Efficient Balance-and-Truncate Model Reduction for Large Scale Systems[1]

V. Balakrishnan, Q. Su and C-K. Koh
School of Electrical and Computer Engineering
Purdue University
West Lafayette, IN 47907–1285
Email: {ragu, qsu,chengkok}@ecn.purdue.edu

Abstract

We present efficient implementations of the balance-and-truncate model reduction technique for large-scale systems. The key observation that distinguishes our approach is that Krylov subspace methods (Arnoldi and Lanczos) directly yield approximate low-rank square roots[2] of the system Gramians; the balancing transformation can then be constructed from these square roots, obviating the need for solving any Lyapunov equations. In addition, the order of the reduced model is not fixed *a priori* as with some existing methods, but is determined from the problem data. Numerical simulations show that our approach performs very well over a range of examples, and offers considerable savings in practice.

1 Introduction

As engineering systems become more and more complex, so do the mathematical models describing them. This is true, for instance, when an engineering system is modeled as an interconnection of a large number of sub-systems, as with VLSI circuit models; the resulting model of the overall system can involve thousands of variables. The analysis and design of large-scale systems can stretch the limits of computing resources. A standard practice that addresses such issues is that of model reduction. Our objective is to present efficient algorithms for the model reduction of large-scale linear time-invariant (LTI) state-space models.

Model reduction of LTI systems is a well-studied topic. One approach is to expand the transfer function as a power series around a suitable point in the complex plane, and obtain a lower order model whose power series coefficients match the first few original coefficients ("moment-matching"). A well-known example of such an approach is Padé approximation [CN92, FF95, GGD94]. Another model-reduction approach involves projecting the state space onto the principal controllable subspace, or the principal observable subspace [Fre98, GG97, GN99]. Some of these approaches use Krylov subspace computation techniques, which are well-conditioned, require only matrix-vector multiplications, and are therefore particularly useful for large-scale systems. A third technique, one that underlies the approach presented in this paper, is the balance-and-truncate method (see for example, [Moo81]). The idea here is to find a state-space coordinate transformation in which the input-to-state map and the state-to-output map are "aligned". Thus, the state-variables that are not easy to reach and not easily observed can be omitted (or the model truncated). The approximation error can be shown to be bounded [Enn84, Glo84]. While the balance-and-truncate method is theoretically attractive and also yields excellent approximate models in practice, its use for the model-reduction of large-scale systems has been hampered by its quickly growing computational demand: Two large-size Lyapunov equations need to be solved, followed by a large-size eigen-decomposition. One approach towards addressing this issue is to obtain low-rank approximate solutions to the large-size Lyapunov equations, for instance, the "Alternate Direction Iteration" or ADI approach [LW91, LW99]. (We note however that in [LW99], the ADI technique is used to find an approximate solution to the system Gramians; the resulting approximate principal controllable and observable subspaces are then used to perform the model reduction, rather than balance-and-truncate.) The drawback of the ADI approach is the requirement that the original system matrix be tridiagonalized first; this step is both computationally demanding and possibly numerically ill-conditioned [GL89, §9.3.6]. Another prevalent approach for balance-and-truncate model reduction is to use Krylov subspace computation methods to first find the principal controllable (or observable) subspace, perform an initial model reduction by projecting the state vector onto the principal subspace, and then solve reduced-order Lyapunov equations to proceed with balancing and truncating; see, for example, [JK94, RP99], and the references therein.

[1] This work was supported in part by NSF under contract number CCR-9984553, and SRC under contract number 99-TJ-689.

[2] We use the term "square root" to mean the not necessarily symmetric square root of a matrix: If $M = M^T = NN^T$, we say N is the square root of M.

Our contribution is an algorithm for balance-and-truncate model reduction, using Krylov methods, where no Lyapunov equations need solution. The distinguishing feature of our approach is the use of Arnoldi and Lanczos iterations to directly compute approximate low-rank square roots of the system Gramians; the balancing transformation is then constructed from these square roots. In Section 2, we introduce the mathematical framework underlying our approach, including an analysis of the approximation error, and the resulting model reduction algorithm. In Section 3, we present a few examples that illustrate that our approach requires greatly reduced computation.

