Model order reduction techniques for linear systems with large numbers of terminals

Peter Feldmann

IBM T.J. Watson Research Center

Yorktown Heights, NY, USA

Abstract

This paper addresses the well known difficulty of applying model order reduction (MOR) to linear circuits with a large number of input-output terminals. Traditional MOR techniques substitute the original large but sparse matrices used in the mathematical modeling of linear circuits by models that approximate the behavior of the circuit at its terminals, and use significantly smaller matrices. Unfortunately these small *MOR matrices become dense as the number of terminals increases, thus canceling the benefits of size reduction. The paper introduces a model reduction technique suitable for circuits with numerous terminals. The technique exploits the correlation that almost always exists between circuit responses at different terminals. The correlation is rendered explicit through an SVD-based algorithm and the result is a substantial sparsification of the MOR matrices. The proposed sparsification technique is applicable to a large class of problems encountered in the analysis and design of interconnect in VLSI circuits. Relevant examples are used to analyze and validate the method.*

1 Introduction

Model order reduction (MOR) has become an established technique to analyze and compress modeling information on linear circuits and systems. MOR relies on the fact that in a vast array of applications we are only interested in the behavior of the circuit at its input and output terminals. Moreover, the number of these inputs and outputs is much smaller than the total size of the state vector (approximated by the number of circuit nodes). For example in the analysis of delay or noise in on-chip interconnect we study the propagation of signals in the wires that connect logic gates. These wires may have numerous features: bends, crossings, vias, etc., and are modeled by circuit extractors in terms of a large number of connected circuit elements: capacitors, resistors and more recently inductors. Nevertheless we are only interested in signal behavior at the inputs and outputs of the gates. MOR techniques generate compact models of the circuit that approximate well circuit behavior at the input and output terminals but renounce 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. 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, $O(m^2)$ 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 $O(m^2)$, which for a circuit with numerous inputs and outputs may approach or even surpass the complexity of working with the original, un-reduced, circuit equations.

This paper proposes a method, SVDMOR, that exploits exactly such a property that occurs in a large class of practical applications: In most cases there is a significant correlation between the $O(m^2)$ entries of the matrix transfer function. This correlation is exploited to produce reduced order models that can be computed and stored with much lower complexity.

The next section formalizes the problem mathematically. Next, Section 2 introduces and justifies the SVDMOR algorithm. Section 3 illustrates the new algorithm on a relevant circuit example.

2 Problem formulation

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

$$C\frac{d}{dt}x + Gx = Mu \qquad (1)$$
$$y = N^T x$$

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

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

and is in fact a $q \times p$ matrix-valued rational function.

Padé-based MOR algorithms [1, 4, 3, 2, 5], operate on the original circuit matrices G, C, M, N, and compute models described by smaller matrices. The transfer function of the reduced-order models approaches the original in the Padé approximation sense

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

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 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 $O(pq)$, while for typical circuits, the system matrices G and C are very sparse, having a number of non-zero entries of order $O(n)$. 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. In the sequel, we introduce a procedure which 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.

3   Algorithm description

We start with some algebraic manipulation, which is common to almost all model-order reduction methods. Let $G = QJQ^T$ be the LDL^T decomposition of the symmetric G matrix, where J is a *simple* matrix, block diagonal with 1×1 and 2×2 diagonal blocks. When G is a symmetric positive definite matrix we perform the Cholesky decomposition and J becomes just the identity matrix.

$$\begin{aligned} H(s) &= N^T (G + sC)^{-1} M \\ &= N^T [Q(J + s\underbrace{Q^{-1}CQ^{-T}}_{A})Q^T]^{-1} M \\ &= N^T Q^{-T} (J + sA)^{-1} Q^{-1} M \end{aligned}$$

