On Symbolic Model Order Reduction[1]

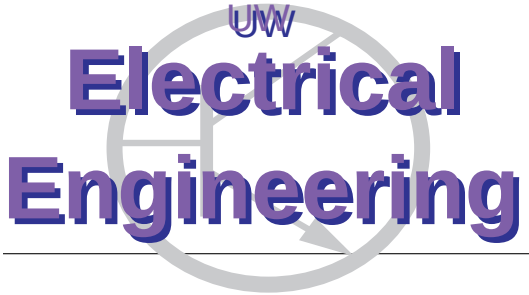
Guoyong Shi, Bo Hu, and C.-J. Richard Shi

Department of Electrical Engineering

University of Washington

Seattle, WA 98195

{gshi,hubo,cjshi}@ee.washington.edu



UWEE Technical Report
Number UWEETR-2004-0031
December 2004

Department of Electrical Engineering
University of Washington
Box 352500
Seattle, Washington 98195-2500
PHN: (206) 543-2150
FAX: (206) 543-3842
URL: http://www.ee.washington.edu

On Symbolic Model Order Reduction*

Guoyong Shi, Bo Hu, and C.-J. Richard Shi

Department of Electrical Engineering
University of Washington
Seattle, WA 98195
{gshi,hubo,cjshi}@ee.washington.edu

University of Washington, Dept. of EE, UWEETR-2004-0031

December 2004

Abstract

Symbolic model order reduction (SMOR) is a macromodeling technique that can be used to create reduced order models while retaining the parameters in the original models. Such symbolic reduced order models can be repeated evaluated (simulated) with greater efficiency for varying model parameters. Although the model order reduction concept has been extensively developed in the literature and widely applied in a variety of problems, model order reduction from a symbolic perspective has not been well studied. Several methods developed in this paper include symbol isolation, nominal projection, and first order approximation. These methods can be applied to models from having only a few parametric elements to many symbolic elements. Of special practical interest are models that have slightly varying parameters such as process related variations, for which efficient reduction procedure can be developed. Each technique proposed in this paper has been tested by circuit examples. Experiments show that the proposed methods are potentially effective for many circuit problems.

[1]This work was sponsored in part by the DARPA NeoCAD Program under Grant No. N66001-01-8920 from Navy Sapce and Naval Warfare Systems Command (SPAWAR), in part by SRC under Contract No. 2001-TJ-921, and in part by the National Science Foundation (NSF) CAREER Award under Grant No. 9985507.

1 Introduction

The model order reduction concept originally developed in control theory has been gaining popularity nowadays in electronic design community. After about ten years of research in its applications to the circuit related problems, model order reduction is now becoming a standard methodology widely used for interconnect modeling [1, 2], timing analysis [3], and in general macromodeling [4]. It provides an efficient means for behavioral modeling and system level design and simulation, and makes the simulation of complex systems possible [5]. In particular, the Krylov subspace based model order reduction approaches are gaining much more popularity because of their numerical efficiency and robustness [6, 7]. A comprehensive review in this regard can be found in the survey paper [8].

While model order reduction techniques for linear models are gaining much popularity and maturity, their extensions to nonlinear models and parametric models are still underdeveloped due to their intrinsically different nature of difficulty. In many electronic design problems, one frequently encounter linear (linearized) models involving parameters, such as process variation parameters, geometric parameters, design parameters, and even artificial parameters for optimization, etc. Such models are typically found in the simulation of parasitics, interconnects, and 3D electromagnetic structures, where models are of high order and parameters are often introduced to facilitate the design and analysis. One might need to carry out Monte Carlo simulation to investigate the parameter-dependent performances or other issues like optimization. In such cases, reduced order models involving the same set of parameters as that in the original models would definitely be a more efficient ways for achieving certain goal than by using the full order models. We formulate such type of problems in the framework of *Symbolic Model Order Reduction* (SMOR), where parameters are treated as symbols.

A few researchers have already attempted to address the problem of parametric reduced order modeling. A multivariate moment matching technique is used in [9] where parameters are assumed to be linearly separable. A variational analysis approach is taken in [10] for RCL interconnect modeling with statistically varying parameters. An interpolation technique is proposed in [11] for parametric interconnect analysis. All these methods pose specific assumptions on the models, hence their potential for general application is very limited. A generic solution for symbolic model order reduction is still not available in the literature.

The primary goal of this paper is to formulate the SMOR problem and propose a set of general-purpose solutions for certain applications. Two basic requirements in the SMOR framework are: (a) The computation procedure for constructing symbolic reduced order models be sufficiently efficient so that it is worthwhile to generate such models; (b) The evaluation of the reduced order symbolic models be sufficiently efficient as well. In other words, a practical SMOR methodology should be able to produce reduced order symbolic models without much effort and the produced models should be simple enough to evaluate. Other mathematical and physical properties such as accuracy, passivity, and stability etc. are naturally required as the ordinary model order reduction methodologies.

Symbolic model order reduction is in general a challenging problem. We do not intend to solve the problem completely in this single paper. The preliminary results developed in this paper are our initial efforts towards an ultimate solution for SMOR. Some partial results of this paper have been reported earlier in [12].

For a self-contained exposition, we first review in Section 2 some basic Krylov subspace techniques widely used in conventional model order reduction algorithms. Then three approaches to symbolic model reduction are developed. The first method for SMOR, introduced in Section 3, is the so called *symbol isolation* method, which is applicable to circuit models with a few symbolic

elements so that they can be isolated by using port connections. While direct extension of Krylov subspace approach to model order reduction to SMOR involves symbolic inversion of matrices, which is computationally expensive, such a difficulty can be overcome for certain special cases such as models with small parameter variations. For such models we attempt to use the nominal models to construct orthonormalized Krylov subspaces, with which approximate symbolic reduced order models are formed. This technique is based on an observation that the nominal orthonormalized Krylov subspaces have certain degree of robustness to warrant an approximate moment matching even the model parameters are perturbed slightly. This nominal projection method is our second method for SMOR and is presented in Section 4, where two methods for creating robust projection matrices are specifically discussed. The third method for SMOR developed in Section 5 is based on the first order approximation for symbolic matrix inversion, which can viewed as an improvement over the nominal projection method. In Section 6 we collects all experimental results to demonstrate the effectiveness of the proposed methods. Future research issues of SMOR are discussed in the conclusion section 7. The feasibility of symbolic inversion of certain special matrices arising from circuit problems is discussed in Section .1 of the appendix.

2 Preliminaries

Consider a linear circuit model that can be described by the following equations:

$$\begin{aligned} C\frac{dx}{dt} + Gx &= Bu \\ y &= Fx \end{aligned} \tag{1}$$

where $u \in \mathbb{R}^m$ is the external stimulus to the system, $B \in \mathbb{R}^{n\times m}$ is the input matrix, $x \in \mathbb{R}^n$ is the state vector, $F \in \mathbb{R}^{p\times n}$ is the output matrix, and $y \in \mathbb{R}^p$ is the output of the model. $C \in \mathbb{R}^{n\times n}$ is the susceptance matrix and $G \in \mathbb{R}^{n\times n}$ is the conductance matrix.

Congruence transformation is commonly used for reduction of circuit models with port formulation because passivity can be preserved [7]. Let $V \in \mathbb{R}^{n\times q}$ be the transformation matrix with $q \ll n$. By defining

$$C_r = V^{\mathrm{T}}CV, \quad G_r = V^{\mathrm{T}}GV, \quad B_r = V^{\mathrm{T}}B, \quad F_r = FV, \tag{2}$$

where the superscript $^{\mathrm{T}}$ indicates transpose of a matrix, the reduced order model of order q can be written as:

$$\begin{aligned} C_r\frac{dz}{dt} + G_r z &= B_r u \\ \widetilde{y} &= F_r x \end{aligned} \tag{3}$$

where $z \in \mathbb{R}^q$ is the new state of the reduced order model.