2 Mathematical Framework

2.1 Balanced Transformation and Truncation

Consider the linear system described by the state-space equations

$$\dot{x} = Ax + Bu, \quad (1a)$$
$$y = Cx + Du, \quad (1b)$$

where $x(t) \in \mathbb{R}^N$, $u(t) \in \mathbb{R}$ and $y(t) \in \mathbb{R}$, and A, B, C and D are real matrices of appropriate sizes. (We will consider only single-input single-output systems in this paper; the extension of the results presented herein to multi-input multi-output systems is straightforward.) We will use ordered quadruple (A, B, C, D) to denote the state-space realization of the system. We assume that A is stable, i.e., all of its eigenvalues have negative real part, and that the realization is minimal.

Balanced truncation is one well-known model reduction scheme. The first step is to compute the controllability and observability Gramians, denoted W_c and W_o respectively, and defined as

$$W_c = \int_0^\infty e^{At}BB^Te^{A^Tt}dt, \quad W_o = \int_0^\infty e^{A^Tt}C^TCe^{At}dt.$$

The Gramians satisfy the Lyapunov equations

$$AW_c + W_cA^T + BB^T = 0, \quad A^TW_o + W_oA + C^TC = 0.$$

With the eigenvalues sorted in decreasing order, the corresponding eigenvectors of W_c yield directions in state-space that are increasingly hard to reach from the input u, and the eigenvectors of W_o yield directions that are increasingly hard to observe from the output y. Define

$$T_b = XU\Sigma^{-\frac{1}{2}} = \left(\Sigma^{-\frac{1}{2}}V^TY^T\right)^{-1},$$

where X and Y are the square roots of the Gramians, i.e., $W_c = XX^T$ and $W_o = YY^T$, and $X^TY = U\Sigma V^T$ is a singular value decomposition (SVD). $\Sigma > 0$ is diagonal, with the diagonal entries in descending order. The diagonal entries of Σ are called the Hankel singular values of the system. In this new, so-called balanced coordinates $\bar{x} = T_b^{-1}x$, the state-space realization is $(\bar{A}, \bar{B}, \bar{C}, D) = (T_b^{-1}AT_b, T_b^{-1}B, CT_b, D)$. It is easily verified that the corresponding controllability and observability Gramians are

$$\bar{W}_c = \bar{W}_o = \Sigma.$$

Thus the state components are as reachable from the input as they are observable at the output, with the corresponding Hankel singular value quantifying their reachability and observability. This motivates the next step, that of "truncation" of the state-vector, i.e., simply "throwing away" state components for which the corresponding diagonal entry σ_i of Σ is small. If

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \gg \sigma_{n+1} \geq \cdots \geq \sigma_N,$$

the state-space realization of the reduced order model is given by $(\bar{A}_{\text{red}}, \bar{B}_{\text{red}}, \bar{C}_{\text{red}}, D)$, where

$$\bar{A}_{\text{red}} = \begin{bmatrix} I_n & 0_{n\times(N-n)} \end{bmatrix} \bar{A} \begin{bmatrix} I_n \\ 0_{(N-n)\times n} \end{bmatrix},$$
$$\bar{B}_{\text{red}} = \begin{bmatrix} I_n & 0_{n\times(N-n)} \end{bmatrix} \bar{B},$$
$$\bar{C}_{\text{red}} = \bar{C} \begin{bmatrix} I_n \\ 0_{(N-n)\times n} \end{bmatrix}.$$

Let $H(s) = C(sI - A)^{-1}B + D$ and $H_{\text{red}}(s) = \bar{C}_{\text{red}}(sI - \bar{A}_{\text{red}})^{-1}\bar{B}_{\text{red}} + D$ denote the transfer functions of the original and the reduced-order system respectively. Then, it can be shown that

$$\sup_{\omega\in\mathbb{R}} \|H(j\omega) - H_{\text{red}}(j\omega)\| \leq 2\sum_{n+1}^{N} \sigma_i.$$

While the approximation properties of the balance-and-truncate model reduction algorithm are excellent, its use for large-scale systems is limited by the heavy computational demand: Two large-size Lyapunov equations need to be solved, followed by one large-size SVD computation. We address this issue with an approximate balance-and-truncate technique that requires far less computation. The idea is to directly compute low-rank square roots of the Gramians; these square-roots can be combined to yield "approximate" balancing transformations that automatically truncate the state space.

