Sparse and efficient reduced order modeling of linear subcircuits with large number of terminals

Peter Feldmann
IBM T.J. Watson Research Center
Yorktown Heights, NY
feldmann@watson.ibm.com

Frank Liu
IBM Austin Research Lab
Austin, TX
frankliu@us.ibm.com

Abstract—**In the process of designing state-of-the art VLSI circuit we often encounter large but highly structured linear subcircuits with large number of terminals. Classical examples are power supply networks, clock distribution networks, large data buses, etc. Various applications would benefit from efficient high level models of such networks. Unfortunately the existing model-order-reduction algorithms are not adapted to handle more than a few tens of terminals. This talk introduces RecMOR, an algorithm for the computation of reduced order models of structured linear circuits with numerous I/O ports. The algorithm exploits certain regularities of the subcircuit response that are typical in numerous applications of interest. When these regularities are present, the normally dense matrix-transfer function of the subcircuit contains sub-blocks that in some sense are significantly low rank and can be compactly modeled by the recently introduced SVDMOR algorithm. The new RecMOR algorithm decomposes the large matrix-transfer function recursively, and applies SVDMOR compression adaptively to the sub-blocks of the transfer function. The result is a reduced order model that is sparse, efficient, and directly usable as an efficient substitute of the subcircuit in circuit simulations. The method is illustrated on several circuit examples.**

I. Introduction

Model order reduction (MOR) has become an established technique for the analysis and compact modeling of linear circuits and systems. In the past decade numerous algorithms have been devised for the computation of reduced-order models [1]–[5]. Model order reduction is useful when only signal behavior at the ports of the linear block is of interest. MOR techniques generate compact models of the circuit that approximate well circuit behavior at the port terminals but sacrifice the modeling of behavior at internal nodes.

Unfortunately the efficiency of model order reduction degrades as the number of external terminals to the circuit increases. The reason for the degradation is fundamental and does not depend on the particular reduction algorithm. A multi-terminal circuit is described by an $m \times m$ matrix-valued transfer function, where m is the number of external terminals. Each entry in the transfer function matrix characterizes the interaction between a pair of two terminals and there are $\mathcal{O}(m^2)$ of such interactions. Moreover, in general, there is no basis to the assumption that any of the interactions is magnitude-wise insignificant, therefore the matrix-valued transfer function must be assumed to be fully populated. Any reduced-order model must approximate in some sense this matrix-valued transfer function. Therefore unless some special properties of the circuit are exploited, the complexity of the reduced order model is at least $\mathcal{O}(m^2)$, which for a circuit with numerous inputs and outputs may approach or even surpass the complexity of working with the original, unreduced, circuit equations.

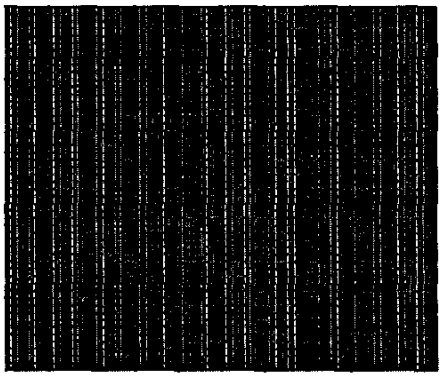

Fig. 1. Power grid mesh section

Recently we introduced a new algorithm, SVDMOR, that can take advantage of situations when the matrix transfer function is numerically close to being low rank and generate for it a compact and sparse reduced order model. While SVDMOR achieves excellent sparsification on this type of transfer functions, a complete matrix transfer function of a subcircuit (relating all port currents and voltages), is very rarely low rank. On the other hand it may contain important low rank sub-blocks. Consider, for example, the power grid section illustrated in Figure 1. Here the inputs and the outputs are the edges of the wire on the right and on the left of the section. The complete matrix transfer function that relates the currents and voltages at all the ports of the section as required to represent its behavior in a higher level simulation, is very large and is also full rank. Therefore SVDMOR can achieve no practical compression on the entire matrix. On the other hand, the submatrix-transfer function that relates the voltages on the left side to the currents on the right is very accurately approximated by a low rank transfer function and SVDMOR achieves a significant sparsification of this off-diagonal sub-

0-7803-8702-3/04/$20.00 ©2004 IEEE.

block.

