Automatic Generation of Accurate Circuit Models of 3-D Interconnect

Mattan Kamon, Nuno Alexandre Marques, *Student Member, IEEE,* Luis Miguel Silveira, *Member, IEEE,* and Jacob White, *Member, IEEE*

***Abstract*— In order to optimize high-speed systems, designers need tools that automatically generate reduced-order SPICE compatible models from geometric descriptions of interconnect and packaging. In this paper, we consider structures small compared to a wavelength, and use a discretized integral formulation combined with an Arnoldi-based model-order reduction strategy to compute efficiently accurate reduced-order models from three-dimensional (3-D) structures. Several issues are addressed including:**

1) **formulation to insure passivity in the reduced-order models;**
2) **efficient reduction using preconditioned inner-loop iterative methods;**
3) **expansion about multiple s-domain points.**

Results are presented on several industrial examples to demonstrate the capabilities and speed of these new methods.

***Index Terms*— Interconnect, packaging, parasitics, passive model order reduction, three-dimensional electromagnetic simulation.**

I. Introduction

IN ORDER TO simulate high-speed systems, designers need accurate models of the interconnect and packaging that can be included in a SPICE-level simulation. For many portions of a design, the significant interconnect may be long and uniform enough to be modeled using a two-dimensional (2-D) approximation and transmission line theory. Unfortunately, discontinuities in this 2-D interconnect, such as vias through planes, chip-to-board-connect and board-to-board connectors, require full three-dimensional (3-D) modeling. In the last few years, much effort has been devoted to computing models of 3-D structures directly from the interconnect or package geometry, usually by combining a Maxwell's equations solver with some post-processing strategy.

One well-known approach for generating accurate circuit models for 3-D structures is the partial element equivalent circuit (PEEC) approach [1]. However, for complex 3-D structures for which skin and proximity effects are important, the number of densely coupled circuit elements can be in the tens of thousands. Such PEEC generated circuits are much too expensive to include in a SPICE-level simulator. Recently, numerically robust model order reduction techniques have been developed to *automatically* generate a low order model from large circuit models [2]–[6]. In their basic form, however, these reduction methods require LU factorization of the original dense circuit matrices whose $O(n^3)$ computational complexity is still too expensive. Additionally, since the interconnect is passive, it is important that the reduced order model preserve this property, but the original PEEC formulation is not in a form to apply the provably passive model order reduction approach of [5].

This paper describes a computationally efficient approach to generating guaranteed passive low order models from large PEEC-like circuit models. To begin, in Section II we review the original PEEC discretization and recent work on passive model reduction. In Section III, we modify the original PEEC methods and follow the circuit solution technique known as mesh analysis. In Section III-C, we derive a first order state-space form of the mesh analysis system which can be used with passive model reduction algorithms. The main contribution of this paper comes in Section IV where we develop iterative approaches for applying the model reduction algorithms in an efficient manner for PEEC models of size exceeding 10^4. In Section V, we present the results of our algorithm followed by our conclusions.

II. Background

In this section, we briefly review the PEEC method derived in [1] and the passive model reduction algorithm from [5].

A. PEEC Formulation

The goal of parasitic interconnect extraction for a set of conductors is to determine the relation between the currents and the voltages at the terminals (or ports) of the conductors. For a k terminal-pair problem in the sinusoidal steady-state at the frequency ω, this relation is described by the admittance matrix, $Y_t(\omega) \in \mathbb{C}^{k\times k}$ where

$$Y_t(\omega)V_t(\omega) = I_t(\omega) \tag{1}$$

where I_t, $V_t \in \mathbb{C}^k$ are vectors of the terminal current and voltage, respectively [7].

This section derives the linear system from Maxwell's equations which must be solved to determine the admittance

Manuscript received January 7, 1998; revised May 7, 1998. This paper was supported by the Defense Advanced Research Projects Agency, the Semiconductor Research Corporation, the National Science Foundation, and grants from IBM and Harris Semiconductor.

M. Kamon is with Microcosm Technologies, Inc., Cambridge, MA 02142 USA (e-mail: matt@memcad.com).

N. A. Marques and L. M. Silveira are with the Instituto de Engenharia de Sistemas e Computadores, Lisboa 1000, Portugal (e-mail: nacm@algos.inesc.pt; lms@inesc.pt).

J. White is with the Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: white@rle-vlsi.mit.edu).

Publisher Item Identifier S 1070-9894(98)05832-0.

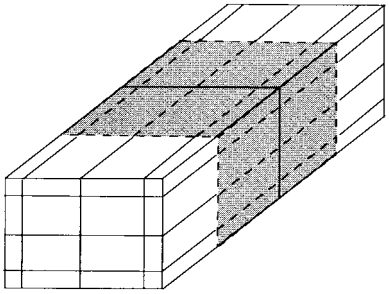


Fig. 1. Discretization of a short section of thin conductor. The volume is discretized into parallel filaments along the length. Surface is discretized into panels shaded in gray.

relation, Y_t. We begin by deriving an integral equation from Maxwell's equations and then discuss the circuit-like discretization to generate a linear system.

1) Integral Equation Formulation: When considering structures small compared to a wavelength, the integral formulation for Maxwell's equations under Laplace transform reduces to

$$\frac{\mathbf{J}(\mathbf{r},s)}{\sigma}+s\frac{\mu}{4\pi}\int_{V'}\frac{\mathbf{J}(\mathbf{r}',s)}{|\mathbf{r}-\mathbf{r}'|}\,dv'=-\nabla\Phi(\mathbf{r},s),\quad \mathbf{r}\in D \tag{2}$$

$$\Phi(\mathbf{r},s)=\frac{1}{4\pi\epsilon}\int_{S}\frac{\rho_s(\mathbf{r}',s)}{|\mathbf{r}-\mathbf{r}'|}\,dv',\quad \mathbf{r}\in\mathbb{R}^3 \tag{3}$$

where s is the Laplace frequency, D is the interior of all conductors, $\mathbf{J}$ is the current density in D, S is the surface of all conductors and ρ_s is the charge density on S. Additionally, the currents and charge obey the conservation equation

$$\nabla\cdot\mathbf{J}(\mathbf{r},s)=0,\quad \mathbf{r}\in D \tag{4}$$

$$\mathbf{n}\cdot\mathbf{J}(\mathbf{r},s)=-s\rho_s(\mathbf{r},s),\quad \mathbf{r}\in S \tag{5}$$

where $\mathbf{n}$ is the inward normal on S.

2) Discretization: The integral equations in the previous section are those used in the original PEEC derivation [1]. We will follow the approach given there to generate a discretization of (2) and (3) and then derive the mesh formulated approach.

To review the modeling of charge in the PEEC discretization, the surface of each conductor is covered with *panels*, each of which hold a constant charge density. To model current flow, the interiors of all conductors are divided into a 3-D grid of *filaments*. Each filament carries a constant current density along its length and this discretization of the interior captures skin and proximity effects. For long, thin wires for which the change in potential in the lateral dimensions is negligible, filaments only along the length of the wire are used. An example for a section of thin wire is shown in Fig. 1.

The approximation to the unknown current distribution can then be written as

$$\mathbf{J}(\mathbf{r})\approx\sum_{i=1}^{b}I_i w_i(\mathbf{r})\mathbf{l_i} \tag{6}$$

where I_i is the current inside filament i, $\mathbf{l_i}$ is a unit vector along the length of the filament and $w_i(\mathbf{r})$ is the weighting function which has a value of zero outside filament i, and $1/a_i$ inside, where a_i is the cross sectional area. By defining the inner product of two vector functions, $\mathbf{a}$ and $\mathbf{b}$, by

$$(\mathbf{a},\mathbf{b})=\int_V\mathbf{a}\cdot\mathbf{b}\,dv \tag{7}$$

and following the method of moments [8], a system of b equations can be generated by taking the inner product of each of the weighting functions with the vector integral equation, (2).

In matrix form, (2) becomes

$$(R+sL)I_b^f=\tilde{\Phi}_A-\tilde{\Phi}_B \tag{8}$$

where $I_b^f\in\mathbb{C}^b$ is the vector of b filament currents, R is the $b\times b$ diagonal matrix of filament DC resistances

$$R_{ii}=\frac{l_i}{\sigma a_i} \tag{9}$$

L is the $b\times b$ dense, symmetric positive definite matrix of partial inductances

$$L_{ij}=\frac{\mu}{4\pi a_i a_j}\int_{V_i}\int_{V_j'}\frac{\mathbf{l_i}\cdot\mathbf{l_j}}{|\mathbf{r}-\mathbf{r}'|}\,dV'\,dV \tag{10}$$

and $\tilde{\Phi}_A$ and $\tilde{\Phi}_B$ are the averages of the potentials over the filament end faces.

For the charge, the approximated charge density can then be written as

$$\rho_s(\mathbf{r})=\sum v_i(\mathbf{r})q_i,\quad \mathbf{r}\in S$$

where q_i is the charge density of panel i and $v_i(\mathbf{r})=1$ if $\mathbf{r}$ is on panel i, zero otherwise.

The filaments are each made branches in a network circuit graph and the junction between filaments are the nodes of the graph. To enforce (5), the panels are added to the circuit at nodes on the surface of the conductors. For the two section example of Fig. 1, a simplified version of the network is shown in Fig. 2. The network in the figure only has filaments which carry current parallel to the length of the wire. As stated before, for long, thin wires, the change in potential in the lateral dimensions is negligible and one can short together the three nodes in the vertical direction of Fig. 2. Note that for the general case of a 3-D grid of filaments, there will be two types of nodes: nodes on the exterior which connect panels and filaments and nodes on the interior which connect only filaments.

The last relation is that of the potential, Φ, to the charge, q, from (3). Approximating the average over the face, $\tilde{\Phi}_A$, by its value at the appropriate node point, the potential becomes

$$\Phi_n=P'q_p \tag{11}$$

where $\Phi_n\in\mathbb{C}^n$ is the vector of the n node voltages, $q_p\in\mathbb{C}^p$ is the charge on each of p panels, and $P'\in\mathbb{R}^{(n_e+n_i)\times p}$ is the potential coefficient matrix given by

$$P'_{ij}=\frac{1}{A_j 4\pi\epsilon_0}\int_{p_j}\frac{1}{|\mathbf{r_i}-\mathbf{r}'|}\,dV' \tag{12}$$

where p_j is the surface of panel j, A_j is its area, $\mathbf{r_i}$ is the ith node location, n_e is the number of node points on the surface

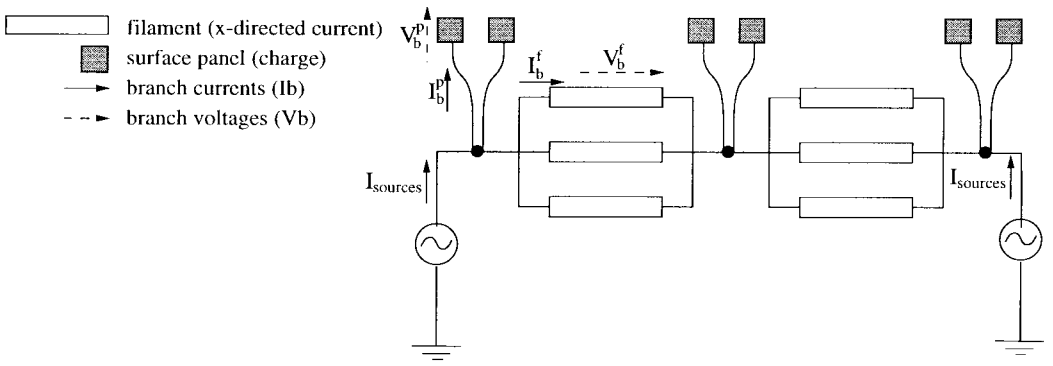


Fig. 2. Circuit for a short section of conductor. The panels are connected at the nodes between sections of filaments, and the sources are connected at the terminals. Only some panels and filaments shown.

and n_i is the number of internal node points with $n = n_e + n_i$. In the original PEEC derivation and for the rest of this section, there is exactly one panel per node, $n_e = p$, unlike what is shown in Fig. 2. This condition will be relaxed in Section III.

