An Efficient Algorithm for Frequency-Weighted Balanced Truncation of VLSI Interconnects in Descriptor form

V.Vasudevan
Department of Electrical Engineering
Indian Institute of Technology-Madras
Chennai-600036, India
vinita@ee.iitm.ac.in

M.Ramakrishna
Department of Aerospace Engineering
Indian Institute of Technology-Madras
Chennai-600036, India
krishna@ae.iitm.ac.in

ABSTRACT

Balanced truncation of descriptor systems requires computation of spectral projectors and solution of the generalized projected Lyapunov equations, both of which have significant complexity. Frequency-weighted balancing methods are more efficient if the response over a specific frequency range is desired. However, a direct extension of these methods to descriptor systems requires the spectral projectors. In this paper, we propose an efficient frequency-weighted balanced truncation algorithm without finding the spectral projectors. Samples of the frequency-domain solution to the system are used to get an accurate estimate of the improper Gramians. The proper Gramians are computed after adjusting for the contribution of the improper subsystem. Low rank factors of these Gramians are used to obtain a basis that includes the contribution of both the proper and improper subsystems. Congruence transform is used to ensure passivity of RLC interconnect models. Results for standard benchmarks show that the method is accurate and efficient.

Categories and Subject Descriptors

B.7.2 [**Integrated Circuits**]: Design Aids

General Terms

Algorithms, Performance

Keywords

Balanced truncation, Descriptor systems

1. INTRODUCTION

Balanced truncation is an attractive method for obtaining reduced order models (ROMs), due to the existence of a global error bound and the guaranteed stability of the resulting ROM. Since the reduced order models are stable, it is also possible to use passivity enforcement techniques to make sure that the resulting ROMs are also passive. However, it requires the solution of the Lyapunov equations to obtain the Gramians, which is an $O(n^3)$ computation. The added complexity in the case of descriptor systems is that the solution of projected generalized continuous and discrete Lyapunov equations is required. This in turn requires the spectral projectors that project the pencil onto the left and right deflating subspaces corresponding to the finite and infinite eigenvalues. One technique to find the projectors is the canonical projector technique proposed by Marz [4] or the modified Marz method proposed in [15,16]. A more efficient technique to evaluate these projectors based on sparse LUQ decomposition has been proposed in [10–12]. It has complexity $O(n^2) - O(n^3)$ depending on the sparsity of the matrices. In some cases, explicit analytical expressions are available for the projectors [7–9, 13, 14]. Even so, its computation is fairly involved and requires the inverse of potentially large and dense matrices. Alternately, the generalized Schur-Hammarling method can be used to compute the Gramians [6], which also has $O(n^3)$ complexity.

An alternative technique to compute the Gramian is to numerically evaluate the corresponding integral in the frequency domain [1–3,17]. This leads to the frequency-weighted balancing methods, which are used when a good match to the transfer function is required over a specified frequency range. This method is computationally much more efficient when the system matrices are sparse, which is typically the case. It is also easily parallelizable. A drawback is that the resulting ROM is not guaranteed to be stable, although it is usually stable if the integral is approximated well. However, if the system matrices are definite matrices, which is the case for VLSI interconnects, a congruence transform can be used to get passive ROMs [1, 2]. In [1],low rank factors of the controllability Gramian is used for reduction of symmetric systems and the cross-Gramian for general systems. In [2, 3], both Gramians are computed and the corresponding Cholesky factors are used for the reduction. However, in both these works, although the authors indicate the method can be used for descriptor systems, the contribution of the improper subsystem is not explicitly taken into account. In the case of singular descriptor systems, for large enough frequencies, the solutions to the equations could tend to a constant or increase with frequency, so that the integral does not exist. A direct extension of methods in [1–3] to include these systems would once again require computation of the spectral projectors [6].

However, if we actually look at the role of these spectral projectors in estimating the Gramian, they essentially zero

DAC '15, June 7–11, 2015, San Francisco, California, USA.

ACM 978-1-4503-3520-1/15/06...$15.00
http://dx.doi.org/10.1145/2744769.2744770.

out the contribution of the improper (proper) subsystem to the Gramian corresponding to the proper (improper) subsystem. In this paper, we show that it is possible to obtain the Gramians in a computationally efficient manner, without explicitly computing the spectral projectors if we look at the problem in the frequency domain. We first obtain an accurate estimate of the contribution of the improper subsystem. The proper Gramians are computed after adjusting for the contribution of the improper subsystem. Low rank factors of these Gramians are used to obtain a basis that includes the contribution of both the proper and improper subsystems. Congruence transform is used to ensure passivity of RLC interconnect models. We demonstrate that the method gives accurate results for various benchmarks [18, 19].