2.2 Approximate Balanced Truncation

We first describe the idea behind an approximate balance-and-truncate method that relies on low-rank square roots of the Gramians. (We will defer a careful analysis of the approximation error to Section 2.4.) Suppose that we have approximate low-rank square roots of the Gramians, i.e., we have full rank $X_k, Y_k \in \mathbb{R}^{N\times k}$ such that

$$W_c \approx X_kX_k^T, \quad W_o \approx Y_kY_k^T.$$

Let

$$X_k^TY_k = \hat{U}\hat{\Sigma}\hat{V}^T$$

be the $k \times k$ SVD. Then, the diagonal entries $\hat{\sigma}_i$ of $\hat{\Sigma}$ approximate the first k Hankel singular values of the system. Suppose that

$$\hat{\sigma}_1 \geq \hat{\sigma}_2 \geq \cdots \geq \hat{\sigma}_n \gg \hat{\sigma}_{n+1} \geq \cdots \geq \hat{\sigma}_k.$$

Define

$$\tilde{T}_b = X_k \hat{U} \hat{\Sigma}^{-\frac{1}{2}} \begin{bmatrix} I_n \\ 0_{(N-n)\times n} \end{bmatrix},$$

and

$$\tilde{T}_b^{\ddagger} = \begin{bmatrix} I_n & 0_{n\times(N-n)} \end{bmatrix} \hat{\Sigma}^{-\frac{1}{2}} \hat{V}^T Y_k^T.$$

Note that $\tilde{T}_b \in \mathbb{R}^{N\times n}$ and $\tilde{T}_b^{\ddagger} \in \mathbb{R}^{n\times N}$, and that $\tilde{T}_b^{\ddagger}\tilde{T}_b = I$.

Consider the nth order system with state-space realization $(\tilde{T}_b^{\ddagger} A \tilde{T}_b, \tilde{T}_b^{\ddagger} B, C\tilde{T}_b, D)$. It can be verified that the controllability and observability Gramians for this realization are

$$\bar{W}_c \approx \mathbf{diag}\,(\hat{\sigma}_1, \ldots, \hat{\sigma}_n) \approx \bar{W}_o.$$

Thus, the matrices $\tilde{T}_b$ and $\tilde{T}_b^{\ddagger}$ directly provide for an approximate balance-and-truncate state-space model reduction.

2.3 Low-rank Square Roots of the Gramians via Krylov Methods

For every real scalar $p < 0$, the equation

$$AW_c + W_c A^T + BB^T = 0$$

is equivalent to

$$A_p W_c A_p^T - W_c + B_p B_p^T = 0$$

where $A_p \triangleq (pI + A)^{-1}(pI - A)$, $B_p \triangleq \sqrt{(-2p)}(pI + A)^{-1}B$. Therefore we have

$$\begin{array}{rcl} W_c & = & \Sigma_{j=0}^{\infty} A_p^j B_p B_p^T (A_p^T)^j \\ & \approx & \Sigma_{j=0}^{k-1} A_p^j B_p B_p^T (A_p^T)^j \\ & = & \mathcal{K}(A_p, B_p, k)\mathcal{K}(A_p, B_p, k)^T \\ & \triangleq & X_k X_k^T, \end{array} \qquad (2)$$

where the kth order Krylov matrix is defined as

$$\mathcal{K}(A_p, B_p, k) \triangleq [B_p \quad A_p B_p \quad A_p^2 B_p \quad \cdots \quad A_p^{k-1} B_p].$$

Similarly, with $C_p \triangleq \sqrt{(-2p)}C(pI + A)^{-1}$, we have

$$\begin{array}{rcl} W_o & = & \Sigma_{j=0}^{\infty} (A_p^T)^j C_p^T C_p A_p^j \\ & \approx & \Sigma_{j=0}^{k-1} (A_p^T)^j C_p^T C_p A_p^j \\ & = & \mathcal{K}(A_p^T, C_p^T, k)\mathcal{K}(A_p^T, C_p^T, k)^T \\ & \triangleq & Y_k Y_k^T. \end{array} \qquad (3)$$

One interpretation of these steps is that we have derived a discrete-time system with state-space realization (A_p, B_p, C_p, D_p), that has the same Gramians as the continuous system, using the conformal mapping $z \mapsto (p - s)/(p + s)$.