P' can be divided separately into its contribution to the internal and external nodes by

$$P' = \begin{bmatrix} P \\ P^i \end{bmatrix} \tag{13}$$

where $P \in \mathbb{R}^{n_e \times n_e}$ and $P^i \in \mathbb{R}^{n_i \times n_e}$.

The above relations, (8), (11), and (5), respectively, give the following linear system:

$$\begin{bmatrix} Z & 0 & A^T \\ 0 & P' & -I \\ A_e & sI & 0 \end{bmatrix} \begin{bmatrix} I_b^f \\ q_p \\ \Phi_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ I_t \end{bmatrix} \tag{14}$$

where $Z = R + sL$, I is the identity matrix of appropriate dimension, I_t are the terminal currents, $A^T = [A_e^T \ A_i^T]$ is the nodal incidence matrix providing the differencing of Φ_n and A_e enforces the boundary condition (5).

This is the PEEC formulation of [1]. In the original work, the elements were created as circuit elements and sent to a circuit simulator which would effectively assemble and solve (14).

It is worth noting that in the interior of conductors, (14) does not *explicitly* enforce current conservation as in (4). Thus, while an exact solution to the integral equation will satisfy current conservation, there is no guarantee that the discrete version will also.

B. Guaranteed Passive Model Order Reduction

In this section we review recent techniques for generating passive reduced order models directly from a large system such as that in (14). These methods require that the original system be written in a state space form such as the nth order system

$$\begin{aligned} s\mathcal{L}x &= -\mathcal{R}x + BV_t \\ I_t &= C^t x \end{aligned} \tag{15}$$

where $x \in \mathbb{C}^n$ is the vector of states, V_t, $I_t \in \mathbb{C}^k$ are the input and output vectors respectively, $\mathcal{L}$, $\mathcal{R} \in \mathbb{R}^{n \times n}$, and B, $C \in \mathbb{R}^{n \times k}$. The idea of model reduction is to derive a much smaller qth order system

$$\begin{aligned} s\tilde{\mathcal{L}}x &= -\tilde{\mathcal{R}}x + \tilde{B}V_t \\ I_t &= \tilde{C}^t x \end{aligned} \tag{16}$$

with $q \ll n$ but which still accurately models the system behavior.

In the area of circuit simulation, asymptotic waveform evaluation (AWE) [9] has popularized the use of model reduction via moment matching. In AWE, the admittance function of the reduced system is chosen to match a number of moments, or terms in the Taylor series about $s = 0$, of the original system. More specifically, the admittance is given by eliminating x in (15)

$$Y_t(s) = C^T(\mathcal{R} + s\mathcal{L})^{-1}B = \sum_{k=0}^{\infty} m_k s^k \tag{17}$$

where the moments are obtained from

$$m_k = -C^T(\mathcal{R}^{-1}\mathcal{L})^k \mathcal{R}^{-1}B.$$

Thus for (16), we seek $\tilde{Y}_t(s) = \sum_{k=0}^{\infty} \tilde{m}_k s^k$, such that $m_k = \tilde{m}_k$, $k = 1, \cdots, q$. The original AWE algorithm suffered from numerical ill-conditioning that prevented its use beyond a few moments. Additionally, since $Y_t(s)$ represents a passive circuit, we require $\tilde{Y}_t(s)$ also be passive. Recently, a numerically robust, guaranteed passive model order reduction algorithm (PRIMA) has been developed in [5]. The idea is to use an Arnoldi algorithm shown in Algorithm 1 to generate a set of orthogonal vectors, V_q [10]. These vectors V_q are applied in a congruence transform [11] to preserve passivity. This corresponds to a reduced system with $\tilde{\mathcal{L}} = V_q^T \mathcal{L} V_q$, $\tilde{\mathcal{R}} = V_q^T \mathcal{R} V_q$, $\tilde{B} = V_q^T B$, and $\tilde{C} = V_q^T C$. In [12], such an approach is shown to match $k-2$ moments and the main result in [5] is that this reduced order model is passive under the following conditions:

1) $\tilde{B} = \tilde{C}$;
2) $z^T(\mathcal{R} + \mathcal{R}^T)z \geq 0$ for all z;
3) $z^T(\mathcal{L} + \mathcal{L}^T)z \geq 0$ for all z.

Note that the PRIMA algorithm generates passive reduced order models which match moments at $s = 0$ for multiple input, multiple output systems. The extension of this algorithm to generate passive models which match moments at multiple values of s is presented in [13].

Algorithm 1 (Block Arnoldi process):

`arnoldi` (input $\mathcal{R}, \mathcal{L}, B, q$; output V_q^b)
{
  Let $A = \mathcal{R}^{-1}\mathcal{L}, D = \mathcal{R}^{-1}B$
  $D = V_1 R_1$ (`QR Factorization`)
  for $(j = 1;\ j <= q-1;\ j++)$ {
    $W = AV_j$
    for $(i = 1;\ i <= j-1;\ i++)$ {
      $H_{i,j} = V_i^T W$
      $W = W - V_j H_{i,j}$
    }
    $W = V_{j+1} H_{j+1,j}$ (`QR Factorization`)
  }
  $V_q^b = [V_1 \cdots V_q]$
  $H_q^b = (H_{i,j}),\ i, j = 1, \cdots, q$
}

III. A Formulation for Model Order Reduction

In order to automatically generate low order models from the large PEEC discretizations, we derive a state-space formulation which obeys the conditions for passivity in Section II-B. Unfortunately, using nodal analysis as in (14) makes passive model order reduction difficult since positive semidefiniteness (conditions 2 and 3 in Section II-B) is difficult to determine. In addition, (14) does not explicitly enforce current conservation on the interior. Also, for efficient quasistatic inductance calculation in [14], it was shown that the nodal formulation is poorly conditioned and this slows iterative algorithm convergence. For these reasons, in this section we derive a mesh analysis version of (14) which also enforces current conservation. The advantages for iterative solution will become apparent for multipoint model-order reduction in Section IV-B.

A. An Alternate PEEC Formulation

In this section, we modify the PEEC formulation before applying mesh analysis. While these changes may be implied by, but not stated, in [1] or have become standard over time, they are included here for completeness.

To begin, we wish to alter (14) to allow multiple panels to be connected to a single surface node. The panel discretization can then be refined independently of the filament discretization to capture, for instance, the sharp changes in charge density at a conductor edge or corner which does not necessarily correspond to changes in the potential. With multiple panels per node, $p \geq n_e$, and we now write $P \in \mathbb{R}^{p \times p}$. To enforce current conservation on the surface, (5), the charge on all the panels at a node must be summed. In addition, the potential of all the panels at a node must equal the node potential. Equation (14) then becomes

$$\begin{bmatrix} Z & 0 & -A_e^T & -A_i^T \\ 0 & P & -A_q^T & 0 \\ 0 & P^i & 0 & -I \\ A_e & sA_q & 0 & 0 \end{bmatrix} \begin{bmatrix} I_b^f \\ q_p \\ \Phi_n^e \\ \Phi_n^i \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ I_t \end{bmatrix} \tag{18}$$

where Φ_n has been divided corresponding to internal and external nodes, $A_q \in \mathbb{R}^{n_e \times p}$ sums the charges at each node and A_q^T copies the potential at a node to all its corresponding panels.

Next, to enforce current conservation on the interior (4), first replace the panel charge with the current into the panel, $\frac{d}{dt} q_p = I_b^p$. Now, enforcing current conservation on the interior involves replacing the set of equations involving P^i with $A_i I_b^f = 0$

$$\begin{bmatrix} Z & 0 & -A_e^T & -A_i^T \\ 0 & P/s & -A_q^T & 0 \\ A_e & A_q & 0 & 0 \\ A_i & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} I_b^f \\ I_b^p \\ \Phi_n^e \\ \Phi_n^i \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ I_t \\ 0 \end{bmatrix}. \tag{19}$$

Combining each of the 2×2 blocks in (19) into single blocks gives

$$\begin{bmatrix} Z_{\mathrm{EM}} & -\tilde{A}^T \\ \tilde{A} & 0 \end{bmatrix} \begin{bmatrix} I_b \\ \Phi_n \end{bmatrix} = \begin{bmatrix} 0 \\ I_t \end{bmatrix} \tag{20}$$

where

$$Z_{\mathrm{EM}} = \begin{bmatrix} Z & 0 \\ 0 & P/s \end{bmatrix}.$$

Additionally, in (12), r_i was chosen to correspond to the node to which the panel is connected. Here, instead, r_i is a collocation point at the center of the panel. Also, a Galerkin scheme could be used for which

$$P_{ij} = \frac{1}{A_j A_i 4\pi\epsilon_0} \int_{p_i} \int_{p_j} \frac{1}{|r - r'|}\, dV'\, dV. \tag{21}$$

B. A Mesh Analysis Formulation

Instead of using a nodal analysis approach to derive circuit equations, consider using mesh analysis as described in [7]. The mesh approach has been used in the context of interconnect analysis in [14] and [15].

To describe mesh analysis, a mesh is a loop of branches in the network graph. Each mesh is assigned an unknown current, I_m, which circulates around its branches. In mesh analysis, these mesh currents are the unknown quantities rather than the node voltages, Φ_n, as in nodal analysis. For a planar graph, a linearly independent set of meshes is exactly all loops which do not enclose any other branches. For nonplanar graphs, algorithms are given in [16].

To derive a mesh analysis version of (20), both the sources and the panels are made explicit branches in the network graph. To assign branch voltages to the panel branches, note that the panel node voltages in (11) are voltages relative to infinity. By adding infinity to the network as a zero volt node, we can view the panel branches as connecting the panel to the zero volt node at infinity. Then the panel branch voltages are given by

$$V_b^p = \Phi_{n_e} - 0 = \Phi_{n_e}. \tag{22}$$

With this definition for the panel branch voltages, Z_{EM} describes the constitutive relations for both filaments and panels

$$V_b = \begin{bmatrix} V_b^f \\ V_b^p \end{bmatrix} = \begin{bmatrix} Z & 0 \\ 0 & P/s \end{bmatrix} \begin{bmatrix} I_b^f \\ I_b^p \end{bmatrix} = Z_{\mathrm{EM}} I_b. \tag{23}$$

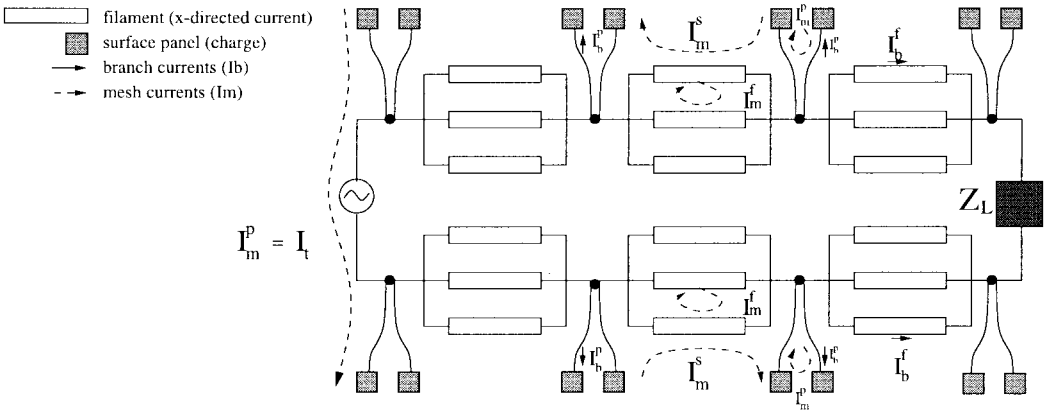


Fig. 3. A circuit describing the mesh quantities for a two-conductor TEM line terminated with a load Z_L.