One popular method for generating the transformation matrix V is by moment matching. Let $X(s) = T(s)U(s)$ be the Laplace transform of the state space model, where $T(s) = (Cs+G)^{-1}B$. For moment matching, we expand $T(s)$ in Taylor expansion at $s = 0$, i.e.

$$\begin{aligned} T(s) &= (Cs+G)^{-1}B \\ &= \sum_{i=0}^{\infty} \left(-G^{-1}C\right)^i G^{-1}Bs^i, \end{aligned}$$

where the coefficients of s^i are called *moments*. Moment matching is directly connected to the Krylov subspace formed by the pair of matrices: $\left(G^{-1}C,\, G^{-1}B\right)$ [6]. The Krylov subspace is

spanned by the column vectors in the following collection of matrices

$$\left\{G^{-1}B,\, \left(G^{-1}C\right)G^{-1}B,\, \cdots,\, \left(G^{-1}C\right)^{i}G^{-1}B, \cdots\right\}$$

which are called the *Krylov vectors*. The qth order Krylov subspace is denoted by

$$\mathcal{K}_q\left(G^{-1}C,\, G^{-1}B\right), \tag{4}$$

which is spanned by the leading q linearly independent Krylov vectors.

Let $V \in \mathbb{R}^{n\times q}$ be any matrix whose columns span the Krylov subspace $\mathcal{K}_q\left(G^{-1}C,\, G^{-1}B\right)$. If the columns of V are orthonormalized, then it can be shown that the following identities hold [7]

$$\left(G^{-1}C\right)^{i}G^{-1}B = V\left(G_r^{-1}C_r\right)^{i}G_r^{-1}B_r \tag{5}$$

for $i = 0,\, 1,\, \cdots,\, q-1$. These identities can be used to verify that a number of the leading moments of the full order and reduced order transfer functions are matched [13].

The block vectors forming the Krylov subspace can be orthonormalized by using the Arnoldi algorithm. A block Arnoldi algorithm for a multicolumn B is described in [7]. Here we list the Arnoldi algorithm for a single column input matrix, i.e. $B = b$, where $b \in \mathbb{R}^n$.

Arnoldi Algorithm:

(i) LU factorize matrix G: $G = LU$.

(ii) Solve $\widetilde{v}_1$ from: $G\widetilde{v}_1 = b$.

(iii) Compute $h_{11} = \|\widetilde{v}_1\|$ and $v_1 = \widetilde{v}_1/h_{11}$.

(iv) For $j = 2, \cdots, q$:

Solve $\widetilde{v}_j$ from: $G\widetilde{v}_j = Cv_{j-1}$.
For $i = 1, \cdots, j-1$: $h_{ij} = v_i^{\mathrm{T}}\widetilde{v}_j$.
$w_j = \widetilde{v}_j - \sum_{i=1}^{j-1} v_i h_{ij}$
$h_{jj} = \|w_j\|$, $v_j = w_j/h_{jj}$.

Note that the Arnoldi algorithm terminates whenever $h_{jj} = 0$, which means that the subsequent vectors belong to the subspace already generated. The Arnoldi algorithm is basically a Gram-Schmidt procedure for orthonormalizing the Krylov vectors.

3 Symbol Isolation Method

In this section we introduce the first SMOR method called *symbol isolation*. This method is motivated by occasions in circuit simulation that a large network has to be simulated by a sweeping analysis over a few critical elements. In this case, the simulation efficiency would be improved greatly if the large network excluding those critical elements is replaced by a compact reduced order model.

The underlying idea of symbol isolation is rather straightforward: Isolate those symbolic elements and treat them as ports (see the illustration in Fig. 1); then reduce the rest of the network by a standard model reduction algorithm. Most model order reduction algorithms in the state space do not change the input/output port structure, the isolated symbolic elements can be merged back with the reduced order model after the subnetwork is reduced.

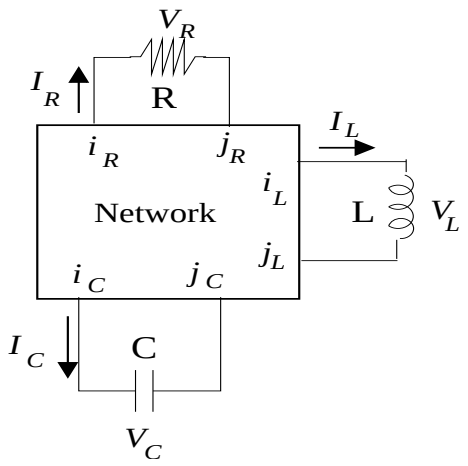


Figure 1: Isolation of symbolic elements from a network.

It would be revealing to formulate this simple idea in formal mathematics. As we shall see shortly, what the symbol isolation method does is equivalent to a block congruence transformation.

We start from considering a network with a single resistor (R), inductor (L), and capacitor (C) isolated for symbolic analysis (see Fig. 1). Each two-terminal element is treated as a port connected to the main network indicated by the block. The following formulation can be extended easily to multiple symbolic RCL elements.

To facilitate the MNA (Modified Nodal Analysis) formulation [14], we introduce some notations for the network model as shown in Fig. 1. Let x be the state of the network inside the block, including nodal voltages and currents inside the block. Together with the variables for the isolated elements, the full state vector becomes

$$x_f^{\mathrm{T}} = [x^{\mathrm{T}},\, \hat{I}_C,\, \hat{v}_C,\, \hat{I}_L,\, \hat{I}_R]^{\mathrm{T}},$$

where $\hat{I}_R$, $\hat{I}_L$, $\hat{I}_C$ are currents through the isolated elements R, L, and C, respectively, and $\hat{v}_C$ is the voltage across the isolated capacitor C. For easy identification, we use the *hat* notation below to indicate a symbolic element and its related variables. The model equations now become

$$\left[\begin{array}{c|cccc} C & & & & \\ 0 & & & & \\ \hline & \hat{C} & & & \\ & & \hat{L} & & \\ & & & 0 & \end{array}\right] \frac{d}{dt} \left[\begin{array}{c} x \\ \hat{I}_C \\ \hline \hat{v}_C \\ \hat{I}_L \\ \hat{I}_R \end{array}\right] = - \left[\begin{array}{cc|ccc} G & E_C & 0 & E_L & E_R \\ -E_C^{\mathrm{T}} & 0 & 1 & 0 & 0 \\ \hline 0 & -1 & 0 & 0 & 0 \\ -E_L^{\mathrm{T}} & 0 & 0 & 0 & 0 \\ -E_R^{\mathrm{T}} & 0 & 0 & 0 & \hat{R} \end{array}\right] \left[\begin{array}{c} x \\ \hat{I}_C \\ \hline \hat{v}_C \\ \hat{I}_L \\ \hat{I}_R \end{array}\right] + \left[\begin{array}{c} B \\ 0 \\ \hline 0 \\ 0 \\ 0 \end{array}\right] u$$

$$y = \left[\begin{array}{cc|ccc} F & 0 & 0 & 0 & 0 \end{array}\right] x_f \tag{6}$$

where $\hat{R}$, $\hat{L}$, $\hat{C}$ are the resistance, inductance, and capacitance of the symbolic elements. In (6) E_R, E_L, and E_C are three column vectors containing all zeros but 1 and -1 at the two locations corresponding to the port node indices with respect to the isolated R, L, and C, respectively. Here we assume that the variables with *hat* are not part of the output. We note from (6) that the symbolic elements have been isolated to the trailing part of the state model by ordering the state variables appropriately.

