A survey of model reduction methods for large-scale systems*†

A.C. Antoulas‡, D.C. Sorensen§, and S. Gugercin‡

e-mail: {aca,sorensen,serkan}@rice.edu

October 27, 2006

Abstract

An overview of model reduction methods and a comparison of the resulting algorithms is presented. These approaches are divided into two broad categories, namely SVD based and moment matching based methods. It turns out that the approximation error in the former case behaves better globally in frequency while in the latter case the local behavior is better.

1 Introduction and problem statement

Direct numerical simulation of dynamical systems has been an extremely successful means for studying complex physical phenomena. However, as more detail is included, the dimensionality of such simulations may increase to unmanageable levels of storage and computational requirements. One approach to overcoming this is through model reduction. The goal is to produce a low dimensional system that has the same response characteristics as the original system with far less storage requirements and much lower evaluation time. The resulting reduced model might be used to replace the original system as a component in a larger simulation or it might be used to develop a low dimensional controller suitable for real time applications.

The *model reduction problem* we are interested in can be stated as follows. Given is a linear dynamical system in state space form:

$$\mathcal{S}: \quad \begin{cases} \sigma x(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) + Du(t) \end{cases} \tag{1.1}$$

where σ is either the derivative operator $\sigma f(t) = \frac{d}{dt} f(t)$, $t \in \mathbb{R}$, or the shift $\sigma f(t) = f(t+1)$, $t \in \mathbb{Z}$, depending on whether the system is continuous- or discrete-time. For simplicity we will use the notation:

$$\mathcal{S} = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathbb{R}^{(n+p)\times(n+m)} \tag{1.2}$$

The *problem* consists in approximating $\mathcal{S}$ with:

$$\hat{\mathcal{S}} = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right] \in \mathbb{R}^{(k+p)\times(k+m)} \tag{1.3}$$

*This work was supported in part by the NSF Cooperative Agreement CCR-9120008 and by the NSF Grant DMS-9972591.

†A preliminary version was presented at the AMS-IMS-SIAM Summer Research Conference on Structured Matrices, Boulder, June 27 - July 1, 1999.

‡Deptartment of Electrical and Computer Engineering, MS 380, Rice University, Houston, Texas 77251-1892.

§Department of Computational and Applied Mathematics, MS 134, Rice University, Houston, Texas 77251-1892.

where $k \ll n$ such that the following properties are satisfied:

1. The approximation error is *small*, and there exists a *global* error bound.

2. System properties, like *stability*, *passivity*, are preserved.

3. The procedure is *computationally stable* and *efficient*.

There are two sets of methods which are currently in use, namely

(a) SVD based methods and

(b) moment matching based methods.

One commonly used approach is the so-called Balanced Model Reduction first introduced by Moore [19], which belongs to the former category. In this method, the system is transformed to a basis where the states which are difficult to reach are simultaneously difficult to observe. Then, the reduced model is obtained simply by truncating the states which have this property. Two other closely related model reduction techniques are Hankel Norm Approximation [20] and the Singular Perturbation Approximation [16], [18]. When applied to stable systems, all of these three approaches are guaranteed to preserve stability and provide bounds on the approximation error. Recently much research has been done to establish connections between Krylov subspace projection methods used in numerical linear algebra and model reduction [8], [10], [12], [13], [15], [24], [11]; consequently The implicit restarting algorithm [23] has been applied to obtain stable reduced models [14].

Issues arising in the approximation of large systems are: **storage**, **computational speed**, and **accuracy**. In general storage and computational speed are finite and problems are *ill-conditioned*. *In addition*: we need global error bounds and preservation of stability/passivity. SVD based methods have provide error bounds and preserve stability, but are computationally not efficient. On the other hand, moment matching based methods can be implemented in a numerically efficient way, but do not automatically preserve stability and have no global error bounds. To remedy this situation, the Approximate Balancing method was introduced in [8]. It attempts to combine all requite properties by iteratively computing a reduced order approximate balanced system.

The paper is organized as follows. After the problem definition, the first part is devoted to approximation methods which are related to the SVD. Subsequently, moment matching methods are reviewed. The third part of the paper is devoted to a comparison of the resulting seven algorithms applied on six dynamical systems of low to moderate complexity. We conclude with unifying features, and complexity considerations.

Contents

1 Introduction and problem statement **1**

2 Approximation in the 2-norm: SVD-based methods **4**
2.1 The singular value decomposition: static systems . . . 4
2.2 SVD methods applied to dynamical systems . . . 5
2.2.1 Proper Orthogonal Decomposition (POD) methods . . . 5
2.2.2 Optimal approximation of linear systems . . . 6
2.2.3 Optimal and suboptimal approximation in the 2-norm . . . 7
2.2.4 Approximation by balanced truncation . . . 7
2.2.5 Singular Perturbation Approximation . . . 8

3 Approximation by moment matching **8**
3.1 The Lanczos procedure . . . 9
3.2 The Arnoldi procedure . . . 10
3.3 The algorithms . . . 11
3.3.1 The Lanczos algorithm: recursive implementation . . . 11
3.3.2 The Arnoldi algorithm: recursive implementation . . . 11
3.3.3 Implicitly restarted Arnoldi and Lanczos methods . . . 12
3.3.4 The Rational Krylov Method . . . 12
3.4 A new approach: The cross grammian . . . 12
3.4.1 Description of the solution . . . 13
3.4.2 A stopping criterion . . . 14
3.4.3 Computation of the solution . . . 15

4 Application of the reduction algorithms **16**
4.1 Structural Model . . . 17
4.2 Building Model . . . 18
4.3 Heat diffusion model . . . 19
4.4 CD Player . . . 21
4.5 Clamped Beam Model . . . 22
4.6 Low-Pass Butterworth Filter . . . 23

5 Projectors and computational complexity **24**

6 Conclusions **25**

List of Figures

1 Approximation of clown image . . . 5
2 Normalized Hankel Singular Values of (a) Heat Model, Butterworth Filter , Clamped Beam Model and Structural Model; (b) Butterworth Filter, Building Example, CD Player. . . . 16
3 Relative degree reduction $\frac{k}{n}$ vs error tolerance $\frac{\sigma_k}{\sigma_1}$ . . . 17
4 $\sigma_{\max}$ of the frequency response of the (a) Reduced and (b) Error Systems of Structural Model . . . 18
5 Nyquist plots of the full order and reduced order models for the (a) Structral (b) Building Model . . . 18
6 $\sigma_{\max}$ of the frequency response of the reduced systems of Building Model . . . 19
7 (a)-(b) $\sigma_{\max}$ of the frequency response of the error systems of Building Model . . . 20
8 $\sigma_{\max}$ of the frequency response of the (a) Reduced and (b) Error Systems of Heat diffusion model . . . 21
9 Nyquist plots of the full order and reduced order models for the (a) Heat Model (b) CD Player . . . 22
10 $\sigma_{\max}$ of the frequency response of the (a) Reduced and (b) Error Systems of CD Player . . . 23
11 $\sigma_{\max}$ of the frequency response of the (a) Reduced and (b) Error Systems of Clamped Beam . . . 24
12 Nyquist plots of the full order and reduced order models for the (a) Clamped Beam Model (b) Butterworth Filter . . . 25
13 $\sigma_{\max}$ of the frequency response of the (a) Reduced and (b) Error Systems of Butterworth Filter . . . 26