For a simple structure such as two thin wires connected through a source at one end and a load Z_L at the other, the network graph illustrating a few sample meshes is shown in Fig. 3. The mesh currents can be divided into three separate types:

- I^f_m those involving only filaments on the interior;
- I^s_m those on the surface which include both a filament and two panels;
- I^p_m and those involving only panels for nodes with multiple panels.

I^s_m and I^p_m both pass through the node at infinity. Note also that the source branch generates a mesh which contains only the source and panels.

In mesh analysis, Kirchoff's voltage law, which implies that the sum of branch voltages around each mesh in the network must be zero, is represented by the matrix M

$$MV_b = \begin{bmatrix} M_{fi} & 0 \\ M_{fs} & M_{ps} \\ 0 & M_{pp} \end{bmatrix} \begin{bmatrix} V^f_b \\ V^p_b \end{bmatrix} = \begin{bmatrix} V^f_s \\ V^s_s \\ V^p_s \end{bmatrix} = V_s \tag{24}$$

where V_s is the mostly zero vector of source voltages. The three block rows of M correspond to the three types of meshes, I^f_m, I^s_m, and I^p_m, respectively. To simplify the notation, let

$$M_f = \begin{bmatrix} M_{fi} \\ M_{fs} \\ 0 \end{bmatrix}, \quad M_p = \begin{bmatrix} 0 \\ M_{ps} \\ M_{pp} \end{bmatrix} \tag{25}$$

so that $M = [M_f \;\; M_p]$. Note that the nonzero terms of V_s correspond to sources applied at the terminals, defined as V_t in (1). The relation between these quantities can be written as

$$V_s = NV_t \tag{26}$$

where N is an easily constructed terminal incidence matrix.

The M matrix also relates the mesh currents to the branch currents via

$$I_b = \begin{bmatrix} I^f_b \\ I^p_b \end{bmatrix} = M^T \begin{bmatrix} I^f_m \\ I^s_m \\ I^p_m \end{bmatrix} = M^T I_m. \tag{27}$$

Equations (23), (24), and (27) combine to give a system in only the unknown mesh currents

$$MZ_{\mathrm{EM}}M^T I_m = V_s. \tag{28}$$

By applying source voltages, the system in (28) could be solved to compute the admittance matrix, $Y_t(s)$, for specified frequencies, s. However the goal of this work is to instead generate reduced order models from this large system.

C. Deriving a State-Space Realization

We seek to apply methods of model order reduction to (28). We thus need to derive a first order state-space realization of the second order system (28). Care must be taken to derive a realization that has appropriate properties for passive model order reduction. Additionally, for computation with $\mathcal{R}$ in (15), we wish the corresponding $\mathcal{R}$ for (28) to be both sparse and nonsingular for expansions about $s = 0$.

1) A Nonsingular $\mathcal{R}$: To derive a first order system, choose new state variables as V^p_b. These new variables can be related to I_m using (23) and (27)

$$sV^p_b = PI^p_b = PM^T_p I_m. \tag{29}$$

Using (23), (24), and (29) in (28) gives a first order system

$$\begin{bmatrix} Z_m & M_p \\ -M^T_p & sP^{-1} \end{bmatrix} \begin{bmatrix} I_m \\ V^p_b \end{bmatrix} = \begin{bmatrix} V_s \\ 0 \end{bmatrix} \tag{30}$$

where $Z_m = M_f Z M^T_f$ and the second row of (30) is (29) multiplied by P^{-1}.

From (26), the terminal currents and voltages are related to their corresponding mesh quantities by $V_s = NV_t$, $I_t = N^T I_m$. Letting $B^T = [N^T \;\; 0]$ and separating out terms multiplying s gives the desired state-space form

$$s\begin{bmatrix} L_m & 0 \\ 0 & P^{-1} \end{bmatrix} \begin{bmatrix} I_m \\ V^p_b \end{bmatrix} = \begin{bmatrix} -R_m & -Mp \\ M^T_p & 0 \end{bmatrix} \begin{bmatrix} I_m \\ V^p_b \end{bmatrix} + \begin{bmatrix} V_s \\ 0 \end{bmatrix}$$
$$s\mathcal{L}x = -\mathcal{R}x + BV_t \tag{31}$$
$$I_t = B^T x$$

where $L_m = M_f L M^T_f$, $R_m = M_f R M^T_f$, and $\mathcal{R}$ and $\mathcal{L}$ are defined as the block matrices of (31). This $\mathcal{R}$ can be shown to be nonsingular under the condition that no node is connected to

the circuit via only panels [17]. In circuit language, there can be no cut-sets of capacitors. Unfortunately, the point at infinity is a such a node. From another point of view, $\mathcal{R}$ represents the DC solution of the system and applying sources between terminals as in Fig. 3 does not specify a DC path to ground (point at infinity).

To make the solution unique, for each conductor of the geometry, Φ must be specified over some terminal. Such a condition is enforced in circuit terms by requiring that at least one source for each conductor is connected to the point at infinity.

Also note that while $\mathcal{R}$ is nonsingular, $\mathcal{L}$ need not be for moment matching about finite s. In fact $\mathcal{L}$ is generally singular since M_f contains one row of zeros for every mesh of I_m^p.

2) A Realization for Practical and Passive Model Order Reduction: With the nonsingular $\mathcal{R}$ of (31), the matrix-vector products of the form $\mathcal{R}^{-1}\mathcal{L}x$ needed for model order reduction are now possible. Note that since the dense P matrix of (29) has been moved into $\mathcal{L}$, $\mathcal{R}$ is now sparse and computing the part of the product $\mathcal{R}^{-1}y$, is inexpensive. Even though computing $\mathcal{L}x$ now involves P^{-1}, this form will have advantages as discussed in the next section.

Additionally, with P^{-1} along the diagonal of $\mathcal{L}$, (31) satisfies the conditions for passive model reduction of Section II-B. In particular, to show that $\mathcal{L}+\mathcal{L}^T$ is positive semidefinite for conditions 3, first note that since L is positive definite, then L_m is also. If P is generated via a Galerkin approach, then it too is positive definite, and since $\mathcal{L}$ is a block diagonal matrix consisting of blocks which are each positive semidefinite, then so is $\mathcal{L}+\mathcal{L}^T$. Since the M_p matrices cancel for $\mathcal{R}+\mathcal{R}^T$, a similar argument holds for condition 2.

Note that the block structure of (31) is similar to the nodal form, (19) and at this point in the development the advantages of pursing a mesh form are not apparent. Even though both are different realizations of the same system, a nodal form may have benefits over a mesh form. One advantage of (19) is that if there are no internal nodes, then the corresponding $\mathcal{L}$ of (19) is nonsingular and smaller than the mesh formulated $\mathcal{L}$. The nonsingular $\mathcal{L}$ could be used for expansions about $s=\infty$. The results pursuing model order reduction for (19) will not be described here, but the interested reader is referred to [18].

Since R is a diagonal matrix, and M_f and M_p are sparse, then $\mathcal{R}$ is now sparse. However, to form the first block of $\mathcal{L}$ requires order f^2 operations and memory since $L \in \mathbb{R}^{f\times f}$ is dense. Similarly, to form the second block, P^{-1} requires p^2 operations and memory to form P, and then order p^3 operations to invert. For complex geometries with tens of thousands of filaments and panels, such growth rates are severely limiting. In the next section we discuss a more efficient technique for generating reduced order models from (31).

IV. Model Reduction for the Full Quasistatic Problem

In this section we describe fast algorithms for computing a reduced order model for (31). In Section IV-A, an approach using the expansion point $s=0$ is presented. This produces a straightforward algorithm, however the models produced are inadequate. To remedy, Section IV-B develops an approach using nonzero expansion points. While significantly more expensive, the generated models are more compact as we will see in Section V.

A. Expansions about $s = 0$

To apply the PRIMA algorithm, or any moment matching scheme about $s=0$, one must compute matrix-vector products with the matrix $(\mathcal{R}^{-1}\mathcal{L})$. For instance, an Arnoldi type algorithm requires $q-1$ such products to produce an order q model. However, because the partial inductance matrix L and potential coefficient matrix, P^{-1}, which appear in $\mathcal{L}$, are both large and dense, many multiplications by $\mathcal{L}$ can be prohibitively expensive. In particular, if done directly, multiplication by P^{-1} would require an initial dense matrix factorization which is $O(p^3)$ operations. For modern packaging structures, for which p exceeds ten thousand, such a factorization is prohibitive.

Efficient model reduction hinges on avoiding this expensive factorization. Fortunately, the expensive factorization can be avoided by noting that the computation $q=P^{-1}v$ is equivalent to solving for the panel charges, q, given a set of voltages, v. It is thus possible to use a preconditioned, Krylov-subspace iterative method to solve $Pq=v$ as outlined in Algorithm 2 [19]. Note that the dominant cost of each iteration is the $O(p^2)$ computation of a dense matrix-vector product, Pw, to acquire the next vector in the subspace.

Algorithm 2 (Iterative Scheme for $Pq=v$)

guess q^0
Initialize the search direction
  $w^0 = v - Pq^0$
for $k = 1, \cdots$ {
  Select $w^k \in$ *span*$\{w^0, Pw^0, \cdots, P^{k-1}w^0\}$
    such that the new solution
      $q^k = q^{k-1} + w^k$
    minimizes $\|r^k\| = \|v - Pq^k\|$
  if $\|r^k\| <$ *tolerance, return solution* q^k
}

In the standard approach, for every product $\mathcal{L}x$, the iterative algorithm would be called to solve $P^{-1}v$, generating a new subspace, $\text{span}\{w^0, Pw^0, P^2w^0, \cdots\}$, and a new set of search directions, w_k. If the number of $\mathcal{L}x$ products is large, the advantage of an iterative method would be degraded by the large number of total Pw products necessary. However, even though w^0 is different for each solve, it may be that the space spanned by $\{w^0, Pw^0, P^2w^0, \cdots\}$ is similar, as is illustrated in Fig. 4. One is thus lead to consider reusing the search directions from the previous solves [20]–[22]. While the recycled vectors are not optimal for the next v, the cost of computing the solution along those directions is negligible compared to a single Pw product. The recycled algorithm using the Krylov-subspace method known as generalized conjugate residual (GCR) [19] can be found in a general form in [20] and specifically for a constant matrix P in [17].

The $O(p^2)$ operations of the iterative algorithm can be reduced further by using a multipole-accelerated iterative algorithm [23] whose cost and memory has been shown to

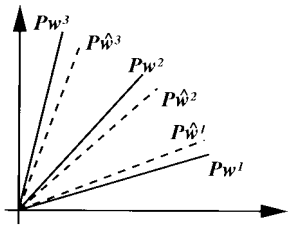


Fig. 4. Two-dimensional illustration of the search direction space for two different calls to the iterative algorithm. Here, the search directions Pw^k and $P\hat{w}^k$ are close so the spaces they span are similar.

grow only as $O(p)$. Similarly, the computation of the product $M_f L M_f^T$ can be performed in $O(f)$ operations also via the multipole-algorithm [14].

B. Efficient Multipoint Approximation

As we shall see in Section V, moment matching about $s = 0$ described above generates poor results. For this reason we wish to consider expanding about some other point or points. Such multipoint expansions have been explored previously for explicit moment matching in [24], and for the Krylov subspace techniques in [25], [12]. Recently, a provably passive multipoint rational Arnoldi algorithm has been derived for the reduction of RLC circuits with multiple inputs and outputs [13].

For expansions about points $s_0 \neq 0$, the moments become

$$m_k = -B^T\{(\mathcal{R} + s_0\mathcal{L})^{-1}\mathcal{L}\}^k\mathcal{R}^{-1}B.$$

Thus, to apply any multipoint scheme for the large dense systems of (31), one must be able to compute not only $\mathcal{L}x$ rapidly, as in the last section, but also

$$(\mathcal{R} + s_0\mathcal{L})^{-1}y. \tag{32}$$

Again, (32) is too large for direct factorization and one is led to iterative solution.