We rewrite the model of (6) in a concise block form as

$$\begin{bmatrix} \bar{C} & \\ & \hat{D} \end{bmatrix} \frac{d}{dt} \begin{bmatrix} \bar{x} \\ \hat{x} \end{bmatrix} = -\begin{bmatrix} \bar{G} & E \\ -E^{\mathrm{T}} & \hat{J} \end{bmatrix} \begin{bmatrix} \bar{x} \\ \hat{x} \end{bmatrix} + \begin{bmatrix} \bar{B} \\ 0 \end{bmatrix} u$$
$$y = \begin{bmatrix} \bar{F} & 0 \end{bmatrix} \begin{bmatrix} \bar{x} \\ \hat{x} \end{bmatrix} \tag{7}$$

where

$$\bar{C} = \begin{bmatrix} C & \\ & 0 \end{bmatrix}, \quad \bar{G} = \begin{bmatrix} G & E_C \\ -E_C^{\mathrm{T}} & 0 \end{bmatrix}, \quad E = \begin{bmatrix} 0 & E_L & E_R \\ 1 & 0 & 0 \end{bmatrix},$$
$$\hat{D} = \begin{bmatrix} \hat{C} & & \\ & \hat{L} & \\ & & 0 \end{bmatrix}, \quad \hat{J} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \hat{R} \end{bmatrix}, \quad \bar{F} = \begin{bmatrix} F, & 0 \end{bmatrix},$$
$$\bar{x} = \begin{bmatrix} x \\ \hat{I}_C \end{bmatrix}, \quad \hat{x} = \begin{bmatrix} \hat{v}_C \\ \hat{I}_L \\ \hat{I}_R \end{bmatrix}.$$

The block model (7) can be split into a nonsymbolic part and a symbolic part.

$$\bar{C}\frac{d\bar{x}}{dt} = -\bar{G}\bar{x} + \begin{bmatrix} -E & \bar{B} \end{bmatrix} \begin{bmatrix} \hat{x} \\ u \end{bmatrix} \tag{8a}$$
$$\hat{D}\frac{d\hat{x}}{dt} = E^{\mathrm{T}}\bar{x} + \hat{J}\hat{x} \tag{8b}$$
$$y = \bar{F}\bar{x}. \tag{8c}$$

Note that in this split model the isolated symbols only appear in equation (8b) which is of order 3. The remaining part consists of a model without symbols, which is duplicated below:

$$\bar{C}\frac{d\bar{x}}{dt} = -\bar{G}\bar{x} + \bar{B}_a u_a$$
$$y = \bar{F}\bar{x} \tag{9}$$

where

$$\bar{B}_a = \begin{bmatrix} E & \bar{B} \end{bmatrix}, \quad u_a = \begin{bmatrix} \hat{x} \\ u \end{bmatrix}$$

are the augmented input matrix and input vector, respectively.

The model in (9) is still a large network but without symbols. The input dimension of this model has been augmented after the variables at the ports are reformulated as part of the input. This model can be reduced by standard reduction methods such as PRIMA [7] or PVL [6]. Suppose we apply PRIMA to the matrix triple ($\bar{C}$, $\bar{G}$, $\bar{B}_a$) and assume $V \in \mathbb{R}^{n_1 \times q}$ is the transformation matrix, where n_1 is the model order of (9) and q is the reduced model order. Let $\bar{x} = Vz$ and premultiply the first equation in (9) by V^{T}. We obtain the following reduced order model from (9):

$$\bar{C}_r\frac{d\bar{x}}{dt} = -\bar{G}_r\bar{x} + \bar{B}_r\bar{u}$$
$$y = \bar{F}_r\bar{x} \tag{10}$$

where

$$\bar{C}_r = V^{\mathrm{T}}\bar{C}V, \quad \bar{G}_r = V^{\mathrm{T}}\bar{G}V, \quad \bar{B}_r = V^{\mathrm{T}}\bar{B}_a, \quad \bar{F}_r = \bar{F}V.$$

Note that after the reduction, the network inside the block in Fig. 1 is replaced by a smaller sized model while the port structure is preserved. As a result, we obtain a reduced order model of the original network with the isolated symbolic elements preserved.

Given the block structure of model (6), the reduction procedure presented so far can be stated equivalently in terms of a block transformation. Using the notation introduced above, the reduced symbolic model is

$$\begin{aligned} \begin{bmatrix} V^{\mathrm{T}}\bar{C}V & \\ & \hat{D} \end{bmatrix} \frac{d}{dt} \begin{bmatrix} z \\ \hat{x} \end{bmatrix} &= -\begin{bmatrix} V^{\mathrm{T}}\bar{G}V & V^{\mathrm{T}}E \\ -E^{\mathrm{T}}V & \hat{J} \end{bmatrix} \begin{bmatrix} z \\ \hat{x} \end{bmatrix} + \begin{bmatrix} V^{\mathrm{T}}\bar{B} \\ 0 \end{bmatrix} u \\ y &= \begin{bmatrix} \bar{F}V & 0 \end{bmatrix} \begin{bmatrix} z \\ \hat{x} \end{bmatrix} \end{aligned} \tag{11}$$

It is easily seen that this reduced order model is obtained from the block model (7) by applying the block transformation $\begin{bmatrix} V & \\ & I_3 \end{bmatrix}$.

One should not have much difficulty to extend the symbol isolation procedure formulated above to models with multiple symbolic RCL elements. However, as the number of ports increases, the input dimension of model (9) will also increase, which may lead to difficulty in reducing the network to a lower order. This limits the application of the symbol isolation method to only a few symbols, regardless of R, L, or C. But the method could be very useful for sweeping analysis of a few critical circuit elements.

4 Nominal Projection Method

Parametric models frequently appear in VLSI circuit design and simulation. For example, RC(L) models for interconnect analysis can have R, C, or L as parameters because of the uncertainty in the process. Geometric parameters as design parameters can also be introduced in models. For many parametric models, directly treating parametric elements as symbolic ones as we did in the previous section would not be efficient, because many elements could depend on only a few parameters and the large number of ports arising from symbol isolation is a problem.

If one attempts to use a projection method for symbolic model order reduction, it would be necessary to obtain symbolic Krylov subspaces. However, since the general procedure of Krylov subspace computation involves a sequence of algebraic operations that forbids efficient symbolic manipulation, in particular the symbolic matrix inversion is not in general computationally feasible, direct construction of symbolic Krylov subspace is in general not a viable solution.

However, there are still certain cases in practice that do not prevent us from developing symbolic approaches to parametric model order reduction. One of such cases is that the model parameters do not change drastically. In this case the parameter variations can be treated as small perturbations to the models. From this perspective, we are interested in developing easily computable approximate Krylov subspaces for symbolic model order reduction. In this section we propose the *second* method for SMOR, called *nominal projection* method. The basic idea is to compute a sufficiently *robust* Krylov subspace so that the same subspace can be used for models with slightly perturbed parameters. Meanwhile we shall investigate the issues on the computation of robust Krylov subspaces and the effectiveness by using nominal projections.

Parametric linear time-invariant models can be described by

$$\begin{aligned} C(p)\frac{dx}{dt} + G(p)x &= B(p)u \\ y &= F(p)x \end{aligned} \tag{12}$$

where $C(p)$, $G(p)$, $B(p)$, and $F(p)$ are model matrices depending on the parameter vector p which contains a number of parameters. With a projection matrix $V(p_0)$ computed from a set of nominal parameters p_0, the reduced parametric matrices are written as

$$\begin{aligned} C_r(p) &= V(p_0)^{\mathrm{T}} C(p) V(p_0), \quad G_r(p) = V(p_0)^{\mathrm{T}} G(p) V(p_0), \\ B_r(p) &= V(p_0)^{\mathrm{T}} B(p), \quad F_r(p) = F(p) V(p_0). \end{aligned} \tag{13}$$

In the following we discuss two methods for computing the nominal projection matrix $V(p_0)$.

We assume that all the model parameters only perturb around some nominal values in a certain design task. The ranges of parameter perturbation are problem dependent. Since parameter perturbation can be viewed as model uncertainty, reduction to a lower dimensional state space in general suppresses such uncertainty, which is a well known fact in the statistical analysis literature [15].

We call the model with all parameters fixed to their nominal values a *nominal model*, from which nominal projections are constructed. A nominal projection matrix should be robust enough to tolerate the model perturbation. Two computation methods are introduced below for robust nominal projection construction. The development is largely heuristic in that no analytical error bound will be established.