2 Approximation in the 2-norm: SVD-based methods

2.1 The singular value decomposition: static systems

Given a matrix $A \in \mathbb{R}^{n\times m}$, its *Singular Value Decomposition* (SVD) is defined as follows:

$$A = U\Sigma V^*,\ \Sigma = \mathrm{diag}\,(\sigma_1, \cdots, \sigma_n) \in \mathbb{R}^{n\times m},$$

where $\sigma_1(A) \geq \cdots \geq \sigma_n(A) \geq 0$: are the *singular values*; recall that $\sigma_1(A)$ is the 2-*induced norm* of A. Furthermore, the columns of the orthogonal matrices $U = (u_1 \quad u_2 \quad \cdots \quad u_n)$, $UU^* = I_n$, $V = (v_1 \quad v_2 \quad \cdots \quad v_m)$, $VV^* = I_m$, are the *left, right singular vectors* of A, respectively. Assuming that $\sigma_r > 0$, $\sigma_{r+1} = 0$, implies that the rank of A is r. Finally, the SVD induces a *dyadic decomposition* of A:

$$A = \sigma_1 u_1 v_1^* + \sigma_2 u_2 v_2^* + \cdots \sigma_r u_r v_r^*$$

Given $A \in \mathbb{R}^{n\times m}$ with $\mathrm{rank}\, A = r \leq n \leq m$, we seek to find $X \in \mathbb{R}^{n\times m}$ with $\mathrm{rank}\, X = k < r$, such that the 2-norm of the *error* $E := A - X$ is minimized.

Theorem 2.1 Schmidt-Mirsky: Optimal approximation in the 2 norm. *Provided that $\sigma_k > \sigma_{k+1}$, there holds:* $\min_{\mathrm{rank} X \leq k} \parallel A - X \parallel_2 = \sigma_{k+1}(A)$. *A (non-unique) minimizer X_* is obtained by truncating the dyadic decomposition: $X_* = \sigma_1 u_1 v_1^* + \sigma_2 u_2 v_2^* + \cdots \sigma_k u_k v_k^*$.*

Next, we address the sub-optimal approximation in the 2 norm, namely: find all matrices X of rank k satisfying

$$\sigma_{k+1}(A) < \parallel A - X \parallel_2 < \epsilon < \sigma_k(A) \tag{2.1}$$

First, we notice that there exist matrices Λ_i, $i = 1, 2$, such that $I_n - \epsilon^{-2} AA^* = \Lambda_1 \Lambda_1^* - \Lambda_2 \Lambda_2^*$, and $\mathrm{rank}\,(\Lambda_1) + \mathrm{rank}\,(\Lambda_2) = \mathrm{rank}\,(A)$. These relationships imply the existence of a J-unitary matrix Θ, such that: $[I_n \quad \epsilon^{-1} T]\Theta = [\Lambda_2 \quad \Lambda_1]$, where

$$\Theta J \Theta^* = J,\ J = \begin{pmatrix} I_n & 0 \\ 0 & -I_n \end{pmatrix},\ \Theta = \begin{pmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & \Theta_{22} \end{pmatrix} \in \mathbb{R}^{2n\times 2n}$$

We now define: $E_i \in \mathbb{R}^{n\times n}$, $i = 1, 2$, $\Delta \in \mathbb{R}^{n\times n}$, as follows:

$$\begin{pmatrix} E_1 \\ E_2 \end{pmatrix} = \Theta \begin{pmatrix} \Delta \\ I_n \end{pmatrix} = \begin{pmatrix} \Theta_{11}\Delta + \Theta_{12} \\ \Theta_{21}\Delta + \Theta_{22} \end{pmatrix}$$

Theorem 2.2 Sub-optimal approximation in the 2 norm. *$\hat{X}$ is a sub-optimal approximant satisfying (2.1) iff there exists a contraction Δ such that: $A - \hat{X} = E := E_1 E_2^{-1}$, where $\parallel \Delta \parallel_2 < 1$, and* $\mathrm{rank}\,(\Lambda_1 + \Lambda_2 \Delta) = k$.

It should be noticed that for a particular choice of Λ_1 and Λ_2, the rank condition above can be converted to the block triangularity of the contraction Δ.

Clown approximation. The above approximation method is applied to the *clown* image shown in figure 2.1, which can be found in matlab. This is a 320×200 pixel image; each pixel has 64 levels of gray. First, the 200 singular values of this 2-dimensional array are computed (see upper right-hand side subplot of the figure); the singular values drop-off rapidly making a low-order approximation with small error, possible. The optimal approximants for rank $k = 1, 3, 10, 30$ are shown. Notice that the storage reduction of a rank k approximant is $(n + m + 1) * k$ compared to $n * m$ for the original image.

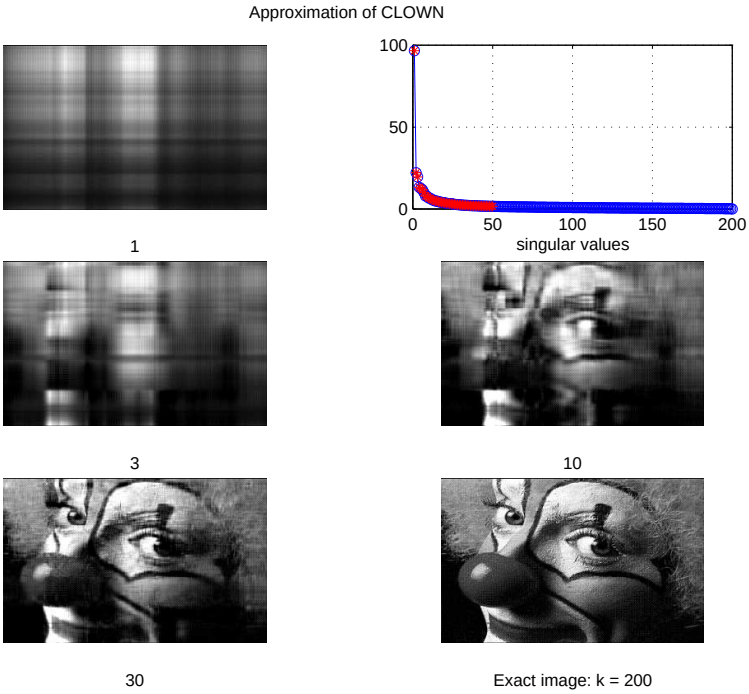


Figure 1: Approximation of clown image

Nonlinear systems	Linear systems
POD methods	Hankel-norm approximation
	Balanced truncation
	Singular perturbation
	New method (Cross grammian)

Table 1: SVD based approximation methods

2.2 SVD methods applied to dynamical systems

There are different ways of applying the SVD to the approximation of dynamical systems. The table below summarizes the different approaches.