2. BACKGROUND

Defining the system under study more precisely, let the differential equations describing the system be written as

$$\begin{aligned} E\frac{d\mathbf{x}}{dt} &= A\mathbf{x} + B\mathbf{u} \\ \mathbf{y} &= C\mathbf{x} + D\mathbf{u} \end{aligned} \tag{1}$$

E and A are $n \times n$ system matrices, $\mathbf{x}$ is the vector of unknowns and $\mathbf{u}$ and $\mathbf{y}$ represent the inputs to and outputs of the system. B is an $n \times m$ matrix corresponding to m inputs and C is an $q \times n$ matrix corresponding to q outputs. We assume that $\lambda E - A$ is a regular pencil and the system is asymptotically stable, i.e., all the finite eigenvalues lie in the open left half plane. In an asymptotically stable system, the matrix A must be nonsingular, but E could be either singular or non-singular. Balanced truncation of systems with non-singular E (state variable form) is a well studied problem. In this paper, we focus on systems for which E is singular i.e., descriptor systems. For these systems, the Gramians corresponding to both the proper (slow) and improper (fast) subsystems are required. Neglecting the improper subsystem could lead to erroneous results [5].

If the pencil is regular, it is possible to make a transformation to the Weierstrass canonical form i.e., there exist nonsingular matrices Q and T such that [6]

$$E = Q\begin{bmatrix} I_f & \mathbf{0} \\ \mathbf{0} & N \end{bmatrix} T \quad \text{and} \quad A = Q\begin{bmatrix} J & \mathbf{0} \\ \mathbf{0} & I_\infty \end{bmatrix} T \tag{2}$$

where I_f and I_∞ are identity matrices with sizes equal to the dimension of the deflating subspaces associated with the finite and infinite eigenvalues. The finite eigenvalues of the pencil are the eigenvalues of the block J and N is a nilpotent matrix. The index of nilpotency of N, denoted by ν, is the index of the system. The equations corresponding to the finite eigenvalues represent the proper part of the system and those corresponding to the infinite eigenvalues represent the improper subsystem. In the discussion that follows, the coordinate system corresponding to the canonical form is referred to as the canonical coordinate system.

If E is singular, the Gramians corresponding to the proper and improper subsystem are the solutions to the following continuous and discrete projected GLYAPs respectively [6].

$$\begin{aligned} EP_{pc}A^T + AP_{pc}E^T + S_lBB^TS_l^T = \mathbf{0}, \quad P_{pc} = S_rP_{pc} \\ E^TP_{po}A + A^TP_{po}E + S_r^TC^TCS_r = \mathbf{0}, \quad P_{po} = P_{po}S_l \end{aligned} \tag{3}$$

$$\begin{aligned} AP_{ic}A^T - EP_{ic}E^T = (I - S_l)BB^T(I - S_l)^T, \quad S_rP_{ic} = \mathbf{0} \\ A^TP_{io}A - E^TP_{io}E = (I - S_r)^TC^TC(I - S_r), \quad P_{io}S_l = \mathbf{0} \end{aligned} \tag{4}$$

where S_l and S_r are the spectral projectors onto the left and right deflating subspaces of the pencil corresponding to the finite eigenvalues, given by

$$S_l = Q\begin{bmatrix} I_f & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} Q^{-1} \quad \text{and} \quad S_r = T^{-1}\begin{bmatrix} I_f & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} T \tag{5}$$

The critical steps in computing the reduced order bases are the computation of the spectral projectors and solution of the projected GLYAPs, both of which have $O(n^3)$ complexity.

In the next two sections, we show that it is possible to perform this computation efficiently without explicitly computing the spectral projectors if we look at the problem in the frequency domain.

3. PROPOSED FREQUENCY WEIGHTED BALANCING METHOD: THEORY

If we look at the role of the spectral projectors S_l and S_r, they essentially project the system so that only the equations corresponding to the finite eigenvalues (proper part) of the system are retained. The improper part is essentially zeroed out. This means that in the time domain, the solutions P_{pc} and P_{po} to the projected GLYAPs can be written as [7]

$$P_{pc} = \int_0^\infty \mathcal{F}(t)BB^T\mathcal{F}^T(t)dt \tag{6}$$

$$P_{po} = \int_0^\infty \mathcal{F}(t)^TC^TC\mathcal{F}(t)dt \tag{7}$$

where

$$\mathcal{F}(t) = T^{-1}\begin{bmatrix} e^{tJ} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} Q^{-1} \tag{8}$$

Similarly, the projection operators used to obtain the improper Gramians are $I - S_l$ and $I - S_r$, which means the equations corresponding to the proper part are nulled out. Consequently, the solutions to the projected discrete-time GLYAPs are given by [7]

$$P_{ic} = \sum_{k=0}^{\nu-1} F_kBB^TF_k^T \tag{9}$$

$$P_{io} = \sum_{k=0}^{\nu-1} F_k^TC^TCF_k \tag{10}$$

where

$$F_k = T^{-1}\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -N^k \end{bmatrix} Q^{-1} \tag{11}$$