4.1 Mixed moment matching

The *first* method for nominal projection computation is to combine the moments from both the zero and infinity frequency points. Given a nominal model as in (1) with $p = p_0$, the Krylov subspace related to moment matching at $s = 0$ is

$$\mathcal{K}_q \left(G(p_0)^{-1} C(p_0),\, G(p_0)^{-1} B(p_0) \right). \tag{14}$$

Similarly, the Krylov subspace related to moment matching at $s = \infty$, i.e. expanding the transfer function $T(s)$ in terms of $1/s^i$, is

$$\mathcal{K}_p \left(C(p_0)^{-1} G(p_0),\, C(p_0)^{-1} B(p_0) \right). \tag{15}$$

Here we assume that both matrices $C(p_0)$ and $G(p_0)$ are nonsingular.

The nominal projection matrix $V(p_0)$ is then constructed from the Krylov vectors partially from (14) and (15). More specifically, we choose the leading q_1 Krylov vectors from (14) and the leading q_2 vectors from (15) to from $q = q_1 + q_2$ vectors for a nominal projection. Since the Krylov subspace formed by the low and high frequency moments captures both steady state and transient behaviors of the response, better robustness is expected than a Krylov subspace formed solely by either low or high frequency components. Its effectiveness is demonstrated by an example in the experiment section.

4.2 Real rational Krylov subspace

Since model order reduction has been widely used in interconnect design and analysis, a few comments on moment matching are worthwhile. As the operating frequency increases, the conventional RC models for interconnect is becoming inadequate, and the inductance effect of interconnect must be explicitly addressed [16]. Since an RC circuit does not have resonance, moment matching at the low frequency can sufficiently capture the frequency response, hence very compact reduced order models can be obtained by moment matching. However, for RCL models with low-loss, due to the resonance, capturing the resonance behavior at the high frequency band usually requires a

high dimensional Krylov subspace. Although the moment matching technique proposed above by matching both the low and high frequency bands can possibly produce a more compact model, we found in our experiment that this method could not help improve the frequency response accuracy in the resonance band. Moreover, the Krylov subspace formed at $s = 0$ often led to unstable or singular reduced order models when the circuit is highly inductive.

To solve this problem, we propose the *second* method for nominal projection computation, where we use real rational Krylov subspace. A real rational Krylov subspace originates from the expansion of the transfer function $T(s) = (Cs + G)^{-1}B$ at a real point σ, i.e.

$$\begin{aligned} T(s) &= (Cs + G)^{-1}B \\ &= \left[C(s - \sigma) + (C\sigma + G)\right]^{-1} B \\ &= \sum_{i=0}^{\infty} \left[-(C\sigma + G)^{-1}C\right]^i (C\sigma + G)^{-1}B(s - \sigma)^i \end{aligned} \tag{16}$$

Similarly to the expansion at $s = 0$, if one would like to have a reduced order model that matches the leading q moments in this expansion, i.e. the coefficients of the terms $(s-\sigma)^i$ for $i = 0, \cdots, q-1$, one can use the Krylov subspace

$$\mathcal{K}_q\left((C\sigma + G)^{-1}C,\, (C\sigma + G)^{-1}B\right). \tag{17}$$

By taking $\sigma = 0$, this Krylov subspace (17) reduces to the one in (4). Here we assume that a real σ has been chosen so that $(C\sigma + G)^{-1}$ exists.

The proper choice of σ has been addressed in many papers. Grimme [13] extensively discussed rational Krylov subspace approach to model order reduction. Chiprout and Nakhla [17] used complex σ for localized moment matching in the frequency domain. Shi and Shi further discussed a new interpretation of real σ from the waveform matching perspective [11] and the dominant subspace computation perspective [18]. The latter new developments motivate us to use real rational Krylov subspaces for robust nominal subspace computation.

An example used in Section 6 will demonstrate that a real point Krylov subspace is quite robust for reducing inductive circuit models. The key feature is that as the parameters are perturbed, the reduced order model obtained by using a nominal projection computed from a real rational Krylov subspace still captures the perturbed resonance modes very well.

5 First Order Approximation Method

The nominal projection method worked well in our experiment for small parameter perturbations. It has the advantage of simple projection matrix computation, numerical stability, and reliability because of orthonormalization. In case the accuracy by using nominal projection cannot meet the requirement, one can make modifications to the nominal projection matrix to achieve higher accuracy. In this section we propose the *third* method for SMOR, called *first order approximation*. The basic idea is to retain the first order terms while generating the Krylov vectors so that the Krylov subspace is approximately updated as the model data are perturbed. The Krylov subspace created this way is partially symbolic.

The main algebraic operations involved in Krylov subspace creation are matrix inversion and matrix vector multiplication. Since direct inversion of a symbolic matrix is computationally non-trivial, we take a partial symbolic approach to symbolic matrix inversion under the condition that *parameter variations are small*. Special circuits that enable fully symbolic matrix inversion are discussed in Section .1 in the appendix.

Let the perturbation of matrices G and C be as follows

$$G' = G + \Delta G, \qquad C' = C + \Delta C$$

where G and C are the matrices with the nominal values. Under the assumption that ΔG is small in certain matrix norm, the inverse of G' can be approximated by the first order expression

$$(G + \Delta G)^{-1} \approx G^{-1} - G^{-1}(\Delta G)G^{-1}. \tag{18}$$

Since the nominal matrix G is numeric, its inverse is also numeric and can be computed in advance. The parameter variations are symbolically characterized by the variation matrix ΔG. Clearly, the formula (18) greatly reduces the symbolic manipulation complexity.

Given the perturbation matrices ΔG and ΔC, the exact basis vectors of the Krylov subspace are computed from

$$\left[(G + \Delta G)^{-1}(C + \Delta C)\right]^k (G + \Delta G)^{-1} B \tag{19}$$

for $k = 0, 1, \cdots, q-1$. Again, for reducing the computation complexity, each expression in (19) can be approximated by keeping the terms up to the first order of ΔG and ΔC. It can be verified by induction that

$$\begin{aligned}&\left[(G + \Delta G)^{-1}(C + \Delta C)\right]^k \\ \approx &(G^{-1}C)^k - \sum_{i=1}^{k-1}(G^{-1}C)^i(G^{-1}\Delta G)(G^{-1}C)^{k-i} + \sum_{j=0}^{k}(G^{-1}C)^j(G^{-1}\Delta C)(G^{-1}C)^{k-j}\end{aligned}$$

by the first order approximation. Note that in this expression all the matrices except ΔG and ΔC are numeric, hence their algebraic multiplications can be carried out *a priori*. Thus it is feasible to obtain an approximate symbolic representation of the Krylov basis vectors. It is not difficult to derive the symbolic first order approximation of all the basis vectors in (19).

Note that, although the symbolic procedure outlined above is feasible for a set of symbolic Krylov basis vectors, the orthonormalization of them is nontrivial, because unlike the numeric case, the memory requirement for symbolic orthonormalization is high. To reduce the computation complexity, the symbolic basis vectors are not orthonormalized in our implementation. However, non-orthonormalized projection matrices could lead to singularity of reduced order models because of bad numerical conditioning when the order becomes high.

6 Experimental Results

Collected in this section are experimental results for testing the three main methods for SMOR presented in the preceding sections.

6.1 Symbol Isolation

The RC ladder circuit shown in Fig. 2 is used to test the symbol isolation method. The circuit has 100 stages with the element (either R, L, or C) between nodes 11 and 12 used as a symbolic element. The full order model is reduced to the 12th order. Shown in Fig. 3 is the transient responses to a two-level stimulus input voltage. The dashed curve indicates the nominal response of the full order model; the dotted solid curves (overlapped) are the responses of the full order and the reduced order models with a new value for R. Other transient responses with the symbolic element replaced by L or C are shown in Fig. 4 (a) and (b). As expected, the symbol isolation method achieves reliable time domain response.