Its application to non-linear systems is known under the name *POD*, that is: Proper Orthogonal Decomposition. Then in the linear case we can make use of additional structure. The result corresponding to a generalization of the Schmidt-Mirsky theorem is known under the name of Hankel norm approximation. Closely related methods are approximation by balanced truncation and approximation by singular perturbation. Finally, the new method proposed in section 3.4 is based on the SVD in a different way.

2.2.1 Proper Orthogonal Decomposition (POD) methods

Consider the nonlinear system described by $\dot{x}(t) = f(x(t), u(t))$; let

$$\mathcal{X} = [x(t_1) \;\; x(t_2) \;\; \cdots \;\; x(t_N)] \in \mathbb{R}^{n \times N}$$

be a collection of snapshots of the solution of this system. We compute the singular value decomposition and truncate depending on how fast the singular values decay:

$$\mathcal{X} = U\Sigma V^* \approx U_k \Sigma_k V_k^*, \quad k \ll n$$

Let $x(t) \approx U_k \xi(t)$, $\xi(t) \in \mathbb{R}^k$. Thus the approximation $\xi(t)$ of the state $x(t)$ evolves in a *low-dimensional* space. Then $U_k \dot{\xi}(t) = f(U_k \xi(t), u(t))$, which implies the reduced order state equation:

$$\dot{\xi}(t) = U_k^* f(U_k \xi(t), u(t))$$

2.2.2 Optimal approximation of linear systems

Consider the following *Hankel operator*:

$$\mathcal{H} = \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \cdots \\ \alpha_2 & \alpha_3 & \alpha_4 & \cdots \\ \alpha_3 & \alpha_4 & \alpha_5 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} : \ \ell(\mathbb{Z}_+) \longrightarrow \ell(\mathbb{Z}_+)$$

It is assumed that $\operatorname{rank} \mathcal{H} = n < \infty$, which is equivalent with the rationality of the (formal) power series:

$$\sum_{t>0} \alpha_t z^{-t} = \frac{\pi(z)}{\chi(z)} =: G_{\mathcal{H}}(z), \ \deg \chi = n > \deg \pi$$

It is well known that in this case $G_{\mathcal{H}}$ possesses a state-space realization denoted by (1.2):

$$G_{\mathcal{H}}(z) = \frac{\pi(z)}{\chi(z)} = C(zI - A)^{-1} B, \ A \in \mathbb{R}^{n \times n}, \ B, C^* \in \mathbb{R}^n$$

This is a discrete-time system; thus the eigenvalues of A (roots of χ) lie inside the unit disc if and only if $\sum_{t>0} \mid \alpha_t \mid^2 < \infty$.

The problem which arises now is to approximate $\mathcal{H}$ by a Hankel operator $\hat{\mathcal{H}}$ of lower rank, *optimally* in the 2-induced norm. The system-theoretic interpretation of this problem is to optimally approximate the linear system described by $G = \frac{\pi}{\chi}$, by a system of lower complexity, $\hat{G} = \frac{\hat{\pi}}{\hat{\chi}}$, $\deg \chi > \deg \hat{\chi}$. This is the problem of *approximation in the Hankel norm*.

First we note that $\mathcal{H}$ is bounded and compact and hence possesses a discrete set of non-zero singular values with an accumulation point at zero:

$$\sigma_1(\mathcal{H}) \geq \sigma_2(\mathcal{H}) \geq \cdots \geq \sigma_n(\mathcal{H}) > 0$$

These are called the *Hankel singular values* of $\mathcal{S}$. By the Schmidt-Mirsky result, any approximant $\mathcal{K}$, not necessarily structured, of rank $k < n$ satisfies:

$$\| \mathcal{H} - \mathcal{K} \|_2 \geq \sigma_{k+1}(\mathcal{H})$$

The question which arises is whether *there exist an approximant of rank k which has Hankel structure and achieves the lower bound*. In system-theoretic terms we seek a low order approximant $\hat{\mathcal{S}}$ to $\mathcal{S}$. The question has an affirmative answer.

Theorem 2.3 Adamjan, Arov, Krein (AAK). *There exists a unique approximant $\hat{\mathcal{H}}$ of rank k, which has Hankel structure and attains the lower bound: $\sigma_1(\mathcal{H} - \hat{\mathcal{H}}) = \sigma_{k+1}(\mathcal{H})$.*

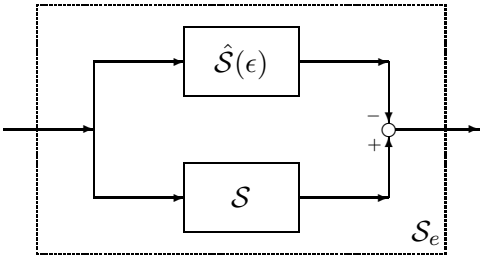
The above result holds for continuous-time systems as well. In this case the discrete-time Hankel operator introduced above is replaced by the continuous-time Hankel operator defined as follows: $y(t) = \mathcal{H}(u)(t) = \int_{-\infty}^{0} h(t - \tau) u(\tau) d\tau$, $t > 0$, where $h(t) = Ce^{At}B$ is the impulse response of the system.

2.2.3 Optimal and suboptimal approximation in the 2-norm

It turns out that both *sub-optimal* and *optimal* approximants of MIMO (multi-input, multi-output) linear continuous- and discrete-time systems can be treated within the same framework. The problem is thus: given a stable system $\mathcal{S}$, we seek approximants $\mathcal{S}_*$ satisfying

$$\sigma_{k+1}(\mathcal{S}) \leq \|\mathcal{S} - \mathcal{S}_*\|_H \leq \epsilon < \sigma_k(\mathcal{S}) \tag{2.2}$$

This is accomplished by the following construction.



Construction of approximants

Given $\mathcal{S}$, construct $\hat{\mathcal{S}}$ such that $\mathcal{S}_e := \mathcal{S} - \hat{\mathcal{S}}$ has norm ϵ, and is all-pass; in this case $\hat{\mathcal{S}}$ is called an *ϵ-all-pass dilation* of $\mathcal{S}$. The following result holds.