The corresponding frequency domain representation of the Gramians can be obtained as follows. Let $Q^{-1} = \begin{bmatrix} Q_{I_1}^T \\ Q_{I_2}^T \end{bmatrix}$ and $T^{-1} = \begin{bmatrix} T_{I_1} & T_{I_2} \end{bmatrix}$, where the matrices are partitioned consistently with (2). Define $K_1 = Q_{I_1}^TB$, $K_2 = Q_{I_2}^TB$. Substituting for $\mathcal{F}(t)$ and using Parseval's theorem, the Gramian

P_{pc} can be obtained from equation (6) as

$$P_{pc} = \frac{1}{2\pi} T_{I_1} \left[\int_{-\infty}^{\infty} (j\omega I - J)^{-1} K_1 K_1^T (j\omega I - J)^{-H} d\omega \right] T_{I_1}^T \tag{12}$$

However, we do not know the matrices Q, T and J, so that this integral cannot be evaluated directly.

To overcome this problem, we look at the system equations (1) in the canonical coordinate system, which can be written as

$$\begin{aligned} \frac{d\boldsymbol{\phi_1}}{dt} &= J\boldsymbol{\phi}_1 + K_1\mathbf{u} \\ \boldsymbol{\phi}_2 &= -K_2\mathbf{u}(t) - \sum_{i=1}^{\nu-1} N^i K_2 \mathbf{u}^{(i)}(t) \end{aligned} \tag{13}$$

where $\mathbf{x} = T^{-1}\boldsymbol{\phi}$, with $\boldsymbol{\phi} = \begin{bmatrix} \boldsymbol{\phi}_1 \\ \boldsymbol{\phi}_2 \end{bmatrix}$. Therefore $\mathbf{x} = T_{I_1}\boldsymbol{\phi}_1 + T_{I_2}\boldsymbol{\phi}_2$. Consequently, if $Z(j\omega)$ denotes the transform of the impulse response matrix, it can be written as

$$\begin{aligned} Z(j\omega) &= (j\omega E - A)^{-1} B \\ &= T_{I_1}(j\omega I - J)^{-1} K_1 - T_{I_2} \sum_{i=0}^{\nu-1} (j\omega)^i N^i K_2 \end{aligned} \tag{14}$$

The proper and improper part of the solutions are clearly decoupled in this frequency domain solution. Let $M_0 = -T_{I_2}K_2$ and $M_1 = -T_{I_2}NK_2$, covering systems of index one and two. The matrices M_i are constant matrices. Also, let $Z_m(j\omega)$ denote the contribution of the improper subsystem. From equation (14), $Z_m(j\omega)$ can be written in terms of the matrices M_i as

$$Z_m(j\omega) = -T_{I_2} \sum_{i=0}^{\nu-1} (j\omega)^i N^i K_2 = M_0 + j\omega M_1 \cdots \tag{15}$$

Hence,

$$Y(j\omega) = Z(j\omega) - Z_m(j\omega) = T_{I_1}(j\omega I - J)^{-1} K_1 \tag{16}$$

which is the strictly proper part of the solution.

P_{ic} and P_{pc} can be constructed using $Z_m(j\omega)$ and $Y(j\omega)$ as follows. From equation (12), it can be seen that the Gramian corresponding to the proper subsystem is given by

$$P_{pc} = \frac{1}{2\pi} \int_{-\infty}^{\infty} Y(j\omega) Y(j\omega)^H d\omega \tag{17}$$

P_{po} can be obtained similarly. From equations (9),(11) and (14), we get

$$\begin{aligned} P_{ic} &= T_{I_2} \left(\sum_{i=0}^{\nu-1} N^i K_2 (N^i K_2)^T \right) T_{I_2}^T \\ &= \sum_{i=0}^{\nu-1} M_i M_i^T \end{aligned} \tag{18}$$

P_{io} can be obtained similarly.

Therefore, instead of computing the spectral projectors, we need to evaluate the constant matrices M_i. The key to obtaining the ROMs efficiently is the efficient computation of these matrices.

4. COMPUTATION OF THE MATRICES OF THE IMPROPER SUBSYSTEM

In this section, we propose an efficient algorithm to obtain accurate estimates of the matrices M_i. In the discussion that follows, we assume that the index of the system is at most two (covering RLC networks, PEEC models and fluid mechanical systems present in the two sets of benchmarks [18, 19]).