The direct computation of the $N \times k$ matrices $\mathcal{K}(A_p, B_p, k)$ and $\mathcal{K}(A_p^T, C_p^T, k)$ is ill-conditioned, as the columns of these matrices quickly converge to the dominant eigenvector of A_p and A_p^T respectively. Krylov methods are natural tools for well-conditioned computation of $\mathcal{K}(A_p, B_p, k)$ and $\mathcal{K}(A_p^T, C_p^T, k)$.

The quality of the approximation of $\mathcal{K}(A_p, B_p, k)$ and $\mathcal{K}(A_p^T, C_p^T, k)$ as square roots of the Gramians depends critically on how fast A_p^k goes to zero with k, or on the spectral radius (i.e., the maximum magnitude of the eigenvalues) $\rho(A_p)$ of A_p. The choice of p to make $\rho(A_p)$ as small as possible is a well-studied problem; see for example [LW91]. The key here is that the eigenvalues of A and A_p are related by $\lambda_i(A_p) = (p - \lambda_i(A))/(p + \lambda_i(A))$. For every i, the value of p that minimizes $|(p - \lambda_i(A))/(p + \lambda_i(A))|$ is $p = |\lambda_i(A)|$. Of course, we need to choose p to minimize

$$\max_i |(p - \lambda_i(A))/(p + \lambda_i(A))|.$$

As discussed in [LW91, Smi68], a good choice for p is simply $-\sqrt{(\max_i |\lambda_i(A)|)\,(\min_i |\lambda_i(A)|)}$. Power iterations requiring only matrix-vector multiplications, and one LU factorization can be used to compute $\max_i |\lambda_i(A)|$ and $\min_i |\lambda_i(A)|$ approximately.

The computation of $\mathcal{K}(A_p, B_p, k)$ and $\mathcal{K}(A_p^T, C_p^T, k)$, can proceed with either the Arnoldi or the Lanczos algorithm. Due to space limitations, we will only describe an implementation of Arnoldi algorithm. The algorithm that we present here is adapted from [GL89, Ch. 9] to compute a QR factorization of both $\mathcal{K}(A_p, B_p, k)$ and $\mathcal{K}(A_p^T, C_p^T, k)$ inside one iterative loop. Note that k is not known *a priori*, but is automatically determined from the stopping criterion. Upon termination, we have QR factorizations $\mathcal{K}(A_p, B_p, k) = Q_k R_k$ and $\mathcal{K}(A_p^T, C_p^T, k) = P_k S_k$; see [BSK00] for details.

$j = 1;$
$q_1 = B_p/\|B_p\|_2;\ \beta = 1;\ Q_1 = q_1;\ R_1 = \|B_p\|_2;$
$p_1 = C_p^T/\|C_p\|_2;\ \gamma = 1;\ P_1 = p_1;\ S_1 = \|C_p\|_2;$
`repeat until stopping criterion is met` {
  `for` $i = 1 : j$
    $h_{ij} = q_i^T A_p q_j;\ f_{ij} = p_i^T A_p^T p_j;$
  `end`
  $r_{j+1} = A_p q_j - \Sigma_{i=1}^{j} h_{ij} q_i;$
  $s_{j+1} = A_p^T p_j - \Sigma_{i=1}^{j} f_{ij} p_i;$

$$H_j = \begin{bmatrix} \begin{matrix} H_{j-1} \\ [0 \ \cdots \ \beta] \end{matrix} & \begin{matrix} \begin{bmatrix} h_{1j} \\ \vdots \\ h_{j-1,j} \end{bmatrix} \\ h_{jj} \end{matrix} \end{bmatrix};$$

