PRIMA: Passive Reduced-order Interconnect Macromodeling Algorithm[*]

Altan Odabasioglu, Mustafa Celik[+], and Lawrence T. Pileggi

Department of Electrical and Computer Engineering
Carnegie Mellon University
Pittsburgh, PA 15213

Abstract

This paper describes PRIMA, an algorithm for generating provably passive reduced order N-port models for RLC interconnect circuits. It is demonstrated that, in addition to requiring macromodel stability, macromodel passivity is needed to guarantee the overall circuit stability once the active and passive driver/load models are connected. PRIMA extends the block Arnoldi technique to include guaranteed passivity. Moreover, it is empirically observed that the accuracy is superior to existing block Arnoldi methods. While the same passivity extension is not possible for MPVL, we observed comparable accuracy in the frequency domain for all examples considered. Additionally, a path tracing algorithm is used to calculate the reduced order macromodel with the utmost efficiency for generalized RLC interconnects.
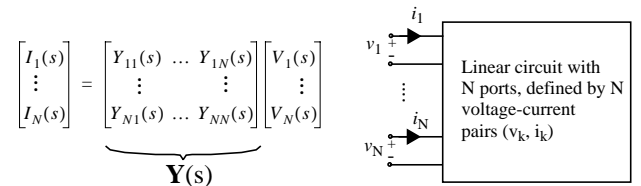
1. Introduction

As integrated circuits and systems are designed with smaller feature sizes and for faster operation, RLC interconnect effects have a more dominant impact on signal propagation than ever before. In addition, parasitic coupling effects and reduced power supply voltage levels make interconnect modeling increasingly important. Since these interconnect models can contain thousands of tightly coupled R-L-C components, reduced order macromodels are imperative [1][2][3][4]. Ideally, a simulator would isolate the large linear portions of the circuit from the nonlinear elements (e.g., transistor models) and preprocess them into reduced order multiport macromodels.

It is well known that an N-port can be fully represented by its admittance parameters in the Laplace domain, however, the objective is to apply model order reduction to produce low order rational approximations for each entry in $\mathbf{Y}(s)$ (see Fig.1). To find $\mathbf{Y}(s)$, voltage sources are connected to the ports and the currents into the ports are measured. The voltage sources are the inputs to the system and the port currents are the outputs. A single-input single-output (SISO) N-port model approach would perform model order reduction on each term Y_{ij} individually. Both Asymptotic Waveform Evaluation (AWE) [1] and Padé via Lanczos (PVL) [2], which are Padé approximations, can perform SISO reduction by matching 2q moments for a qth order approximation of each Y_{ij} term. The Arnoldi Algorithm [4] can also be used to obtain SISO approximations, however it matches only q moments for a qth order approximation. MPVL (Matrix Padé via Lanczos) [5] and Block Arnoldi [6] are multi-input multi-output (MIMO) versions of PVL and Arnoldi respectively. In the block techniques, the system Modified Nodal Analysis (MNA) matrices are directly reduced by matrix transformations.

Regardless of the reduction method used in all of the approaches cited above, the reduced order model of an RLC circuit can have unstable poles. It is always possible to obtain an asymptotically stable model by simply discarding the unstable poles, however, passivity is not guaranteed. In addition, discarding unstable poles requires re-adjustment of the residues to improve the quality of the approximation. Passivity uncertainty is problematic since even the test for N-port passivity can be very costly for a large number of ports [7]. The coordinate transformed Arnoldi Algorithm [8] was introduced as a remedy for the instability problem, but it does not guarantee passivity. The passivity extension of this stable Arnoldi algorithm was recently developed in [9], however its applicability is limited to RC circuits only. The PACT algorithm [3] proposed a new direction for passive reduced-order model for RC circuits based on congruence transformations. The same authors proposed Split Congruence Transformations [10] for passive reductions of RLC circuits, producing equivalent circuit realizations. In [10], however, the extra steps required to split the transformation matrix can result in a decrease in accuracy and efficiency. Moreover, the passivity proof is somewhat controversial, and we will consider a more complete proof in this paper.

$$\begin{bmatrix} I_1(s) \\ \vdots \\ I_N(s) \end{bmatrix} = \underbrace{\begin{bmatrix} Y_{11}(s) & \dots & Y_{1N}(s) \\ \vdots & & \vdots \\ Y_{N1}(s) & \dots & Y_{NN}(s) \end{bmatrix}}_{\mathbf{Y}(s)} \begin{bmatrix} V_1(s) \\ \vdots \\ V_N(s) \end{bmatrix}$$

FIGURE 1: The multiport representation of a linear circuit.

1

A passive system denotes a system that is incapable of generating energy, and hence one that can only absorb energy from the sources used to excite it [11]. As we will show in Section 2.2., passivity is an important property to satisfy because stable, but not passive macromodels can produce unstable systems when connected to other stable, even passive, loads. A property in classical circuit theory states that: interconnections of stable systems may not necessarily be stable; but (strictly) passive circuits are (asymptotically) stable; and arbitrary interconnections of (strictly) passive circuits are (strictly) passive, and, therefore, (asymptotically) stable [12].

In this paper, we propose a Passive Reduced-order Interconnect Macromodeling Algorithm, PRIMA, based on the Block Arnoldi Algorithm but with congruence transformations that produce provably passive reduced order macromodels for arbitrary RLC circuits. PRIMA has accuracy comparable to MPVL and superior to Block Arnoldi. Furthermore, the block Arnoldi vectors are generated with the utmost efficiency following the algorithms in RICE [13] that are used to calculate moments. This includes efficient handling of interconnect trees and meshes, as in RICE, but with renewed focus on efficient handling of large problems with a huge number of mutual inductances.

2. Background

To obtain the admittance matrix of a multiport, voltage sources are connected to the ports. The multiport, along with these sources, constitutes the Modified Nodal Analysis (MNA) equations:

$$\mathbf{C}\,\dot{\mathbf{x}}_n = -\mathbf{G}\,\mathbf{x}_n + \mathbf{B}\,\mathbf{u}_p$$
$$\mathbf{i}_p = \mathbf{L}^T\,\mathbf{x}_n \tag{1}$$

The $\mathbf{i}_p$ and $\mathbf{u}_p$ vectors denote the port currents and voltages respectively and

$$\mathbf{G} \equiv \begin{bmatrix} \mathbf{N} & \mathbf{E} \\ -\mathbf{E}^T & \mathbf{0} \end{bmatrix} \qquad \mathbf{C} \equiv \begin{bmatrix} \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{H} \end{bmatrix} \qquad \mathbf{x}_n \equiv \begin{bmatrix} \mathbf{v} \\ \mathbf{i} \end{bmatrix} \tag{2}$$

where $\mathbf{v}$ and $\mathbf{i}$ are the MNA variables corresponding to the node voltages, inductor and voltage source currents respectively. The $n \times n$ matrices $\mathbf{G}$ and $\mathbf{C}$ represent the conductance and susceptance matrices (except that the rows corresponding to the current variables are negated as in [7]). $\mathbf{N}$, $\mathbf{Q}$ and $\mathbf{H}$ are the matrices containing the stamps for resistors, capacitors and inductors respectively. $\mathbf{E}$ consists of ones, minus ones and zeros, which represent the current variables in KCL equations. Provided that the original N-port is composed of passive linear elements only, $\mathbf{Q}$, $\mathbf{H}$ and $\mathbf{N}$ are symmetric nonnegative definite matrices. This implies $\mathbf{C}$ is also symmetric and nonnegative definite. Since this is an N-port formulation, whereby the only sources are the voltage sources at the N port nodes, $\mathbf{B}=\mathbf{L}$. But we maintain the separate $\mathbf{B}$ and $\mathbf{L}$ notation for generality of the equations.