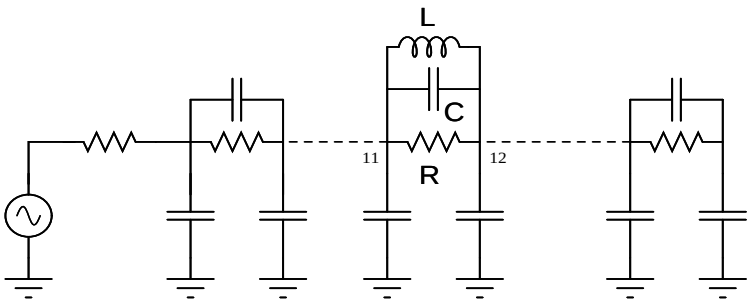


Figure 2: Test circuit for symbol isolation.

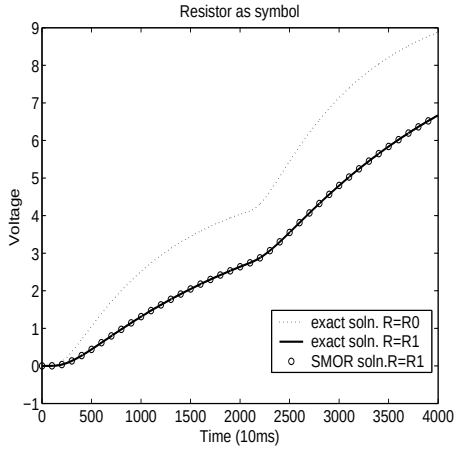
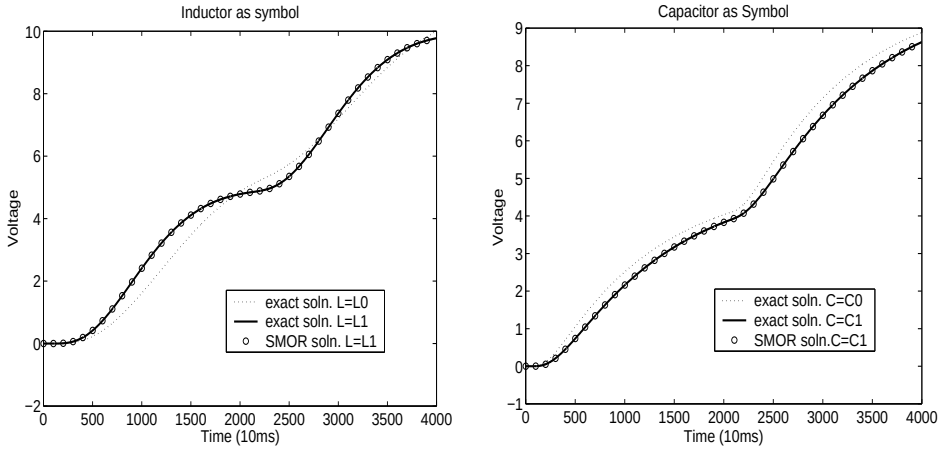


Figure 3: Transient responses with the resistor between nodes 11 and 12 as a symbolic element.



(a) (b)

Figure 4: Transient responses with a symbolic element between nodes 11 and 12: (a) inductor L, (b) capacitor C.

6.2 Nominal Projection

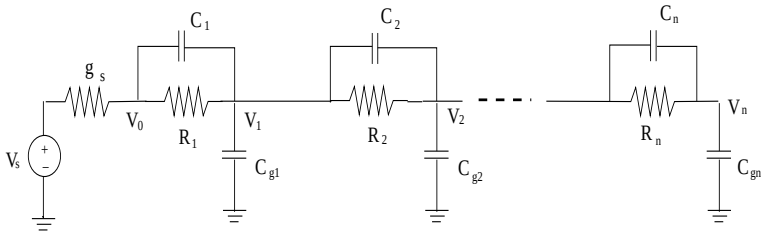
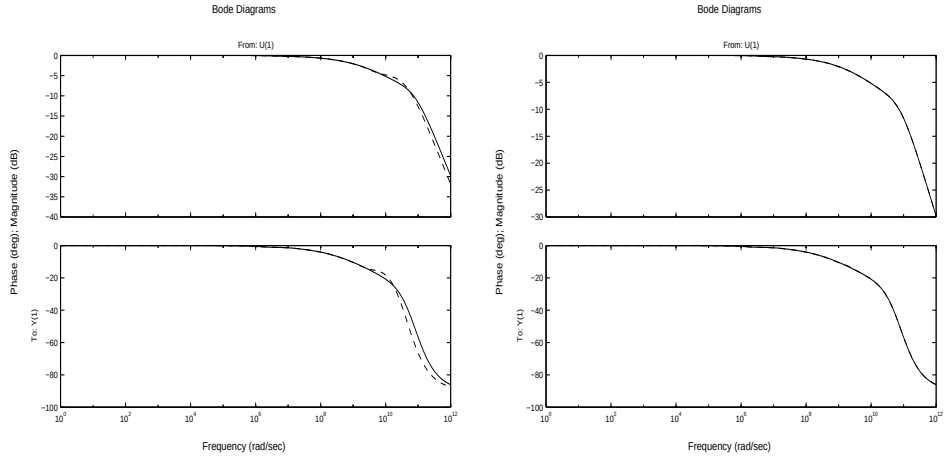


Figure 5: An RC ladder circuit.

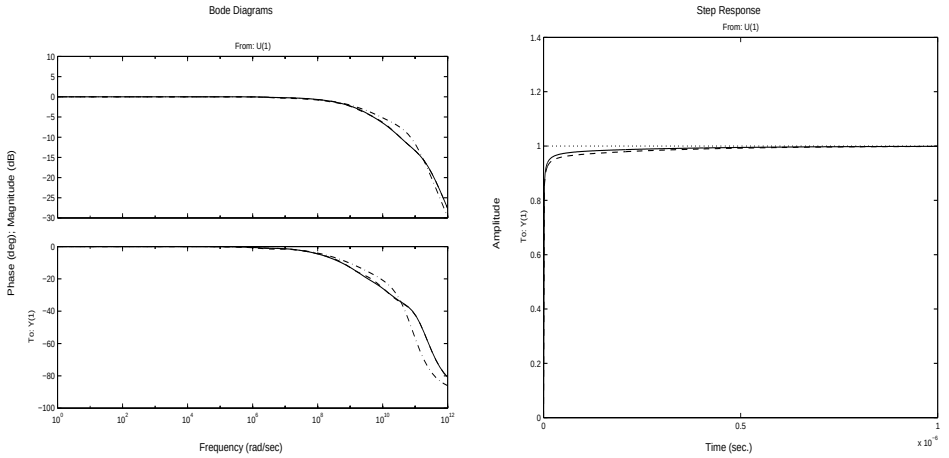
The RC ladder circuit in Fig. 5 is used for testing the nominal projection methods. V_s is the input voltage and V_1 is the output voltage. For demonstration purpose, we choose all $R_i = 20\,\Omega$, all $C_i = C_{gi} = 1\,pF$, and $g_s = 1/50\,S$ for nominal values. These values are perturbed up to certain level to test the robustness of the nominal projection. In this experiment the full order model is of the 200th order and is reduced to the 10th order. We choose five Krylov vectors each from the Krylov subspaces in (14) and (15) to form a 200×10 nominal projection matrix V.



(a) (b)

Figure 6: Comparison of the reduction effect in the frequency domain: (a) By moment matching at $s = 0$. (b) By mixed moment matching at both low and high frequencies.