The transfer function reduces to the standard form

$$H(s) = L^T (J + sA)^{-1} R \qquad (4)$$

where $A = Q^{-1}CQ^{-T}$, $L = Q^{-1}N$, and $R = Q^{-1}M$. In general we are interested to maintain the symmetry of the formulation. Therefore we introduce the $n \times (m = p + q)$ matrix B obtained from the juxtaposition of matrices L and R, $B = [LR]$. Using selection matrices E_L and E_R we can recover the original L and R matrices

$$\begin{aligned} L &= BE_L \qquad (5) \\ R &= BE_R \end{aligned}$$

The transfer function can now be expressed in terms of the juxtaposed matrix B

$$H(s) = E_L^T \underline{B^T (J + sA)^{-1} B} E_R \qquad (6)$$

where the underlined middle part of the expression is symmetrical.

The matrix B encodes all the input/output port definitions. Obviously in many applications all the inputs and outputs are not independent. On the contrary, typically there is a large degree of correlation between the various inputs and outputs. Such a correlation would manifest itself in the matrix B having highly dependent entries, or in other words with B being well approximated by a lower rank matrix. Note that in our formulation, B only contains DC (zero frequency) information on the system and the sparsification will be based on correlation that manifests itself at DC. The algorithm can be extended to use more complicated response correlations.

The low-rank approximation to B is computed through the *singular value decomposition*, (SVD),

$$B = U\Sigma V^T \qquad (7)$$

where $\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_m)$, and U and V are orthogonal matrices. 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
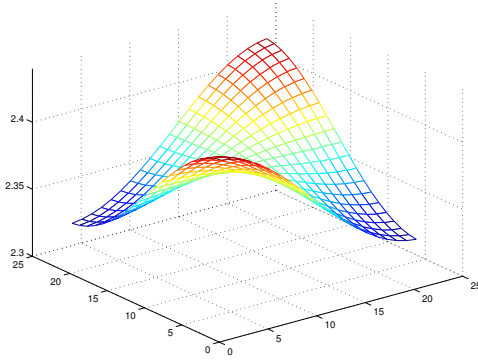
$$B = U\Sigma V^T \approx U_r V_r^T \qquad (8)$$

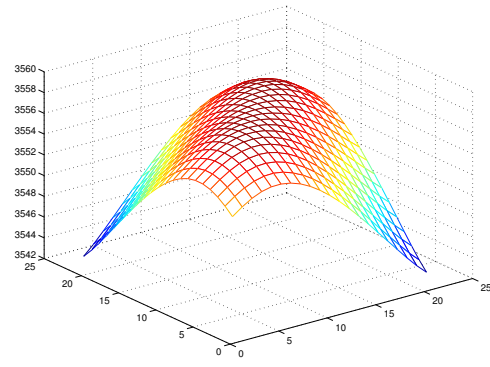**Figure 1. DC response of mesh**

Figure 2. Elmore delay of mesh

and U_r and V_r are $n \times r$ and $r \times r$ matrices respectively. The transfer function becomes

$$H(s) \approx E_L^T V_r \underbrace{U_r^T (J + sA)^{-1} U_r}_{H_r(s)} V_r^T E_R \qquad (9)$$

The standard model order reduction technique [1, 4, 3, 2, 5] can now be applied just to

$$H_r(s) = U_r^T (J + sA)^{-1} U_r \qquad (10)$$

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

$$H(s) \approx E_L^T V_r \tilde{H}_r(s) V_r^T E_R \qquad (11)$$

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

4 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 nodes are on the left side of the mesh and the output on the right side of the mesh. Assuming the mesh is of size 20×50 the transfer function that the reduced-order model needs to capture will be a 20×20 matrix-valued transfer function. The zeroth order moment (the DC component) and the first order moment (the Elmore delay) have the following expressions

$$\begin{aligned} \mathcal{M}_0 &= L^T G^{-1} R \\ \mathcal{M}_1 &= L^T G^{-1} C G^{-1} R \end{aligned} \qquad (12)$$