Theorem 2.4 *Let $\hat{\mathcal{S}}$ be an ϵ-all-pass dilation of the linear, stable, discrete- or continuous-time system $\mathcal{S}$. The stable part $\hat{\mathcal{S}}_+$ of $\mathcal{S}$ has exactly k stable poles and the inequalities (2.2) hold. Furthermore, if $\sigma_{k+1}(\mathcal{S}) = \epsilon$, $\sigma_{k+1}(\mathcal{S}) = \|\mathcal{S} - \hat{\mathcal{S}}\|_H$.*

Computation. The Hankel singular values of $\mathcal{S}$ given by (1.2), can be computed by solving two Lyapunov equations in finite dimensions. For continuous-time systems, let $\mathcal{P}$, $\mathcal{Q}$ be the system grammians:

$$A\mathcal{P} + \mathcal{P}A^* + BB^* = 0, \quad A^*\mathcal{Q} + \mathcal{Q}A + C^*C = 0 \tag{2.3}$$

It follows that

$$\sigma_i(\mathcal{S}) = \sqrt{\lambda_i(\mathcal{P}\mathcal{Q})} \tag{2.4}$$

The *error bound* for optimal approximants, in the 2-norm of the *convolution operator*:

$$\sigma_{k+1} \leq \| \mathcal{S} - \hat{\mathcal{S}} \|_\infty \leq 2(\sigma_{k+1} + \cdots + \sigma_n) \tag{2.5}$$

where the $\mathcal{H}_\infty$ norm is maximum of the largest singular value of the frequency response, or alternatively, the 2-induced norm of the *convolution operator*, namely: $y(t) = \mathcal{S}(u)(t) = \int_{-\infty}^{\infty} h(t-\tau)u(\tau)d\tau$, $t \in \mathbb{R}$, where $h(t) = Ce^{At}B$. For details on these issues we refer to [7].

2.2.4 Approximation by balanced truncation

A linear system $\mathcal{S}$ in state space form is called *balanced* if the solutions of the two grammians (2.3) are equal and diagonal:

$$\mathcal{P} = \mathcal{Q} = \Sigma = \mathrm{diag}\,(\sigma_1, \cdots, \sigma_n) \tag{2.6}$$

It turns our that every controllable and observable system can be transformed to balanced form by means of a basis change $\hat{x} = Tx$. Let $\mathcal{P} = UU^*$ and $\mathcal{Q} = LL^*$ where U and L are upper and lower triangular matrices respectively. Let also $U^*L = ZSY^*$ be the singular value decomposition (SVD) of U^*L. A the balancing transformation is $T = S^{\frac{1}{2}}Z^*U^{-1} = S^{-\frac{1}{2}}Y^*L^*$. Let $\mathcal{S}$ be balanced with grammians equal to

$\Sigma = \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{pmatrix}$, where $\Sigma_1 \in \mathbb{R}^{k\times k}$, and Σ_2 contains all the small Hankel singular values. Partition conformally the system matrices:

$$\mathcal{S} = \left[\begin{array}{cc|c} A_{11} & A_{12} & B_1 \\ A_{21} & A_{22} & B_2 \\ \hline C_1 & C_2 & D \end{array}\right] \text{ where } A_{11} \in \mathbb{R}^{k\times k},\ B_1 \in \mathbb{R}^{k\times m},\ C_1 \in \mathbb{R}^{p\times k} \tag{2.7}$$

The system $\hat{\mathcal{S}} := \left[\begin{array}{c|c} A_{11} & B_1 \\ \hline C_1 & \end{array}\right]$, is a reduced order system obtained by *balanced truncation*. This system has the following guaranteed properties: (a) stability is preserved, and (b) the same error bound (2.5) holds as in Hankel-norm approximation.

2.2.5 Singular Perturbation Approximation

A closely related approximation method, is the so-called *singular perturbation* approximation. It is based on the balanced form presented above. Thus, let (2.7) hold; the reduced order model is given by

$$\hat{\mathcal{S}} = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right] = \left[\begin{array}{c|c} A_{11} - A_{12}A_{22}^{-1}A_{21} & B_1 - A_{12}A_{22}^{-1}B_2 \\ \hline C_1 - C_2A_{22}^{-1}A_{21} & D - C_2A_{22}^{-1}B_2 \end{array}\right]. \tag{2.8}$$

Again, the same guaranteed properties as for the approximation by balanced truncation, are satisfied.

3 Approximation by moment matching

Given a linear system $\mathcal{S}$ in state space form (1.2), its transfer function $G(s) = C(sI - A)^{-1}B + D$, is expanded in a Laurent series around a given point $s_0 \in \mathbb{C}$ in the complex plane:

$$G(s_0 + \sigma) = \eta_0 + \eta_1\sigma + \eta_2\sigma^2 + \eta_3\sigma^3 + \cdots$$

The η_t are called the *moments* of $\mathcal{S}$ at s_0. We seek a reduced order system $\hat{\mathcal{S}}$ as in (1.3), such that the Laurent expansion of the corresponding transfer function at s_0 has the form

$$\hat{G}(s_0 + \sigma) = \hat{\eta}_0 + \hat{\eta}_1\sigma + \hat{\eta}_2\sigma^2 + \hat{\eta}_3\sigma^3 + \cdots$$

where k moments are matched:

$$\eta_j = \hat{\eta}_j,\ j = 1, 2, \cdots, k$$

for appropriate $k \ll n$. If s_0 is infinity, the moments are called *Markov parameters*; the corresponding problem is known as *partial realization*, or *Padé approximation*; the solution of this problem can be found in [1], [4]. Importantly, the solution of these problems can be implemented in a numerically stable and efficient way, by means of the *Lanczos* and *Arnoldi* procedures. For arbitrary $s_0 \in \mathbb{C}$, the problem is known as *rational interpolation*, see e.g. [2], [3]. A numerically efficient solution is given by means of the *rational Lanczos/Arnoldi* procedures.

Recently, there has been renewed interest in moment matching and projection methods for model reduction in LTI systems. Three leading efforts in this area are Padé via Lanczos (PVL) [10], multi-point rational interpolation [12], and implicitly restarted dual Arnoldi [15].

The PVL approach exploits the deep connection between the (nonsymmetric) Lanczos process and classic moment matching techniques. The multi-point rational interpolation approach utilizes the rational Krylov method of Ruhe [22] to provide moment matching of the transfer function at selected frequencies

and hence to obtain enhanced approximation of the transfer function over a broad frequency range. These techniques have proven to be very effective. PVL has enjoyed considerable success in circuit simulation applications. Rational interpolation achieves remarkable approximation of the transfer function with very low order models. Nevertheless, there are shortcomings to both approaches. In particular, since the methods are local in nature, it is difficult to establish rigorous error bounds. Heuristics have been developed that appear to work, but no global results exist. Secondly, the rational interpolation method requires selection of interpolation points. At present, this is not an automated process

and relies on ad-hoc specification by the user.

