Efficient Model Reduction of Interconnects Via Double Gramians Approximation *

Boyuan Yan†, Sheldon X.-D. Tan†, Gengsheng Chen‡ and Yici Cai§
†Department of Electrical Engineering, University of California, Riverside, CA 92521
‡State Key Laboratory of ASIC & System, Fudan Univeristy, Shanghai, China, 200433
§Computer Science Dept., Tsinghua University, Beijing, China, 100084

ABSTRACT

The gramian approximation methods have been proposed recently to overcome the high computing costs of classical balanced truncation based reduction methods. But those methods typically gain efficiency by projecting the original system only onto one dominant subspace of the approximate system gramian (for instance using only controllability gramian). This single gramian reduction method can lead to large errors as the subspaces of controllability and observability can be quite different for general interconnects with unsymmetric system matrices. In this paper, we propose a fast balanced truncation method where the system is balanced in terms of two approximate gramians as achieved in the classical balanced truncation method. The novelty of the new method is that we can keep the similar computing costs of the single gramian method. The proposed algorithm is based on a generalized SVD-based balancing scheme such that the dominant subspace of the approximate gramian product can be obtained in a very efficient way without explicitly forming the gramians. Experimental results on a number of published benchmarks show that the proposed method is much more accurate than the single gramian method with similar computing costs.

1. INTRODUCTION

Model order reduction (MOR) is an efficient technique to reduce the interconnect complexity at small errors. Existing methods mainly project the original system onto a subspace. These approaches are divided into two broad categories, namely moment matching based methods (Krylov subspace methods) and balanced truncation based methods. In the former case, the system is projected onto a subspace to match dominant moments while in the latter case the system is projected onto a subspace both easily controllable and easily observable.

Moment matching based approaches have been a great success in the past due to its efficiency and scalability [4,5,7, 13,19]. Due to the introduction of Krylov subspace [4,7,19], implicit moment matching can be performed in a projection framework with very good numerical stability. As a result, moment matching idea [17] can be applied to very large-scale problems. Another reason for the success is that, when applied to a special class of systems (like interconnect circuits in MNA formulation), those methods can be modified to preserve passivity [5, 13].

While suitable for reduction of large-scale circuits, Krylov subspace methods do not offer a *priori* error bound and may generate models not as compact as desired. Therefore, another approach, truncated balanced realization (TBR), or balanced truncation (BT), which has been well developed in the control community [1, 6, 8, 12], has been studied intensively recently [9–11, 14–16, 18, 20].

The classical balanced truncation approaches [8, 12] produce nearly optimal models with controllable a *priori* global error bound [6]. However, those classical methods are too expensive to directly apply to large-scale problems due to the cubic cost to solve two Lyapunov equations. There has been significant effort devoted to mitigate this difficulty recently, which has led to two classes of approximate balanced truncation methods. The first class is based on the approximate balancing by iterative low rank solution of Lyapunov equation [9–11, 20, 22]. The second class is based on the low rank gramian approximation [15, 16, 18].

The low rank gramian approximation methods were proposed in [15, 21]. In [21], both controllability and observability gramians are computed in a sampling based method and eigendecomposition is performed on the product of the two gramians to compute the projection matrix, which is still very expensive. In [15], which is called the PMTBR method, only controllability gramian is computed in a similar sampling way and the system is reduced by projecting onto the approximate dominant controllable subspace only, which can be obtained much more cheaply by using singular value decomposition (SVD). Although no rigorous global error bounds exist as the classical method, those methods often produce a better approximation over a wide frequency range than Krylov subspace methods at the similar cost.

As shown in [20], considering only one gramian can lead to large errors as both controllability and observability gramians and their corresponding subspaces can be quite different for general interconnect circuits. Considering both gramians requires eigendecomposition of the product of two gramians, which is a $n \times n$ matrix, where n is the size of the problem. Hence the resulting method will have the similar computational cost, which is $O(n^3)$, of solving Lyapunov equations in the standard TBR method.

In this paper, we propose a fast balanced truncation method where the system is balanced in terms of two approximate gramians as achieved in the classical balanced truncation method. The novelty of the new method is that we can keep the similar computing costs of the single gramian method. The proposed algorithm is based on a generalized SVD-based balancing scheme, which is the extension of the classical balanced truncation method [8], where the balancing transformation is determined through the SVD of the product of Cholesky factors of gramians without explicitly forming the gramian product. In the proposed method, instead of Cholesky factors, different factors are applied such that the dominant invariant subspace of the approximate gramian product can be obtained in a very efficient way without explicitly forming the gramians. Experimental results on a number of published benchmarks show that the proposed method is much more accurate than the single gramian approximation method at the similar computing cost.