$Y(j\omega)$ denotes the strictly proper part of the solution and hence eventually decays with frequency. Therefore, at large enough frequencies, the improper part of the solution dominates and $Z(j\omega) \approx Z_m(j\omega) = M_0 + j\omega M_1$. At these frequencies, therefore, $M_0 \approx Z_r$ and $M_1 \approx \frac{dZ_I}{d\omega}$ where $Z_r = \mathrm{Re}(Z(j\omega))$ and $Z_I = \mathrm{Im}(Z(j\omega))$. The question is what is a large enough frequency. To determine this we set up an iteration. The steps are outlined in Algorithm 1. We start the iteration at a frequency ω_{max} and define $\omega_{min} = \omega_{max}^{\gamma}$, $\gamma < 1$. Samples of $Z(j\omega)$ are obtained by varying ω uniformly on a log scale between ω_{min} and ω_{max}. We then estimate the first derivative of Z_r (dz_r) and the second derivative of Z_I ($d2z_I$) with respect to ω at each of these frequencies using finite differences. At frequency ω_i, the two derivatives can be written as

$$dz_r[i] = \frac{Z_r[i] - Z_r[i-1]}{\omega_i - \omega_{i-1}}, \quad d2z_I[i] = \frac{2(dz_I[i+1] - dz_I[i])}{\omega_{i+1} - \omega_{i-1}}$$

where dz_I is defined similarly as dz_r. Note that both $dz_r[i]$ and $d2z_I[i]$ are $n \times m$ or $n \times q$ matrices depending on whether we are computing matrices for the controllability or observability Gramian. If the solution is dominated by the improper part, then both these derivatives should be approximately zero. To check this, we evaluate $\{\max_{j,k}\{|dz_{r_{jk}}|\}\}$ and $\{\max_{j,k}\{|d2z_{I_{jk}}|\}\}$ at each frequency ω_i. If both measures are less than a threshold (typically 10^{-15}) for all ω_i between ω_{min} and ω_{max}, then M_0 and M_1 are estimated as

$$M_0 \approx \frac{1}{r+1} \sum_{i=0}^{r} Z_r[i], \quad M_1 \approx \frac{1}{r} \sum_{i=1}^{r} \frac{Z_I[i] - Z_I[i-1]}{\omega_i - \omega_{i-1}}$$

i.e. the average value of the real part of $Z(j\omega_i)$ and the first derivative of the imaginary part of $Z(j\omega_i)$ respectively over the samples obtained at $r+1$ frequency points between ω_{min} and ω_{max}. If the threshold condition is not satisfied, ω_{max} is increased and the derivatives are once again estimated. For all the benchmarks, we have found that if we start at a high enough frequency, typically between $10^{12} - 10^{16}$rad/s for most benchmarks, the matrices can be obtained in one or two iterations. Moreover, in many cases, the additional samples computed are not wasted, since they are used to estimate the proper Gramians.

Once these matrices are evaluated, $Z_m(j\omega)$ is obtained as $M_0 + j\omega M_1$ and $Y(j\omega) = Z(j\omega) - Z_m(j\omega)$. It is possible now compute both Gramians - the improper Gramian as a simple outer product and the proper Gramian using a quadrature. For the quadrature, we require samples between $\begin{bmatrix} 0 & \omega_{max_q} \end{bmatrix}$. If ω_{max_q} is larger than ω_{min}, the samples obtained for estimation of the constant matrices are reused. For many of the benchmarks, especially the MNA benchmarks, therefore, there is no real overhead for computation of the improper Gramians, other than the cost of estimating the derivatives. We have tried both trapezoidal and Simpson's rule for the quadrature. Both these methods were found to give reasonable results.

Algorithm 1 Algorithm to compute M_0 and M_1 for index 1 and 2 systems

Input: ω_{max}, γ, β, $\begin{bmatrix} E & A & B & C & D \end{bmatrix}$, Threshold
$\omega_{min} = \omega_{max}^{\gamma}$; $Z = [\ ]$; flag = 0;
While flag == 0
  For ω_i between ω_{min} and ω_{max}:
    $Z[i] \leftarrow (j\omega_i E - A)^{-1} B$
    $Z_r = \mathrm{Re}\{Z[i]\}$; $Z_I = \mathrm{Im}\{Z[i]\}$
    $dz_r[i] = \frac{Z_r[i] - Z_r[i-1]}{\omega_i - \omega_{i-1}}$
    $d2z_I[i] = \frac{2}{\omega_{i+1} - \omega_{i-1}}(dz_I[i+1] - dz_I[i])$
    $Ndz_r[i] = \max_{j,k}\{|dz_r[i]_{jk}|\}$
    $Nd2z_I[i] = \max_{j,k}\{|d2z_I[i]_{jk}|\}$
  If $(\max_i\{Ndz_r[i]\}$ or $\max_i\{Nd2z_I[i]\}) >$ Threshold:
    $\omega_{min} = \omega_{max}$; $\omega_{max} = \beta\ \omega_{max}$ //$\beta > 1$//
  Else
    flag = 1
$M_0 = \frac{1}{r+1}\sum_{i=0}^{r} Z_r[i]$; Average over the $r+1$ samples between ω_{min} and ω_{max}
$M_1 = \frac{1}{r}\sum_{i=1}^{r} \frac{Z_I[i] - Z_I[i-1]}{\omega_i - \omega_{i-1}}$;
$Z_m(j\omega_i) = M_0 + j\omega_i M_1$;