In [15] an implicitly restarted dual Arnoldi approach is described. The dual Arnoldi method runs two separate Arnoldi processes, one for the controllability subspace, and the other for the observability subspace and then constructs an oblique projection from the two orthogonal Arnoldi basis sets. The basis sets and the reduced model are updated using a generalized notion of implicit restarting. The updating process is designed to iteratively improve the approximation properties of the model. Essentially, the reduced model is reduced further, keeping the best features, and then expanded via the dual Arnoldi processes to include new information. The goal is to achieve approximation properties similar to those of balanced truncation. Other related approaches [9, 17, 21] work directly with projected forms of the two Lyapunov equations (2.3) to obtain low rank approximations to the system Grammians.

In the sequel we will review the Lanczos and Arnoldi procedures. We will also review the concept of implicit restarting. For simplicity only the scalar (SISO) versions will be discussed.

3.1 The Lanczos procedure

Given is the scalar system $\mathcal{S}$ as in (1.2) with $m = p = 1$, we seek to find $\hat{\mathcal{S}}$ as in (1.3), $k < n$, such that the first $2k$ *Markov parameters* $\eta_i = CA^{i-1}B$, of $\mathcal{S}$, and $\hat{\eta}_i := \hat{C}\hat{A}^{i-1}\hat{B}$, of $\hat{\mathcal{S}}$, are matched: $\eta_i = \hat{\eta}_i,\ i = 1,\ \cdots, 2k$. We will solve this problem following a non-conventional path with system-theoretic flavor; the Lanczos factorization in numerical analysis is introduced using a different set of arguments. First, the *observability* matrix $\mathcal{O}_t$, and the *reachability* matrix $\mathcal{R}_t$ are defined:

$$\mathcal{O}_t = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{t-1} \end{bmatrix} \in \mathbb{R}^{t\times n},\ \mathcal{R}_t = \begin{bmatrix} B & AB & \cdots & A^{t-1}B \end{bmatrix} \in \mathbb{R}^{n\times t}$$

Secondly, we define the $t \times t$ Hankel matrix, and its shift:

$$\mathcal{H}_t := \begin{bmatrix} \eta_1 & \eta_2 & \cdots & \eta_t \\ \eta_2 & \eta_3 & \cdots & \eta_{t+1} \\ \vdots & & \ddots & \\ \eta_t & \eta_{t+1} & \cdots & \eta_{2t-1} \end{bmatrix},\ \sigma\mathcal{H}_t := \begin{bmatrix} \eta_2 & \eta_3 & \cdots & \eta_{t+1} \\ \eta_3 & \eta_4 & \cdots & \eta_{t+2} \\ \vdots & & \ddots & \\ \eta_{t+1} & \eta_{t+2} & \cdots & \eta_{2t} \end{bmatrix}$$

It follows that $\mathcal{H}_t = \mathcal{O}_t\mathcal{R}_t$ and $\sigma\mathcal{H}_t = \mathcal{O}_tA\mathcal{R}_t$. The *key step* is as follows: assuming that $\det\mathcal{H}_i \neq 0$, $i = 1, \cdots, k$, we compute the LU factorization of $\mathcal{H}_k$:

$$\mathcal{H}_k = LU$$

with $L(i,j) = 0$, $i < j$, $U(i,j) = 0$, $i > j$, and $L(i,i) = \pm U(i,i)$. Define the maps:

$$\pi_L := L^{-1}\mathcal{O}_k \quad \text{and} \quad \pi_U := \mathcal{R}_kU^{-1} \tag{3.1}$$

Clearly, the following properties hold: (a) $\pi_L\pi_U = 1$, and (b) $\pi_U\pi_L$: orthogonal projection. The **reduced order system** $\hat{\mathcal{S}}$, is now defined as follows:

$$\hat{A} := \pi_LA\pi_U,\ \hat{B} = \pi_LB,\ \hat{C} = C\pi_U \tag{3.2}$$

Theorem 3.1 *$\hat{\mathcal{S}}$ as defined above matches $2k$ Markov parameters. Furthermore, $\hat{A}$ is tridiagonal, and $\hat{B}, \hat{C}^*$ are multiples of the unit vector e_1.*

3.2 The Arnoldi procedure

As in the Lanczos case, we will derive the Arnoldi factorization following a non-conventional path, which is different from the path usually adopted by numerical analysts in this case. Let $\mathcal{R}_n = [B \;\; AB \;\cdots\; A^{n-1}B]$ with $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^n$. Then:

$$A\mathcal{R}_n = \mathcal{R}_n F \quad \text{where} \quad F = \begin{pmatrix} 0 & 0 & \cdots & 0 & -\alpha_0 \\ 1 & 0 & \cdots & 0 & -\alpha_1 \\ 0 & 1 & \cdots & 0 & -\alpha_2 \\ & & \ddots & & \\ 0 & 0 & \cdots & 1 & -\alpha_{n-1} \end{pmatrix}$$

and $\chi_A(s) = \det(sI - A) = s^n + \alpha_{n-1}s^{n-1} + \cdots + \alpha_1 s + \alpha_0$. The *key step* in this case consists in computing the QR factorization of $\mathcal{R}_n$:

$$\mathcal{R}_n = VU$$

where V is orthogonal and U is upper triangular. It follows that $AVU = VUF$, $AV = VUFU^{-1}$, which in turn with $\bar{F} = UFU^{-1}$, implies that $AV = V\bar{F}$; thereby U, U^{-1} are upper triangular, F is upper Hessenberg, and therefore $\bar{F}$: is *upper Hessenberg*. This yields the k-step Arnoldi factorization:

$$[AV]_k = \left[V\bar{F}\right]_k \quad \Rightarrow \quad A[V]_k = [V]_k \bar{F}_{kk} + fe_k^*$$

where f is a multiple of the $k+1$-st column of V; $\bar{F}_{kk}$, is still upper Hessenberg, and the columns of $[V]_k$ provide an orthonormal basis for the space spanned by the first k columns of $\mathcal{R}_n$.

Recall that $\mathcal{H}_k = \mathcal{O}_k\mathcal{R}_k$; a projection π can be attached to the QR factorization of $\mathcal{R}_k$:

$$\pi := \mathcal{R}_k U^{-1} = V, \quad V \in \mathbb{R}^{n\times k}, \quad V^*V = I_k, \quad U: \text{ upper triangular} \tag{3.3}$$

The *reduced order system* $\hat{\mathcal{S}}$ is now defined as follows:

$$\hat{A} := \pi^* A \pi, \;\; \hat{B} = \pi^* B, \;\; \hat{C} = C\pi \tag{3.4}$$

Theorem 3.2 *$\hat{\mathcal{S}}$ matches k Markov parameters: $\hat{\eta}_i = \eta_i$, $i = 1, \cdots, k$. Furthermore, $\hat{A}$ is in Hessenberg form, and $\hat{B}$ is a multiple of e_1.*

Remarks. **(a)** Number of operations needed to compute $\hat{\mathcal{S}}$ using Lanczos or Arnoldi is $O(k^2 n)$, vs. $O(n^3)$ operations needed for the other methods. Only *matrix-vector* multiplications are required as opposed to matrix factorizations and/or inversions.