$$R_j = \left[\begin{array}{c|c} \dfrac{R_{j-1}}{[0 \cdots 0]} & H_j \begin{bmatrix} R_{j-1}(:, j-1) \\ 0 \end{bmatrix} \end{array} \right];$$

$$F_j = \begin{bmatrix} F_{j-1} & \begin{bmatrix} f_{1j} \\ \vdots \\ f_{j-1,j} \end{bmatrix} \\ [0 \quad \cdots \quad \gamma] & f_{jj} \end{bmatrix};$$

$$S_j = \left[\begin{array}{c|c} \dfrac{S_{j-1}}{[0 \cdots 0]} & F_j \begin{bmatrix} S_{j-1}(:, j-1) \\ 0 \end{bmatrix} \end{array} \right];$$

$\beta = \|r_{j+1}\|_2; \quad \gamma = \|s_{j+1}\|_2;$
$q_{j+1} = r_{j+1}/\beta; \quad p_{j+1} = s_{j+1}/\gamma;$
$Q_{j+1} = [Q_j \quad q_{j+1}]; \ P_{j+1} = [P_j \quad p_{j+1}];$
j=j+1;

}

2.4 Error Analysis and Stopping Criterion

We first analyze the error in approximating the Gramians. Consider the error $E_{k,c} = W_c - X_k X_k^T$. Clearly $E_{k,c} \geq 0$ for all k, so that $\mathbf{Tr}\,(E_{k,c})$ serves as a norm of the error. Now,

$$\begin{aligned} \mathbf{Tr}\,(E_{k,c}) &= \mathbf{Tr}\,\left(\textstyle\sum_{i=k}^{\infty} A_p^i B_p B_p^T (A_p^T)^i\right) \\ &= \textstyle\sum_{i=k}^{\infty} \|A_p^i B_p\|_2^2. \end{aligned}$$

Thus, the error converges monotonically to zero with k. Moreover,

$$\begin{aligned} \textstyle\sum_{i=k}^{\infty} \|A_p^i B_p\|_2^2 &\leq K_c \rho(A_p)^k \textstyle\sum_{i=0}^{\infty} \|A_p^i B_p\|_2^2 \\ &= K_c \rho(A_p)^k \mathbf{Tr}\,(W_c) \end{aligned}$$

for some constant K_c. Thus,

$$\mathbf{Tr}\,(W_c - X_k X_k^T) \leq K_c \rho(A_p)^k \mathbf{Tr}\,(W_c),$$

and similarly $\mathbf{Tr}\,(W_o - Y_k Y_k^T) \leq K_o \rho(A_p)^k \mathbf{Tr}\,(W_o)$,

or the relative error in the approximation depends critically on the spectral radius of A_p.

Finally, we may derive an expression for the error in the approximation of the Hankel singular values themselves. Recall that with $W_c = XX^T$ and $W_o = YY^T$, the Hankel singular values are simply the singular values σ_i of $X^T Y$. Our algorithm yields k approximate Hankel singular values $\hat{\sigma}_i$, via an SVD of $X_k^T Y_k$. Then,

$$\|X^T Y\|_F^2 = \mathbf{Tr}\,(Y^T X X^T Y) = \sum_{i=1}^{N} \sigma_i^2,$$

and $$\|X_k^T Y_k\|_F^2 = \mathbf{Tr}\,(Y_k^T X_k X_k^T Y_k) = \sum_{i=1}^{k} \hat{\sigma}_i^2.$$

Moreover, it can be shown that

$$\begin{aligned} &\mathbf{Tr}\,(Y^T X X^T Y - Y_k^T X_k X_k^T Y_k) \\ \leq\ &K \rho(A_p)^{2k} \|W_c\|_F \|W_o\|_F \end{aligned}$$

for some constant K. Once again, the approximation error depends critically on the spectral radius of A_p.

Stopping Criterion: One practical stopping criterion with both the Arnoldi and Lanczos iterations is to monitor the Frobenius norm of the product $X_k^T Y_k$, and to stop when the change is smaller than some tolerance. The quantity $\|X_k^T Y_k\|_F$ can be computed iteratively, with only matrix-vector multiplications.