*This work is funded by NSF grant under No. CCF-0448534, in part by NSF grant under No. OISE-0929699, in part by National Natural Science Foundation of China (NSFC) grant under No. 60828008

978-1-4244-5766-3/10/$26.00 ©2010 IEEE

2. BACKGROUND

2.1 Classic balanced truncation

Classical balanced truncation was introduced to the system and control society by [12]. Given a stable minimal linear time invariant (LTI) system in standard state-space form (A, B, C, D)

$$\begin{array}{rcl} \dot{x}(t) & = & Ax(t) + Bu(t) \\ y(t) & = & Cx(t) + Du(t) \end{array} \quad (1)$$

where $A \in R^{n\times n}$, $B \in R^{n\times p}$, $C \in R^{q\times n}$, $y(t) \in R^q$, $u(t) \in R^p$, the controllability gramian X and observability gramian Y are the unique symmetric positive definite solutions to the Lyapunov equations

$$\begin{array}{l} AX + XA^T + BB^T = 0 \\ A^TY + YA + C^TC = 0 \end{array} \quad (2)$$

and they are related to the energy demanded to control and observe the system, respectively. Notice that solving the two Lyapunov equations takes $O(n^3)$ even for sparse matrices, which makes the BT methods not suitable for large scale problems.

Given a dynamic system, the state-space representation is not unique. Any nonsingular linear transformation $x = T\tilde{x}$ can be applied to the system (A, B, C, D) to obtain a new state-space representation $(\tilde{A}, \tilde{B}, \tilde{C}, D)$

$$\begin{array}{rcl} \dot{\tilde{x}}(t) & = & \tilde{A}\tilde{x}(t) + \tilde{B}u(t) \\ y(t) & = & \tilde{C}\tilde{x}(t) + Du(t) \end{array} \quad (3)$$

where

$$\tilde{A} = T^{-1}AT \quad \tilde{B} = T^{-1}B \quad \tilde{C} = CT \quad (4)$$

Such a transformation is known as similarity transformation, which does not change the input-output behavior of the system. It is easy to see both representations (A, B, C, D) and $(\tilde{A}, \tilde{B}, \tilde{C}, D)$ have the same transfer function $H(s)$.

A *balanced* realization is a special state-space representation, where the controllability and observability gramians are diagonal and equal. The balancing transformation can be computed by calculating the eigenmodes of the gramian product XY

$$XY = T\Lambda T^{-1} \quad (5)$$

It can be seen that the eigenvetors of XY are the basis vectors that describe the balancing transformation as follows. It is easy to show the gramians of the transformed system have the following expressions

$$\tilde{X} = T^{-1}XT^{-T} \quad \tilde{Y} = T^TYT \quad (6)$$

For a balanced system, we require $\tilde{X} = \tilde{Y} = \Sigma$, where Σ is a diagonal matrix. From (6), we can write

$$T^{-1}X = \Sigma T^T \quad YT = T^{-T}\Sigma \quad (7)$$

or

$$T^{-1}XYT = \Sigma^2 \quad (8)$$

which means the transformation T that balances the system is that containing the eigenvectors of the gramian product XY as its columns.

From the gramian expression (5) and (8), it can be seen that the eigenvalues λ_i contained in the diagonal matrix Λ are positive real numbers, and $\sigma_i = \sqrt{\lambda_i}$ are known as the Hankel singular values of the system. The eigenvectors of XY correspond to states through which the input is transmitted to the output. The magnitudes of the Hankel singular values describe the relative importance of these states and are independent of the particular realization of the system.

Therefore, a general idea of balanced truncation is to transform the system into a balanced form $(\tilde{A}, \tilde{B}, \tilde{C}, D)$, where the states which are difficult to control are also difficult to observe, and to discard the parts of the dynamics that correspond to small Hankel singular values. This truncation leads to a balanced reduced-order system (A_r, B_r, C_r, D). The error in the transfer function of the order r approximation is bounded by [6]

$$\|H(s) - H_r(s)\| = 2\sum_{i=r+1}^{n} \sigma_k \quad (9)$$

2.2 Gramian approximation method

To mitigate high computing costs of classic BT methods for solving large problems in VLSI design, gramian approximation methods have been proposed, where the approximate dominant subspace of a gramian can be obtained in a very efficient way.

2.2.1 *Gramian expression in frequency domain*

Given the state-space model (A, B, C, D), in frequency domain, the controllability gramian X can be computed from the expression

$$X = \frac{1}{2\pi}\int_{-\infty}^{+\infty} (j\omega I - A)^{-1}BB^T(j\omega I - A)^{-H}dw \quad (10)$$

and the observability gramian Y can be computed from the expression

$$Y = \frac{1}{2\pi}\int_{-\infty}^{+\infty} (j\omega I - A^T)^{-1}C^TC(j\omega I - A^T)^{-H}dw \quad (11)$$

where superscript H denotes Hermitian transpose.