Computationally, this is much more efficient than finding the spectral projectors and using them to solve the Lyapunov equations. For p frequency points, the evaluation of the integral for the Gramian corresponding to the proper subsystem requires p complex matrix factorizations and p solves, corresponding to a computational complexity of $O(pn^{\alpha} + pn^{\beta}), 1 \leq \alpha \leq 1.2, 1.1 \leq \beta \leq 1.5$ [1]. For large sparse systems, this complexity is significantly lower than solution of the Lyapunov equations [1, 3]. An additional advantage is that this computation is very easily parallellizable and is, in fact, embarrassingly parallel.

5. COMPUTATION OF THE TBR

Once the matrices M_i and the samples of $Y(j\omega)$ required for the quadrature are obtained, the Gramians corresponding to the proper and improper subsystems can be evaluated using an outer product and quadrature. But this is expensive and we do not actually require the Gramians themselves. What is required are Cholesky factors of the Gramians L_p, R_p, L_i and R_i where $P_{ic} = R_i R_i^T$; $P_{io} = L_i^T L_i$ and $P_{pc} = R_p R_p^T$; $P_{po} = L_p^T L_p$. These factors can be obtained without the building the Gramian. The steps for computing the ROM are detailed in Algorithm 2. It is based on the algorithm proposed in [6]. The differences are we do not solve the projected GLYAPs, we use low rank factors and when applicable, we use congruence transform. The inputs M_c and M_o are matrices containing the constant matrices corresponding to the improper controllability and the observability Gramian. First, low rank factors L_i and R_i are obtained through a "thin" singular value decomposition (SVD) of the matrices M_c and M_o. The low rank factor R_p (L_p) can be obtained via a thin SVD of the matrix Y_c (Y_o), the columns of which are the real and imaginary parts of $Y_i = Y(j\omega_i)$ multiplied by the square root of the weighting factor used in the quadrature. If we use Y_i directly, the SVD will result in a complex valued R_p and L_p, which is not desired. To find a real valued R_p and L_p, we can use the fact that both Y_i and Y_i^* are required to build the Gramians.

Algorithm 2 Algorithm used to find the reduced order model

Input: $\begin{bmatrix} E & A & B & C & D \end{bmatrix}$
Input: Y_c and Y_o; Samples of $Y(j\omega)$ corresponding to the proper controllability and observability Gramian; $Y_c = \begin{bmatrix} \beta_0 Y_0 & \beta_1 \mathrm{Re}(Y_1) & \beta_1 \mathrm{Im}(Y_1) & \cdots \end{bmatrix}$; $\beta_0 = \sqrt{\alpha_i/(2\pi)}$, $\beta_i = \sqrt{\alpha_i/(\pi)}$, $i > 0$ // α_i: quadrature weights;
Input M_c and M_o; Constant matrices corresponding to the controllability and observability Gramian ($M_c = \begin{bmatrix} M_0 & M_1 \end{bmatrix}$);
Input Sym: If true, $B = C^T$ and A and E are symmetric matrices
Input CT: If true, perform reduction using a congruence transform
l_p, l_i: Orders of the proper and improper part of the reduced system; $l = l_p + l_i$
Output: Reduced model $\begin{bmatrix} E_r & A_r & B_r & C_r \end{bmatrix}$
$M_c = U\Lambda V^T$; //Perform an SVD of M_c
$R_i = U\Lambda$. //Compute L_i in a similar fashion from M_o if Sym is false;
$Y_c = U_y \Lambda_y V_y^T$; $R_p = U_y \Lambda_y$. //Compute L_p from Y_o in a similar fashion if Sym is false
If (Sym):
  $L_p = R_p^T$; $L_i = R_i^T$
$L_p E R_p = U_p \Sigma V_p^T$; $L_i A R_i = U_i \Theta V_i^T$;
If (CT):
  $V_l = \begin{bmatrix} R_p V_{p_1} & R_i V_{i_1} \end{bmatrix}$; where $V_{p_1} = V_p[:, 1 : l_p]$, $V_{i_1} = V_i[:, 1 : l_i]$;
  $V_l = Q_r R$ //QR decomposition
  $E_r = Q_r^T E Q_r$; $A_r = Q_r^T A Q_r$; $B_r = Q_r^T B$; $C_r = C Q_r$
Else:
  $W_l = \begin{bmatrix} L_p^T U_{p_1} & L_i^T U_{i_1} \end{bmatrix}$; where $U_{p_1} = U_p[:, 1 : l_p]$, $U_{i_1} = U_i[:, 1 : l_i]$
  $V_l = \begin{bmatrix} R_p V_{p_1} & R_i V_{i_1} \end{bmatrix}$; where $V_{p_1} = V_p[:, 1 : l_p]$, $V_{i_1} = V_i[:, 1 : l_i]$;
  $W_l = Q_l L$; $V_l = Q_r R$; // QR decompositions for numerical stability
  $E_r = Q_l^T E Q_r$; $A_r = Q_l^T A Q_r$; $B_r = Q_l^T B$; $C_r = C Q_r$