Iterative Solution:

$$x = (\mathcal{R} + s_0\mathcal{L})^{-1}y = \begin{bmatrix} R_m + s_0 L_m & M_p \\ -M_p^T & s_0 P^{-1} \end{bmatrix}^{-1} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{33}$$

is particularly expensive because each matrix vector product of an iterative algorithm requires an inner solve for P^{-1}. Thus, the model reduction algorithm would have three levels of nested loops:

1) Arnoldi iteration;
2) for each Arnoldi iteration, an iterative solve of (32);
3) at each iteration for (32), a solution with P^{-1}.

Fortunately, this last inner iteration can be avoided by realizing that computation of (32) is very close to solving the mesh formulated circuit of (31) at a single frequency s_0 given an input vector y. Thus we can return to solving the second order form, (28), to compute $x_1 = I_m$, and then compute V_b^p separately. However, (33) differs from (31) since y_2 is not generally zero, yet the last entry in the input for (31) is always zero. Thus, to eliminate V_b^p, we use

$$sV_b^p = PM_p^T I_m + Py_2 \tag{34}$$

in place of (29) and arrive at

$$(MZ_{\mathrm{EM}}M^T)I_m = y_1 - M_p P y_2. \tag{35}$$

After solving (35), $x_2 = V_b^p$ can be directly computed from (34). This approach avoids the inner computation of $P^{-1}x$ but more importantly, it returns us to the familiar mesh matrix, (28), to which we can apply the effective preconditioning techniques developed for inductance extraction in [14].

In general, Krylov-subspace iterative methods applied to solving (28) can be significantly accelerated by *preconditioning* if there is an easily computed good approximation to the inverse of $MZ_{\mathrm{EM}}M^T$. We denote the approximation to $(MZ_{\mathrm{EM}}M^T)^{-1}$ by Q, in which case preconditioning the iterative algorithm is equivalent to solving

$$(MZ_{\mathrm{EM}}M^T)Qx = V_s \tag{36}$$

for the unknown vector x. The mesh currents are then computed with $I_m{=}Qx$. Clearly, if Q is precisely $(MZ_{\mathrm{EM}}M^T)^{-1}$, then (36) is trivial to solve, but then Q will be very expensive to compute.

To follow the approach of [14], consider preconditioning with a block diagonal version of Z_{EM}. Thus, the preconditioner will be an LU factored version of

$$Q^{-1} = M\begin{bmatrix} R + s_0\hat{L} & 0 \\ 0 & \hat{P}/s_0 \end{bmatrix}M^T \tag{37}$$

where $\hat{L}$ and $\hat{P}$ are block diagonal.

One can improve this preconditioner by noting that for fast capacitance extraction in [23], [26], it was found that block diagonal preconditioning for P is not adequate to capture the strong coupling involved in charge interaction. For that reason, in [23], [26] a local-inversion preconditioner was developed. Since we know this preconditioner works well for P, we wish to use it in (37). Each row, i, of the local inversion preconditioner is formed by directly inverting a small P matrix corresponding only to those panels near panel i. The row in this inverse corresponding to panel i is then used as the row i of the preconditioner, C.

For the local inversion preconditioner to be effective inside the mesh formulated preconditioner of (37), it must be positive definite. While this preconditioner is not guaranteed to be positive definite, we have found from experiment that it produces good results implying it must be "close" to positive definite. This can be explained by realizing that the inverse of each of the small P matrices described above is close to a capacitance matrix and capacitance matrices are diagonally dominant. This implies that each row of C is likely to be diagonally dominant, so C is likely to be positive definite.

The preconditioner of the form (37) requires an approximation for P, but C is an approximation of P^{-1}. Thus, (37) would become

$$M\begin{bmatrix} R + s_0\hat{L} & 0 \\ 0 & \frac{1}{s_0}C^{-1} \end{bmatrix}M^T \tag{38}$$

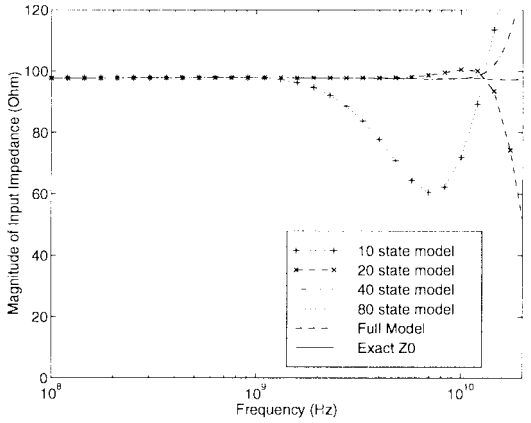


Fig. 5. Impedance looking into a matched transmission line for various reduced order models and the original discretized system.

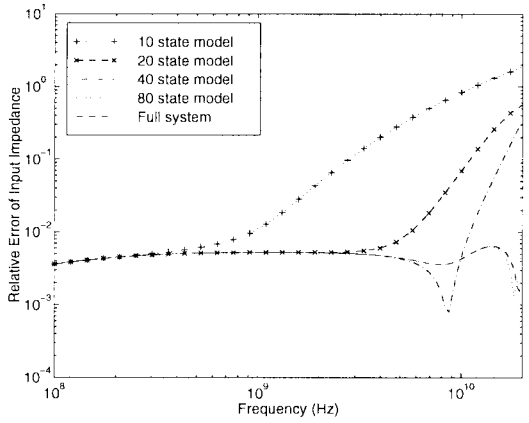


Fig. 6. Relative error for models for matched transmission line.

requiring an *inversion* of C, destroying its sparsity. C^{-1} would be too dense to LU factor efficiently as a submatrix of (37).

Fortunately this inversion can be avoided. Recall that the form of (28) was desirable because it avoided the inverse of P. Now, to avoid the inversion of C we can return to the first order form, (31), for the preconditioning step. More specifically, the preconditioning step requires computing x

$$\left(M\begin{bmatrix} R+s_0\hat{L} & 0 \\ 0 & \frac{1}{s_0}C^{-1}\end{bmatrix}M^T\right)^{-1}y_1 = x_1. \qquad (39)$$

Instead solve

$$\begin{bmatrix} \hat{R}_m + s_0\hat{L}_m & M_p \\ -M_p^T & s_0 C\end{bmatrix}^{-1}\begin{bmatrix} y_1 \\ 0\end{bmatrix} = \begin{bmatrix} x_1 \\ x_2\end{bmatrix} \qquad (40)$$

and discard x_2. The matrix is sparse and can be computed rapidly with LU factorization.

C. Recap

To recap, we noted that the dominant cost of applying Arnoldi-based model order reduction was in the repeated computation of Ax where, for a nonzero expansion point s_0, $A = (\mathcal{R}+s_0\mathcal{L})^{-1}\mathcal{L}$. Since the submatrices of $\mathcal{L}$ are L and P^{-1} which are dense matrices of dimension 10^4 to 10^5 a practical algorithm must avoid $O(n^3)$ computation such as the explicit formation of P^{-1} or $(\mathcal{R}+s_0\mathcal{L})^{-1}$. In Section IV-A for $s_0 = 0$, we used an iterative algorithm to avoid forming P^{-1} in the computation of $\mathcal{L}x$. Applying a preconditioned iterative algorithm was straightforward since it was identical to capacitance computation as explored in [23]. Since many solves must be performed, a recycling algorithm was employed for further speedup. In Section IV-B, for $s_0 \neq 0$, iterative solution is not as straightforward to compute the $(\mathcal{R}+s_0\mathcal{L})^{-1}y$ portion of the Ax product. Since P^{-1} is contained within $(\mathcal{R}+s_0\mathcal{L})$, iterative solution would require two levels of nested solve. Fortunately, we can return to a pure mesh form for this computation and then the iterative solve is instead for $(MZ_{\mathrm{EM}}M^T)^{-1}$. Not only does this form avoid the nested solve, but also provides a good method of preconditioning from the previous work for inductance extraction. To apply these preconditioning ideas to include the added capacitive portion of Z_{EM}, we saw that the best preconditioning approach for capacitance required us to return to the $(\mathcal{R}+s_0\mathcal{L})$ block form for just the preconditioning step.

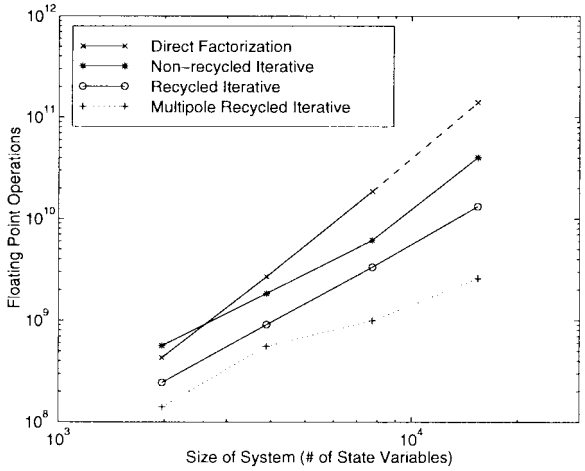


Fig. 7. Flop count for different methods of computing $P^{-1}x$.

V. Results

In this section, we present results from model order reduction. First we investigate the models produced and then observe the efficiency of the iterative solution algorithms.[1]

A. Expansions about $s = 0$

Consider generating a reduced order models of order 10, 20, 40, and 80, for a matched 2-D transmission line. The lines are copper and have a rectangular cross section with a width of 37 μm and height 15 μm. They are vertically spaced with a 42 μm center-to-center spacing. From Figs. 5 and 6, as the model order is increased, a model which matches to higher frequency is generated. For a reduced order model with 1% error, a twentieth order model is valid up to 6 GHz, a fortieth order model up to 12 GHz, and an eightieth order model past 20 GHz (to about 26 GHz).

[1] Note that, due to its simplicity of implementation, our implementation uses a collocation approach rather than a Galerkin approach to compute the entries of P. Such an approach is not guaranteed to give a positive semidefinite $P + P^T$ but has yet to cause nonpassive models.

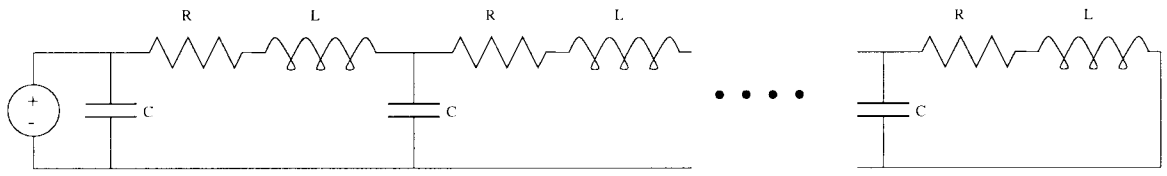
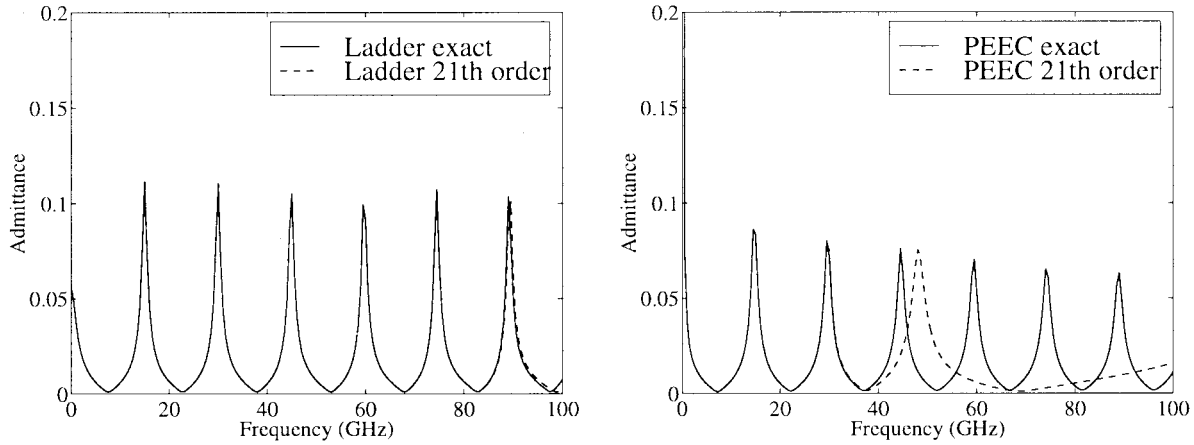


Fig. 8. A ladder circuit to model 2-D TEM transmission line.



