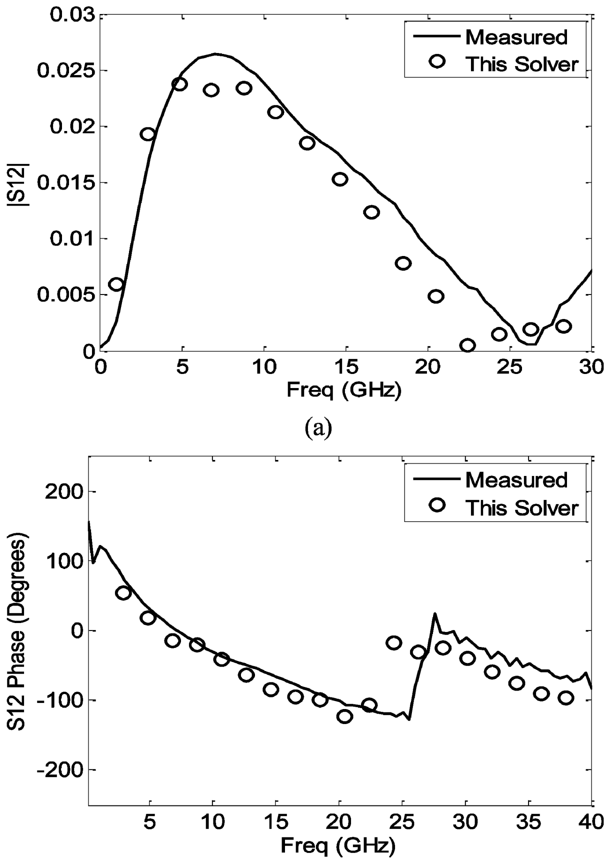


Fig. 9. Test-chip interconnect (cross-sectional view).



(b)

Fig. 10. Simulation of a test-chip interconnect. (a) $|S12|$. (b) $S12$ phase (degrees).

eigenvalue solver. The proposed solver clearly outperforms a conventional solver, and its linear complexity can be observed.

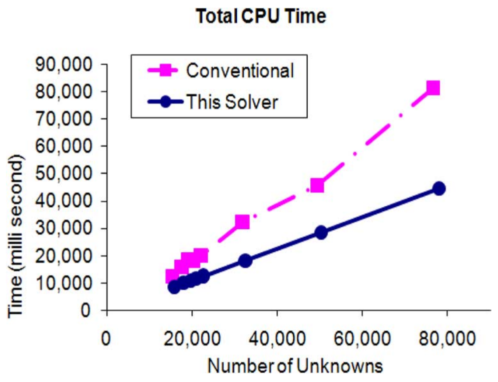
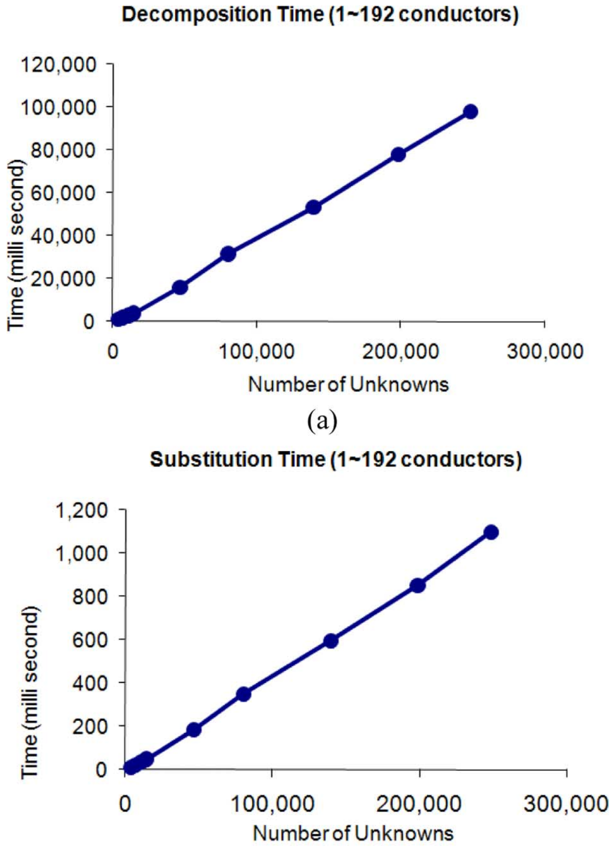


Fig. 11. Total CPU time comparison for the simulation of a test-chip interconnect example shown in Fig. 9.