This is done as follows.

$$\begin{aligned} \begin{bmatrix} Y_i & Y_i^* \end{bmatrix} &= \begin{bmatrix} \sqrt{2}\mathrm{Re}(Y_i) & \sqrt{2}\mathrm{Im}(Y_i) \end{bmatrix} \times \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix} \\ &= \begin{bmatrix} \sqrt{2}\mathrm{Re}(Y_i) & \sqrt{2}\mathrm{Im}(Y_i) \end{bmatrix} F_y \end{aligned}$$

Note that the multiplying matrix F_y is a unitary matrix. Therefore if $U_y \Lambda_y V_y^T = \begin{bmatrix} \sqrt{2}\mathrm{Re}(Y_i) & \sqrt{2}\mathrm{Im}(Y_i) \end{bmatrix}$, then

$$\begin{bmatrix} Y_i & Y_i^* \end{bmatrix} = U_y \Lambda_y (V_y^T F_y)$$

$V_y^T F_y$ is clearly a unitary matrix. As seen in Algorithm 2, U_y and Λ_y are used to construct the low rank factors R_p and L_p of the proper controllability and observability Gramians.

Algorithm 2 has two additional inputs Sym and CT. The input Sym is an indicator of a symmetric system i.e. $A^T = A$, $E^T = E$ and $C = B^T$. In this case $L_p^T = R_p$ and $L_i^T = R_i$ which means $W_l = V_l$. If E and A are definite matrices, CT is set to one, indicating a congruence transform can be used. While it is not possible to guarantee passivity for the general case, it is possible to use a congruence transform as proposed in [2] to guarantee passivity when the matrices E and A are

definite matrices, which is the case for VLSI interconnects. In this case, as explained in Algorithm 2, we can perform an additional QR decomposition of the right basis V_l and use the resulting orthogonal basis to perform a congruence transform resulting in passive reduced order models [1, 2]. The left basis W_l is not used. Therefore, it is not equivalent to a balanced realization unless the system is symmetric in addition to A and E being definite matrices.

The RLC network is not a symmetric system. However, we have observed that good reduced order models are obtained if we perform an orthogonal projection onto the subspace spanned by the right basis V_l (containing the eigenvectors of $P_{pc}E^TP_{po}E$ and $P_{ic}A^TP_{po}A$) and use congruence transform. It is also seen in [2]. In fact, we have found that in some cases, like the benchmark MNA1, it gives much better results and is numerically more stable. This is explained as follows. For RLC networks, $B = C^T$. Based on the structure of the matrices A and E [18], it can be shown that the Cholesky factors of the observability and controllability Gramian are related as

$$L_p^T = \pm \begin{bmatrix} -I & \mathbf{0} \\ \mathbf{0} & I \end{bmatrix} R_p \tag{19}$$

depending on whether the input to the network consists of voltage sources or current sources. A similar result holds for L_i and R_i. Since E is symmetric and block diagonal, L_pER_p is also symmetric. The structure of A ensures L_iAR_i is also symmetric. From Algorithm 2, it is clear that $U_p = V_p$ and $U_i = V_i$ so that W_l and V_l differ due to a sign change of some of the rows in R_p and L_p. Due to this, it turns out that there is a substantial overlap between the two subspaces, leading to good ROMs when congruence transform is used. For the supersonic inlet benchmark [19], the matrices do not have this property and the model obtained using a congruence transform is very poor.

6. RESULTS

The algorithms were implemented in PYTHON and SCIPY was used for the computations. The parameter *Threshold* in Algorithm 2 is set to 10^{-16}. The experiments were run on a 3.5GHz core-i7 machine. We obtained TBRs of a number of benchmarks in the sets [18, 19]. The CPU times taken to find the ROM for various benchmarks are given in Table 1. For each case a ROM was obtained so that its frequency response matched that of the full system over a specific range (in most cases, specified in the documentation of the benchmarks). The order of the ROM obtained compares well with the order obtained using other methods. The CPU times indicate that the method is suitable even for relatively large systems.