Returning to equation (1), following the notation in [2] we define

$$\mathbf{A} \equiv -\mathbf{G}^{-1}\mathbf{C} \text{ and } \mathbf{R} \equiv \mathbf{G}^{-1}\mathbf{B}. \tag{3}$$

With unit voltages at the ports, taking the Laplace transformation of (1) and solving for the port current variables, the y-parameter matrix is given as

$$\mathbf{Y}(s) = \mathbf{L}^T(\mathbf{I}_n - s\mathbf{A})^{-1}\mathbf{R} \tag{4}$$

where $\mathbf{I}_n$ is the $n \times n$ identity matrix. It is apparent from (4) that the eigenvalues of $\mathbf{A}$ represent the inverses of the poles of $\mathbf{Y}(s)$.

Using any of the aforementioned model-order reduction techniques, we can find reduced order rational approximations to $Y_{jk}(s)$ terms, for all j, k ≤ N. The reduced-order $\mathbf{Y}(s)$ can then be simulated along with other nonlinear and linear portions of the complete circuit using a simulator that employs either recursive convolution [14] or state-space realization [7], both of which have linear complexity. If the reduction is block, the reduced order multi-input multi-output circuit can also be realized using linear circuit elements.

2.1. Block Arnoldi Algorithm

The Block Arnoldi algorithm reduces the system matrix $\mathbf{A}$ in (3) to a small block upper Hessenberg matrix $\mathbf{H}_q$. To do so requires an orthonormal basis, $\mathbf{X}$, for the corresponding Krylov space which satisfies the following:*

$$colsp(\mathbf{X}) = Kr\left(\mathbf{A}, \mathbf{R}, \left\lfloor \frac{q}{N} \right\rfloor\right)$$
$$\mathbf{X}^T \mathbf{A}\,\mathbf{X} = \mathbf{H}_q \tag{5}$$
$$\mathbf{X}^T \mathbf{X} = \mathbf{I}_q$$

where N is the number of ports and $\mathbf{I}_q$ is a $q \times q$ identity matrix. The Krylov space is defined as

$$Kr(\mathbf{A}, \mathbf{R}, k) \equiv colsp[\mathbf{R}, \mathbf{AR}, \mathbf{A}^2\mathbf{R}, ..., \mathbf{A}^k\mathbf{R}]. \tag{6}$$

Finding the reduced order admittance matrix can be explained by a change of variable,

$$\mathbf{x}_n = \mathbf{X}_{\{n \times q\}}\,\mathbf{z}_q. \tag{7}$$

where $\mathbf{z}_q$ is now the reduced order system variable, which reduces the number of unknowns in the system (q is generally much smaller than n). Substituting (7) into (1), then multiplying the first equation by $\mathbf{X}^T\mathbf{G}^{-1}$ yields

$$\mathbf{H}_q\dot{\mathbf{z}}_q = \mathbf{z}_q - \mathbf{X}^T\,\mathbf{R}\,\mathbf{u}_p$$
$$\mathbf{i}_p = \mathbf{L}^T\mathbf{X}\,\mathbf{z}_q \tag{8}$$

Therefore, in the Laplace domain,

$$\hat{\mathbf{Y}}(s) = \mathbf{L}^T\mathbf{X}(\mathbf{I}_q - s\mathbf{H}_q)^{-1}\mathbf{X}^T\,\mathbf{R} \tag{9}$$

where $\mathbf{I}_q$ is a $q \times q$ identity matrix.

The reduced order system equations and admittance matrix are given by (8) and (9) respectively. The poles of

---

*The $\lfloor . \rfloor$ operator is the truncation to the nearest integer towards zero.

$$Y_1(s) = \frac{s+4}{s^2+2s+5}$$

$$Y_{dr}(s) = 0.06 + 0.056s$$

$$Z_c(s) = \frac{s^2+2s+5}{0.056s^3+0.172s^2+1.4s+4.3} \Rightarrow \text{ Poles at: } -3.074, \ \pm 5j$$
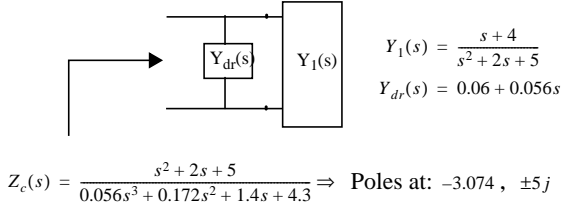
FIGURE 2: A non-passive system example demonstrating potential instability.

the reduced order system are the reciprocal eigenvalues of $\mathbf{H}_q$. A complete pole/residue decomposition can be obtained by eigendecomposing $\mathbf{H}_q$. Using the information in [6], it can be shown that the first $\left\lfloor \frac{q}{N} \right\rfloor$ moments of $\hat{\mathbf{Y}}(s)$ in (9) match those of $\mathbf{Y}(s)$ in (4).

2.2. Importance of Passivity

It is always possible to come up with stable reduced order macromodels by utilizing a number of heuristics, however, none of these tricks can be used to obtain provably passive approximations. Moreover, in [7] it was shown that the test for passivity of an AWE-reduced N-port macromodel is prohibitive in terms of CPU run time cost. Fig.2 is a numerical example generated in [7] that demonstrates the passivity problem. $Y_1(s)$ in this figure represents a reduced order transfer function which has all poles and zeros in the left half plane. $Y_{dr}(s)$ represents a capacitor and resistor in parallel. If we drive this circuit with a current source, it will oscillate at $2.5/\pi$ Hz. To show that it is also a practical problem, we took a simple interconnect and connected the load and the nonlinear driver as shown in the examples section in Fig.5. The interconnect is represented by a fifth order approximation obtained by PVL [2]. The figure clearly shows the growing oscillations at the output (instability) although all of the poles obtained from PVL were stable. A Thevenin equivalent linear driver with a resistance of 2 ohms generates a similar instability for this 2-port example.