(a) (b)

Fig. 9. Responses of various models for the transmission line. (a) Exact and reduced ladder circuit. (b) Exact and reduced PEEC.

Next, to demonstrate the efficiency of the recycled iterative scheme, consider refining the discretization of the transmission line of the previous example and extracting a fiftieth order model. Fig. 7 shows the number of floating point operations (flops) required for direct factorization with back substitution, a nonrecycled Krylov-subspace method, a recycled Krylov method, and multipole-accelerated recycled Krylov method, for various levels of discretization. Our implementation, called FASTPEP, uses direct matrix-vector products and thus the multipole-accelerated times are projected based on flop counts from multipole-accelerated capacitance and inductance codes [23], [14]. The residual error tolerance of the iterative algorithm had to be chosen as 10^{-6} so that the difference between models produced by the iterative scheme versus direct factorization differed by less than 1% up to 100 GHz. As can be seen from Fig. 7 for an original 15 409 state system, the recycled scheme performs an order of magnitude faster than direct factorization, and similarly, the multipole algorithm would provide another order of magnitude speed up. Note that the CPU time comparison would be similar to the flop count comparison for the direct factorization and direct recycled iterative scheme, however the overhead in arranging the multipole computation would shift its curve slightly upward.

B. Quality of the Models

The methods of the last section provide efficient generation of a reduced order model, and in this section we investigate the quality of the generated models.

To compare to an analytic result, consider the 2-D transmission line of uniform cross section and length 1 cm described in the last section. The line is divided into 40 sections along its length. Each section has a nine filament bundle, and each node has 12 panels leading to a 1704 element circuit. The line is shorted at the far end instead of being matched in order to emphasize the resonances. The admittance is then computed through a number of the resonant frequencies of the line using both the full model and a twenty-first order reduced model.

For comparison, given the exact per unit length 2-D line parameters, L' and C', a similar 40 section ladder circuit is constructed as shown in Fig. 8. The resistance per unit length is chosen to roughly match the actual resistance at the first resonance (and thus not at DC!). The admittance is computed for both this eightieth order model and also a twenty-first order reduced model. The four admittance functions are shown in Fig. 9. The solid lines in each figure represent the full PEEC model and the full ladder model. The solid lines show qualitatively the same results: there is a periodic resonance with the first occurring when the 1 cm structure is half a wavelength long. The resonant peaks show a decay for the PEEC model since it captures skin effects. Similarly, one might hope that because the frequency behavior of the two is roughly the same, the model order reduction results for similar order would be the same. However, the twenty-first order reduced models are very different. The PEEC twenty-first order model loses accuracy before the third resonance, however the ladder model does not begin to degrade until the sixth.

To understand this phenomenon, Fig. 10(a) and (b) plot the poles of the admittance function for the four cases. In Fig. 10(a), the poles of the exact admittance lie evenly spaced on a vertical line in the s plane. Since the model order reduction was performed about $s = 0$, one would expect a trend of pole matching starting at the origin and moving outward as shown. In Fig. 10(b), the full PEEC model has

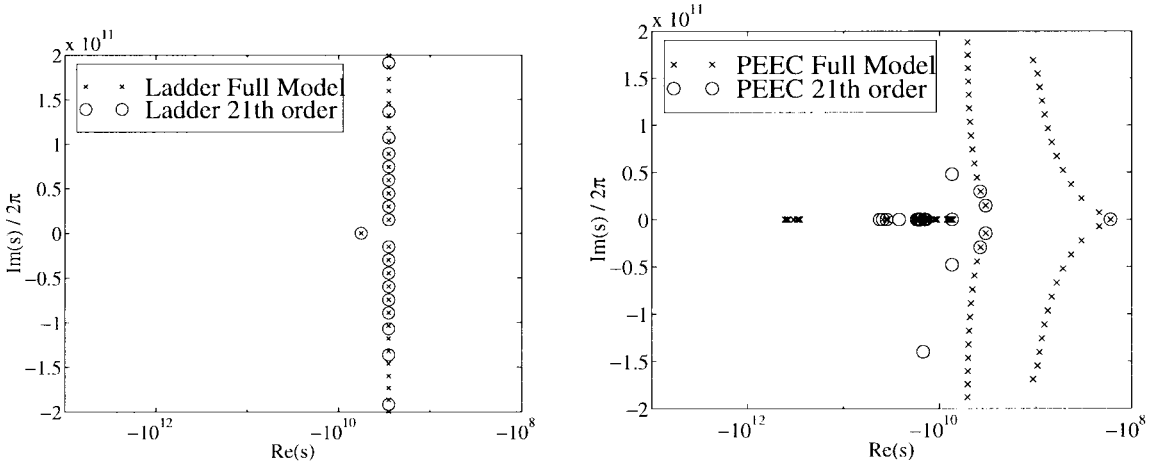


Fig. 10. Pole locations for (a) RLC ladder model and (b) the PEEC models. Note axis scaling.

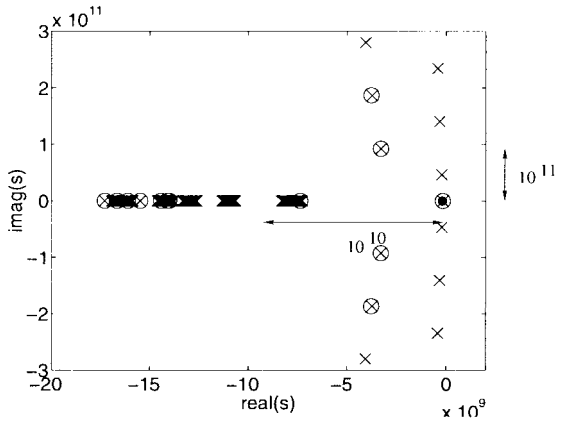


Fig. 11. Closeup of origin on linear scale for the pole locations of the PEEC models. Solid black circle marks origin.

two sets of vertically spaced poles. The complex conjugate pairs with real part in $[-10^9, -10^8]$ are exactly pole-zero cancelled and do not affect the frequency response perhaps due to symmetry in the two conductor geometry. The conjugate pairs with real part near -5×10^{-10} correspond to the vertical line of poles in Fig. 10(a) and are the dominant poles of the system, responsible for the resonances in Fig. 9. Note that they do not lie along a vertical line due to skin effects. In addition, there are also a large number of purely real poles. By noting the scaling of the plot, these real poles are the closest to the origin (see closeup in Fig. 11). For this reason, moment matching about the origin captures these poles first instead of the poles responsible for the resonant behavior.

The real poles result from the discretization of the conductor into bundles of filaments in order to capture skin and proximity effects. Because the system is driven from external terminals, it is difficult to excite the modes corresponding to the many interior filaments. For this reason, the effect on the frequency response of the large cluster of real poles near the origin is weak; in fact all of these poles are nearly or exactly cancelled by zeros. In addition, as the order is raised beyond 21, most additional poles are matched near the origin resulting in very slow convergence to the full model.

C. Improving the Model

The cluster of real, weak poles near the origin is not limited to the transmission line example and has been observed in most examples. We thus seek a geometry independent means of avoiding the matching of the weak poles near the origin. One approach is to expand about $s = \infty$ which would select the weak poles last and model generation would involve $\mathcal{L}^{-1}\mathcal{R}$. But for most discretizations, there are multiple panels at the nodes, and thus $\mathcal{L}$ is singular. The nodal formulation approach could be used for such expansions as pursued in [18].

Another approach is to use some multipoint scheme [24], [25] as developed in Section IV-B. Results of this approach are described next.

D. Results of Multipoint Model Order Reduction

With multipoint models computationally tractable, we can begin to investigate generating low order models. In this section we give results of expansions about nonzero s to give insight into the properties of the resulting models specific to interconnect analysis. The general topic of multipoint model generation via Krylov subspace methods is addressed in detail in [12].

1) Expansion Points: In this section we explore choosing a single nonzero s as an expansion point and discuss the quality of the resultant reduced order models. We then further improve the models by using multiple expansion points.

From Section V-B where $s_0 = 0$, the poles, s_i, closest to the origin were captured first. This was a natural conclusion of matching moments about the origin. One can also explain this occurrence by viewing the Arnoldi process as it was originally developed as method of eigenvalue computation on $\mathcal{R}^{-1}\mathcal{L}$. It is well known that the Arnoldi algorithm will converge fastest to eigenvalues, $\lambda_i = \frac{1}{s_i}$, which are well separated from other eigenvalues, and slower to clustered eigenvalues [27]. Since poles clustered close to the origin are very well separated under the map $\frac{1}{s}$ they will be captured first.

Now consider choosing $s_0 \in \mathbb{R}$ and $s_0 > 0$. For $s_0 \neq 0$, poles for which $\frac{1}{s_i - s_0}$ is well separated come first. Since all the poles have $\mathrm{Re}(s) < 0$, no pole will be closer than the distance

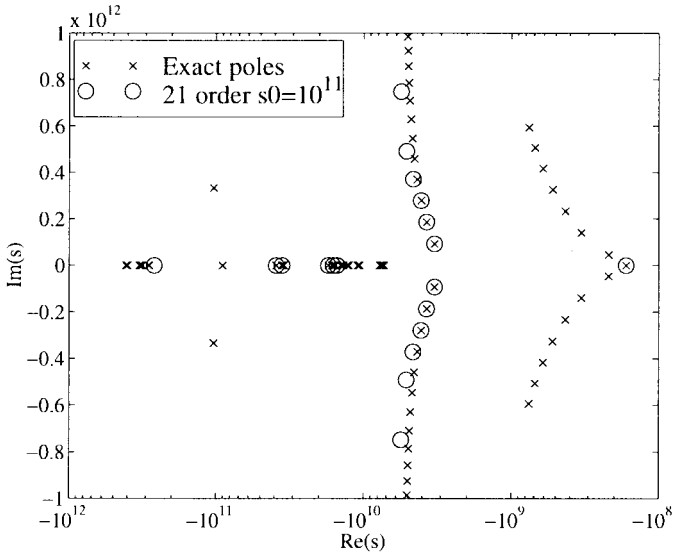


Fig. 12. Poles of reduced model for 2-D two conductor TEM line for $s_0 = 10^{11}$.

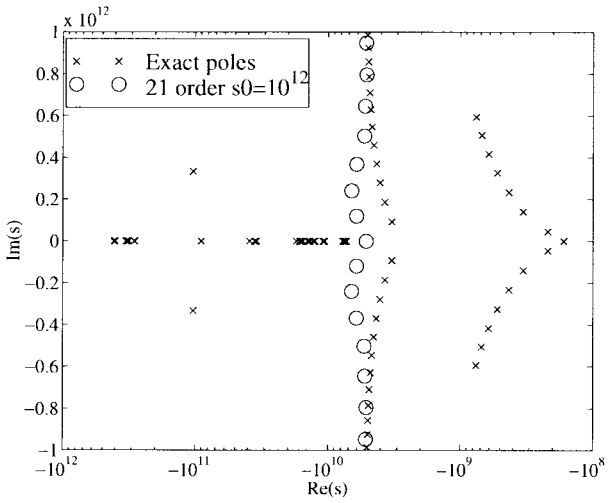


Fig. 13. Poles of reduced model for 2-D two conductor TEM line for $s_0 = 10^{12}$.