Table 1: Comparison of flop counts.

Original model order	100	200	400
Average flop count with SBT	$2.25e8$	$19.3e8$	$164e8$
Average savings with ABT-Arnoldi	46×	108×	173×
Average savings with ABT-Lanczos	47×	109×	174×

3 Numerical Results

We now demonstrate the performance of the model reduction schemes described thus far on some numerical examples.

3.1 Damped Systems

We considered randomly generated stable LTI systems with twenty pairs of eigenvalues with a real part of −1, twenty pairs of eigenvalues with a real part of −2, with the remaining eigenvalues having smaller (more negative) real parts. We considered full-order models of three different sizes: 100, 200 and 400. For each size, we generated thirty different test cases, and applied our model reduction schemes to obtain an approximately balanced-and-truncated reduced order model for each of them. Table 1 shows statistics describing the performance of our algorithm. (The term "savings" in the table is the ratio of the flop count of the standard balance and truncate (SBT) model reduction scheme to the flop count of our algorithm, approximate balanced truncation with Arnoldi (ABT-Arnoldi) or approximate balanced truncation with Lanczos (ABT-Lanczos). All simulations were performed with MATLAB.) It is evident that with our algorithm, considerable computational savings accrue as compared with the standard balance-and-truncate model reduction.

In order to illustrate the error in approximation, we consider a typical test case of a full-order model with 100 states. Our algorithm yielded a reduced-order model with 21 states. Figure 1 shows the relative approximation error of the 21-state reduced-order models obtained with the standard balance-and-truncate method, balance-and-truncate with Arnoldi iterations, and balance-and-truncate with Lanczos iterations respectively. It is evident that error performance of our algorithms are comparable with that of the standard balance-and-truncate method. Figure 2 show the magnitude and phase of the system response of the original system, and that of the reduced-order systems, once again illustrating that the reduced-order model obtained from our algorithms are virtually indistinguishable from those obtained by the standard balance-and-truncate method.

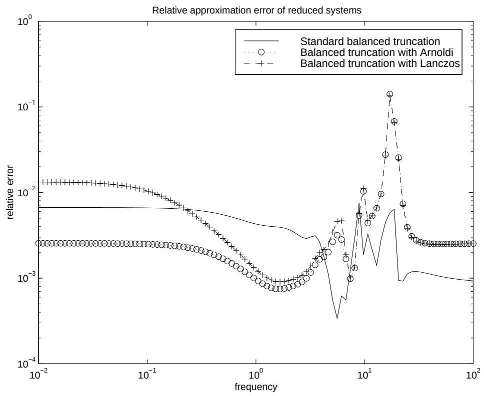

Figure 1: Relative approximation error of 21-state reduced-order models (original model order is 100.

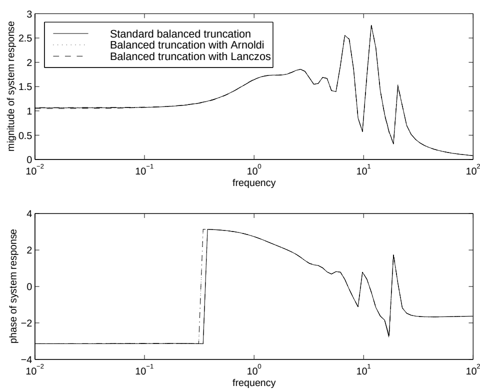

Figure 2: Relative approximation error of 21-state reduced-order models (original model order is 100).