3. PRIMA: Passive Reduced-order Interconnect Macromodeling Algorithm

The Block Arnoldi Algorithm is employed in PRIMA to generate the orthonormal basis for a congruence transformation matrix. After $\left\lfloor \frac{q}{N} \right\rfloor + 1$ (the extra step is not necessary when $\frac{q}{N}$ is an integer) iterations of PRIMA, the $n \times q$ matrix $\mathbf{X}$ is found such that:

$$colsp(\mathbf{X}) = Kr\left(\mathbf{A}, \mathbf{R}, \left\lfloor \frac{q}{N} \right\rfloor\right) \tag{10}$$

$$\mathbf{X}^T \mathbf{X} = \mathbf{I}_q$$

FIGURE 3: The Passive Reduction Algorithm

- Connect voltage sources to the multiport & obtain the MNA matrices as in (2).
- Set $[\, \mathbf{b}_1 \mid \mathbf{b}_2 \mid ... \mid \mathbf{b}_p \,] = \mathbf{B}$ and $[\, \mathbf{l}_1 \mid \mathbf{l}_2 \mid ... \mid \mathbf{l}_p \,] = \mathbf{L}$
- Solve $\mathbf{GR} = \mathbf{B}$ for $\mathbf{R}$
- $(\mathbf{X}_0, \mathbf{K}) = qr(\mathbf{R})$; qr factorization of $\mathbf{R}$
- Set $n = int\left(\frac{q}{N}\right) + 1$
- For $k = 1, 2, ..., n$
   Set $\mathbf{V} = \mathbf{CX}_{k-1}$
   Solve $\mathbf{G}\,\mathbf{X}_k^{(0)} = \mathbf{V}$ for $\mathbf{X}_k^{(0)}$
   For $j = 1, .., k$
      $\mathbf{H} = \mathbf{X}_{k-j}^T \mathbf{X}_k^{(j-1)}$
      $\mathbf{X}_k^{(j)} = \mathbf{X}_k^{(j-1)} - \mathbf{X}_{k-j}\mathbf{H}$
   $(\mathbf{X}_k, \mathbf{K}) = qr(\mathbf{X}_k^{(k)})$; qr factorization of $\mathbf{X}_k^{(k)}$
- Set $\mathbf{X} = \begin{bmatrix} \mathbf{X}_0 \ \mathbf{X}_1 \ ... \ \mathbf{X}_{k-1} \end{bmatrix}$ and truncate $\mathbf{X}$ so that it has q columns only.
- Compute $\tilde{\mathbf{C}} = \mathbf{X}^T \mathbf{C}\, \mathbf{X}$, $\tilde{\mathbf{G}} = \mathbf{X}^T \mathbf{G}\, \mathbf{X}$
- Find eigendecomposition of $\tilde{\mathbf{G}}^{-1}\tilde{\mathbf{C}}$: $\tilde{\mathbf{G}}^{-1}\tilde{\mathbf{C}} = \mathbf{S}\Lambda\mathbf{S}^{-1}$ *
   $\Lambda = diag(\lambda_1, \lambda_2, ..., \lambda_q)$
- To find poles and residues for $\hat{Y}_{i,j}(s)$:
   Solve $\tilde{\mathbf{G}}\mathbf{w} = \mathbf{X}^T\mathbf{b}_j$ for $\mathbf{w}$ *
   Set $\mu = \mathbf{S}^T\mathbf{X}^T\mathbf{l}_i$ and $\nu = \mathbf{S}^{-1}\mathbf{w}$
   $\hat{Y}_{i,j}(s) = \sum_{n=1}^{q} \frac{\mu_n \nu_n}{1 + s\lambda_n}$
- Set $\hat{\mathbf{Y}}(s) = \begin{bmatrix} \hat{Y}_{1,1} & ... & \hat{Y}_{1,p} \\ \vdots & & \vdots \\ \hat{Y}_{p,1} & ... & \hat{Y}_{p,p} \end{bmatrix}$

   *Inversion of $\tilde{\mathbf{G}}$ can be avoided.

In the classical Arnoldi approach [4], the reduced order $\mathbf{Y}(s)$ is calculated using the $q \times q$ upper Hessenberg matrix in (5) as shown in (9):

$$\hat{\mathbf{Y}}(s) = \mathbf{L}^T\mathbf{X}(\mathbf{I} - s\mathbf{H})^{-1}\mathbf{X}^T\mathbf{R} \tag{11}$$

In our variation, the conductance and susceptance matrices are directly reduced so that passivity is preserved during reduction. Applying the change of variable $\mathbf{x}_n = \mathbf{X}_{\{n \times q\}}\tilde{\mathbf{x}}_q$ in (1), and multiplying the first row by $\mathbf{X}^T$ from (10) yields

$$(\mathbf{X}^T\mathbf{C}\,\mathbf{X})\dot{\tilde{\mathbf{x}}}_q = -(\mathbf{X}^T\mathbf{G}\,\mathbf{X})\tilde{\mathbf{x}}_q + (\mathbf{X}^T\mathbf{B})\mathbf{u}_p \tag{12}$$

$$\mathbf{i}_p = (\mathbf{L}^T\mathbf{X})\,\dot{\tilde{\mathbf{x}}}_q$$

The reduced order MNA matrices are, therefore,

$$\tilde{\mathbf{C}} = \mathbf{X}^T\mathbf{C}\,\mathbf{X} \qquad \tilde{\mathbf{G}} = \mathbf{X}^T\mathbf{G}\,\mathbf{X}$$
$$\tilde{\mathbf{B}} = \mathbf{X}^T\mathbf{B} \qquad \tilde{\mathbf{L}} = \mathbf{X}^T\mathbf{L} \tag{13}$$

These types of transformations are known as congruence transformations. Congruence transformations were first introduced by [3] for order reduction of circuits. From (12) and (13), the reduced $\mathbf{Y}(s)$, namely $\hat{\mathbf{Y}}(s)$, is now

$$\hat{\mathbf{Y}}(s) = \tilde{\mathbf{L}}^T(\tilde{\mathbf{G}} + s\tilde{\mathbf{C}})^{-1}\tilde{\mathbf{B}} \tag{14}$$

Since the size of $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{C}}$ is typically very small, it is easy to find the poles and zeros of $\hat{\mathbf{Y}}(s)$ by eigendecomposition. The complete algorithm is given in Fig.3. It employs the Block Arnoldi Algorithm using modified Gram-Schmidt orthogonalization [6], which is mathematically equivalent to ordinary Gram-Schmidt process, but behaves better numerically [15]. In addition, it is possible to avoid the inversion of $\tilde{\mathbf{G}}$ to find the poles and residues by using a generalized eigendecomposition. In this case, the computation of $\tilde{\mathbf{G}}^{-1}\tilde{\mathbf{B}}$ can be avoided by using (31) and replacing it by $\mathbf{X}^T\mathbf{R}$. It is observed that this scheme is numerically much better.