We now discuss the results for two of the benchmarks MNA5 from the set [18] and the supersonic inlet [19] in more detail. MNA5 is an index 2 system and the supersonic inlet has index 1. ROMs for these two benchmarks have been obtained in [11, 20] using technques that require explicit construction of the spectral projectors, allowing for comparison of results. MNA5 has nine ports which results in an improper subsystem of size 18. Figure 1(a) shows the magnitude response at port 1 for a ROM that has a proper subsystem of size 450. The relatively large order is due to a gradual decay of singular values of the proper subsystem. The high and low frequency portion can be captured with a relatively small model order, but capturing all the oscillations requires a higher order proper subsystem.

Benchmark	size	index	Freq. Rng.(Hz)	ROM order	Time(s)
PEEC Sp.Ind	1434	0	$0 - 10^{10}$	10	29
MNA1	578	2	$0 - 10^{15}$	164	9
MNA5	10913	2	$0 - 10^{10}$	450	69
Supersonic Inlet	11,730	1	0-20	27	40
xingo3012	20,944	1	0-1000	36	8

Table 1: Frequency range, ROM order and CPU times for various benchmarks.

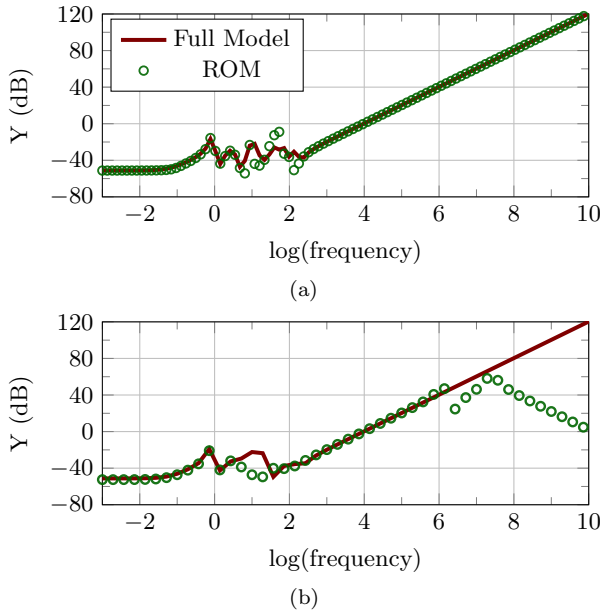

Figure 1: Magnitude response at port 1 for the benchmark MNA5. The order of the proper subsystem is 450. The order of the ROM is 468. (b) Order 450 model without explicitly finding the improper subsystem.

Figure 1(b) shows a comparison of the full model and ROM of size 450 obtained without explicitly accounting for the improper subsystem. The method is similar to the one used in [3] (and in spirit, to PMTBR). The resulting ROM has several finite high frequency poles, but no poles at infinity. This results in a poor match beyond 1MHz. Since methods like multi-point PRIMA and PMTBR do not explicitly obtain the polynomial part of the transfer function, the resulting ROMs have high frequency errors [21].

Figure 2(a) contains the absolute value of the error as a function of frequency. The relatively large error between 0 and 100Hz is due to the reduced size of the proper subsystem. This error reduces as the size of the proper subsystem is increased. The high frequency error is between 10^{-10} to 10^{-8}, though there is a tendency to increase beyond 10^9Hz. We believe this increase is mainly due to round-off errors involved in the computation of the reduced order matrices. Also, the condition number of the matrix $(j\omega E_r - Ar)$ increases rapidly beyond 10^{11}Hz, making the computation of the reduced order transfer function susceptible to round-off

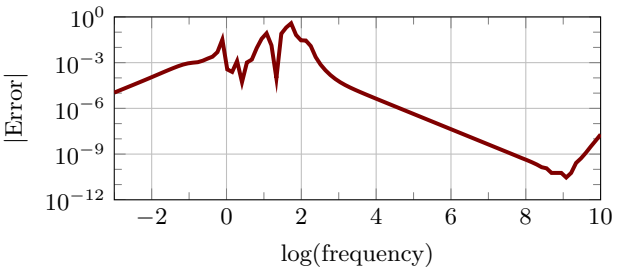

Figure 2: Absolute value of the error for MNA5.

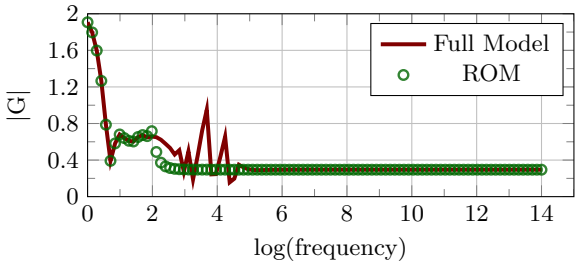

Figure 3: Magnitude response of the supersonic inlet benchmark. The order of the ROM is 27.

errors. A steady increase in the error is also seen [11].