(b) Drawback Lanczos: it breaks down if $\det \mathcal{H}_i = 0$, for some $1 \le i \le n$. The remedy in this case are *look-ahead* methods. Arnoldi breaks down if $\mathcal{R}_i$ does not have full rank; this happens less frequently.

(c) $\hat{\mathcal{S}}$ tends to approximate the high frequency poles of $\mathcal{S}$. Hence the steady-state error may be significant. Remedy: match expansions around other frequencies. This leads to *rational Lanczos*.

(d) $\hat{\mathcal{S}}$ may not be stable, even if $\mathcal{S}$ is stable. Remedy: *implicit restart* of Lanczos and Arnoldi. ■

3.3 The algorithms

3.3.1 The Lanczos algorithm: recursive implementation

Given: the triple $A \in \mathbb{R}^{n\times n}$, $B, C^* \in \mathbb{R}^n$, **find**: $V, W \in \mathbb{R}^{n\times k}$, $f, g \in \mathbb{R}^n$, and $K \in \mathbb{R}^{k\times k}$, such that

$$AV = VK + fe_k^*,\ A^*W = WK^* + ge_k^*,\ \text{where}$$
$$K = V^*AW,\ V^*W = I_k,\ W^*f = 0,\ V^*g = 0$$

where e_k denotes the k^{th} unit vector in $\mathbb{R}^n$. The projections π_L, π_U defined above are given by V^*, W.

Two-sided Lanczos algorithm

1. $\beta_1 := \sqrt{|CB|}$, $\gamma_1 := \operatorname{sgn}(CB)\beta_1$
   $v_1 := B/\beta_1$, $w_1 := C^*/\gamma_1$
2. For $j = 1, \cdots, k$, set
   (a) $\alpha_j := w_j^* A v_j$
   (b) $r_j := Av_j - \alpha_j v_j - \gamma_j v_{j-1}$, $q_j := A^* w_j - \alpha_j w_j - \beta_j w_{j-1}$
   (c) $\beta_{j+1} = \sqrt{|r_j^* q_j|}$, $\gamma_{j+1} = \operatorname{sgn}(r_j^* q_j)\beta_{j+1}$
   (d) $v_{j+1} = r_j/\beta_{j+1}$, $w_{j+1} = q_j/\gamma_{j+1}$

The following relationships hold: $V_k = (v_1\ \ v_2\ \cdots\ v_k)$, $W_k = (w_1\ \ w_2\ \cdots\ w_k)$, where $AV_k = V_k K_k + \beta_{k+1} v_{k+1} e_k^*$, $A^*W_k = W_k K_k^* + \gamma_{k+1} w_{k+1} e_k^*$, and $K_k = \begin{pmatrix} \alpha_1 & \gamma_2 & & \\ \beta_2 & \alpha_2 & \ddots & \\ & \ddots & \ddots & \gamma_k \\ & & \beta_k & \alpha_k \end{pmatrix}$, $r_k \in \mathcal{R}_{k+1}(A, B)$, $q_k^* \in \mathcal{O}_{k+1}(C, A)$.

3.3.2 The Arnoldi algorithm: recursive implementation

Given: the triple $A \in \mathbb{R}^{n\times n}$, $B, C^* \in \mathbb{R}^n$, **find**: $V \in \mathbb{R}^{n\times k}$, $f \in \mathbb{R}^n$, and $K \in \mathbb{R}^{k\times k}$, such that

$$AV = VK + fe_k^*,\ \text{where}$$
$$K = V^*AV,\ V^*V = I_k,\ V^*f = 0$$

where K is in *upper Hessenberg* form. The projection π defined above is given by V.

The Arnoldi algorithm

1. $v_1 := \frac{v}{\|v\|}$, $w := Av_1$; $\alpha_1 := v_1^* w$
   $f_1 := w - v_1\alpha_1$; $V_1 := (v_1)$; $K_1 := (\alpha_1)$
2. For $j = 1, 2,\ \cdots\ , k-1$
   $\beta_j :=\| f_j \|$, $v_{j+1} := \frac{f_j}{\beta_j}$
   $V_{j+1} := (V_j\ \ v_{j+1})$, $\hat{K}_j = \begin{pmatrix} K_j \\ \beta_j e_j^* \end{pmatrix}$
   $w := Av_{j+1}$, $h := V_{j+1}^* w$, $f_{j+1} = w - V_{j+1}h$
   $H_{j+1} := \begin{pmatrix} \hat{H}_j & h \end{pmatrix}$

Remarks. **(a)** The residual $f_j := Av_j - V_j h_j$, where h_j is chosen so that the norm $\| f_j \|$ is minimized. It turns out that $V_j^* h_j = 0$, and $h_j = V_j^* A v_j$, where $w = Av_j$, that is $f_j = (I - V_j V_j^*)Av_j$.

(b) If A is symmetric, then H_j is tridiagonal, and the Arnoldi algorithm coincides with the Lanczos algorithm. ■

3.3.3 Implicitly restarted Arnoldi and Lanczos methods

The goal of *restarting* the Lanczos and Arnoldi factorizations is to get a better approximation of some desired set of preferred eigenvalues, for example, those eigenvalues that have

- Largest modulus
- Largest real part
- Positive or negative real part

Let A have eigenvalues in the left half-plane, and let the approximant K_m obtained through Lanczos or Arnoldi, have an eigenvalue μ in the right half-plane. To eliminate this unwanted eigenvalue the reduced order system obtained at the m-th step $AV_m = V_m K_m + f_m e_m^*$, is projected onto an $(m-1)$-st order system. First, compute the QR-factorization of $K_m - \mu I_m = Q_m R_m$:

$$A\bar{V}_m = \bar{V}_m \bar{K}_m + f_m e_m^* Q_m \quad \text{where} \quad \bar{V}_m = V_m Q_m, \ \bar{K}_m = Q_m^* K_m Q_m$$

We now truncate the above relationship to contain $m-1$ columns; let $\bar{K}_{m-1}$ denote the principal submatrix of $\bar{K}_m$, containing the leading $m-1$ rows and columns.

Theorem 3.3 *Given the above set-up, $\bar{K}_{m-1}$ can be obtained through an $m-1$ step Arnoldi process with A unchanged, and the new starting vector $\bar{B} := (\mu I_n - A)B$: $A\bar{V}_{m-1} = \bar{V}_{m-1}\bar{K}_{m-1} + \bar{f} e_{m-1}^*$.*

This process can be repeated to eliminate other unwanted eigenvalues (poles) from the reduced order system.