The complexity of the algorithm to produce q poles for an N-port is slightly less than AWE, PVL and MPVL. It requires 1 LU factorization (or path tracing equivalent as explained in Section 4.) of the $\mathbf{G}$ (MNA conductance) matrix, which dominates all the other computational costs and is common in all reduction techniques. However, to find q poles, only q backward-forward substitutions are needed, whereas in MPVL, PVL and AWE, twice as many are required. As in MPVL, there will be only one eigendecomposition to find the poles and residues, whereas PVL requires N^2 eigendecompositions, since for each $Y_{ij}(s)$, there will be a different T_q. AWE will solve the N^2 different Hankel matrices to get to the poles.

3.1. Preservation of Passivity

If the system described by (1) and (2) is reduced by the transformations in (13), it can be shown that the reduced system is always passive. In [16], necessary and sufficient conditions for the system admittance matrix $\hat{\mathbf{Y}}(s)$ (eqn. (14)) to be passive are:

1. $\hat{\mathbf{Y}}(s*) = \hat{\mathbf{Y}}^*(s)$ for all complex s, where * is the complex conjugate operator.

2. $\hat{\mathbf{Y}}(s)$ is a positive matrix, that is $\mathbf{z}^{*T}(\hat{\mathbf{Y}}(s) + \hat{\mathbf{Y}}^{*T}(s))\mathbf{z} \geq 0$ for all complex values of s satisfying Re(s) > 0 and for any complex vector $\mathbf{z}$.

The second condition also implies the analyticity of $\hat{\mathbf{Y}}(s)$ for Re(s) > 0, since $\hat{\mathbf{Y}}(s)$ is a rational function of s (details in [16]). Therefore, the test of analyticity is unnecessary.

Due to the fact that the reduced matrices, $\tilde{\mathbf{G}}$, $\tilde{\mathbf{C}}$, $\tilde{\mathbf{B}}$ and $\tilde{\mathbf{L}}$ are all real since the transformation matrix, $\mathbf{X}$, is real, condition 1 is automatically satisfied. To show that condition 2 is

satisfied, we first set $\mathbf{Y}_h(s) = \hat{\mathbf{Y}}(s) + \hat{\mathbf{Y}}^T(s*)$ and use the property $\tilde{\mathbf{B}} = \tilde{\mathbf{L}}$ (since $\mathbf{B} = \mathbf{L}$ in our formulation when there are no sources inside the N-port, $\mathbf{X}^T\mathbf{B} = \mathbf{X}^T\mathbf{L}$) and some algebra to obtain,

$$
\begin{aligned}
\mathbf{z}^{*T}\mathbf{Y}_h(s)\mathbf{z} &= \mathbf{z}^{*T}(\tilde{\mathbf{B}}^T(\tilde{\mathbf{G}} + s\tilde{\mathbf{C}})^{-1}\tilde{\mathbf{B}} + \tilde{\mathbf{B}}^T(\tilde{\mathbf{G}} + s^*\tilde{\mathbf{C}})^{-T}\tilde{\mathbf{B}})\mathbf{z} \\
&= \mathbf{z}^{*T}\tilde{\mathbf{B}}^T(\tilde{\mathbf{G}} + s\tilde{\mathbf{C}})^{-1}[(\tilde{\mathbf{G}} + s\tilde{\mathbf{C}}) + (\tilde{\mathbf{G}} + s^*\tilde{\mathbf{C}})^T](\tilde{\mathbf{G}} + s^*\tilde{\mathbf{C}})^{-T}\tilde{\mathbf{B}}\mathbf{z}
\end{aligned} \tag{15}
$$

Setting $\mathbf{w} = (\tilde{\mathbf{G}} + s^*\tilde{\mathbf{C}})^{-T}\tilde{\mathbf{B}}\mathbf{z}$ and $s = j\omega + \sigma$ yields,

$$
\begin{aligned}
\mathbf{z}^{*T}\mathbf{Y}_h(s)\mathbf{z} &= \mathbf{w}^{*T}[(\tilde{\mathbf{G}} + (\sigma + j\omega)\tilde{\mathbf{C}}) + (\tilde{\mathbf{G}} + (\sigma - j\omega)\tilde{\mathbf{C}})^T]\mathbf{w} \\
&= \mathbf{w}^{*T}[\tilde{\mathbf{G}} + \tilde{\mathbf{G}}^T + \sigma(\tilde{\mathbf{C}} + \tilde{\mathbf{C}}^T)]\mathbf{w} \\
&= \mathbf{w}^{*T}\mathbf{X}^T[\mathbf{G} + \mathbf{G}^T + \sigma(\mathbf{C} + \mathbf{C}^T)]\mathbf{X}\mathbf{w}
\end{aligned} \tag{16}
$$

Similarly, let $\mathbf{y} = \mathbf{X}\mathbf{w}$ to get

$$\mathbf{z}^{*T}\mathbf{Y}_h(s)\mathbf{z} = \mathbf{y}^{*T}[\mathbf{G} + \mathbf{G}^T + \sigma(\mathbf{C} + \mathbf{C}^T)]\mathbf{y} \tag{17}$$

Since $\mathbf{C}$ is symmetric, $\mathbf{C}^T + \mathbf{C} = 2\mathbf{C}$. $\mathbf{C}$ is known to be non-negative definite (since we negate the rows corresponding to current variables as in (2)), so

$$\mathbf{y}^{*T}\sigma(\mathbf{C}^T + \mathbf{C})\mathbf{y} = 2\sigma\mathbf{y}^{*T}\mathbf{C}\mathbf{y} \geq 0 \tag{18}$$

for any complex vector $\mathbf{y}$ and $\sigma = \text{Re}(s) > 0$. $\mathbf{N}$ (the resistor stamps) is a symmetric nonnegative definite matrix, therefore

$$
\begin{aligned}
\mathbf{y}^{*T}(\mathbf{G}^T + \mathbf{G})\mathbf{y} &= \mathbf{y}^{*T}\left(\begin{bmatrix} \mathbf{N} & \mathbf{E} \\ -\mathbf{E}^T & \mathbf{0} \end{bmatrix}^T + \begin{bmatrix} \mathbf{N} & \mathbf{E} \\ -\mathbf{E}^T & \mathbf{0} \end{bmatrix}\right)\mathbf{y} \\
&= \mathbf{y}^{*T}\begin{bmatrix} 2\mathbf{N} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\mathbf{y} \geq 0
\end{aligned} \tag{19}
$$

is also nonnegative definite for any complex vector $\mathbf{y}$. From (17), (18) and (19), it follows that the second passivity condition is satisfied.