$|s_0|$. In particular, the weak poles at the origin appear more as a cluster and we would expect to capture the poles along the imaginary axis more readily. Additionally, the choice of a real expansion point implies solution of (35) will not involve complex arithmetic.

The results for various s_0 for a PEEC model for a 2-D TEM transmission line are shown in Figs. 12 and 13. We see that $s_0 = 10^{11}$ does slightly better compared to $s_0 = 0$ of Fig. 10 but still places multiple poles near the origin, and $s_0 = 10^{12}$ does not seem to capture any single pole accurately. To explain $s_0 = 10^{12}$, the expansion point is too far away, and the entire region of poles appears as a distance cluster for which convergence to any one pole is slow. While it may be possible to find a better choice for $10^{11} < s_0 < 10^{12}$, these results demonstrate that choosing an appropriate real expansion point *a priori* is difficult.

Instead, consider choosing $s_0 \in \mathbb{C}$. In fact, choosing purely imaginary expansion points is the common choice since it is the response along the $j\omega$ axis which is of interest. As with choosing the origin, the poles nearest the expansion point will tend to appear in the model first. For model reduction for interconnect analysis, we can exploit that the only large dense cluster of poles is at the origin, and thus by choosing an imaginary expansion point away from the origin, the algorithm will not stagnate as for $s_0 = 0$.

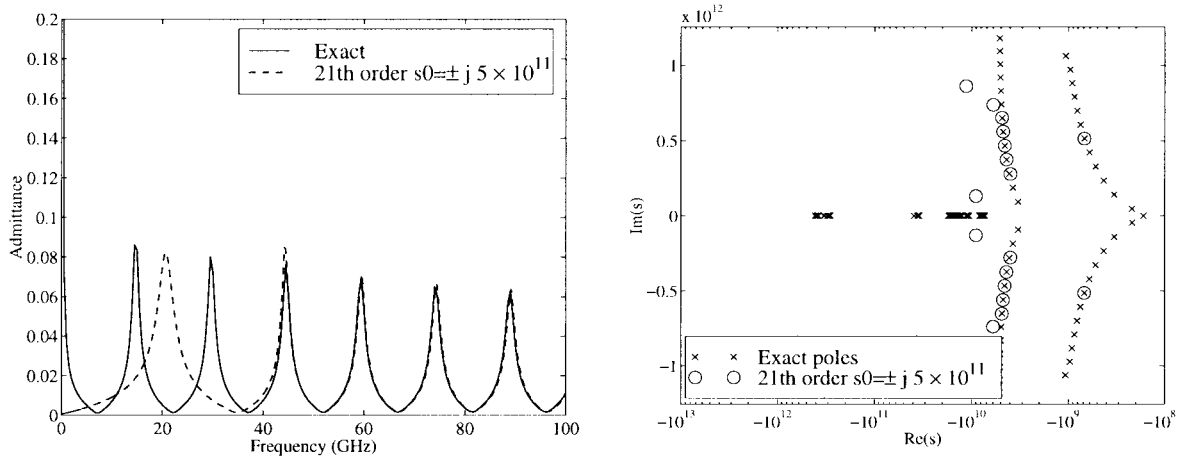
The advantage of complex s_0 comes at a cost. Solving (35) will involve complex arithmetic and is thus four times as expensive as the real case. However, to maintain a real reduced order model, moments must be matched at conjugate pair points, s_0 and s_0^*. It was pointed out in [28] that the $x = x_r + jx_i$ and $x^* = x_r - jx_i$ resulting from solving (35) at s_0 and s_0^*, respectively, generate only two independent vectors, x_r and x_i, for the Arnoldi algorithm. These directions can be computed with a single solve of (35) and thus the cost of a complex expansion point is only double that of a real one.

For the TEM example, we choose $s_0 = \pm j5 \times 10^{11}$ and in Fig. 14 we see that model order reduction matches the dominant poles starting at $\pm j5\times 10^{11}$ and moving outward as desired. Note that the model is close to optimal in the sense that almost every pole matched corresponds to a pole responsible for the resonant peaks.

To take advantage of the efficient approach developed here for multipoint expansions, consider matching many moments at multiple points along the imaginary axis. Consider choosing six points, $s_0 = 0, \pm j1\times 10^{11}, \pm j2\times 10^{11}, \pm j3\times 10^{11}, \pm j4\times 10^{11}, \pm j5\times 10^{11}$ and matching 1, 4, 4, 4, 4, 4 moments at each, respectively. The results for this 21st order model are shown in Fig. 15 where we see that this approximation accurately captured the dominant poles to a frequency comparable to that for the RLC ladder network of Figs. 10(a) and 9(a).

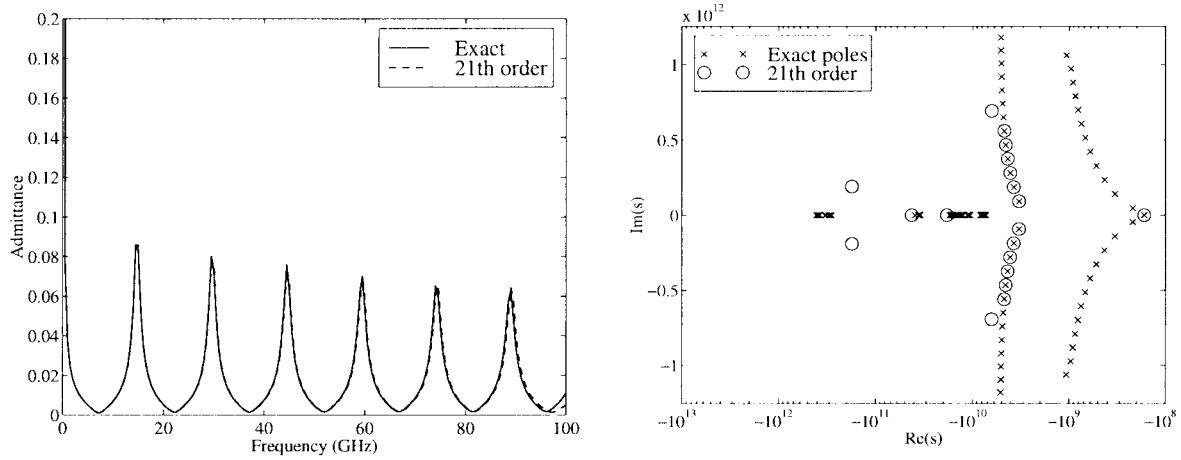
While the results in Fig. 15 were the underlying goal of this section, the choice of expansion points and number of moments to match at each did not come without trial and error. For instance, consider an 11th order model matching two moments at each point instead of four in the previous example. The results in Fig. 16 show that even though the model roughly captured the poles, it did not capture their magnitude well and thus did not give a very accurate frequency response. Comparing to Fig. 15, perhaps the influence of *all* the weak poles near the origin has an effect at these higher frequencies and more moments should be matched at zero. This is not the case as shown in Fig. 17. We now match five moments at $s = 0$ instead of 1, and while the response for the first two resonances near $s = 0$ improved, near the third around 40 GHz, it has worsened. Also, the pole for the second resonance has moved away from its exact value but the magnitude of the resonance is close to the exact. Both of these examples indicate that strictly observing pole locations is not a direct measure of error in the frequency response.

These examples illustrate the need for sophisticated methods of error analysis and expansion point and order selection. For the provably passive multipoint Krylov-subspace Arnoldi schemes used here, some of ideas of Grimme [12] might work here.



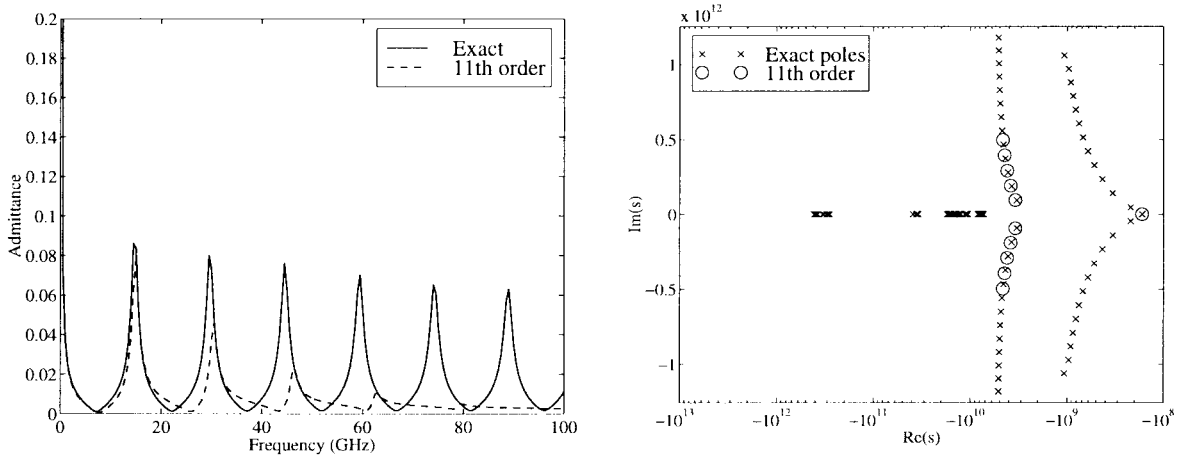
(a) (b)

Fig. 14. Admittance and poles of reduced model for 2-D two conductor TEM line for complex expansion point.



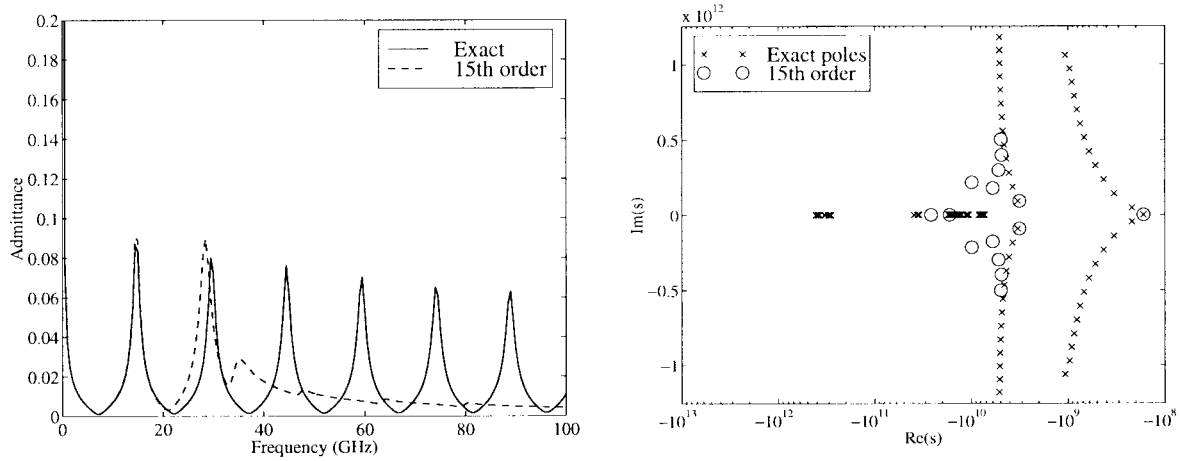
(a) (b)

Fig. 15. Admittance and poles of reduced model for 2-D two conductor TEM line for complex expansion points $s_0 = 0, \pm j1 \times 10^{11}, \pm j2 \times 10^{11}, \pm j3 \times 10^{11}, \pm j4 \times 10^{11}, \pm j5 \times 10^{11}$ and matching 1, 4, 4, 4, 4, 4 moments at each.



(a) (b)

Fig. 16. Admittance and poles of reduced model for 2-D two conductor TEM line for complex expansion points $s_0 = 0, \pm j1 \times 10^{11}, \pm j2 \times 10^{11}, \pm j3 \times 10^{11}, \pm j4 \times 10^{11}, \pm j5 \times 10^{11}$ and matching 1, 2, 2, 2, 2, 2 moments at each.