(b)

Fig. 12. Simulation of a suite of on-chip interconnects consisting of three wires to 192 wires. (a) CPU cost for evaluating $(\mathbf{A}' - \tau \mathbf{B}')^{-1}$. (b) CPU cost for evaluating $(\mathbf{A}' - \tau \mathbf{B}')^{-1}\mathbf{B}x$.

The third example is a test-chip interconnect structure, as shown in Fig. 9. The structure was 2000-μm long, consisting of 11 inhomogeneous layers. It involves 12 parallel returns in the M1 and M3 layers, respectively. These returns were 1.05-μm wide and 1 μm apart. They were shorted to the ground at the near and far ends. Two wires were placed in the center of M2. One was of 1.1-μm wide, and the other was of a 2.07-μm width. The spacing between these two wires was 2.0 μm. The distance to M2 returns at the left- and right-hand sides was 10.1 μm. The reference ground is located at the bottom of the silicon substrate. The far ends of the two center wires in M2 were left open. The

S-parameters at the near ends of the two M2 wires were extracted by using the proposed eigenvalue solver and compared with the measured data. Very good agreement can be observed as can be seen from Fig. 10. In Fig. 11, the total CPU time cost by the proposed eigenvalue solver at one frequency point is plot against that of a conventional Arnoldi-based eigenvalue solver; the advantage of the proposed solver is evident.

In the last example, we simulated a suite of on-chip interconnect structures containing from three wires to 192 wires. The structures were discretized with up to 250K unknowns. Fig. 12(a) shows the decomposition time, i.e., the time for evaluating $(\mathbf{A}' - \tau\mathbf{B}')^{-1}$ in (14) of the proposed eigenvalue solver as a function of the number of unknowns, and Fig. 12(b) shows the time complexity of evaluating the dense-matrix multiplication in step 2.1 of (11). In both figures, linear complexity can be observed.

V. CONCLUSIONS

In this paper, a linear-time complex-valued eigenvalue solver was developed to solve large-scale on-chip interconnect problems. Numerical and experimental results have demonstrated the accuracy and efficiency of the proposed method.

ACKNOWLEDGMENT

The authors would like to thank Dr. M. J. Kobrinsky and Dr. S. Chakravarty, both with the Intel Corporation, Hillsboro, OR, for providing measured data.

REFERENCES

[1] D. Jiao, C. Dai, S.-W. Lee, T. R. Arabi, and G. Taylor, “Computational electromagnetics for high-frequency IC design,” in *IEEE Int. AP-S Symp.*, 2004, pp. 3317–3320.

[2] J. Ihm and A. C. Cangellaris, “Modeling of semiconductor substrate on on-chip power grid switching,” in *IEEE 13th Elect. Performance Electron. Packag. Topical Meeting*, 2004, pp. 265–268.

[3] J. Ihm and A. C. Cangellaris, “Distributed on-chip power grid modeling: An electromagnetic alternative to RLC extraction-based models,” in *IEEE 12th Elect. Performance Electron. Packag. Top. Meeting*, 2003, pp. 37–40.

[4] C. C. Chen, T. Lee, N. Murugesan, and S. C. Hagness, “Generalized FDTD-ADI: An unconditionally stable full-wave Maxwell’s equations solver for VLSI interconnect modeling,” *Int. Comput.-Aided Design Conf.*, pp. 156–163, 2000.

[5] A. Rong and A. C. Cangellaris, “Generalized PEEC models for three-dimensional interconnect structures and integrated passives of arbitrary shapes,” in *IEEE 10th Elect. Performance Electron. Packag. Top. Meeting*, Oct. 2001, pp. 29–31.

[6] A. E. Ruehli, G. Antonini, J. Esch, J. Ekman, A. Mayo, and A. Orlandi, “Nonorthogonal PEEC formulation for time- and frequency-domain EM and circuit modeling,” *IEEE Trans. Electromagn. Compat.*, vol. 45, no. 2, pp. 167–176, May 2003.

[7] D. Gope, A. E. Ruehli, C. Yang, and V. Jandhyala, “(S)PEEC: Time- and frequency-domain surface formulation for modeling conductors and dielectrics in combined circuit electromagnetic simulations,” *IEEE Trans. Microw. Theory Tech.*, vol. 54, no. 6, pp. 2453–2464, Jun. 2006.