Since moment matching at $s = 0$ only does not guarantee a good matching at the high frequency, we compare the reduction effects by using the Krylov subspace (14) with $q = 10$ and the mixed Krylov subspace consisting of (14) with $q_1 = 5$ and (15) with $q_2 = 5$. The nominal reduction effects are compared in Fig. 6 in Bode plots, where the dashed curve indicates the full order model. Clearly the mixed moment matching method gives a better frequency response matching at all frequencies of interest. The robustness of the nominal projection matrix formed by mixed moment matching is further demonstrated by the plots in Fig. 7, where Fig. 7(a) shows the frequency responses of full and reduced order models at the nominal and perturbed parameter values, with the circuit parameters perturbed by 60%. Note that the frequency responses of the full and reduced order perturbed models overlap in the plot. The dashed curve indicates the nominal full order model for reference. The transient response is shown in Fig. 7(b).



(a) (b)

Figure 7: Normal projection by mixed moment matching: (a) Bode plots of the full and reduced order models with perturbation up to 60%. (b) Step responses of the perturbed full and reduced order models (solid: full order; dashed: reduced order.)

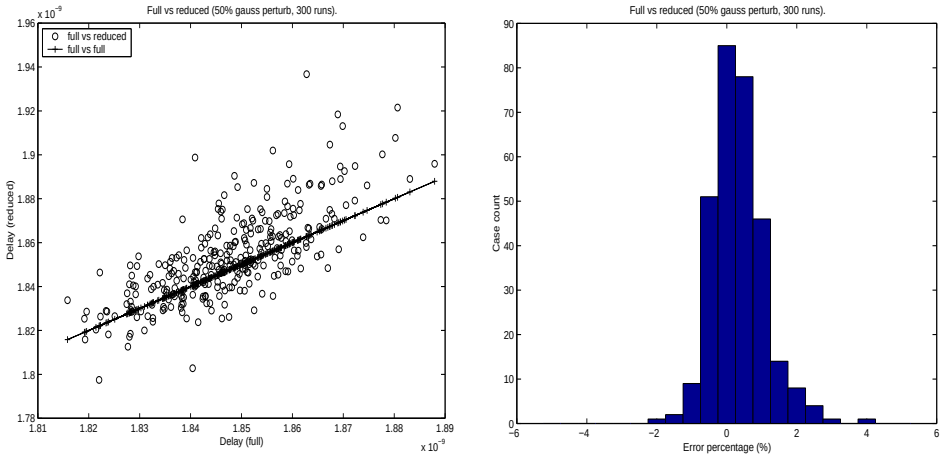
The reliability of the nominal projection method is further demonstrated by Monte Carlo tests. Shown in Fig. 8 are the Monte Carlo test results for the RC ladder circuit with 300 runs. The nominal parameters are perturbed individually by Gaussian distribution up to 50%. We use the rise time up to 0.8 of a step response for delay measurement. Fig. 8(a) shows the distribution of the measured delays out of 300 runs by using a nominal projection matrix computed by the mixed moment matching method, with the x-axis for the delay of the full order model and the y-axis for the delay of the reduced order model. The diagonal line marked '+' is the equal-delay line for reference. The plot shows the clustering of the delay measurements surrounding the reference line. In view of the small axes range used in this plot, 10^{-9}, the delays measured from the reduced order models are mostly quite accurate. Fig. 8(b) shows a histogram of the *relative delay error percentage* computed from 300 runs. The relative delay error is defined by

$$\frac{\textit{Delay of reduced model} - \textit{Delay of full model}}{\textit{Delay of full model}}.$$

The relative delay error percentage mostly falls within 2%, meaning that the nominal projection has adequate robustness given the fact that the parameters have been perturbed up to 50%.

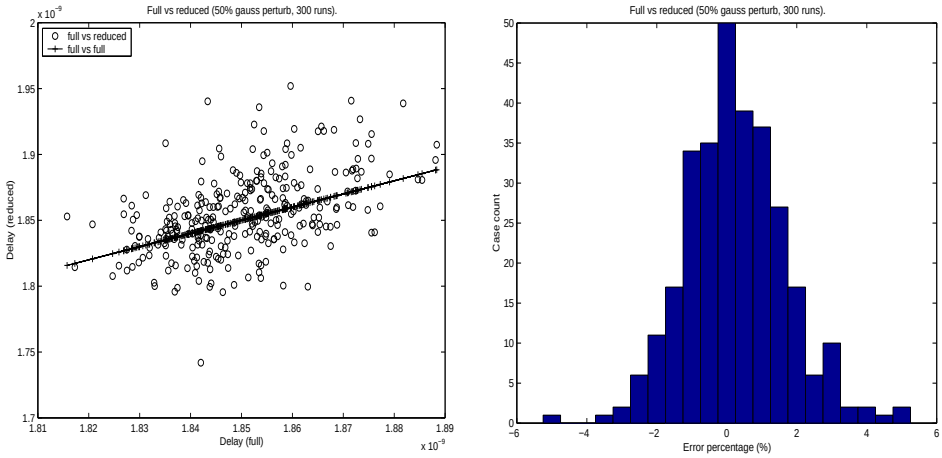
To compare whether the mixed moment matching has a better performance in delay measuring, we ran another Monte Carlo test with 300 runs, but using the 10th order Krylov subspace at $s = 0$. The test results are shown in Fig. 9. The histogram in Fig. 9(b) shows a wider deviation of the relative delay error with a lower count at the center, comparing to that in Fig. 8(b). Hence, the nominal projection computed from mixed moment matching has a better robustness. We note that among the 300 runs both methods did not encounter an unstable reduced order model.

The second method for nominal projection computation is to use real rational Krylov subspace. The robustness of such a nominal projection is tested by the inductive ladder circuit shown in Fig. 10. In this example, the nodal voltages and the currents passing the inductors are the state variables. We choose a model with the 320th order and reduce it to the 50th order with a real rational Krylov subspace at $\sigma = 10^9$. Uniform nominal RCL values are chosen with $R = 0.2\,\Omega$, $L = 1.0\,nH$, and $C = 0.5\,pF$. Figure 11 shows the frequency responses of both the full order and



(a) (b)

Figure 8: Monte Carlo test of the RC ladder circuit reduction using the mixed moment matching. (a) Delay distribution of 300 cases with Gaussian perturbation up to 50%. (b) Histogram of the measured delay error percentage.



(a) (b)

Figure 9: Monte Carlo test of the RC ladder circuit reduction using the moment matching at $s = 0$. (a) Delay distribution of 300 cases with Gaussian perturbation up to 50%. (b) Histogram of the measured delay error percentage.

reduced order models together with the error plot. The real rational Krylov subspace has achieved good approximation at the resonance frequency band.

To test whether the nominal projection still works for perturbed parameters, we add random perturbations to each RLC value up to $\pm 50\%$ and then perform the reduction by using the nominal projection again. Shown in Fig. 12 is the frequency response result and the error plot. The frequency response of the nominal full order is also plotted (the dotted curve) for reference. Clearly, the frequency response of the reduced order model still captures the full order frequency response quite well. Although the error becomes larger, it is still at an acceptable level.

The transient response is shown in Fig. 13, which displays the output of the perturbed full and

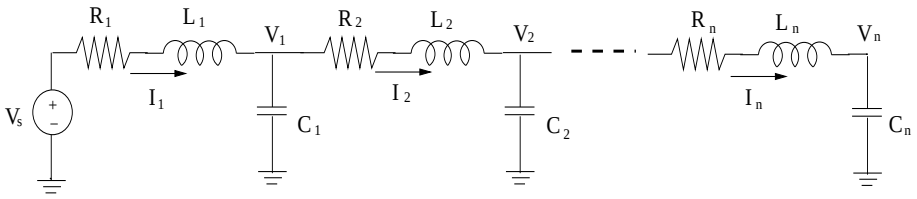


Figure 10: An RCL ladder circuit.