2.2.2 *Single gramian approximation*

Willcox proposed the sampling-based method to compute the approximate gramians to avoid the Lyapunov equations [21]. But the method can not avoid the expensive eigendecomposition of the gramian product. To mitigate this problem, Phillips proposed single gramian approximation method (PMTBR) [16], where the approximate dominant subspace of controllability gramian (10) can be obtained in a very efficient way.

Specifically, let ω_k be kth sampling point. If we define

$$z_{c_k} = (j\omega_k I - A)^{-1}B \quad (12)$$

then X can be approximated as

$$\hat{X} = \sum w_k z_{c_k} z_{c_k}^H = Z_c W^2 Z_c^H \quad (13)$$

where $Z_c = [z_{c_1}, z_{c_2}, \ldots, z_{c_N}]$ and W a diagonal matrix with diagonal entries $W_{kk} = \sqrt{w_k}$. The weight w_k may come from a specific numerical quadrature method. In fact, we can set $w_k = 1$. Since $\hat{X}$ is symmetric, it is orthogonally diagonalizable. If we perform the SVD on

$$Z_c = U\Sigma V^T \quad (14)$$

then we have

$$\hat{X} = Z_c {Z_c}^H = U\Sigma^2U^T = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1^2 & 0 \\ 0 & \Sigma_2^2 \end{bmatrix} \begin{bmatrix} U_1^T \\ U_2^T \end{bmatrix} \quad (15)$$

where $U^TU = I$. If the quadrature rule is accurate, $\hat{X}$ will converge to X and the dominant eigenspace of $\hat{X}$ converges to the dominant eigenspace of X by perturbation analysis of invariant subspaces. As a result, the dominant eigenvectors U_1 can be used as the projection matrix and the reduced model (A_r, B_r, C_r, D) can be obtained as

$$A_r = U_1^TAU_1 \quad B_r = U_1^TB \quad C_r = CU_1 \quad (16)$$

Note that the method does not need to form the $n \times n$ gramian $Z_c Z_c^H$ explicitly and neither does it need to perform the eigenvalue decomposition at the cost of $O(n^3)$. Instead, it only performs the SVD on a $n \times Np$ thin matrix Z_c instead ($Np \ll n$).

Given N sampling points and p inputs, the cost of SVD on matrix $Z_{n \times Np}$ is $O(n(Np)^2)$. In addition, it takes N matrix factorizations and Np matrix solves. The total cost is $O(n(Np)^2 + Nn^\beta + Npn^\alpha)$ (typically, $1.1 \leq \beta \leq 1.5$ and $1 \leq \alpha \leq 1.2$ for circuits) [16], which is dominated by $O(Nn^\beta) < O(n^2)$.

Note that, the same procedure can be performed based on observability gramian (11) as well if we define $z_{o_k} = (j\omega_k I - A^T)^{-1} C^T$ and perform an SVD on $Z_o = [z_{o_1}, \ldots, z_{o_N}]$.

2.2.3 Existing problems

PMTBR only uses controllability gramian and it works well for symmetric systems where both controllability and observability gramians are the same. But it may not work well for general unsymmetrical systems like RLC interconnect systems as shown in [20].

To consider two gramians, the most straightforward way is to compute $\hat{X} = Z_c Z_c^H$ and $\hat{Y} = Z_o Z_o^H$ respectively and perform an eigendecomposition on the product $\hat{X}\hat{Y}$ [21]. However, the eigendecomposition has to be performed on a $n \times n$ full matrix $\hat{X}\hat{Y}$, which still has the computational order of $O(n^3)$.

Another possible way is to use cross-gramian X_{CG}, which contains both controllability and observability information in a single matrix. In the frequency domain, X_{CG} is expressed as

$$X_{CG} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} (j\omega I - A)^{-1} BC (j\omega I - A)^{-1} dw \qquad (17)$$

which can be approximated as $\hat{X}_{CG} = \sum z_{c_k} z_{o_k}^H = Z_c Z_o^H$. In this case, we do not need to compute $\hat{X}$, $\hat{Y}$, and $\hat{X}\hat{Y}$. However, to determine the dominant subspace of X_{CG}, we still need to perform an eigendecomposition on a $n \times n$ full matrix $Z_c Z_o^H$.

3. NEW DOUBLE GRAMIAN APPROXIMATION METHOD

In this section, we present the new double gramians approximation method.

3.1 Classical SVD-based balancing algorithm

In classical balanced truncation [12], the most straightforward way to determine the balancing transformation T is to perform an eigendecomposition of the gramian product XY and T is the invariant subspace.

However, in practice, an SVD-based method was proposed in [8], where the balancing transformation T is determined through computing the singular value decomposition (SVD) of a certain product of matrices without explicitly forming the gramian product. The algorithm is shown in *Algorithm* 1.