3.2. Preservation of Moments

The transformation in (13) preserves $\left\lfloor \dfrac{q}{N} \right\rfloor$ moments of the original system, which is the same as the classical Block Arnoldi reduction and half of that in MPVL. The proof is as follows. The exact (block) moments, $\mathbf{M}_i$, of the circuit are given as:

$$\mathbf{M}_i = \mathbf{L}^T\mathbf{A}^i\mathbf{R} \tag{20}$$

where $\mathbf{A} \equiv -\mathbf{G}^{-1}\mathbf{C}$, $\mathbf{R} \equiv \mathbf{G}^{-1}\mathbf{B}$ and $\mathbf{G}, \mathbf{C}, \mathbf{B}, \mathbf{L}$ are the system matrices as defined in (1).

Likewise, the moments of the PRIMA reduced order system are given by

$$\hat{\mathbf{M}}_i = \tilde{\mathbf{L}}^T\tilde{\mathbf{A}}^i\tilde{\mathbf{R}} \tag{21}$$

where $\tilde{\mathbf{A}} \equiv -\tilde{\mathbf{G}}^{-1}\tilde{\mathbf{C}}$, $\tilde{\mathbf{R}} \equiv \tilde{\mathbf{G}}^{-1}\tilde{\mathbf{B}}$ and $\tilde{\mathbf{G}}, \tilde{\mathbf{C}}, \tilde{\mathbf{B}}, \tilde{\mathbf{L}}$ are as defined in (13). Substitution of (13) in (21) yields:

$$\hat{\mathbf{M}}_i = \mathbf{L}^T \mathbf{X}[-(\mathbf{X}^T \mathbf{G} \mathbf{X})^{-1}(\mathbf{X}^T \mathbf{C} \mathbf{X})]^i (\mathbf{X}^T \mathbf{G} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{B} \qquad (22)$$

It is shown in [6] that the Arnoldi algorithm yields

$$\mathbf{A}^i \mathbf{R} = \mathbf{X} \, \mathbf{H}_q^i \, \mathbf{X}^T \, \mathbf{R}, \qquad 0 \le i < \left\lfloor \frac{q}{N} \right\rfloor. \qquad (23)$$

Rearranging the terms and using the definitions from (13):

$$\mathbf{A} \, \mathbf{A}^{i-1} \mathbf{R} = \mathbf{X} \, \mathbf{H}_q^i \, \mathbf{X}^T \, \mathbf{R}$$
$$-\mathbf{G}^{-1}\mathbf{C} \, \mathbf{A}^{i-1}\mathbf{R} = \mathbf{X} \, \mathbf{H}_q^i \, \mathbf{X}^T \, \mathbf{R}$$
$$-\mathbf{C} \, \mathbf{A}^{i-1}\mathbf{R} = \mathbf{G} \, \mathbf{X} \, \mathbf{H}_q^i \, \mathbf{X}^T \, \mathbf{R} \qquad (24)$$
$$-\mathbf{X}^T \mathbf{C} \, \mathbf{A}^{i-1}\mathbf{R} = \mathbf{X}^T \mathbf{G} \, \mathbf{X} \, \mathbf{H}_q^i \, \mathbf{X}^T \, \mathbf{R}$$

$$-\mathbf{X} \, (\mathbf{X}^T \mathbf{G} \, \mathbf{X})^{-1} \mathbf{X}^T \mathbf{C} \, \mathbf{A}^{i-1}\mathbf{R} = \mathbf{X} \, \mathbf{H}_q^i \, \mathbf{X}^T \, \mathbf{R} \qquad (25)$$

Inserting (23) in (25) results in:

$$\mathbf{K} \, \mathbf{A}^{i-1}\mathbf{R} = \mathbf{A}^i \mathbf{R}, \qquad 0 \le i < \left\lfloor \frac{q}{N} \right\rfloor \qquad (26)$$

where

$$\mathbf{K} = -\mathbf{X} \, (\mathbf{X}^T \mathbf{G} \, \mathbf{X})^{-1} \mathbf{X}^T \mathbf{C} . \qquad (27)$$

From (26), it can be shown by recursion that

$$\mathbf{K}^i \mathbf{R} = \mathbf{A}^i \mathbf{R}, \qquad 0 \le i < \left\lfloor \frac{q}{N} \right\rfloor \qquad (28)$$

Therefore, using (27) it follows that

$$\mathbf{X}[-(\mathbf{X}^T \mathbf{G} \, \mathbf{X})^{-1}(\mathbf{X}^T \mathbf{C} \, \mathbf{X})]^i = \mathbf{K}^i \, \mathbf{X} \qquad (29)$$

Replacing $\mathbf{X}[-(\mathbf{X}^T \mathbf{G} \, \mathbf{X})^{-1}(\mathbf{X}^T \mathbf{C} \, \mathbf{X})]^i$ in (22) with $\mathbf{K}^i \, \mathbf{X}$ yields

$$\hat{\mathbf{M}}_i = \mathbf{L}^T \mathbf{K}^i \, \mathbf{X}(\mathbf{X}^T \mathbf{G} \, \mathbf{X})^{-1} \mathbf{X}^T \mathbf{B} \qquad (30)$$

Evaluating (26) when i=0 gives

$$\mathbf{X} \, (\mathbf{X}^T \mathbf{G} \, \mathbf{X})^{-1} \mathbf{X}^T \mathbf{B} = \mathbf{R} \qquad (31)$$

Then from (30) and (31),

$$\hat{\mathbf{M}}_i = \mathbf{L}^T \mathbf{K}^i \, \mathbf{R}, \qquad 0 \le i < \left\lfloor \frac{q}{N} \right\rfloor, \qquad (32)$$

Finally, combining (32) and (28) with (20), it follows that

$$\hat{\mathbf{M}}_i = \mathbf{M}_i, \qquad 0 \le i < \left\lfloor \frac{q}{N} \right\rfloor \qquad (33)$$

Note that the number of poles in each entry of $\mathbf{Y}(s)$ is q, and we have matched the first $\left\lfloor \frac{q}{N} \right\rfloor$ moments at all N ports, yielding a total of q moments. The number of moments matched in PRIMA is, therefore, the same as that for the Block Arnoldi algorithm and half as many as matched by MPVL.

4. Integration of PRIMA within RICE

For all of the model order reduction schemes, the LU decomposition of the MNA conductance matrix ($\mathbf{G}$ in (2)) dominates the run time. In [13], RICE (Rapid Interconnect Circuit Evaluation) was described as a general path tracing algorithm to obtain moments with optimal efficiency for interconnect trees and mesh structures. Using RICE to calculate moments, the explicit construction and inversion of $\mathbf{G}$ is avoided, and the moments are more accurate than those obtained via matrix factorization.