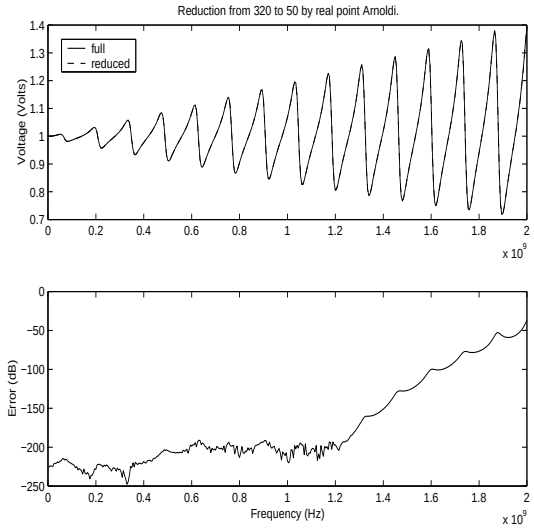


Figure 11: Nominal reduction.

reduced order models in response to a sinusoidal input at the frequency 1GHz. Except some slight distortion at the transient, the two waveforms match quite well.

Shown in Figs. 14 and 15 are reduction results for another set of perturbed parameters, with Fig. 15 showing more cycles of the sinusoidal response.

6.3 First Order Approximation

The first order approximation method is an improvement over the nominal projection method for better accuracy with the price of higher algebraic operation complexity. The RC circuit in Fig. 5 is used again to test the first order approximation method for approximate symbolic Krylov subspace construction. The circuit consists of 100 stages of RC blocks with the nodal voltages as the state variables with the 100th order.

For demonstration purpose, we used the mathematical software Maple to do symbolic algebra. The circuit is divided into three sections: nodes 1 to 30 for section 1, nodes 31 to 60 for section 2, and the rest for section 3. Four parameters are introduced as the symbols, which are respectively the perturbations added to the capacitors in the first section, to the resistors in the first section, to the capacitors in the second section, and to the resistors in the second section. The resistors and capacitors in the third section are not perturbed. The model is reduced from the 100th order to the 4th order. The effectiveness of the first order approximation method is tested by running Monte Carlo test with the four parameters perturbed up to 30% with normal distribution. We

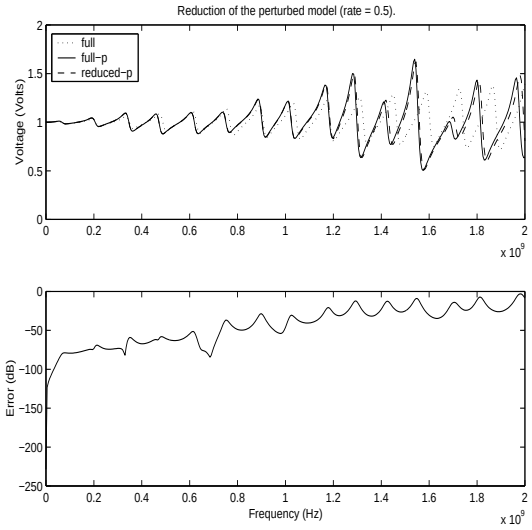


Figure 12: Reduction of a perturbed model by nominal reduction.

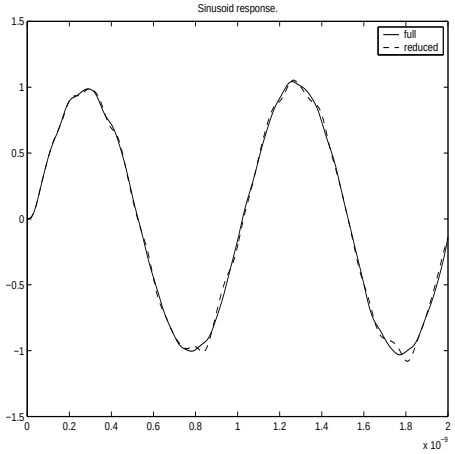


Figure 13: Sinusoidal response at $f = 1e9\,Hz$.

simulated the unit step responses of both the full and reduced order models, and compared the 50% rise times. The statistical test results are shown in Fig. 16, with the delay pairs shown in Fig. 16(a) and a histogram in Fig. 16(b). These figures show that the first order approximation method produced very accurate reduced order models measured by the rise time, more accurate than the nominal projection method, but with more complex reduced order models because of the symbolic representation.

7 Conclusions

Motivated by the practical needs for behavioral modeling in parametric form, we have explored the possibility and potential of symbolic model order reduction in this paper. Three methods are proposed, symbol isolation, nominal projection, and first order approximation, with each being

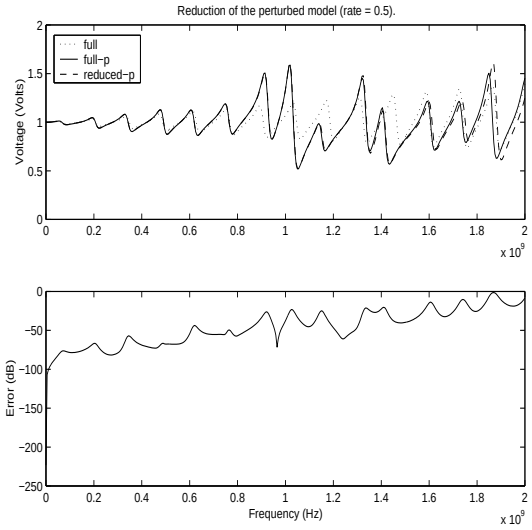


Figure 14: Reduction of another perturbed model by nominal reduction.

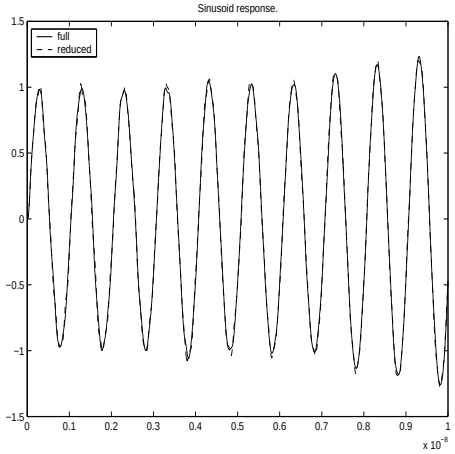
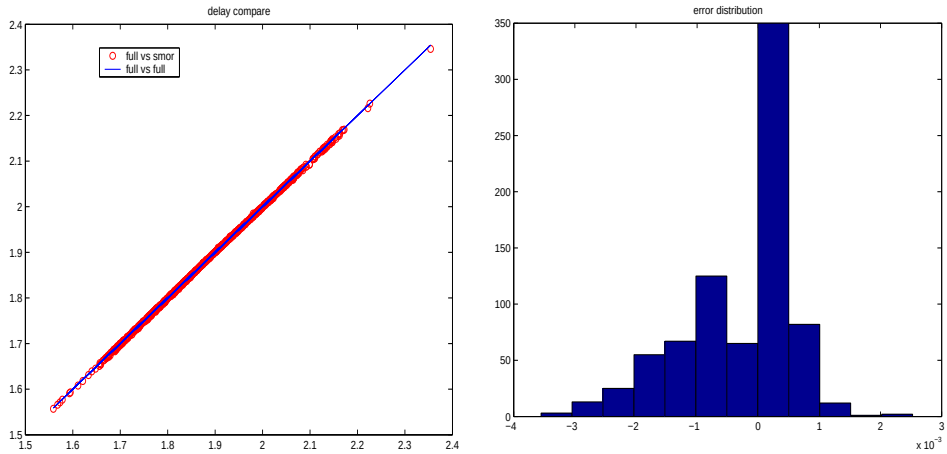


Figure 15: Another sinusoidal response at $f = 1\,GHz$.

applicable for circuit models with certain features. The potential effectiveness of these methods have been demonstrated by experiments on some typical circuits that are representative models for interconnects.

Although numerous model order reduction algorithms are available in the literature for numeric linear model order reduction, their extension to symbolic model order reduction is in general not trivial. The experimental study carried out in this paper has revealed that many other potential solutions for SMOR are worth further investigation. We anticipate that a general SMOR theory might not be available in the near future. Nevertheless, solutions with potential generic applications still can be developed. Other directions for further research include efficient model order reduction methods for uncertain models, non-subspace based approaches, and other identification-based approaches to SMOR.