This paper introduces RecMOR, an algorithm for the computation of reduced order models of large linear subcircuits with a large number of input/output ports. The algorithm depends of and exploits certain regularities of the subcircuit response that are typical in numerous applications of interest such as power meshes, buses, and clock networks. When these regularities are present, the normally large and dense matrix transfer function of the subcircuit has significant low rank subblocks. The RecMOR algorithm recursively and adaptively applies the SVDMOR [6] sparsification to these portions of the transfer function. The result is a reduced order model that is sparse, efficient, and directly usable as an efficient substitute of the subcircuit in higher level simulations.

II. Overview of the SVDMOR algorithm

First, we briefly summarize the essence of MOR methods. We are interested to compute the reduced-order model for a linear circuit characterized by a *large* number of input/output terminals. The general state-space formulation of the circuit is

$$\begin{aligned} C\frac{d}{dt}x + Gx &= Mu \\ y &= N^T x \end{aligned} \tag{1}$$

Here C, and G, are $n \times n$ matrices describing the reactive, and dissipative parts of the circuit respectively. M is a $n \times p$ matrix that defines the input ports, and N is a $n \times q$ matrix that defines the outputs. For most circuits these matrices are quite sparse.

A large class of MOR methods operate on the Laplace-domain transfer function of the multi-port circuit. The Laplace transform of the input output transfer function has the expression

$$H(s) = N^T(G + sC)^{-1}M, \tag{2}$$

and is in fact a $q \times p$ matrix-valued rational function. Padé-based MOR algorithms [1]–[5], operate on the original circuit matrices G, C, M, N, and compute models described by smaller matrices, which usually are just projections of the circuit matrices in well chosen subspaces. The transfer function of the reduced-order models approaches the original in the Padé approximation sense

$$H(s) \approx H_l(s) = N_l^T(G_l + sC_l)^{-1}M_l. \tag{3}$$

In this reduced-order model (3) G_l and C_l are $l \times l$ matrices. where l depends on the number of I/O ports and the order of approximation. Typically, l is much smaller than n, the size of the original system matrices and, therefore the reduced-order model is expressed in terms of significantly smaller matrices.

However, reduced-order model matrices may also be much denser. The number of non-zero entries in the reduced-order model matrices is increasing rapidly with the number of I/O ports and is of order $\mathcal{O}(pq)$, while for typical circuits, the original system matrices G and C are very sparse, having only an order $\mathcal{O}(n)$ of non-zero entries. This situation causes the benefits of model-order reduction (compactness and computational efficiency) to vanish rapidly as the number of I/O ports is increased.

The SVDMOR algorithm [6] exploits the structure of a wide class of transfer functions and can often result in compact reduced-order models even for circuits with large numbers of I/O ports. Matrix N and M encode all the input/output port definitions. Obviously in many applications, the responses at the circuit inputs and outputs are not independent. On the contrary, typically there is a large degree of correlation between circuit responses at various ports. Such a correlation will manifest itself in the matrix H having highly dependent entries, or in other words H can be well approximated by a lower rank matrix.

Note that in our formulation, M and N only contain topology information. SVDMOR first transforms the equations in a way that reveals circuit response correlations. One possible transformation focuses on the DC response matrix of the circuit H_{DC}. For a highly regular circuit we expect this response to be highly correlated and therefore to be well approximate by a low rank matrix. We determine this approximation by performing SVD and keeping only the important singular values.

$$H_{\text{DC}} = N^T G^{-1} M = U\Sigma W^T \tag{4}$$

where $\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_m)$, and U and W are orthogonal matrices. Note that we chose the unusual notation W instead of the traditional V in order to avoid confusion with the voltage variables. In many important situations there will be a relatively small number of dominant singular values, say $\sigma_1 \ldots \sigma_r$, $r << m$, and the error caused by setting the remaining singular values to zero, will be relatively small. In these cases

$$B = U\Sigma W^T \approx U_r \Sigma_r W_r^T \tag{5}$$

Other good matrices for revealing correlations that are the first moment of the response $\mathcal{M}_1 = M^T G^{-1} C G^{-1} N$, frequency shifted moments $\mathcal{M}_{s_0} = M^T(G + s_0 C)^{-1}N$ or even combinations of these. We approximate

$$M = b_m U_r^T \quad \text{and} \quad N = b_n W_r^T \tag{6}$$

where b_m and b_n are obtained using the Moore-Penrose pseudoinverse.

$$b_m = M U_r (U_r^T U_r)^{-1} \quad \text{and} \quad b_n = N W_r (W_r^T W_r)^{-1} \tag{7}$$