[8] P. J. Restle, A. E. Ruehli, S. G. Walker, and G. Papadopou-los, “Full-wave PEEC time-domain method for the modeling of on-chip interconnects,” *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 20, no. 7, pp. 877–887, Jul. 2001.

[9] Z. H. Zhu, B. Song, and J. K. White, “Algorithm in FastImp: A fast and wideband impedance extraction program for complicated 3D geometries,” in *40th ACM/IEEE Design Automat. Conf.*, 2003, pp. 981–988.

[10] W. C. Chew, “Toward a more robust and accurate fast integral solver for microchip applications,” in *IEEE 12th Elect. Performance Electron. Packag. Top. Meeting*, 2003, p. 333.

[11] S. Kapur and D. E. Long, “Large-scale full-wave simulation,” *DAC*, pp. 806–809, 2004.

[12] A. E. Yilmaz, J. M. Jin, and E. Michielssen, “A parallel FFT-accelerated transient field-circuit simulator,” *IEEE Trans. Microw. Theory Tech.*, vol. 53, no. 9, pp. 2851–2865, Sep. 2005.

[13] F. Ling, V. I. Okhamtovski, W. Harris, S. McCracken, and A. Dengi, “Large-scale broad-band parasitic extraction for fast layout verification of 3D RF and mixed-signal on-chip structures,” *IEEE Trans. Microw. Theory Tech.*, vol. 53, no. 1, pp. 264–273, Jan. 2005.

[14] Y. Wang, V. Jandhyala, and C. J. Shi, “Coupled electromag-netic-circuit simulation of arbitrarily-shaped conducting structures,” in *IEEE 12th Elect. Performance Electron. Packag. Top. Meeting*, 2001, pp. 233–236.

[15] Z. Cendes and A. Yen, “Mixed electromagnetic and elec-trical circuit simulation for RFIC characterization,” in *IEEE AP-S Int. Symp.*, 2004, vol. 3, pp. 3289–3292.

[16] D. Jiao, M. Mazumder, S. Chakravarty, C. Dai, M. Ko-brinsky, M. Harmes, and S. List, “A novel technique for full-wave modeling of large-scale three-dimensional high-speed on/off-chip interconnect structures,” in *Int. Simulation Semiconduct. Processes and Devices Conf.* , 2003, pp. 39–42.

[17] D. Jiao, S. Chakravarty, and C. Dai, “A layered finite-element method for electromagnetic analysis of large-scale high-frequency integrated circuits,” *IEEE Trans. Antennas Propag.*, vol. 55, no. 2, pp. 422–432, Feb. 2007.

[18] ARPACK. Rice Univ., Houston, TX, 2008. [Online]. Available: http://www.caam.rice.edu/software/ARPACK/

[19] J. M. Jin*, The Finite Element Method in Electromagnetics*, 2nd ed. New York: Wiley, 2002.

[20] G. Meurant, “A review on the inverse of symmetric tridia-gonal and block tridiagonal matrices,” *SIAM J. Matrix Anal. Appl.*, vol. 13, no. 3, pp. 707–728, Jul. 1992.

[21] S. Cauley, J. Jain, C.-K. Koh, and V. Balakrishnan, “A scal-able distributed method for quantum-scale device simulation,” *J. Appl. Phys.*, vol. 101, pp. 12–12, 2007, Art. ID 123715.

[22] M. J. Kobrinsky, S. Chakravarty, D. Jiao, M. C. Harmes, S. List, and M. Mazumder, “Experimental validation of cross-talk simulations for on-chip interconnects using S-parameters,” *IEEE Trans. Adv. Packag.*, vol. 28, no. 1, pp. 57–62, Feb. 2005.

Jongwon Lee (S’09) received the B.S. degree in electrical engineering from Seoul National University, Seoul, Korea, in 2002, the M.S. degree in electrical and computer engineering from Purdue University, West Lafayette, IN, in 2007, and is currently working toward the Ph.D. degree in electrical and computer engineering from Purdue University.

He was a System Programmer with Chosun-Ilbo, Seoul, Korea, for three years. He is currently with the On-Chip Electromagnetics Research Group, School of Electrical and Computer Engineering, Purdue University. His current research interest is computational electromagnetics for large-scale high-frequency integrated circuit design.