3.3.4 The Rational Krylov Method

The rational Krylov Method is a generalized version of the standard Arnoldi and Lanczos methods. Given a dynamical system Σ, a set of interpolation points $w_1, \cdots, w_l$, and an integer N, the Rational Krylov Algorithm produces a reduced order system Σ_k that matches N moments of Σ at $w_1, \cdots, w_l$. The reduced system is not guaranteed to be stable and no global error bounds exist. Moreover the selection of interpolation points which determines the reduced model is not an automated process and has to be figured out by the user using trial and error.

3.4 A new approach: The cross grammian

The approach to model reduction proposed below is related to the implicitly restarted dual Arnoldi approach developed in [15]; although it is not a moment matching method it belongs to the general set of Krylov projection methods. Its main feature is that is based on *one Sylvester equation* instead of *two Lyapunov equations*. One problem with prior attempts at working with the two Lyapunov equations separately and then applying dense methods to the reduced equations, is consistency. One cannot be certain that the two separate basis sets are the ones that would have been selected if the full system Grammians had been available. Since our method actually provides best rank k approximations to the system Grammians with a computable error bound, we are assured to obtain a valid approximation to the balanced reduction.

Given $\mathcal{S}$ as in (1.2) with $m = p = 1$, the *cross grammian* $X \in \mathbb{R}^{n \times n}$ is the solution to the following Sylvester equation:

$$AX + XA + BC = 0 \tag{3.5}$$

Recall the definition of the controllability and observability grammians (2.3). The relationship between these three grammians is:

$$X^2 = \mathcal{P}\mathcal{Q} \tag{3.6}$$

Moreover, the eigenvalues of X are equal to the non-zero eigenvalues of the Hankel operator $\mathcal{H}$. If A is stable

$$X = \int_0^\infty e^{At}BCe^{At}dt = \frac{1}{2\pi}\int_{-\infty}^{\infty}(j\omega - A)^{-1}BC(-j\omega - A)^{-1}$$

Furthermore, in this case the $\mathcal{H}_2$ norm of the system is given by $\| \mathcal{S} \|^2_{\mathcal{H}_2} = CXB$. Often, the *singular values* of X drop off very rapidly and X can be well approximated by a low rank matrix. Therefore the idea is to capture most of the energy with X_k, the best rank k approximation to X:

$$\| \mathcal{S} \|^2_{\mathcal{H}_2} = CX_kB + \mathcal{O}(\sigma_{k+1}(X))$$

The Approximate Balanced Method [8] solves a Sylvester Equation to obtain a reduced order almost balanced system iteratively without computing the full order balanced realization $\mathcal{S}$ in (2.7).

3.4.1 Description of the solution

We now return to the study of the *Sylvester equation* (3.5). It is well known that X is a solution iff

$$\begin{pmatrix} A & BC \\ 0 & -A \end{pmatrix}\begin{pmatrix} I & X \\ 0 & I \end{pmatrix} = \begin{pmatrix} I & X \\ 0 & I \end{pmatrix}\begin{pmatrix} A & 0 \\ 0 & -A \end{pmatrix}$$

This suggests that X can be computed using a projection method:

$$\begin{pmatrix} A & BC \\ 0 & -A \end{pmatrix}\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} H$$

where V is orthogonal: $V_1^*V_1 + V_2^*V_2 = I$. If V_2 is non-singular, then $AV_1 + BCV_2 = V_1H$, $-AV_2 = V_2H$, which implies $A(V_1V_2^{-1}) + BC = (V_1V_2^{-1})\hat{H}$, $H = -A$, where $\hat{H} = V_2HV_2^{-1}$. Therefore the solution is:

$$X = V_1V_2^{-1}$$

The best rank k approximation to X is related to the *C-S decomposition*. Let $V = [V_1^* \ V_2^*]^* \in \mathbb{R}^{2n\times k}$, with $V^*V = I_k$; then we have $V_1 = U_1\Gamma W^*$, $V_2 = U_2\Delta W^*$, where U_1, U_2 are orthogonal, W nonsingular and $\Gamma^2 + \Delta^2 = I_k$. Assuming A stable, V_2 has full rank iff the eigenvalues of A include those of H. It follows that the SVD of X can be expressed as $X = U_1(\Gamma/\Delta)U_2^*$. To compute the best rank k approximation to X, we begin with the full (n-step) decomposition: $V_1, V_2 \in \mathbb{R}^{n\times n}$, V_2 full rank:

$$\begin{aligned} AV_1 + BCV_2 &= V_1H \\ -AV_2 &= V_2H \end{aligned}$$

Let $W_k := W(:, 1:k)$, $\Gamma_k := \Gamma(1:k, 1:k)$, $\Delta_k := \Delta(1:k, 1:k)$. Then

$$\begin{aligned} A(V_1W_k) + BC(V_2W_k) &= (V_1W)(W^*HW_k) \\ -A(V_2W_k) &= (V_2W)(W^*HW_k) \end{aligned}$$

Therefore

$$\begin{aligned} A(U_{1k}\Gamma_k) + BC(U_{2k}\Delta_k) &= (U_{1k}\Gamma_k)H_k + E_k \\ -U_{2k}^*AU_{2k}\Delta_k &= \Delta_kH_k \end{aligned}$$

where $U_{1k}^* E_k = 0$ and $H_k = W_k^* H W_k$. We thus obtain the *projected* Sylvester equation

$$U_{1k}^*(AX_k + X_k A + BC)U_{2k} = 0$$

This implies the error equation

$$AX_k + X_k A + BC = -A(X - X_k) - (X - X_k)A = \mathcal{O}(\gamma_{k+1}/\delta_{k+1})$$

where

$$X_k = U_{1k}(\Gamma_k/\Delta_k)U_{2k}^*$$

is the best rank k approximation of the cross grammian X. The *Reduced order system* is now defined as follows: let the SVD of the cross grammian be $X = U\Sigma V^*$, and the best rank k approximant be $X_k = U_k \Sigma_k V_k^*$. Then $\hat{\mathcal{S}}$ is given by

$$\hat{\mathcal{S}} = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \end{array}\right] = \left[\begin{array}{c|c} V_k^* A V_k & V_k^* B \\ \hline C V_k & \end{array}\right] \tag{3.7}$$

A closely related alternative method of defining a reduced order model is the following. Let $\hat{X}$ be the best rank k approximation to X. Compute a partial eigenvalue decomposition of $\hat{X} = Z_k D_k W_k^*$ where $W_k^* Z_k = I_k$. Then an approximately balanced system is obtained as

$$\hat{\mathcal{S}} = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \end{array}\right] = \left[\begin{array}{c|c} W_k^* A Z_k & W_k^* B_k \\ \hline C Z_k & \end{array}\right] \tag{3.8}$$

The advantage of the this method which will be referred to as *Approximate Balancing* in the sequel, is that it computes an *almost* balanced reduced system *iteratively* without computing a balanced realization of the full order system first, and then truncating. Some details on the implementation of this algorithm are provided in section 3.4.3; more details can be found in [8].