The circuit transfer function now becomes

$$H(s) \approx U_r \underbrace{b_m^T (G + sC)^{-1} b_n}_{H_r(s)} W_r^T \tag{8}$$

The standard model order reduction technique [1]–[5] can now be applied just to

$$H_r(s) = b_m^T (G + sC)^{-1} b_n \tag{9}$$

which is just a $r \times r$ matrix transfer function, and obtain $\tilde{H}_r(s)$. The complete transfer function is approximated by

$$H(s) \approx U_r \tilde{H}_r(s) W_r^T \tag{10}$$

where all the matrices involved have $\mathcal{O}(r^2)$ non-zero entries.

Note also that SVDMOR compression works at the problem formulation level, therefore it is not tied to a particular choice of reduction method. As such, SVDMOR compression can be applied in conjunction with any of the quoted model reduction algorithms.

III. Hierarchical decomposition of complete matrix transfer function

By complete matrix-transfer function we mean a transfer function that completely characterizes the port behavior of a circuit block. It can be, for example, the full Z-parameter matrix relating all the port currents to all port voltages, or the S-parameter matrix, or any other representation.

We assume that the input ports are sorted, to the extent it is possible, according to some *electrical distance* metric. In other words, ports with indices that are close to each other are likely to be electrically close, and widely separated indices are likely to correspond to *electrically distant* ports. This indexing policy does not have to be absolutely enforced but can significantly impact the effectiveness of model reduction.

For example, we consider the rectangular grid network in Figure 1, modeling a portion of the power grid, that has ports on the right and on the left edges. The Laplace transform of the state equations of this system are

$$\begin{aligned} (G+sC)\,X &= BI \\ V &= B^T X \end{aligned} \tag{11}$$

where G and C are matrices that describe the resistive and the reactive elements comprising the mesh and B is the vector that defines the I/O ports to the subcircuit. The equations determine the following transfer function that establishes the relationship between port currents and port voltages

$$H(s) = B^T\,(G+sC)^{-1}\,B$$

The vector B which defines the ports is partitioned in two components: $B = \begin{bmatrix} B_l & B_r \end{bmatrix}$, with B_l and B_r selecting the ports on the left and right sides of the grid, respectively. The partitioning of the ports results in the partitioning of the transfer function of as follows

$$\begin{aligned} H(s) &= B^T\,(G+sC)^{-1}\,B \\ &= \begin{bmatrix} B_l^T\,(G+sC)^{-1}\,B_l & B_l^T\,(G+sC)^{-1}\,B_r \\ B_r^T\,(G+sC)^{-1}\,B_l & B_r^T\,(G+sC)^{-1}\,B_r \end{bmatrix} \end{aligned} \tag{12}$$

We assume that model order reduction is performed separately on the four components of the transfer function matrix resulting in the following *reduced* matrix

$$\tilde{H}(s) = \begin{bmatrix} b_l^T \tilde{Y}_{ll}^{-1} b_l & d_l^T \tilde{Y}_{lr}^{-1} d_r \\ d_r^T \tilde{Y}_{lr}^{-T} d_l & b_r^T \tilde{Y}_{rr}^{-1} b_r \end{bmatrix} \tag{13}$$

where for notational brevity we denote $\tilde{Y}_{pq} = \tilde{G}_{pq} + s\tilde{C}_{pq}$. Note also that the off-diagonal blocks are just transposes of each other.

This transfer function approximates the port current to port voltage relationship

$$\begin{bmatrix} V_l \\ V_r \end{bmatrix} = \begin{bmatrix} b_l^T \tilde{Y}_{ll}^{-1} b_l & d_l^T \tilde{Y}_{lr}^{-1} d_r \\ d_r^T \tilde{Y}_{lr}^{-T} d_l & b_r^T \tilde{Y}_{rr}^{-1} b_r \end{bmatrix} \begin{bmatrix} I_l \\ I_r \end{bmatrix} \tag{14}$$

where for notational convenience we replaced the $\approx$ relationship with $=$. We now rewrite this reduced system in state variable form using $\left[X_{ll}, X_{rl}, X_{lr}, X_{rr}\right]^T$ as the state of the reduced system

$$\begin{array}{lll} \tilde{Y}_{ll} X_{ll} = b_l I_l, & V_{ll} = b_l^T X_{ll} & \\ \tilde{Y}_{lr} X_{lr} = d_r I_r, & V_{lr} = d_l^T X_{lr}, & V_l = V_{ll} + V_{lr} \\ \tilde{Y}_{lr}^T X_{rl} = d_l I_l, & V_{rl} = d_r^T X_{rl} & \\ \tilde{Y}_{rr} X_{rr} = b_r I_r, & V_{rr} = b_r^T X_{rr}, & V_r = V_{rl} + V_{rr} \end{array}$$