Since we could not find an interconnect benchmark with index 1, we illustrate the method for the supersonic inlet. Figure 3 shows the frequency response for the supersonic inlet. The proper subsystem has order 27 and the improper subsystem has order one. The match at low frequencies about upto 100Hz and at very high frequencies is very good. For a good match at intermediate frequencies, a higher order proper subsystem is required. The model obtained is much better than the one obtained in [20].

7. CONCLUSIONS

In this paper, we have presented an efficient algorithm for obtaining frequency-weighted TBRs including the contribution of both the proper and improper subsystem. Results indicate it is accurate and can be used for large systems. It can be used to get passive models for definite systems. It is also easily parallelizable. The key step is to estimate the contribution of the improper subsystem efficiently and accurately. Possible future directions include better sampling techniques and reduction of round-off errors.

8. REFERENCES

[1] J.R.Phillips and L.M.Silveira, " Poor man's TBR: A simple model reduction scheme *IEEE Trans. Computer-Aided Design*, vol. 24, no.1, pp 43-55, Jan. 2005.

[2] B.Yan, S.Tan, P.Liu, B.McGaughy, "Passive Interconnect Macromodelling via balanced truncation of linear systems in the descriptor form", *Proc. Asia South Pacific Design Autom.Conf.(ASPDAC)* pp355-360, 2007.

[3] B.Yan, S.Tan, G.Chen and Y.Cai, "Efficient model reduction of Interconnects via double Gramians approximation", *Proc. ASPDAC*, pp25-30, 2010.

[4] R.Marz, " Canonical projectors for linear differential algebraic equations", *Computers. Math. Applic.*, vol. 31, no.4, pp121-135, Feb.1995.

[5] W.Q.Liu and V.Sreeram, "Model reduction of singular systems", *Int.J.Syst Sci.*, vol.32, pp1205-1215, 2001.

[6] T.Stykel, "Gramian based model reduction for descriptor systems", *Math. Contol Sig. Syst.*, vol. 16, pp. 297-319, 2004.

[7] T.Stykel, "Balanced Truncation model reduction for semidiscretized Stokes equation", *Linear Algebra Appl.*, vol. 415, pp262-289, 2006.

[8] F.Freitas, J.Rommes and N.Martins, "Gramian based reduction method applied to large sparse power system descriptor models", *IEEE Trans. Power Syst,*, vol.23, no.3, pp1258-1270, Aug.2008

[9] M.Heinkenschloss, D.Sorenson and K.Sun, "Balanced truncation model reduction for a class of descriptor systems with application to the Oseen equation", *SIAM J.Sci.Comput*, vol.30, no.2, pp 1038-1063.

[10] N.Wong, "An efficient passivity test for descriptor systems via canonical projector techniques", *Proc. Design Automation Conf.,*, pp957-962, July 2009.

[11] Z.Zhang and N.Wong, "An efficient projector based passivity test for descriptor systems",*IEEE Trans. Computer-Aided Design*, vol. 29, no.8, pp 1202-1210, Aug.2010.

[12] Z.Zhang, Q.Wang, N.Wong and L.Daniel, "A moment-matching scheme for the passivity preserving model order reduction of indefinite descriptor systems with possible polynomial parts",*Proc. ASPDAC*, pp49-54, 2011.

[13] T.Reis and T.Stykel, "PABTEC: Passivity preserving balanced truncation for electrical circuits", *IEEE Trans. Computer-Aided Design*, vol. 29, no.9, pp 1354-1367, Sept.2010.

[14] T.Reis and T.Stykel, "Lyapunov balancing for passivity-preserving model reduction of RC circuits", *SIAM J. App. Dynamical Syst.*,vol.10, no.1, pp1-34, 2011.

[15] G.Ali, N.Banagaaya, W.H.A.Schilders and C.Tischendorf, "Index-aware model order reduction for linear index-2 DAEs with constant coefficients", *SIAM J.Sci.Comput*, vol.35, no.3, ppA1487-A1510.

[16] N.Banagaaya, G.Ali, W.H.A.Schilders and C.Tischendorf, "Implicit index-aware model order reduction for RLC/RC networks", *Proc. Int. Conf. Design, Automation and Test in Europe*, 2014.

[17] W.Gawronski and J.N.Juang, "Model reduction in limited time and frequency intervals", *Int.J.Syst Sci.*, vol.21, pp349-376, 1990.

[18] Available online at http://www.icm.tu-bs.de/NICONET/benchmodred.html

[19] Available online at http://portal.uni-freiburg.de/imteksimulation/downloads/benchmark

[20] P.Benner and V.I.Sokolov, "Partial realization of descriptor systems", *Systems Control Lett.,*, vol.55, no.11, pp 929-938,2006.

[21] S.Gugercin, T.Stykel and S.Wyatt, "Model reduction of Descriptor systems by Interpolatory Projection Methods", Preprint, available at http://www.math.uni-augsburg.de, Jan. 2013.