(a) (b)

Fig. 17. Admittance and poles of reduced model for 2-D two conductor TEM line for complex expansion points $s_0 = 0, \pm j1 \times 10^{11}, \pm j2 \times 10^{11}, \pm j3 \times 10^{11}, \pm j4 \times 10^{11}, \pm j5 \times 10^{11}$ and matching 5, 2, 2, 2, 2, 2 moments at each.

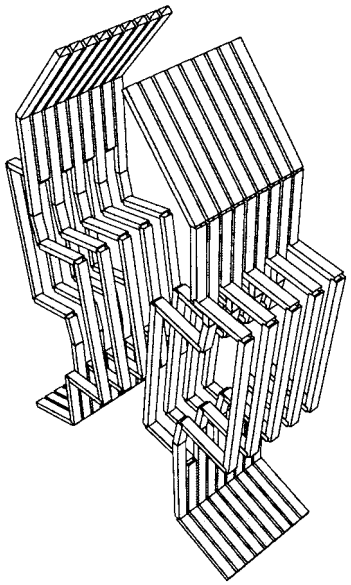


Fig. 18. An 18-pin backplane connector.

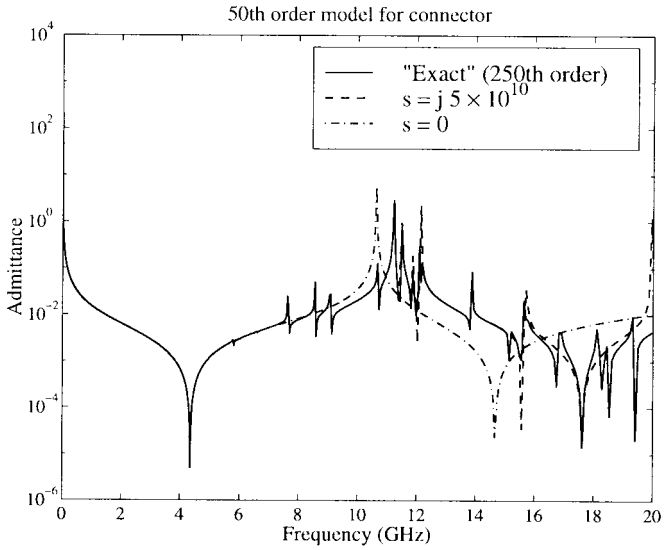


Fig. 19. Various reduced order models for the self-admittance of a middle pin from part of the back plane connector.

The point of the above discussion was to describe the need for automatic multipoint methods, but such methods could not be applied unless the algorithms of this paper make problems for $n > 10^4$ computationally possible. Next we show that indeed they do.

2) Results for a Practical Example: Finally, consider a real example to show the computational efficiency of multipoint expansions via the algorithms of this chapter. We generate a fiftieth order model for half of the backplane connector of Fig. 18. The discretization generated 1560 panels and 480 filaments. For simplicity, only one input and one output is modeled, $B = b$, corresponding to exciting only one of the middle pins. Assume we desire an accurate frequency response up to $\omega = 10^{11}$. Since poles tend to be matched outward from the expansion point, we match 48 moments about a midpoint $s_0 = j5\times 10^{10}$ and then two at $s_0 = 0$ to insure accurate capture of the DC behavior. To match moments at multiple points, as in the previous section, we use the algorithm in [13]. The self-admittance of the excited pin is shown in Fig. 19. The result is compared to a much higher order model of size 250. The fiftieth order model matches well up to around 15 GHz, compared to the poor response of a fiftieth order model for which all moments were matched at $s = 0$. The poles captured in the reduced model are shown in Fig. 20. The many real poles in the original model are greater than 10^7 and are out of range of the plot.

To observe the computational efficiency of the preconditioned iterative algorithm, consider one solve of (35) using the simple block preconditioner in (37) for a finer discretization of the connector resulting in 5112 panels and 2592 filaments. The results shown in Fig. 21 are good compared to no preconditioning.

One can improve the block preconditioner by using the preconditioner of (40) which uses a local inversion preconditioner for the capacitive part. For (40), the iteration count is smaller as shown in Fig. 22 and also the number of nonzeros

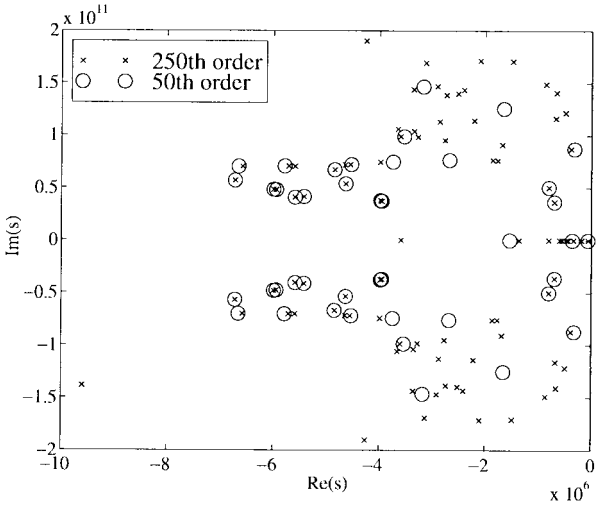


Fig. 20. Poles of reduced model and higher order model for part of the back plane connector.

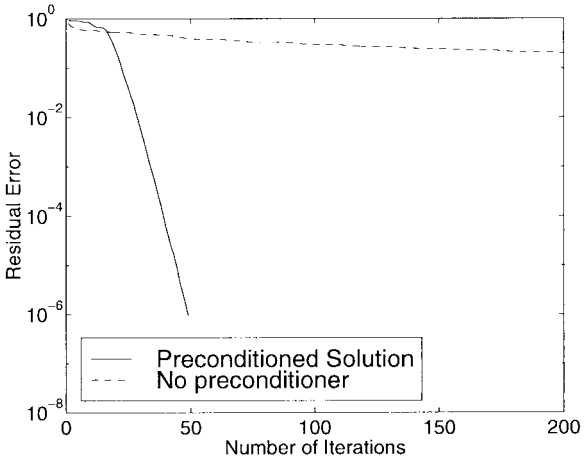


Fig. 21. Convergence of iterative solver for one solve of $MZ_{\rm EM}M^T$ using a block diagonal preconditioner.

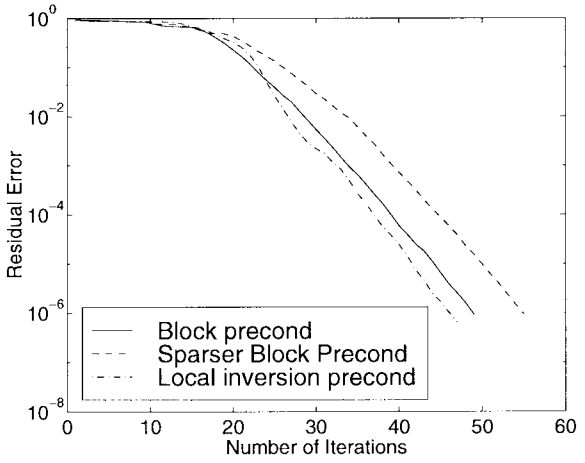


Fig. 22. Convergence of iterative solver for one solve of $MZ_{\rm EM}M^T$ using different preconditioners.

is considerably less as shown in Table I. The table and figure also include a block preconditioner with smaller blocks than in Fig. 21. This smaller block preconditioner had roughly the same number of nonzeros in the unfactored matrix as the local

TABLE I
NONZEROS AND CPU TIME FOR FACTORIZATION OF ONLY THE PRECONDITIONER. LOCAL INVERSION IS FOR A BLOCK DIAGONAL $\hat{L}$ AND LOCAL INVERSION BASED C

Preconditioner	Number of iterations	non-zeros before factor	non-zeros after factor	Factor time (CPU secs)
Block Diag 1	49	288622	894175	14.2
Block Diag 2	55	152974	675278	9.4
Local Inversion	47	143438	276561	4.5

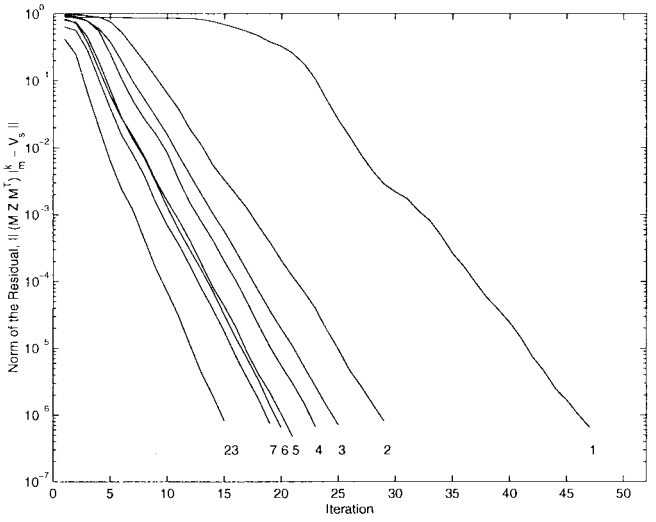


Fig. 23. Convergence of iterative solver for $MZ_{\rm EM}M^T$ using recycling. Numbers for each line correspond to solve number. With more vectors from recycling, the later solves converge faster.

inversion preconditioner, yet it was worse in all respects, as shown in the table.

The local inversion preconditioner required less time, fewer iterations and had many fewer nonzeros in the factored matrix. Since the CPU time to factor each of these is so small, choosing a denser preconditioner could significantly improve results. Such an optimization will not be pursed here. Note that for a denser preconditioner, the CPU time advantage of the local inversion preconditioner would become considerably more significant.

Just as for $s = 0$, if many moments are to be matched at a given s_0, then the Krylov subspace from previous moments can be reused. Consider now computing multiple moments at $s_0 = j * 5\times\ 10^{11}$.

For each iterative solve, we see that the number of iterations decreases as shown in Fig. 23. The 23 solves required a total of 422 matrix-vector products, compared to the roughly $23{*}47 = 1081$ that would be required without recycling, representing over a factor of two speedup. This speedup is counterbalanced by the memory consumption in storing the back vectors. For the above problem, storage of the dense L and P matrices with 5112 panels and 2592 filaments requires 263 MB. Storage of the back vectors with 6858 complex entries each requires 46 MB, which is over 17% of the total storage. For this small problem, such memory is acceptable, but for larger problems which require multipole or precorrected-FFT acceleration, such consumption is unacceptable. By noting that the most

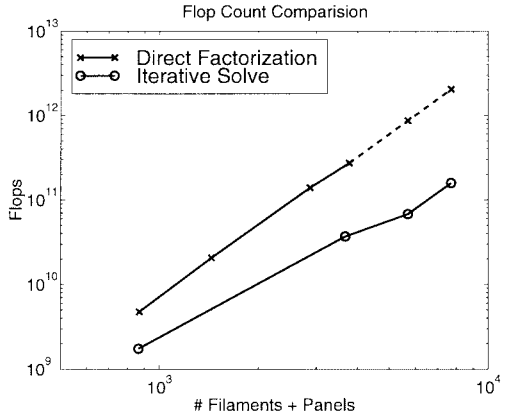


Fig. 24. CPU flops required to generate a fiftieth order model.

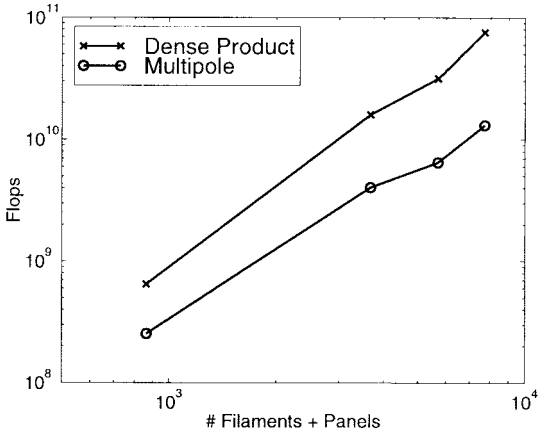


Fig. 25. Flop count projections for use of a multipole algorithm for the dense matrix vector product for the fiftieth order connector model.

benefit from recycling comes from the first few solves in Fig. 23, the storage of these back vectors could be stopped when memory consumption is a concern.