Venkataramanan Balakrishnan (M’94–SM’06) received the B.Tech. degree in electronics and communication from the Indian Institute of Technology, Madras, India, in 1985, and the M.S. degree in statistics and Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1992.

Since 1994, he has been a faculty member with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, where he is currenlty Professor and Interim Head. His primary research interests are the application of numerical techniques, especially those based on convex optimization, to problems in engineering. He coauthored the monograph LINEAR MATRIX INEQUALITIES IN SYSTEM AND CONTROL THEORY (SIAM, 1994).

Dr. Balakrishnan was the recipient of the 1985 President of India Gold Medal presented by the Indian Institute of Technology, the 1997 Young Investigator Award presented by the Office of Naval Research, the 1998 Ruth and Joel Spira Outstanding Teacher Award, and the 2001 Honeywell Award for Excellence in Teaching presented by the School of Electrical and Computer Engineering, Purdue University. He was named a Purdue University Faculty Scholar in 2008.

Cheng-Kok Koh (S'92–M'98–SM'06) received the B.S. degree in computer science (with first class honors) and M.S. degree in computer science from the National University of Singapore, Singapore, in 1992 and 1996, respectively, and the Ph.D. degree in computer science from the University of California at Los Angeles (UCLA), in 1998.

He is currently an Associate Professor of electrical and computer engineering with Purdue University, West Lafayette, IN. His research interests include physical design of VLSI circuits and modeling and analysis of large-scale systems.

Dr. Koh was the recipient of the 1990 Lim Soo Peng Book Prize for Best Computer Science Student presented by the National University of Singapore, and the Tan Kah Kee Foundation Postgraduate Scholarship (1993 and 1994), the GTE Fellowship and the Chorafas Foundation Prize presented by UCLA (995 and 1996), the 1998 ACM Special Interest Group on Design Automation (SIGDA) Meritorious Service Award and Distinguished Service Award, the 1999 Chicago Alumni Award presented by Purdue University, the 2000 National Science Foundation CAREER Award, and the 2002 ACM/SIGDA Distinguished Service Award.

Dan Jiao (S'00–M'02–SM'06) received the Ph.D. degree in electrical engineering from the University of Illinois at Urbana-Champaign, in 2001.

She then joined the Technology Computer-Aided Design (CAD) Division, Intel Corporation, until September 2005, as a Senior CAD Engineer, Staff Engineer, and Senior Staff Engineer. In September 2005, she joined Purdue University, West Lafayette, IN, as an Assistant Professor with the School of Electrical and Computer Engineering. She is currently an Associate Professor with Purdue Univeristy. She has authored two book chapters and over 90 papers in refereed journals and international conferences. Her current research interests include computational electromagnetics, high-frequency digital, analog, mixed-signal, and RF integrated circuit (IC) design and analysis, high-performance VLSI CAD, modeling of microscale and nanoscale circuits, applied electromagnetics, fast and high-capacity numerical methods, fast time-domain analysis, scattering and antenna analysis, RF, microwave, and millimeter-wave circuits, wireless communication, and bio-electromagnetics.

Dr. Jiao has served as the reviewer for many IEEE journals and conferences. She was the recipient of the 2008 National Science Foundation (NSF) CAREER Award, the 2006 Jack and Cathie Kozik Faculty Start up Award (which recognizes an outstanding new faculty member of the School of Electrical and Computer Engineering, Purdue University), a 2006 Office of Naval Research (ONR) Award under the Young Investigator Program, the 2004 Best Paper Award presented at the Intel Corporation's annual corporate-wide technology conference (Design and Test Technology Conference) for her work on generic broadband model of high-speed circuits, the 2003 Intel Corporation's Logic Technology Development (LTD) Divisional Achievement Award in recognition of her work on the industry-leading BroadSpice modeling/simulation capability for designing high-speed microprocessors, packages, and circuit boards, the Intel Corporation's Technology CAD Divisional Achievement Award for the development of innovative full-wave solvers for high frequency IC design, the 2002 Intel Corporation's Components Research the Intel Hero Award (Intel-wide she was the tenth recipient) for the timely and accurate 2-D and 3-D full-wave simulations, the Intel Corporation's LTD Team Quality Award for her outstanding contribution to the development of the measurement capability and simulation tools for high frequency on-chip crosstalk, and the 2000 Raj Mittra Outstanding Research Award presented by the University of Illinois at Urbana-Champaign.