We gather these equation in matrix form and arrange them in a way that results in a symmetric system

$$\begin{bmatrix} \tilde{Y}_{ll} & & & \\ & 0 & \tilde{Y}_{lr} & \\ & \tilde{Y}_{lr}^T & 0 & \\ & & & \tilde{Y}_{rr} \end{bmatrix} \begin{bmatrix} X_{ll} \\ X_{rl} \\ X_{lr} \\ X_{rr} \end{bmatrix} = \begin{bmatrix} b_l & & \\ & d_r & \\ d_l & & \\ & & b_r \end{bmatrix} \begin{bmatrix} I_l \\ I_r \end{bmatrix}$$

$$\begin{bmatrix} V_l \\ V_r \end{bmatrix} = \begin{bmatrix} b_l & & \\ & d_r & \\ d_l & & \\ & & b_r \end{bmatrix}^T \begin{bmatrix} X_{ll} \\ X_{rl} \\ X_{lr} \\ X_{rr} \end{bmatrix}$$

When any of these transfer function sub-matrices can be modeled by low-rank models as shown in the previous section, or its number of ports is sufficiently small, it is approximated by a sparse reduced-order model. Otherwise the sub-block is further subdivided and the process continues recursively until all the sub-block are modeled by sparse approximations.

IV. Combining hierarchical decomposition with SVDMOR compression

We continue the example and assume that the off-diagonal transfer function sub-matrix

$$H_{lr} = B_l^T\,(G+sC)^{-1}\,B_r \tag{15}$$

can be approximated by the low rank reduced order model. We also observe that, G and C, being symmetric matrices, $H_{rl} = H_{lr}^T$, so it is sufficient to work on only one of the blocks. Applying the SVDMOR method B_l and B_r are approximated

$$B_l \approx b_l U_l^T \;\text{ and }\; B_r \approx b_r W_r^T \tag{16}$$

Standard model reduction is than applied to the inner part of the expression resulting in a sparse multi-port model

$$\begin{aligned} H_{lr} &\approx U_l \underbrace{b_l^T\,(G+sC)^{-1}\,b_r}_{\text{MOR}} W_r^T \\ &\approx U_l\, u_l^T \tilde{Y}_{lr}^{\;-1} u_r\, W_r^T \end{aligned} \tag{17}$$

In other words the current to voltage transfer function $V_{lr} = H_{lr} I_{lr}$ can be approximated in terms of fewer state variables since

$$V_{lr} = U_l\, u_l^T \tilde{Y}_{lr}^{\;-1} u_r\, W_r^T I_{lr} \tag{18}$$

can be expressed as

$$\tilde{Y_{lr}} X_{lr} = u_r W_r^T I_{lr} \quad (19)$$
$$V_{lr} = U_l u_l^T X_{lr} \quad (20)$$

We replace the expression in the hierarchical decomposition obtained in the previous section

$$\begin{bmatrix} \tilde{Y_{ll}} & & & \\ & 0 & \tilde{Y_{lr}} & \\ & \tilde{Y_{lr}}^T & 0 & \\ & & & \tilde{Y_{rr}} \end{bmatrix} \begin{bmatrix} X_{ll} \\ X_{rl} \\ X_{lr} \\ X_{rr} \end{bmatrix} = \begin{bmatrix} b_l & & \\ & u_r W_r^T & \\ u_l U_l^T & & \\ & & b_r \end{bmatrix} \begin{bmatrix} I_l \\ I_r \end{bmatrix}$$

$$\begin{bmatrix} V_l \\ V_r \end{bmatrix} = \begin{bmatrix} b_l & & \\ & u_r W_r^T & \\ u_l U_l^T & & \\ & & b_r \end{bmatrix}^T \begin{bmatrix} X_{ll} \\ X_{rl} \\ X_{lr} \\ X_{rr} \end{bmatrix}$$

The resulting submatrices that correspond to off-diagonal terms are small and sparse when the correlation assumptions hold well. The diagonal terms Y_{ll} and Y_{rr} had their port count reduced by half. Furthermore, this process can be repeated recursively on the diagonal terms until the off-diagonal elements no longer yield significant reductions through correlation exploitation.