In this algorithm, given the controllability gramian $X > 0$ and observability gramian $Y > 0$, the Cholesky factors are computed first. Let L_c and L_o denote the lower triangular Cholesky factors of the gramians X and Y

$$X = L_c L_c^T \quad Y = L_o L_o^T \qquad (18)$$

Then the singular value decomposition of the product of the Cholesky factors is computed as

$$L_o^T L_c = U \Sigma V^T \qquad (19)$$

where $U^T U = I$ and $V^T V = I$. The balancing transformation T and T^{-1} are given as

$$T = L_c V \Sigma^{-1/2} \quad T^{-1} = \Sigma^{-1/2} U^T L_o^T \qquad (20)$$

Therefore, instead of explicitly forming the gramian product XY and performing an eigendecomposition, the invariant subspace T can be determined in a less expensive and more efficient way in the classical balanced truncation algorithm.

Algorithm 1: Classical SVD-based balancing method

Input: $H : (A, B, C, D)$
Output: $H_r : (A_r, B_r, C_r, D)$

1. Compute $X > 0$ and $Y > 0$
2. Cholesky factorization $X = L_c L_c^T$ and $Y = L_o L_o^T$
3. Compute SVD of $U\Sigma V = L_o^T L_c$
4. Compute $T = L_c V \Sigma^{-1/2}$ and $T^{-1} = \Sigma^{-1/2} U^T L_o^T$
5. Compute the balanced realizations $\tilde{A} = T^{-1} A T$,$\tilde{B} = T^{-1} B$,$\tilde{C} = CT$
6. Truncate to form the reduced system (A_r, B_r, C_r, D)

3.2 Generalized SVD-based balancing algorithm

In this paper, we propose a generalized SVD-based balancing scheme. In the classical SVD-based balancing algorithm, the Cholesky factors are used. However, we show that Cholesky factors are not the only choice to compute the balancing transformation. In fact, there is no restriction on the structure of matrix factors at all. We first show the following result:

THEOREM 1. *Assume the gramians X and Y can be factorized as*

$$X = Z_c Z_c^T \quad Y = Z_o Z_o^T \qquad (21)$$

where Z_c and Z_o are matrix factors with arbitrary structure. Then the singular value decomposition of the product of the factors is computed as

$$Z_o^T Z_c = U \Sigma V^T \qquad (22)$$

where $U^T U = I$ and $V^T V = I$.

In this case, the balancing transformation T and T^{-1} are given as

$$T = Z_c V \Sigma^{-1/2} \quad T^{-1} = \Sigma^{-1/2} U^T Z_o^T \qquad (23)$$

PROOF. Given (21) and (22), it can be shown that T^{-1} is an inverse matrix of T in (23)

$$\begin{aligned} T^{-1}T &= (\Sigma^{-1/2} U^T Z_o^T)(Z_c V \Sigma^{-1/2}) \\ &= \Sigma^{-1/2} U^T (Z_o^T Z_c) V \Sigma^{-1/2} \\ &= \Sigma^{-1/2} U^T (U \Sigma V^T) V \Sigma^{-1/2} \\ &= \Sigma^{-1/2} (U^T U) \Sigma (V^T V) \Sigma^{-1/2} \\ &= \Sigma^{-1/2} \Sigma \Sigma^{-1/2} \\ &= I \end{aligned} \qquad (24)$$

and T is exactly the invariant subspace of gramian product XY (8)

$$\begin{aligned} T^{-1}XYT &= (\Sigma^{-1/2} U^T Z_o^T)(XY)(Z_c V \Sigma^{-1/2}) \\ &= (\Sigma^{-1/2} U^T Z_o^T)(Z_c Z_c^T)(Z_o Z_o^T)(Z_c V \Sigma^{-1/2}) \\ &= \Sigma^{-1/2} U^T (Z_o^T Z_c)(Z_c^T Z_o)(Z_o^T Z_c) V \Sigma^{-1/2} \\ &= \Sigma^{-1/2} U^T (U \Sigma V^T)(V \Sigma U^T)(U \Sigma V^T) V \Sigma^{-1/2} \\ &= \Sigma^{-1/2} \Sigma \Sigma \Sigma \Sigma^{-1/2} \\ &= \Sigma^2 \end{aligned} \qquad (25)$$

□

Therefore, the factors of gramians are no longer limited to Cholesky factors.

3.3 New double gramians approximation method

Now, we apply the proposed generalized SVD-based balancing algorithm for double gramians approximation.