(a) (b)

Figure 16: Monte Carlo test of the symbolic first order approximation method. (a) Delay distribution of 800 cases with normal perturbation up to 30%. (b) Histogram of the measured delay error.

.1 Symbolic Inversion of the G Matrix

Symbolic inversion of a symbolic matrix is computationally expensive in general. Here we briefly discuss that for certain circuit problems, symbolic inversion of the G matrix is relatively simple. Although this is a quite restricted type of matrices, it actually covers a large class of circuits.

There is a type of circuits whose models can be written in MNA formulation as

$$\begin{bmatrix} \boldsymbol{C} & \\ & \boldsymbol{L} \end{bmatrix} \frac{d}{dt} \begin{bmatrix} \boldsymbol{v}_C \\ \boldsymbol{v}_L \end{bmatrix} + \begin{bmatrix} \boldsymbol{G} & -\boldsymbol{A}^{\mathrm{T}} \\ \boldsymbol{A} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_C \\ \boldsymbol{v}_L \end{bmatrix} = \begin{bmatrix} \boldsymbol{B}_1 \\ \boldsymbol{B}_2 \end{bmatrix} \tag{20}$$

where $\boldsymbol{A}$ is the incidence matrix with only 1 and -1 as its nonzero elements. Assume that the nodes are ordered appropriately that $\boldsymbol{A}$ takes the following form

$$\boldsymbol{A} = \begin{bmatrix} 1 & & & \\ -1 & \ddots & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix}.$$

Then

$$\boldsymbol{A}^{-1} = \begin{bmatrix} 1 & & & \\ \vdots & \ddots & & \\ \vdots & & \ddots & \\ 1 & \cdots & \cdots & 1 \end{bmatrix}$$

and

$$\begin{bmatrix} \boldsymbol{G} & -\boldsymbol{A}^{\mathrm{T}} \\ \boldsymbol{A} & \boldsymbol{0} \end{bmatrix}^{-1} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{A}^{-1} \\ -\boldsymbol{A}^{-\mathrm{T}} & \boldsymbol{A}^{-\mathrm{T}} \boldsymbol{G} \boldsymbol{A}^{-1} \end{bmatrix}.$$

The circuit model (20) represents RCL models for interconnects, buses, and discretized networks for power ground analysis.

However, one can also easily find circuit examples that the symbolic inversion of a matrix is computationally intensive. This can be illustrated by the inversion of a tridiagonal or Hessenberg matrix, which arises from a resistor ladder circuit as shown in Fig. 17. The inverse of G is a dense matrix and the symbolic expression for each entry of G^{-1} is rather complicated.

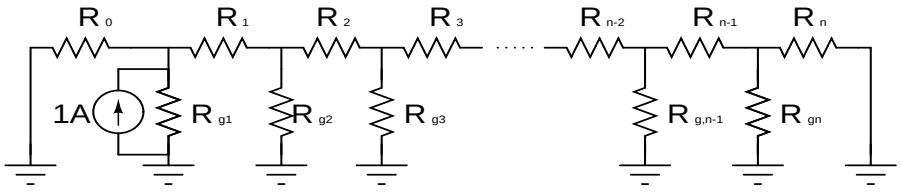


Figure 17: A resistor ladder circuit.

Determinant decision diagrams (DDDs) are means for finding the cofactors of a matrix [19]. It is a potential tool for finding the symbolic inverse of a highly sparse matrix. The key problem is to find an efficient way to manage the memory in the process of symbolic Krylov subspace generation, which will be a subject for future research.

References

[1] L. Silveira, I. Elfadel, J. White, M. Chilukuri, and K. Kundert, "Efficient frequency-domain modeling and circuit simulation of transmission lines," *IEEE Trans. on Components, Packaging, and Manufacturing Technology – Part B*, vol. 17, no. 4, pp. 505–513, Nov. 1994.

[2] L. Silveira, M. Kamon, and J. White, "Efficient reduced-order modeling of frequency-dependent coupling inductances associated with 3-D interconnect structures," *IEEE Trans. on Component, Packaging, and Manufacturing Technology – Part B*, vol. 19, no. 2, pp. 283–288, May 1996.

[3] L. Pillage and R. Rohrer, "Asymptotic waveform evaluation for timing analysis," *IEEE Trans. on Computer-Aided Design*, vol. 9, pp. 352–366, April 1990.

[4] J. Roychowdhury, "Automated macromodel generation for electronic systems," in *IEEE Int'l Behavioral Modeling and Simulation (BMAS) Workshop*, San Jose, CA, 2003, pp. 11–16.

[5] E. Hung and S. Senturia, "Generating efficient dynamical models for MEMS systems from a few finite-element simulation runs," *IEEE Journal of MEMS*, vol. 8, no. 3, pp. 280–289, 1999.

[6] P. Feldmann and R. Freund, "Efficient linear circuit analysis by Padé approximation via the Lanczos process," *IEEE Trans. on Computer-Aided Design*, vol. 14, no. 5, pp. 639–649, May 1995.

[7] A. Odabasioglu, M. Celik, and L. Pileggi, "PRIMA: Passive reduced-order interconnect macromodeling algorithm," *IEEE Trans. on Computer-Aided Design*, vol. 17, no. 8, pp. 645–654, August 1998.

[8] Z. Bai, "Krylov subspace techniques for reduced-order modeling of large-scale dynamical systems," *Applied Numerical Mathematics*, vol. 43, pp. 9–44, 2002.

[9] L. Daniel, C. S. Ong, S. C. Low, K. H. Lee, and J. White, "Geometrically parameterized interconnect performance models for interconnect synthesis," in *Proc. Int'l Symosium on Physical Design*, San Diego, CA, 2002, pp. 202–207.

[10] Y. Liu, L. Pileggi, and A. Strojwas, "Model order-reduction of RC(L) interconnect including variational analysis," in *Proc. 36th Design Automation Conference*, New Orleans, LA, 1999, pp. 201–206.

[11] G. Shi and C.-J. Shi, "Parametric reduced order modeling for interconnect analysis," in *Proc. Asia South Pacific Design Automation Conference (ASPDAC)*, Yokohama, Japan, 2004, pp. 774–779.

[12] B. Hu, G. Shi, and C.-J. R. Shi, "Symbolic model order reduction," in *Proc. IEEE Int'l Workshop on Behavioral Modeling and Simulation (BMAS)*, San Jose, CA, 2003, pp. 34–40.

[13] E. Grimme, "Krylov projection methods for model reduction," Ph.D. dissertation, ECE Dept., University of Illinoise at Urbana-Champaign, 1997.

[14] C.-W. Ho, A. E. Ruehli, and P. A. Brennan, "The modified nodal approach to network analysis," *IEEE Trans. on Circuits and Systems*, vol. CAS-22, no. 6, pp. 504–509, 1975.

[15] T. Kohonen, *Self-Organization and Associative Memory.* Berlin, Germany: Springer-Verlag, 1988.

[16] Y. Ismail and E. Friedman, "Effects of inductance on the propagation delay and repeater insertion in VLSI circuits," *IEEE Trans. on Very Large Scale Integration Systems*, vol. 8, no. 2, pp. 195–206, April 2002.

[17] E. Chiprout and M. Nakhla, "Analysis of interconnect networks using complex frequency hopping (CFH)," *IEEE Trans. on Computer-Aided Design*, vol. 14, no. 2, pp. 186–200, February 1995.

[18] G. Shi and C.-J. Shi, "Model order reduction by dominant subspace projection: Error bound, subspace computation and circuit application," *IEEE Trans. on Circuits and Systems – Part I*, 2005, in press.

[19] C.-J. Shi and X.-D. Tan, "Canonical symbolic analysis of large analog circuits with determinant decision diagrams," *IEEE Trans. on Computer-Aided Design*, vol. 19, no. 1, pp. 1–18, January 2000.