Note that this sub-matrix based sparsification procedure is similar to the method used in [7] for matrix-vector multiplication in a fast Poisson equation solver context.

V. Examples

As an example we analyze an RC rectangular mesh such as would result from the modeling of the on-chip power-grid. The grid is quite regular, therefore we expect the responses of the signals to be highly correlated. We assume that all the input/output ports are on the left and right side of the mesh. Assuming the mesh is of size 50×60 the transfer function that the reduced-order model needs to capture will be a 120×120 matrix-valued rational function. The original network has over 3000 nodes. The SVDMOR compression is guided by the correlations revealed by

$$\mathcal{M}_{\mathrm{DC}} = L^T G^{-1} R \quad (21)$$
$$\mathcal{M}_{s_0} = L^T (G + s_0 C)^{-1} R,$$

the DC component, and a shifted moment around the normalized frequency $s_0 = 0.1$.

Figure 2 shows 7 transfer functions out of the 120 matrix generated through the application of the RecMOR algorithm on the mesh. We observe that the accuracy of the RecMOR model below that the brute force application of a standard MOR procedure, however it is above the preimposed error limit of 10^{-3}. The brute force application of MOR would have produced as part of the "reduced" model a dense matrix of sizes several times 120, the number of I/O ports. Such a matrix is much more expensive to employ in a circuit simulation (if at all possible) than the original 3000×3000 circuit matrices. RecMOR produces instead a much sparser, block-structured matrix shown in Figure 3. It is larger than the matrix that would have been produced by the brute force model reduction but significantly sparser and therefore compatible with efficient use in circuit simulators. Moreover one can take advantage of its block diagonal structure to extract additional efficiency.

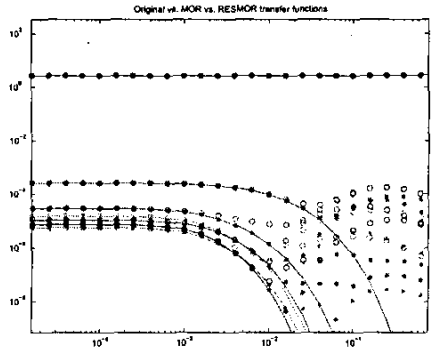

Fig. 2. 7 representative responses

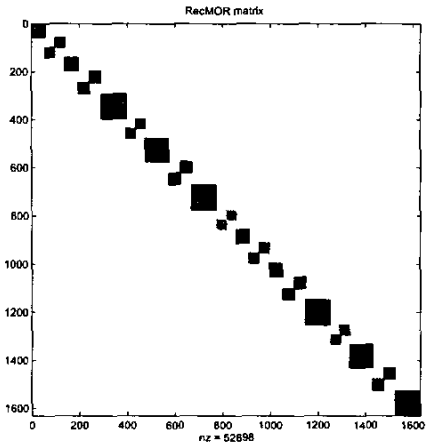

Fig. 3. Reduced system matrix structure

The next example is a realistic power grid section implemented on three layers of metal. The topology of the grid together with the current response to a unit excitation applied to one of the nodes in the top layer are shown in Figure 4. The grid consists of interlaced power and ground wires modeled with just under 3000 nodes. The ports are assumed to be all the nodes of the top layer, 50 in number. Note that the voltage-to-current response is very localized in the X and Y coordinates. Figure 5 shows the first moment response of the network.

RecMOR is applied to this circuit and the structure of the sparsified model is shown in Figure 6. RecMOR achieves a very significant reduction in model complexity at a small accuracy cost. This is shown in Figure 7.

Conclusion

The paper introduced a new method, RecMOR, for model-order reduction of *complete* linear subcircuits characterized by a very large number of terminals. Previously, such systems were not amenable to reduction, because traditional methods