For the overall computational efficiency, Fig. 24 shows the total CPU floating point operations (flops) required to generate a fiftieth order model for the *full* connector for various levels of discretization. As can be seen, if (35) were solved by direct factorization, the flops would grow as $O(n^3)$, but with the iterative solver, the growth is only $O(n^2)$. Note that for a modest problem size still under 10^4 elements, the iterative algorithm is an order of magnitude faster than direct factorization.

With efficient iterative solution in place, the Multipole algorithm could be directly applied to bring the operation count and memory growth to $O(n)$. The benefits of such an approach are shown in Figs. 25 and 26. We see that using multipole acceleration is roughly a factor of five improvement in both time and memory over dense matrix vector products.

VI. Conclusion

In this paper, we developed a mesh formulated approach for passive model order reduction of the full quasistatic Maxwell's equations. We found that model reduction about $s = 0$ stagnates, that is, the reduced order transfer function along

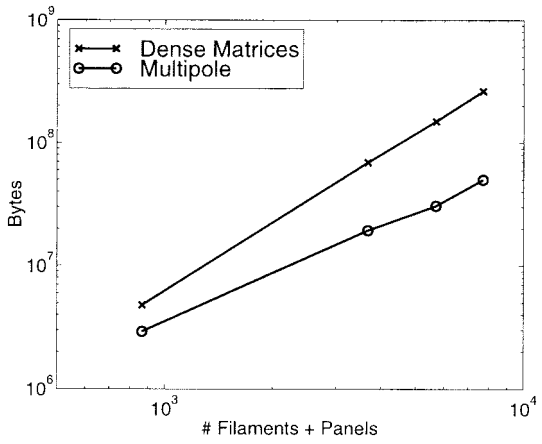


Fig. 26. Memory use projections for use of a multipole algorithm for the dense matrix vector product for the fiftieth order connector model.

the $j\omega$ axis converges slowly to the exact transfer function. This behavior was attributed to the large clusters of poles near the origin, a common feature of the PEEC RLC method. A common solution to such a problem is to match moments about some $s \neq 0$. Computations for the Arnoldi algorithm for $s \neq 0$ however, would be extremely expensive due to the required nested dense iterative solve.

The advantage of the mesh formulation became apparent for computing multipoint expansions because the first order state-space form could be reformulated in a pure mesh form. From a pure mesh form, effective preconditioning and multipole acceleration could be applied to give an algorithm of nearly $O(n)$ flop and memory growth which, for $n \approx 10^4$, was 50 times faster and consumed 5 times less memory than direct factorization. Such growth rates make such algorithms essential as problems near $n = 10^5$.

While we saw that an accurate low order model could be generated for the backplane connector, the method is not fully automatic in regard to error control. An accurate model was generated only after comparing the response of the reduced model to the exact model. Unfortunately, the exact response is generally not available for comparison. It is thus not clear how to fully automate the generation of a good model. Methods to investigate include [24] and recently [12].

Acknowledgment

The authors would like to thank M. Tsuk, Digital Equipment Corporation, for the connector example.

References

[1] A. E. Ruehli, "Equivalent circuit models for three-dimensional multiconductor systems," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-22, pp. 216–221, Mar. 1974.

[2] P. Feldmann and R. W. Freund, "Efficient linear circuit analysis by Padé approximation via the Lanczos process," in *EURO-DAC'94 with EURO-VHDL'94*, Sept. 1994.

[3] W. B. Gragg and A. Lindquist, "On the partial realization problem," *Linear Algebra Applicat.*, vol. 50, pp. 277–319, 1983.

[4] K. Gallivan, E. Grimme, and P. Van Dooren, "Asymptotic waveform evaluation via a Lanczos method," *Appl. Math Lett.*, vol. 7, no. 5, pp. 75–80, 1994.

[5] A. Odabasioglu, M. Celik, and L. Pileggi, "Prima: Passive reduced-order interconnect macromodeling algorithm," in *Int. Conf. Computer Aided-Design*, San Jose, CA, Nov. 1997.
[6] L. M. Silveira, M. Kamon, I. M. Elfadel, and J. K. White, "A coordinate-transformed Arnoldi algorithm for generating guaranteed stable reduced order models of RLC circuits," in *IEEE/ACM Int. Conf. Computer Aided-Design*, San Jose, CA, Nov. 1996.
[7] C. Desoer and E. Kuh, *Basic Circuit Theory*. New York: McGraw-Hill, 1969.
[8] R. F. Harrington, *Field Computation by Moment Methods*. New York: MacMillan, 1968.
[9] L. T. Pillage and R. A. Rohrer, "Asymptotic waveform evaluation for timing analysis," *IEEE Trans. Computer-Aided Design*, vol. 9, pp. 352–366, Apr. 1990.
[10] L. M. Silveira, M. Kamon, and J. K. White, "Algorithms for coupled transient simulation of circuits and complicated 3-D packaging," *IEEE Trans. Comp., Packag., Manufact. Technol.*, vol. 18, pp. 92–98, Feb. 1995.
[11] K. J. Kerns, I. L. Wemple, and A. T. Yang, "Stable and efficient reduction of substrate model networks using congruence transforms," in *IEEE/ACM Int. Conf. Computer-Aided Design*, San Jose, CA, Nov. 1995, pp. 207–214.
[12] E. J. Grimme, "Krylov projection methods for model reduction," Ph.D. dissertation, Univ. Illinois, Urbana-Champaign, 1997.
[13] I. M. Elfadel and D. D. Ling, "A block rational Arnoldi algorithm for multipoint passive model-order reduction of multiport RLC networks," in *Int. Conf. Computer-Aided Design*, San Jose, CA, Nov. 1997.
[14] M. Kamon, M. J. Tsuk, and J. White, "Fasthenry: A multipole-accelerated 3-d inductance extraction program," *IEEE Trans. Microwave Theory Tech.*, vol. 42, pp. 1750–1758, Sept. 1994.
[15] J. Garrett, A. Ruehli, and C. Paul, "Efficient frequency domain solutions for sPEEC EFIE for modeling 3d geometries," in *Proc. Zurich Symp. Electromag. Compat.*, Zurich, Switzerland, Mar. 1995, pp. 179–184.
[16] N. Deo, *Graph Theory with Applications to Engineering and Computer Science*. Englewood Cliffs, NJ: Prentice-Hall, 1974.
[17] M. Kamon, "Fast parasitic extraction and simulation of three-dimensional interconnect via quasistatic analysis," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, Feb. 1998.
[18] N. Marques, M. Kamon, J. White, and L. M. Silveira, "An efficient algorithm for fast parasitic extraction and passive order reduction of 3d interconnect models," in *Proc. DATE'98, Design Automat. Test Euro.*, Feb. 1998.
[19] Y. Saad, *Iterative Methods for Sparse Linear Systems*. Boston, MA: PWS, 1996.
[20] R. Telichevesky, K. Kundert, and J. White, "Efficient AC and noise analysis of two-tone RF circuits," in *Proc. 33rd Design Automat. Conf.*, Las Vegas, NV, June 1996.
[21] J. R. Phillips, E. Chiprout, and D. D. Ling, "Efficient full-wave electromagnetic analysis via model-order reduction of fast integral transforms," in *Proc. 33rd Design Automat. Conf.*, Las Vegas, NV, June 1996.
[22] M. Kamon, N. Marques, and J. White, "FastPep: A fast parasitic extraction program for complex three-dimensional geometries," in *IEEE/ACM Int. Conf. Computer Aided-Design*, San Jose, CA, Nov. 1997.
[23] K. Nabors and J. White, "Fast capacitance extraction of general three-dimensional structures," *IEEE Trans. Microwave Theory Tech.*, June 1992.
[24] E. Chiprout and M. S. Nakhla, "Analysis of interconnect networks using complex frequency hopping (CFH)," *IEEE Trans. Computer-Aided Design*, vol. 14, pp. 186–200, Feb. 1995.
[25] K. Gallivan, E. Grimme, and P. Van Dooren, "Multi-point Padé approximants of large-scale systems via a two-sided rational Krylov algorithm," in Proc. 33rd *IEEE Conf. Decision Contr.*, Lake Buena Vista, FL, Dec. 1994.
[26] S. A. Vavasis, "Preconditioning for boundary integral equations," *SIAM J. Matrix Anal. Appl.*, vol. 13, no. 3, pp. 905–925, July 1992.
[27] L. N. Trefethen and D. Bau, *Numerical Linear Algebra*. Philadelphia, PA: Soc. Ind. Appl. Math., 1997.
[28] A. Ruhe, "The rational Krylov algorithm for nonsymmetric eigenvalue problems III: Complex shifts for real matrices," *BIT*, vol. 34, pp. 165–176, 1994.

Mattan Kamon received the B.S. degree in engineering science and the M.A. degree in mathematics, both from Pennsylvania State University, University Park, in 1991 and the M.S. degree in electrical engineering and the Ph.D. degree in computer science, both from the Massachusetts Institute of Technology, Cambridge, in 1994 and 1998, respectively.

For his graduate work he developed efficient algorithms for 3-D interconnect parameter extraction and simulation. He joined Microcosm Technologies, Cambridge, after graduation, where he has developed CAD tools for fast inductance and magnetic force calculation for package analysis and microelectromechanical system (MEMS) design.

Nuno Alexandre Marques (S'97) was born in Lisbon, Portugal. He received the Engineer's degree in electrical and computer engineering from the Instituto Superior Técnico, Technical University of Lisbon, Portugal, in 1995 and is currently pursuing the M.S. degree in efficient algorithms for extraction and passive order-reduction of 3-D interconnect models at the Instituto de Engenharia de Sistemas e Computadores, Lisbon.

His research interests are in developing algorithms and techniques for three-dimensional modeling and simulation of electromagnetic field effects in interconnects and package structures.

Luís Miguel Silveira (S'86–A'86–M'95) was born in Lisbon, Portugal. He received the Engineer's (summa cum laude) and Master's degrees in electrical and computer engineering from the Instituto Superior Técnico, Technical University of Lisbon in 1986 and 1989, respectively, and the M.S., E.E., and Ph.D. degrees from the Massachusetts Institute of Technology, Cambridge, in 1990, 1991, and 1994, respectively.

From 1992 to 1994, he was supported by an IBM doctoral fellowship. He is currently an Associate Professor at the Instituto Superior Técnico, Technical University of Lisbon, a Researcher in the Electronics Division, INESC, Instituto de Engenharia de Sistemas e Computadores, Lisbon, Portugal, and a founding member of the Lisbon Center of the Cadence European Laboratories. His research interests are in various aspects of computer-aided design of integrated circuits with emphasis on parallel computer algorithms and the theoretical and practical issues concerning numerical simulation methods for circuit design problems.

Dr. Silveira is a member of Sigma Xi.

Jacob White (S'80–M'83) received the B.S. degree in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, and the S.M. and Ph.D. degrees in electrical engineering and computer science from the University of California, Berkeley.

He worked at the IBM T. J. Watson Research Center, Yorktown Heights, NY, from 1985 to 1987, was the Analog Devices Career Development Assistant Professor at the Massachusetts Institute of Technology, from 1987 to 1989, he is currently a Professor at MIT and his research interests are in serial and parallel numerical algorithms for problems in circuit, interconnect, device, and microelectromechanical system design.

Dr. White was a 1988 Presidential Young Investigator and was an Associate Editor of the IEEE TRANSACTIONS ON COMPUTER-AIDED DESIGN, from 1992 until 1996.