Let ω_k be kth sampling point. If we define

$$\begin{aligned} z_{c_k} &= (j\omega_k I - A)^{-1} B \\ z_{o_k} &= (j\omega_k I - A^T)^{-1} C^T \end{aligned} \tag{26}$$

then X and Y can be approximated as

$$\begin{aligned} \hat{X} &= \sum z_{c_k} z_{c_k}^H = Z_c Z_c^H \\ \hat{Y} &= \sum z_{o_k} z_{o_k}^H = Z_o Z_o^H \end{aligned} \tag{27}$$

where $Z_c = [z_{c_1}, z_{c_2}, \ldots, z_{c_N}]$ and $Z_o = [z_{o_1}, z_{o_2}, \ldots, z_{o_N}]$.

According to the generalized SVD-based balancing algorithm, Z_c and Z_o can be used as the factors to compute the balancing transformation T. We perform a singular value decomposition on

$$Z_o^H Z_c = U \Sigma V^H \tag{28}$$

Notice that, given N sampling points and p inputs and q outputs, Z_c and Z_o are $n \times Np$ and $n \times Nq$ matrices, respectively, where $Np \ll n$ and $Nq \ll n$. Assume $m = max(Np, Nq)$, the dimension of matrix $Z_o^T Z_c$ is smaller than $m \times m$ and the cost of SVD is $O(m^3)$, which is much smaller than $O(n^3)$ when $m \ll n$. Then we have

$${Z_o}^H Z_c = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1^2 & 0 \\ 0 & \Sigma_2^2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \tag{29}$$

where U_1 and V_1 are dominant subspace corresponding to the first r largest singular values Σ_1. Then the right projection matrix and left projection matrix are given by

$$T_r = Z_c V_1 \Sigma_1^{-1/2} \quad T_l = \Sigma_1^{-1/2} U_1^T Z_o^T \tag{30}$$

T_r and T_l are $n \times r$ and $r \times n$ matrices respectively, corresponding to the dominant invariant subspace of the approximate gramian product $\hat{X}\hat{Y}$. The reduced model (A_r, B_r, C_r, D) can be obtained as

$$A_r = T_l A T_r \quad B_r = T_l B \quad C_r = C T_r \tag{31}$$

The resulting double gramians approximation algorithm, called DGA, is given in *Algorithm* 2.

Given N sampling points and p inputs and q outputs, assume $m = max(Np, Nq)$, the cost of SVD on matrix $Z_o^T Z_c$ is $O(m^3)$ as analyzed before. In addition, it takes $2N$ matrix factorizations and $pN + qN$ matrix solves to obtain Z_c and Z_o, and $O(pqn)$ to obtain $Z_o^T Z_c$.

The total cost is $O(m^3 + 2Nn^\beta + (p+q)Nn^\alpha + pqn)$ (typically, $1.1 \le \beta \le 1.5$ and $1 \le \alpha \le 1.2$ for circuits). As $m \ll n$, the cost is still dominated by $O(2Nn^\beta)$, which is about twice the cost of single gramian approximation method PMTBR [16] but still in the same growth order. In fact, the cost of DGA is less than twice the cost of PMTBR as the cost of SVD process in DGA is $O(m^3)$, which is much less than $O(n(Np)^2)$ in PMTBR, where $Np \le m$ and $m \ll n$.

Practically, we notice that PMTBR uses incremental QR on Z_c to find the project matrix. But it will have the same computational costs of SVD on Z_c. Incremental QR can be used for the proposed method on $Z_o^T Z_c$ also.

We want to emphasize that although the proposed method is slower than the PMTBR as DGA computes two approximate gramians, *it does not mean that PMTBR will achieve the same accuracy of the proposed method if both have the exactly same computing costs (for instance PMTBR samples twice of the DGA method).* The proposed method addresses the fundamental problem of *using only single gramians* in the PMTBR-like method.

3.4 Practical Implementation

Algorithm 2: New double gramians approximation method (DGA)

Input: $H : (A, B, C, D)$
Output: $H_r : (A_r, B_r, C_r, D)$

1. Select N sampling points $z_k (k = 1, 2, \ldots, N)$
2. Compute $Z_c = [z_{c_1}, z_{c_2}, \ldots, z_{c_N}]$ where $z_{c_k} = (j\omega_k I - A)^{-1} B$
3. Compute $Z_o = [z_{o_1}, z_{o_2}, \ldots, z_{o_N}]$ where $z_{o_k} = (j\omega_k I - A^T)^{-1} C^T$
4. Perform SVD on matrix ${Z_o}^H Z_c = U \Sigma V^T$
5. Compute right projection matrix T_r and left projection matrix T_l as $T_r = Z_c V_1 \Sigma_1^{-1/2}$ and $T_l^T = \Sigma_1^{-1/2} U_1^T Z_o^T$
6. Project onto the dominant invariant subspace of the approximate gramian product $A_r = T_l A T_r, B_r = T_l B, C_r = C T_r$