The moments of the circuit can be obtained recursively from:

$$\mathbf{M}_0 = \mathbf{G}^{-1}\mathbf{B}$$
$$\mathbf{M}_k = \mathbf{G}^{-1}\mathbf{C}\mathbf{M}_{k-1} \qquad k > 0 \qquad (34)$$

where the matrices $\mathbf{G}$, $\mathbf{C}$ and $\mathbf{B}$ are as defined in (2). As shown in [1], this can be viewed as recursive dc circuit solutions, when capacitors and inductors are replaced by current and voltage sources respectively, with the values derived from the columns of $\mathbf{C}\mathbf{M}_{k-1}$. The Krylov vectors, which can be viewed as well conditioned moments, can be obtained from a very similar recursive scheme:

1. Obtain zero[th] moment and orthonormalize it:

$$\text{Solve } \mathbf{M}_0 \text{ from } \mathbf{G} \, \mathbf{M}_0 = \mathbf{B} \qquad (35)$$
$$\mathbf{X}_0 = orth(\mathbf{M}_0) \qquad (36)$$

2. Recursively obtain higher order Krylov vectors:

$$\text{Solve } \mathbf{M}_k \text{ from } \mathbf{G} \, \mathbf{M}_k = \mathbf{C}\mathbf{X}_{k-1} \qquad (37)$$
$$\mathbf{X}_k^\zeta = \mathbf{M}_k - \mathbf{X}_{k-1}(\mathbf{X}_{k-1}^T\mathbf{M}_k) - ... - \mathbf{X}_0(\mathbf{X}_0^T\mathbf{M}_k) \qquad (38)$$
$$\mathbf{X}_k = orth(\mathbf{X}_k^\zeta) \qquad (39)$$

The "*orth*" operator can be implemented as a simple Gram-Schmidt orthonormalization procedure. The space spanned by the block Krylov terms $(\mathbf{X}_k, \mathbf{X}_{k-1}, ..., \mathbf{X}_0)$ is called the Krylov space. Therefore, the Krylov vectors can be obtained via a path tracing procedure using RICE-like routines to solve for equations (35) and (37).

The Krylov space constitutes the congruence transformation matrix, $\mathbf{X}$ in PRIMA. The reduced MNA matrices $\tilde{\mathbf{G}}$ and $\tilde{\mathbf{C}}$ are

$$\tilde{\mathbf{C}} = \mathbf{X}^T \mathbf{C} \, \mathbf{X} \qquad \tilde{\mathbf{G}} = \mathbf{X}^T \mathbf{G} \, \mathbf{X}. \qquad (40)$$

Note, however, that the matrices $\mathbf{C} \, \mathbf{X}$ and $\mathbf{G} \, \mathbf{X}$ are obtained using RICE without explicitly constructing $\mathbf{G}$ and $\mathbf{C}$. The columns of $\mathbf{C} \, \mathbf{X}$ are the values of current and voltage sources that are used to replace capacitors and inductors at each moment computation stage. This information is easily obtained during a path trace [13]. The kth block of $\mathbf{G} \, \mathbf{X}$ (i.e. $\mathbf{G}\mathbf{X}_k$) is a function of previous blocks of $\mathbf{G} \, \mathbf{X}$ and $\mathbf{C}\mathbf{X}_{k-1}$ since from (38),

$$\mathbf{G}\mathbf{X}_k^\zeta = \mathbf{G}\mathbf{M}_k - \mathbf{G}\mathbf{X}_{k-1}(\mathbf{X}_{k-1}^T\mathbf{M}_k) - ... - \mathbf{G}\mathbf{X}_0(\mathbf{X}_0^T\mathbf{M}_k) \qquad (41)$$

and using $\mathbf{G} \, \mathbf{M}_k = \mathbf{C}\mathbf{X}_{k-1}$,

$$\mathbf{GX}_k^\zeta = \mathbf{CX}_{k-1} - \mathbf{GX}_{k-1}(\mathbf{X}_{k-1}^T \mathbf{M}_k) - \ldots - \mathbf{GX}_0(\mathbf{X}_0^T \mathbf{M}_k) \ . (42)$$

5. Time-domain simulation of the macromodels

For a complete circuit simulation, the nonlinear elements should be simulated along with the reduced order macro-models. There are two ways to include the PRIMA macro-models into circuit simulators such as SPICE [17]. One approach is in terms of the frequency domain y-parameters. Combining the nonlinear time domain analysis in SPICE with the frequency dependent y-parameters requires convolution of $O(T^2)$ complexity, where T is the number of simulation time-points. For this reason, recursive convolution [14] and time-domain y-parameter macromodels [7] were developed, where the complexity is linear with the number of time-points. The second method is the direct stamping (i.e. circuit model realization). Since the reduction method we use is block, the reduced matrices can be directly stamped into the SPICE MNA matrices. Noticing that the reduced order *q*-variable system has the equation shown in (12) and (13), and recognizing that it is possible to introduce $\tilde{\mathbf{x}}_q$ as a circuit variable into the MNA matrix, the direct stamps for the macromodel can be generated as below:

$$\begin{bmatrix} \text{Stamps for} & \mathbf{0} & \mathbf{0} & \\ f(x_{NL},u_p) & \mathbf{I}_N & \mathbf{0} & \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_N & -\tilde{\mathbf{L}}^T \\ \mathbf{0} & -\tilde{\mathbf{B}} & \mathbf{0} & \left(\tilde{\mathbf{G}} + \tilde{\mathbf{C}}\dfrac{d}{dt}\right) \end{bmatrix} \begin{bmatrix} \mathbf{x}_{NL} \\ \mathbf{u}_p \\ \mathbf{i}_p \\ \tilde{\mathbf{x}}_q \end{bmatrix} = \begin{bmatrix} \mathbf{v}_{NL} \\ \mathbf{v}_p \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \quad (43)$$

In (43), $\mathbf{x}_{NL}$ denotes the other variables of the circuit (other node voltages and currents) $\mathbf{u}_p$ and $\mathbf{i}_p$ are port voltages and currents respectively, and $\tilde{\mathbf{x}}_q$ denotes the extra variables that are introduced from the inclusion of realized macromodel into the circuit. Since $\tilde{\mathbf{C}}$ is a symmetric and real matrix, it can be diagonalized using singular value decomposition. In this case, all the capacitance values will be real and positive, since they will be the singular values of $\tilde{\mathbf{C}}$. Note that it is also possible to come up with a realization scheme similar to (43) starting from $\hat{\mathbf{Y}}(s)$.