3.2 Lightly Damped Systems

Recall that the analysis of the approximation error in Section 2.4 revealed that the quality of our low-rank approximation of the square root of the Gramian depended critically on how small the spectral radius $\rho(A_p) = \rho((pI+A)^{-1}(pI-A))$ is. When the eigenvalues of A are well-damped, as with the test cases presented thus far, the spectral radius of $\rho(A_p)$ can be made significantly less than one with an appropriate choice of p. This is one of the reasons for the remarkably good performance of our approximate balance-and-truncate schemes. For very lightly damped systems, for every choice of p, the value of $\rho(A_p)$ will be very close to one, implying that the quality of approximation with our methods should be poor with the same number of iterations. To explore this issue further, we considered a typical test case of a full order LTI system with 100 states, where we randomly generated twenty pairs of eigenvalues with a real part of $-.001$, twenty pairs of eigenvalues with a real part of $-.002$, and the remaining eigenvalues with smaller (more negative) real parts. Our algorithm yielded a reduced-order model with 44 states. An examination of the quality of approximation, shown in Figure 3, reveals the remarkable fact that over a large range of frequencies, our approximate balance-and-truncate schemes perform better than the standard balance-and-truncate scheme. A possible explanation for this is that for very lightly-damped systems, the Gramians themselves are ill-conditioned (the Lyapunov operator $\mathcal{L}(\cdot) \triangleq A^T(\cdot) + (\cdot)A$ is close to being singular), and therefore numerical errors lead to the poor performance of the standard balance-and-truncate method. In contrast our algorithms, especially the Arnoldi method, are numerically more stable. Figures 4, 5 and 6 show the magnitude and phase of the system response of the reduced systems. From these plots, it is once again evident that our algorithms perform better than the standard balanced truncation method.

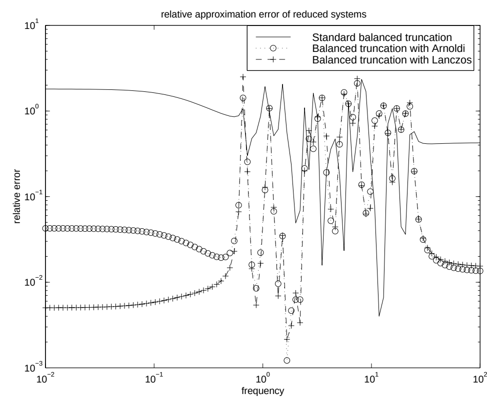

Figure 3: Relative approximation error of 44-state reduced-order models (original model order is 100).

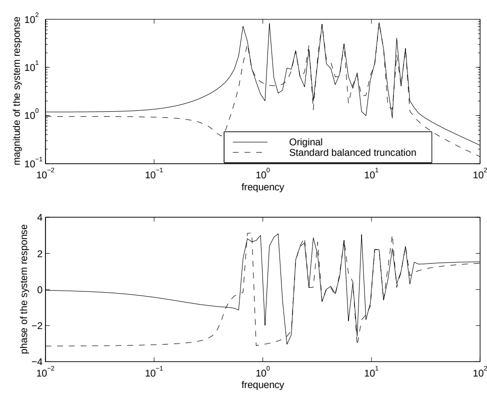

Figure 4: System response of the original 100-state lightly-damped system and the 44-state reduced-order system generated by the standard balance-and-truncate method.

4 Conclusion

We have presented efficient implementations of the balance-and-truncate model reduction technique for large-scale systems, using Krylov subspace methods.

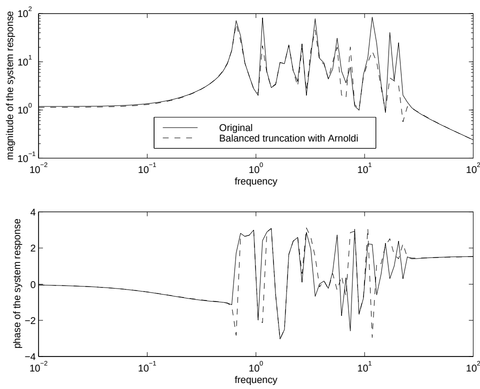

Figure 5: System response of the original 100-state lightly-damped system and the 44-state reduced-order system generated by balance-and-truncate with Arnoldi.