3.4.1 Descriptor systems

A special class of dynamic systems is the RLC interconnect circuit described by state-space equations in descriptor form

$$\begin{aligned} C\dot{x}(t) &= -Gx(t) + Bu(t) \\ y(t) &= Lx(t) \end{aligned} \tag{32}$$

In this case, the controllability gramian and observability gramian can be obtained from generalized Lyapunov equations [16, 20] and the frequency domain expressions are [16]

$$\begin{aligned} X &= \tfrac{1}{2\pi} \textstyle\int_{-\infty}^{+\infty} (j\omega C + G)^{-1} B B^T (j\omega C + G)^{-H} dw \\ Y &= \tfrac{1}{2\pi} \textstyle\int_{-\infty}^{+\infty} (j\omega C + G^T)^{-1} L^T L (j\omega C + G^T)^{-H} dw \end{aligned} \tag{33}$$

As a result, z_{c_k}'s and z_{o_k}'s in *Algorithm* 2 will be replaced by

$$z_{c_k} = (j\omega_k C + G)^{-1} B \quad z_{o_k} = (j\omega_k C + G^T)^{-1} L^T \tag{34}$$

Just as PMTBR, the complications present in applying standard balanced truncation to problems with singular descriptor matrix C vanish in the proposed method.

3.4.2 Passivity

Similar to classical balanced truncation [8], the proposed double gramians approximation method does not preserve the passivity for general dynamical systems. But post-passivity-enforcement process can be carried out to ensure the passivity [3, 20], which is out of scope of this paper.

4. EXPERIMENTAL RESULTS

In this section, we present experimental results on four benchmark examples used in published papers, which are NOT symmetric. The proposed double gramians approximation method, called DGA, is compared with existing single gramian approximation method PMTBR and Krylov subspace method PRIMA at the same reduced order. Note that, the sampling points (total number and their locations) for both DGA and PMTBR are exactly the same.

4.1 Two RLC lines

The first and second examples are RLC lines used in [20] (in Fig.2 and Fig.4). The two RLC lines are of the same order 1502 but with different topologies. In both examples, input signal $u(t)$ is the voltage at the first node and output is the current flowing through the voltage source. The state vector consists of node voltages, inductor currents, and currents through the voltage source. The MNA formulation for the two lines results in two systems (G, C, B, L) with unsymmetric matrices G. In both examples, the parameters are $R = 0.1, L = 2, C = 15$ and the reduced orders are set to be 10.

Table 1: Reduction CPU time comparison of PMTBR and DGA (seconds).

n	152	1520	150002
DGA	0.014597	0.06298	8.810724
PMTBR	0.008164	0.03228	5.070467

The results for the first and second RLC lines are shown in Fig. 1(above) and Fig. 1(below), respectively. Clearly, we see that DGA is much more accurate that PMTBR and PRIMA in both examples. The reason why PRIMA's results are quite off is that it approximates only dominant controllable states [20]. We then compare the CPU time of DGA and

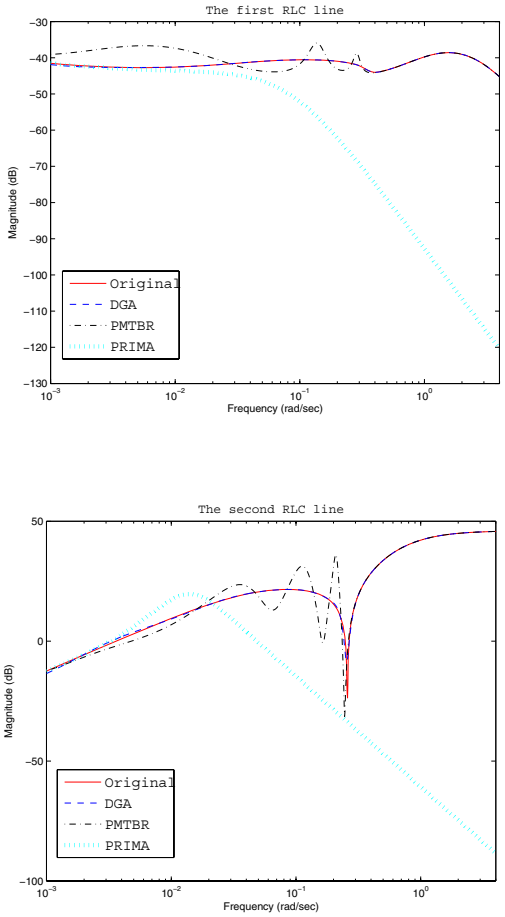

Figure 1: Comparison results on the two RLC line examples.

PMTBR. The reduction CPU times are shown in Table 1, where the n is the order of the RLC line and the reduced order is 10. From Table 1, we can see, the reduction time of DGA is less than twice the reduction time of PMTBR. As analyzed in Section 3, this is because DGA is much less expensive than PMTBR in the SVD process although DGA has to take twice matrix factorizations and solving.