6. Results

In this section, PRIMA is demonstrated and compared with other approaches. All reductions are done using RICE v5.0, a program which integrates the PRIMA algorithm with the RICE moment calculation routines. For the frequency domain examples, the y-parameters are compared with the reduced order models from different reduction methods. Time domain results via recursive convolution are obtained using a modified version of SPICE3f4 [18]. For all the exam-
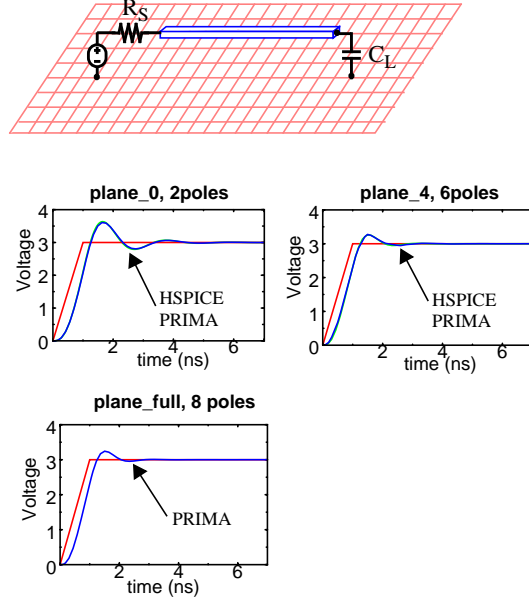
Circuit complexity

Circuit name	# of R	# of L	# of K
plane_0	859	858	0
plane_4	859	858	17,892
plane_full	859	858	183,693

Run time comparisons

Circuit name	HSPICE	PRIMA	RecConv
plane_0	39.97 secs	0.25 secs	0.03 secs
plane_4	17,343 secs	1.23 secs	0.05 secs
plane_full	can not run	11.73 secs	0.05 secs

FIGURE 4: Mesh ground plane example

ples, the poles obtained via PRIMA were observed to be *stable*.

6.1. Mesh ground plane

With the ability to calculate a large number of poles accurately, PRIMA can be applied to analysis problems which include complex, high frequency responses. One such application is the R-L mesh plane encountered in MCM and packaging problems. Since such a problem is strongly coupled, the L-matrix is dense and thereby destroys the matrix sparsity in a classical SPICE simulation. In this example, the ground plane is modeled by a 20x20 mesh, and each square is modeled as a resistor and an inductor. The coupling can be adjusted to make the induc-tance matrix sparse as described in [19]. There is an RL line over the ground plane that is terminated with a capacitor load, as shown in Fig.4. Also shown in the figure are time domain results from PRIMA and HSPICE for various levels

of L-matrix sparsity. The full matrix response is also shown for PRIMA, but the HSPICE simulation would not complete its run due to memory and run-time limits.

This circuit is a worst-case interconnect topology for a path tracing algorithm [13] (all loops), however, RICE v5, our path tracing implementation of PRIMA, showed excellent speed-up over HSPICE, a commercial circuit simulation tool. The table in this figure also includes the time required for recursive convolution of the reduced-order model in SPICE3f, denoted by RecConv.

6.2. Nonlinear driver driving a transmission line

Fig.5 shows a lossy transmission line represented by 40 lumped RLC sections and reduced to 5 poles using both PVL and PRIMA. Although all of the poles from PVL were stable (i.e. negative real parts), the overall PVL response was clearly unstable as shown in Fig.5. The 5th order approximation from PRIMA is indistinguishable from the exact response, which was obtained by an HSPICE simulation for this example.

6.3. Coupled noise for a two-bit bus

Next consider the two-bit bus driven by CMOS inverters in Fig.6. One of the drivers is switching while the other is quiet. The interconnect, consisting of 40 coupled RLC sections, is modeled as a 4-port and reduced by PRIMA. Transient analysis is done using recursive convolution. The time domain waveforms at the load end are compared for various order of approximations. Since this is a 4-port, an 8 pole approximation corresponds to matching only m_0 and m_1 generated by four different sources. The plot shows that in the time domain, even the coupled noise can be accurately simulated using the 8 poles from PRIMA. Although the interconnect inductance was exaggerated in this example to make the approximation more difficult, it is observed that an 8th order approximation is sufficient to capture the coupled noise from the active driver to the quiet load end.
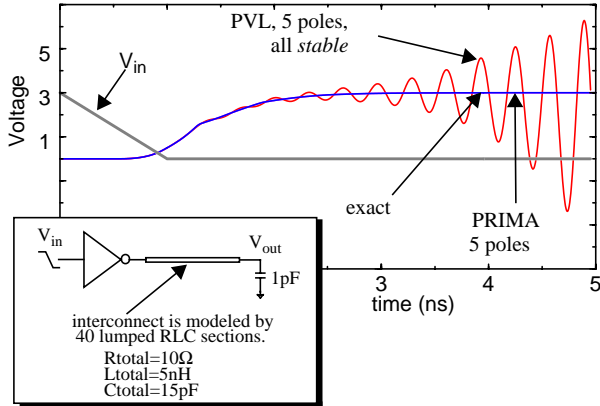

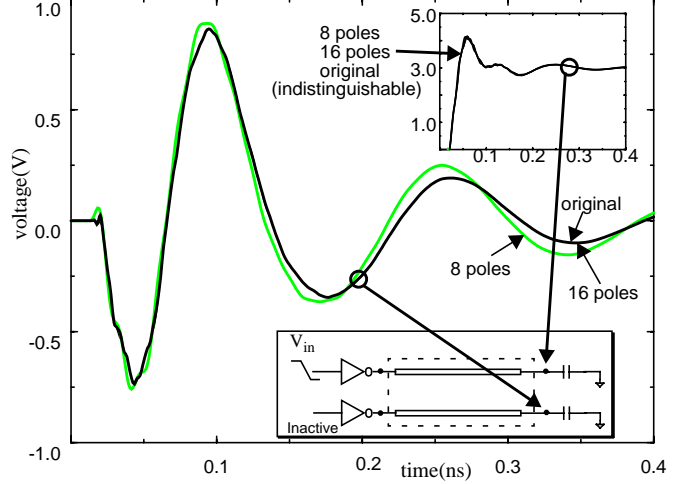FIGURE 5: Instability in the time domain for non-passive model

FIGURE 6: Waveform comparisons for a four port.

To compare the difference between direct realization and y-parameter based simulation (i.e. recursive convolution here), the reduced order circuit (via PRIMA) is simulated using both techniques. In the direct realization, $\tilde{\mathbf{C}}$ is diagonalized to increase the speed. The run times are given in Table 1. Although the circuit is relatively small (i.e. $\mathbf{G}$ is only *300x300*), the gain in using a PRIMA reduced macromodel and y-parameter based simulation is about 50x over direct realization. For larger circuits such as the mesh plane example, this gain is expected to be much larger. Direct realization is inferior when the order of approximation gets bigger, mainly because the dense $\tilde{\mathbf{G}}$ matrix gets larger.