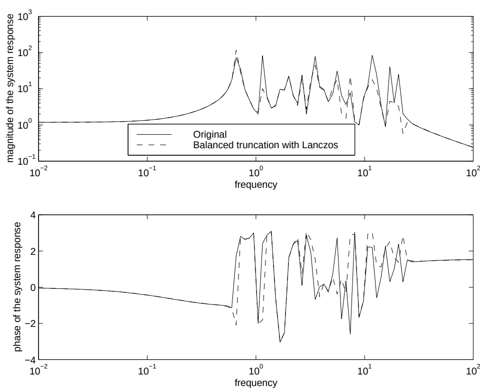

Figure 6: System response of the original 100-state lightly-damped system and the 44-state reduced-order system generated by balance-and-truncate with Lanczos.

The two distinguishing features of our algorithms are: (i) We directly compute state coordinate transformations that approximately balance-and-truncate the state vector. (ii) The coordinate transformations are computed directly from Krylov subspace methods and a small-size SVD, without the need for solving any Lyapunov equations. Numerical simulations show that our approach holds promise in the balance-and-truncate model reduction of large-scale systems.

References

[BSK00] V. Balakrishnan, Q. Su, and C-K. Koh. Efficient balance-and-truncate model reduction for large scale systems. Technical Report TR-ECE 00-15, School of Electrical and Computer Engineering, Purdue University, November 2000.

[CN92] E. Chiprout and M. Nakhla. Generalized moment-matching methods for transient analysis of interconnect networks. In *Proc. 29th ACM/IEEE Design Automation Conference*, pages 201–206, Anaheim, CA, June 1992.

[Enn84] D. F. Enns. Model reduction with balanced realizations: An error bound and a frequency weighted generalization. In *Proc. IEEE Conf. on Decision and Control*, pages 127–132, Las Vegas, NV, December 1984.

[FF95] P. Feldmann and R. W. Freund. Efficient linear circuit analysis by Padé approximation via the Lanczos process. *IEEE Trans. Computer-Aided Design*, 14:639–649, May 1995.

[Fre98] R. W. Freund. Reduced-order modeling techniques based on Krylov subspaces and their use in circuit simulation. Technical Report 98-3-02, Numerical Analysis Manuscript, Bell Laboritories, Murray Hill, New Jersey, February 1998.

[GG97] E. Grimme and K. Gallivan. Approximate solves in Krylov-based modeling methods. In *Proc. IEEE Conf. on Decision and Control*, pages 3849–3854, December 1997.

[GGD94] K. Gallivan, E. Grimme, and P. Van Dooren. Padé approximation of large-scale dynamic systems with Lanczos methods. In *Proc. IEEE Conf. on Decision and Control*, pages 443–448, December 1994.

[GL89] G. Golub and C. Van Loan. *Matrix Computations*. Johns Hopkins Univ. Press, Baltimore, second edition, 1989.

[Glo84] K. Glover. All optimal Hankel-norm approximations of linear multivariable systems and their $\mathbf{L}_\infty$-error bounds. *Int. J. Control*, 39(6):1115–1193, 1984.

[GN99] P. K. Gunupudi and M. S. Nakhla. Model-reduction of nonlinear circuits using Krylov-space techniques. In *Proc. 29th ACM/IEEE Design Automation Conference*, pages 13–16, 1999.

[JK94] I. M. Jaimoukha and E. M. Kasenally. Krylov subspace methods for solving large Lyapunov equations. *SIAM J. on Numerical Analysis*, 31(1):227–251, February 1994.

[LW91] A. Lu and E. L. Wachspress. Solution of Lyapunov equations by alternating direction implicit iteration. *Computers Math. Appl.*, 21(9):43–58, 1991.

[LW99] J. Li and J. White. Efficient model reduction of interconnnect via approximate system Gramians. In *Proc. Int. Conf. on Computer-Aided Design*, pages 380–383, 1999.

[Moo81] B. C. Moore. Principal component analysis in linear systems: Controllability, observability, and model reduction. *IEEE Trans. Aut. Control*, AC-26(1):17–32, February 1981.

[RP99] P. Rabiei and M. Pedram. Model order reduction of large circuits using balanced truncation. In *Design Automation Conference, Asia and South Pacific, In Proc. ASP-DAC, 99*, pages 237–240, 1999.

[Smi68] R. A. Smith. Matrix equation $XA + BX = C$. *SIAM Journal on Applied Mathematics*, 16(1):198–201, 1968.