4.2 A transmission line model

The third example is a transmission line model from [2, 10], which is not symmetric and has an order of 256. There are 2 inputs and 2 outputs in this model, which results in a 2×2 transfer matrix $H(s)$. The reduced orders are 30.

The results for the diagonal terms $h_{11}(s)$ and $h_{22}(s)$ of $H(s)$ are shown in Fig. 2(above) and Fig. 2(below), respectively. For this unsymmetric example, the proposed DGA method produces the best wideband approximation again. PRIMA is very accurate in low frequency range but not accurate beyond 10^{11} rad/sec. DGA is much more accurate than PMTBR also in this case.

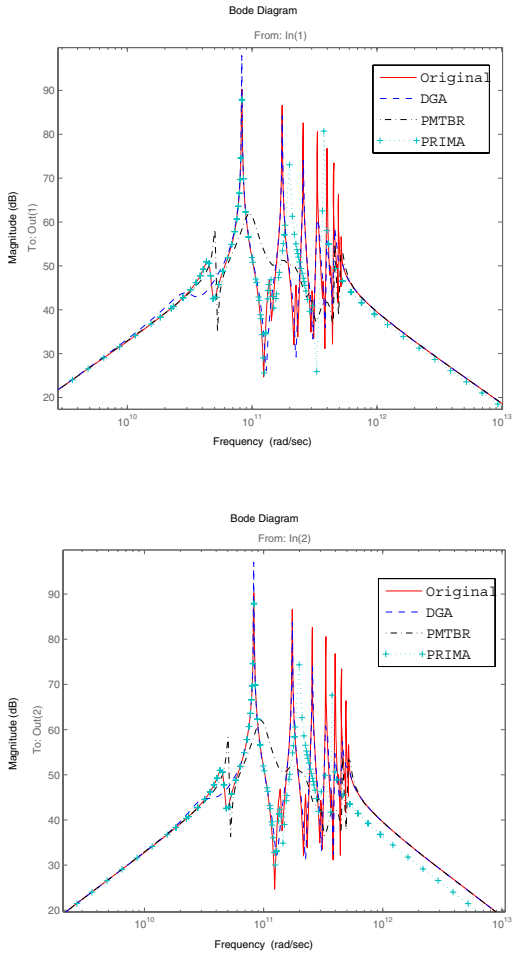

Figure 2: Comparison result on the transmission line example.

4.3 An international space station example

In addition to interconnect modeling, model reduction is being used to generate compact models of various dynamic systems [20]. The forth example is a structural model of component 1r (Russian service module) of the International Space Station from [2]. This example is also not symmetric and has an order of 270. There are 3 inputs and 3 outputs in this model, which results in a 3×3 transfer matrix $H(s)$. The reduced orders are 25.

The results for the diagonal term $h_{22}(s)$ and the off diagonal term $h_{32}(s)$ are shown in Fig. 3(above) and Fig. 3(below), respectively. The same conclusion can be drawn here. PRIMA is only accurate in low frequency range, which can match up to about 10 rad/sec for $h_{22}(s)$ and about 1 rad/sec for $h_{32}(s)$. However, DGA still has the excellent wideband accuracy and the performance is much better than PMTBR.

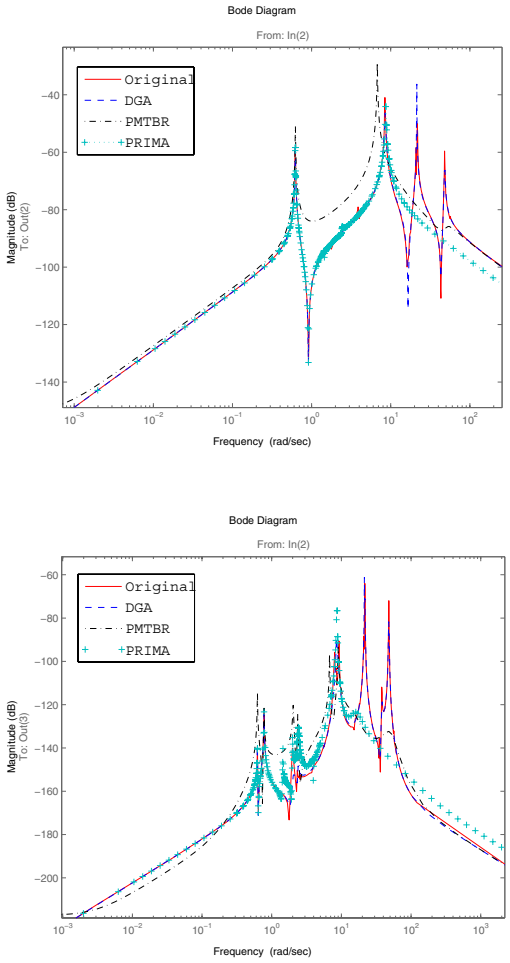

Figure 3: Comparison result on the ISS example.