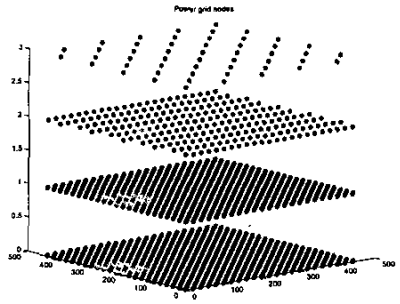

Fig. 4. Power grid and voltage-to-current response

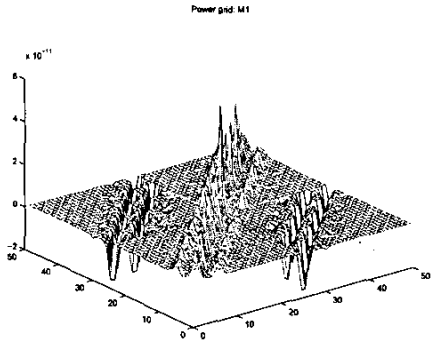

Fig. 5. Power grid first moment response

would result in *reduced* models that are more complex to store and evaluate than the original circuit. This apparent paradox is explained by the fact that reduced-order models for systems with large number of terminals are based on dense matrices while the original circuit equations are written in terms of sparse matrices, albeit much larger.

The RecMOR method restores the sparsity of the reduced-order model even in the cases when the number of terminals is very large. The method employs the recent SVDMOR algorithm which takes advantage of the correlations between circuit responses at various network terminals The models resulting from these algorithms become more efficient as the correlation between circuit responses is more pronounced.

While not a universal property of electrical circuits, such correlations are characteristic to large number of practical applications. As the examples analyzed in the paper indicate, the RecMOR and SVDMOR algorithms are particularly powerful in the analysis of regularly structured circuits, often used in modeling of power grids, clock networks, and wide buses.

SVDMOR and RecMOR compreson are compatible with practically any of the existing model order reduction algorithms.

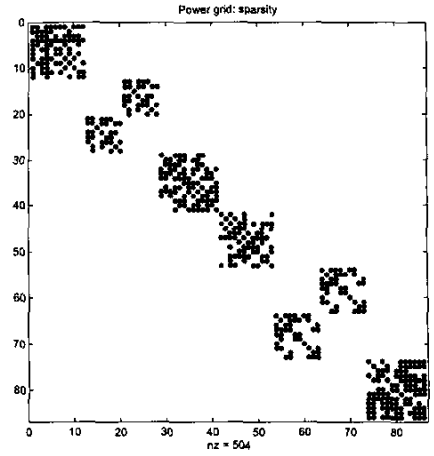

Fig. 6. Power grid model sparsity structure

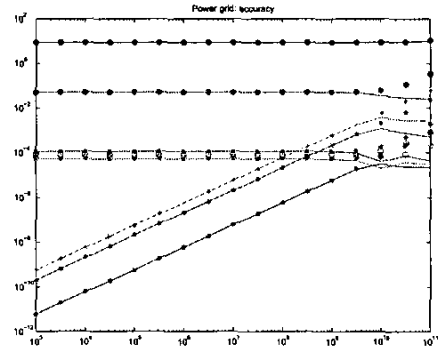

Fig. 7. Power grid model sparsity structure

References

[1] P. Feldmann and R. W. Freund, "Reduced-order modeling of large linear subcircuits via a block Lanczos algorithm," in *Proceedings of the 32nd Design Automation Conference, San Francisco, California.* Association for Computing Machinery, Inc., 1995, pp. 474–479.

[2] K. J. Kerns and A. T. Yang, "Preservation of passivity during RLC network reduction via split congruence transformations," in *Design Automation Conference*, 1997, pp. 34–39.

[3] K. Gallivan, E. Grimme, and P. Van Dooren, "A rational Lanczos algorithm for model reduction," *Numerical Algorithms*, vol. 12, no. 1–2, pp. 33–63, 1996.

[4] R. W. Freund and P. Feldmann, "The SyMPVL algorithm and its applications to interconnect simulation," in *Proceedings of the 1997 International Conference on Simulation of Semiconductor Processes and Devices.* IEEE, 1997, pp. 113–116.

[5] A. Odabasioglu, M. Celik, and L. Pileggi, "PRIMA: Passive reduced-order interconnect macromodeling algorithm," *IEEE trans. on CAD*, vol. 17, no. 8, pp. 645–654, 1998.

[6] P. Feldmann, "Model order reduction techniques for linear systems with large numbers of terminals," in *Proceedings of the Design, Automation and Test in Europe Conference 2004.* IEEE Computer Society Press, 2003.

[7] S. Kapur and D. E. Long, "IES3: a fast integral equation solver for efficient 3-dimensional extraction," in *Proceedings of the 1997 International Conference on Computer-aided Design.* IEEE, 1997, pp. 448–455.