Figures 1 and 2 plot the entries of the two moment matrices and shows that their entries, far from being random, exhibit a high degree of correlation.
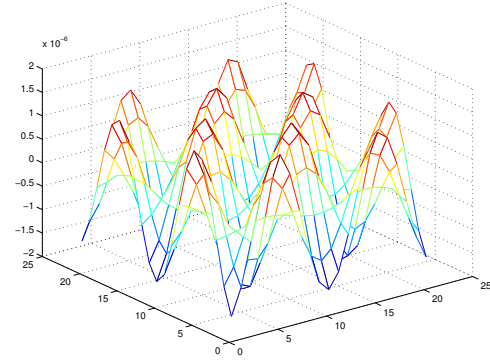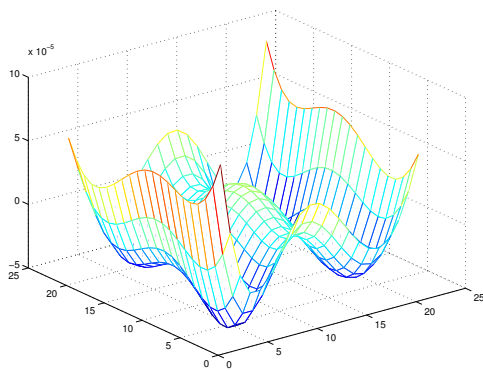

Figure 3. Error of sparsification: DC

Figure 3 and 4 show the relative error incurred by a *low-rank* approximation of the two moments. It turns out that a rank-4 sparsification is very accurate

The algorithm described in the previous section produces exactly such a low-rank approximation of matrix-moments. Figure 5 shows the s-domain transfer function of 4 entries entries of the transfer matrix, chosen to be as different as possible. The solid lines represent the exact response as obtained from solving the 1000-node circuit. The discrete points represent the approximations of the same transfer function by SVDMOR sparsification and by SyMPVL [1] model-order reduction. As expected the SVDMOR sparsified model matches the original transfer function quite accurately.

Conclusion

The paper introduced a new method, SVDMOR, for model-order reduction of linear systems characterized by a very large number of terminals. Previously, such systems were not amenable to reduction, since their so-called reduced-order model, could become as complex to store and evaluate as the original un-reduced model. This apparent paradox is explained by the fact that reduced-order models for systems with large number of terminals are based

Figure 4. Error of sparsification: Elmore
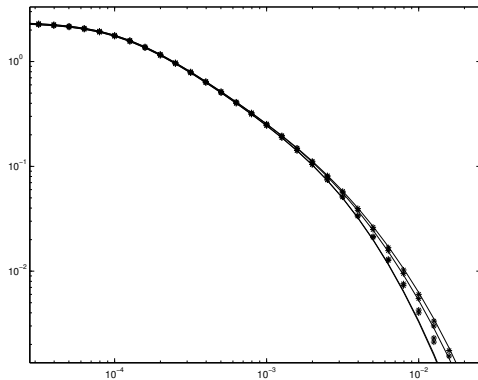
Figure 5. SVDMOR vs. exact responses

on dense matrices while the original circuit equations are written in terms of sparse matrices albeit much larger.

The SVDMOR method restores the sparsity of the reduced-order model even in the cases when the number of terminals is very large. The method exploits the correlation between circuit responses at various network terminals, and becomes 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 SVDMOR method is particularly powerful in the analysis of regularly structured circuits, often used in modeling of power grids and buses.

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*, pages 474–479. Association for Computing Machinery, Inc., 1995.

[2] 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*, pages 113–116. IEEE, 1997.

[3] K. Gallivan, E. Grimme, and P. Van Dooren. A rational Lanczos algorithm for model reduction. *Numerical Algorithms*, 12(1–2):33–63, 1996.

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

[5] A. Odabasioglu, M. Celik, and L. Pileggi. PRIMA: Passive reduced-order interconnect macromodeling algorithm. *IEEE trans. on CAD*, 17(8):645–654, 1998.