TABLE 1: Run time comparisons

Exact	Reduced	
Full simulation	**Simulated after Direct Realization**	**Simulated by Y-parameter based**
17.78 sec	0.6 with 8 poles	0.18 s. with 8 poles
	3.98 s. with 16 poles	0.28 s. with 16 poles
	10.29 s. with 24 poles	0.32 s. with 24 poles

6.4. Six coupled transmission lines

The second example is a 12-port containing six coupled transmission lines modeled by 40 coupled RLC sections. The input admittance ($Y_{11}(s)$), reduced by Block Arnoldi, MPVL and PRIMA are compared with the exact input admittance in Fig.7 using 48th order approximations in all cases. Block Arnoldi captures the exact response up to 16 GHz, while MPVL and PRIMA match up to 28 GHz. When the order of approximation is increased to 72 poles, it is observed that the frequency spectrum is captured up to 60 GHz by MPVL and PRIMA.
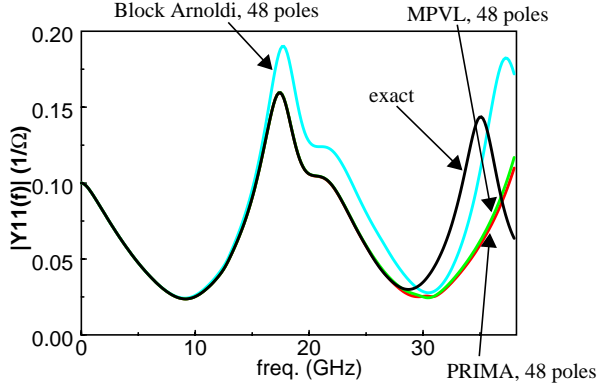
FIGURE 7: $Y_{11}(s)$ in frequency domain for six coupled TR. lines.

6.5. Large Coupled RLC Circuit

The third example in Fig.8 displays the responses for a 3-port composed of densely coupled RLC circuits. Approximations are done using 25 poles for the three methods. As can be observed from the figure, both PRIMA and MPVL capture the entire frequency spectrum.
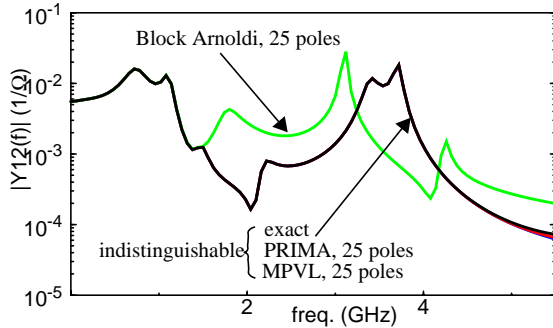

FIGURE 8: 3-Port consisting of a large lumped RLC circuit.

7. Conclusions

This paper presented a novel algorithm for producing provably passive macromodels for arbitrary RLC circuits. The method uses a Block Arnoldi algorithm to generate the vectors needed for applying a transformation to the macromodel MNA matrices. Empirical results show that PRIMA produces comparable or superior results in terms of accuracy with respect to all other known reduction techniques, but superior in that it guarantees the passivity that is critical for time domain analyses. The implementation of PRIMA with path tracing algorithms from RICE enables extremely accurate high frequency response approximations of enormous, complex, RLC circuits with excellent efficiency.

The PRIMA algorithm presented in this paper can be easily extended to implement a number of heuristics such as moment shifting [13] and frequency shifting [20]. However, these heuristics are unnecessary and merely increase the complexity.

BIBLIOGRAPHY

[1] L. T. Pillage and R. A. Rohrer, "Asymptotic waveform evaluation for timing analysis", *IEEE Trans. Computer-Aided Design*, vol. 9, no. 4, pp. 352-366, Apr. 1990.

[2] P. Feldmann and R. W. Freund, "Efficient linear circuit analysis by Padé approximation via the Lanczos process", *IEEE Trans. on CAD*, vol. CAD-14, pp. 639-649, May 1995

[3] K. J. Kerns, I. L. Wemple, and A. T. Yang, "Stable and efficient reduction of substrate model networks using congruence transforms" *IEEE/ACM Proc. ICCAD*, pp. 207-214, Nov. 1995

[4] L. M. Silveira, M. Kamon and J. White, "Efficient reduced-order modeling of frequency-dependent coupling inductances associated with 3-D interconnect structures", *IEEE/ACM Proc. DAC*, pp. 376-380, Jun. 1995

[5] P. Feldmann and R. W. Freund, "Reduced-order modeling of large linear subcircuits via a block Lanczos algorithm", *IEEE/ACM Proc. DAC*, pp. 474-479, Jun.. 1995

[6] D. L. Boley, "Krylov space methods on state-space control models", *Circuits Syst. Signal Process*. vol.13, no.6, pp. 733-758, 1994

[7] S. Y. Kim, N. Gopal and L. T. Pillage, "Time-Domain Macromodels for VLSI Interconnect Analysis", *IEEE Trans. on CAD*, vol 13, No. 10, pp. 1257-1270, Oct. 1994

[8] L. M. Silveira, M. Kamon, I. Elfadel and J. White, "A coordinate-transformed Arnoldi algorithm for generating guaranteed stable reduced-order models of arbitrary RLC circuits", *IEEE/ACM Proc. ICCAD*, pp. 288-294, Nov. 1996

[9] I. M. Elfadel and David D. Ling, "Zeros and passivity of Arnoldi-reduced-order models for interconnect networks", *IEEE/ACM Proc. DAC*, pp. 28-33, Jun. 1997

[10] K. J. Kerns and A. T. Yang, "Preservation of Passivity During RLC Network Reduction via Split Congruence Transformations", *IEEE/ ACM Proc. DAC*, pp. 34-39, Jun. 1997

[11] B. D. Anderson and S. Vongpanitlerd, *Network analysis and synthesis*. Prentice-Hall, Inc., 1973

[12] R. A. Rohrer and H. Nosrati, "Passivity considerations in stability studies of numerical integration algorithms", *IEEE Trans. on Circuits and Systems*, vol. CAS-28, no. 9, pp. 857-866, Sep. 1981

[13] C. L. Ratzlaff and L. T. Pillage, "RICE: Rapid interconnect circuit evaluation using AWE", *IEEE Trans. CAD*, vol. 13, no. 6, pp. 763-776, Jun. 1994

[14] V. Raghavan, J. E. Bracken and R. A. Rohrer, "AWESpice: A general tool for the accurate and efficient simulation of interconnect problems", *IEEE/ACM Proc. DAC*, Jun. 1992

[15] G. H. Golub and C. F. Van Loan, *Matrix Computations*. John Hopkins Univ. Press, 2nd ed., 1989

[16] E. S. Kuh and R. A. Rohrer, *Theory of Linear Active Networks*. Holden-Day Inc., 1967

[17] L. W. Nagel, "SPICE2, a computer program to simulate semiconductor circuits", Technical Report ERL-M520, UC-Berkeley, May 1975

[18] T. L. Quarles, "The SPICE3 implementation guide", Tech. Rep. Memo ERL-M89/44, U. of California, Berkeley, 1989

[19] B. Krauter and L. T. Pileggi, "Generating sparse partial inductance matrices with guaranteed stability", *IEEE/ACM Proc. ICCAD*, pp. 45-52, Nov. 1995

[20] X. Huang, "Padé approximation of linear(ized) circuit responses", Ph.D. Thesis, Carnegie Mellon Univ., Nov. 1990