5. CONCLUSION

In this paper, a generalized SVD-based balancing scheme has been proposed to balance and reduce interconnect circuits modeled as unsymmetric systems, where the dominant controllable subspace and the dominant observable subspace do not coincide. We demonstrated that Cholesky factors are not the only choice to determine the balancing transformation in the classical SVD-based method. In the generalized balancing framework, different matrix factors are applied and the system can be balanced and truncated in terms of the product of two approximate gramians in an efficient way without explicitly computing the gramians. Experimental results on a number of published benchmarks show that the proposed method is much more accurate than the single gramian method at the similar computing costs.

6. REFERENCES

[1] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*. The Society for Industrial and Applied Mathematics (SIAM), 2005.

[2] Y. Chahlaoui and P. V. Dooren, "A collection of benchmark examples for model reduction of linear time invariant dynamical systems," SLICOT Working Note 2002-2, February 2002.

[3] C. P. Coelho, J. Phillips, and L. M. Silveira, "A convex programming approach for generating guaranteed passive approximations to tabulated frequency-data," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 23, no. 2, pp. 293–301, Feb. 2004.

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

[5] R. W. Freund, "SPRIM: structure-preserving reduced-order interconnect macromodeling," in *Proc. Int. Conf. on Computer Aided Design (ICCAD)*, 2004, pp. 80–87.

[6] K. Glover, "All optimal Hankel-norm approximations of linear multi-variable systems and their L_∞ error bounds"," *Int. J. Control*, vol. 36, pp. 1115–1193, 1984.

[7] E. J. Grimme, *Krylov projection methods for model reduction (Ph.D. Thesis)*. University of Illinois at Urbana-Champaign, 1997.

[8] A. J. Laub, M. T. Heath, C. C. Paige, and R. C. Ward, "Computation of system balancing transformations and other applications of simultaneous diagonalization algorithms," *IEEE Trans. Automat. Contr.*, vol. 32, pp. 115–122, 1987.

[9] J. R. Li, *Model reduction of large linear systems via low rank system gramians (Ph.D. Thesis)*. MIT, 2002.

[10] J. R. Li, F. Wang, and J. White, "An efficient Lyapunov equation-based approach for generating reduced-order models of interconnect," in *Proc. Design Automation Conf. (DAC)*, 1999, pp. 1–6.

[11] J. R. Li and J. White, "Efficient model reduction of interconnect via approximate system gramians," in *Proc. Int. Conf. on Computer Aided Design (ICCAD)*, 1999, pp. 380–383.

[12] B. Moore, "Principal component analysis in linear systems: Controllability, and observability, and model reduction," *IEEE Trans. Automat. Contr.*, vol. 26, no. 1, pp. 17–32, 1981.

[13] A. Odabasioglu, M. Celik, and L. Pileggi, "PRIMA: Passive reduced-order interconnect macromodeling algorithm," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, pp. 645–654, 1998.

[14] J. R. Phillips, L. Daniel, and L. M. Silveira, "Guaranteed passive balancing transformation for model order reduction," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 22, no. 8, pp. 1027–1041, 2003.

[15] J. R. Phillips and L. M. Silveira, "Poor man's TBR: a simple model reduction scheme," in *Proc. European Design and Test Conf. (DATE)*, 2004, pp. 938–943.

[16] ——, "Poor man's TBR: a simple model reduction scheme," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 24, no. 1, pp. 43–55, 2005.

[17] L. T. Pillage and R. A. Rohrer, "Asymptotic waveform evaluation for timing analysis," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, pp. 352–366, April 1990.

[18] L. M. Silveira and J. R. Phillips, "Exploiting input information in a model reduction algorithm for massively coupled parasitic networks," in *Proc. Design Automation Conf. (DAC)*, 2004, pp. 385–388.

[19] M. Silveira, M. Kamon, I. Elfadel, and J. White, "A coordinate-transformed Arnoldi algorithm for generating guaranteed stable reduced-order models of RLC circuits," in *Proc. Int. Conf. on Computer Aided Design (ICCAD)*, 1996, pp. 288–294.

[20] D. Vasilyev and J. White, "A more reliable reduction algorithm for behavioral model extraction," in *Proc. Int. Conf. on Computer Aided Design (ICCAD)*, 2005, pp. 813–820.

[21] K. Willcox and J. Peraire, "Balanced model reduction via the proper orthogonal decomposition," *AIAA Journal*, 2002.

[22] Y. Zhou, *Numerical methods for large scale matrix equations with applications in LTI system model reduction (Ph.D. Thesis)*. Rice University, 2002.