Finally, a word about the MIMO case. Recall that (3.5) is not defined unless $m = p$. Hence, we to apply this method to MIMO systems, we proceed by embedding the system $\mathcal{S}$ in a system $\tilde{\mathcal{S}}$ which has the same order, is square, and symmetric:

$$\tilde{\mathcal{S}} = \left[\begin{array}{c|c} A & \tilde{B} \\ \hline \tilde{C} & \end{array}\right] \left[\begin{array}{c|cc} A & JC^* & B \\ \hline C & & \\ B^* J^{-1} & & \end{array}\right], \; J = J^* \tag{3.9}$$

The symmetrizer J is chosen so that $AJ = JA^*$ and $\lambda_i(\mathcal{PQ}) \approx \lambda_i(\tilde{\mathcal{P}}\tilde{\mathcal{Q}}) = \lambda(\tilde{X})^2$ where $\tilde{X}$ is the cross grammian of $\tilde{\mathcal{S}}$; for details see [8].

3.4.2 A stopping criterion

As *stopping criterion*, we look at the $\mathcal{L}_\infty$-norm of the following *residual*:

$$R(s) = BC - (sI - A)V_k(sI - \hat{A})^{-1}\hat{B}\hat{C}(-sI - \hat{A})^{-1}V_k^*(-sI - A)$$

Notice that *projected residual* is zero: $V_k^* R(s) V_k = 0$. Consider:

$$(sI - A)V(sI - V^*AV)^{-1}V^*B = (sI - A)VV^*(sI - AVV^*)^{-1}B$$

Assume that we change basis so that $V^* = [I_k \;\; 0]$; let in this basis

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \; B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \; C = \begin{pmatrix} C_1 & C_2 \end{pmatrix}$$

Then this last expression becomes:

$$\begin{pmatrix} I \\ -A_{21}(sI_k - A_{11})^{-1} \end{pmatrix} B_1$$

Similarly, $CV(-sI - V^*AV)^{-1}V^*(-sI - A) = C_1[I \quad (-sI - A_{11})^{-1}A_{12}]$. Hence

$$R(s) = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \begin{pmatrix} C_1 & C_2 \end{pmatrix} - \begin{pmatrix} I \\ -A_{21}(sI_k - A_{11})^{-1} \end{pmatrix} B_1C_1 \begin{pmatrix} I & (-sI - A_{11})^{-1}A_{12} \end{pmatrix}$$

In state space form we have

$$R = \left[\begin{array}{cc|cc} A_{11} & B_1C_1 & B_1C_1 & 0 \\ 0 & -A_{11} & 0 & A_{12} \\ \hline 0 & -B_1C_1 & 0 & B_1C_2 \\ A_{21} & 0 & B_2C_1 & B_2C_2 \end{array}\right] \in \mathbb{R}^{(2k+n)\times(2k+n)}$$

The implication is that $R(s)$ is $n \times n$ proper rational matrix with McMillan degree $2k$; its $\mathcal{L}_\infty$-norm $\| R(s) \|_\infty$ can be readily computed.

3.4.3 Computation of the solution

The following points are important to keep in mind. First, we give up Krylov (Arnoldi requires special starting vector), and second, an iterative scheme related to the Jacobi-Davidson algorithm together with implicit restarting will be used. Given is

$$(-A)V = VH + F,\ V^*V = I,\ V^*F = 0$$

1. Solve the projected Sylvester equation in the Controllability space and compute the SVD of the solution:
$$AY + BCV = YH \ \Rightarrow\ [Q, S, W] = \text{svd}\,(Y)$$

2. Project onto the space of the largest singular values
$$\left|\begin{array}{l} S_k = S(1:k, 1:k) \\ Q_k = Q(:, 1:k) \\ W_k = W(1:k, :) \end{array}\right. \qquad \left|\begin{array}{l} V \leftarrow VW_k \\ \bar{H} = Q_k^* A Q_k \\ H \leftarrow W_k^* H W_k \end{array}\right.$$

3. Correct the projected Sylvester equation in the observability space:
$$E := A^*VW_kS_k + VW_kS_k\bar{H}^* + C^*B^*Q_k$$
Solve $A^*Z + Z\bar{H}^* = -E$.

4. Adjoin Correction and project:
$$\left|\begin{array}{l} [V, R] = \text{qr}\,([VW_k,\ Z]) \\ H \leftarrow -(V^*AV) \\ F \leftarrow (-I + VV^*)AV \end{array}\right.$$

Remark. It should be stressed that the equations in 1. and 3.

$$\begin{pmatrix} A & BCV \\ 0 & -H \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} R, \quad \begin{pmatrix} A^* & E \\ 0 & -H^* \end{pmatrix} \begin{pmatrix} W_1 \\ W_2 \end{pmatrix} = \begin{pmatrix} W_1 \\ W_2 \end{pmatrix} Q$$

above are solved by IRAM. No inversions are required. However convergence may be accelerated with a single sparse direct factorization. ■

4 Application of the reduction algorithms

In this section we apply the algorithms mentioned above to six different dynamical systems: A Structural Model, Building Model, Heat Transfer Model, CD Player, Clamped Beam, Low-Pass Butterworth Filter. We reduce the order of models with a *tolerance*[1] value, ρ, of 1×10^{-3}. Table-2 shows the order of the systems, n; the number of inputs, m; and outputs, p; and the order of reduced system, k. Moreover, the Normalized[2] Hankel Singular Values of each model are depicted in Figure 2-a and 2-b. To make a better comparison between the systems, in Figure 3 we also show relative degree reduction $\frac{k}{n}$ vs a given error tolerance $\frac{\sigma_k}{\sigma_1}$. This figure shows how much the order can be reduced for the given tolerance: the lower the curve, the easier to approximate. It can be seen from Figure 3 that among all models for a fixed tolerance value less than 1.0×10^{-1}, the building model is the hardest one to approximate. One should notice that specification of the tolerance value ρ determines everything in all of the methods except the Rational Krylov Method. The order of the reduced model and the eigenvalue placements are completely automatic. On the other hand, in the Rational Krylov Method, one has to choose the interpolation points and the integer N which determines the number of moments matched per point.

	n	m	p	k
Structural Model	270	3	3	37
Building Model	48	1	1	31
Heat Model	197	2	2	5
CD Player	120	1	1	12
Clamped Beam	348	1	1	13
Butterworth Filter	100	1	1	35

Table 2: The systems used for comparing the model reduction algorithms

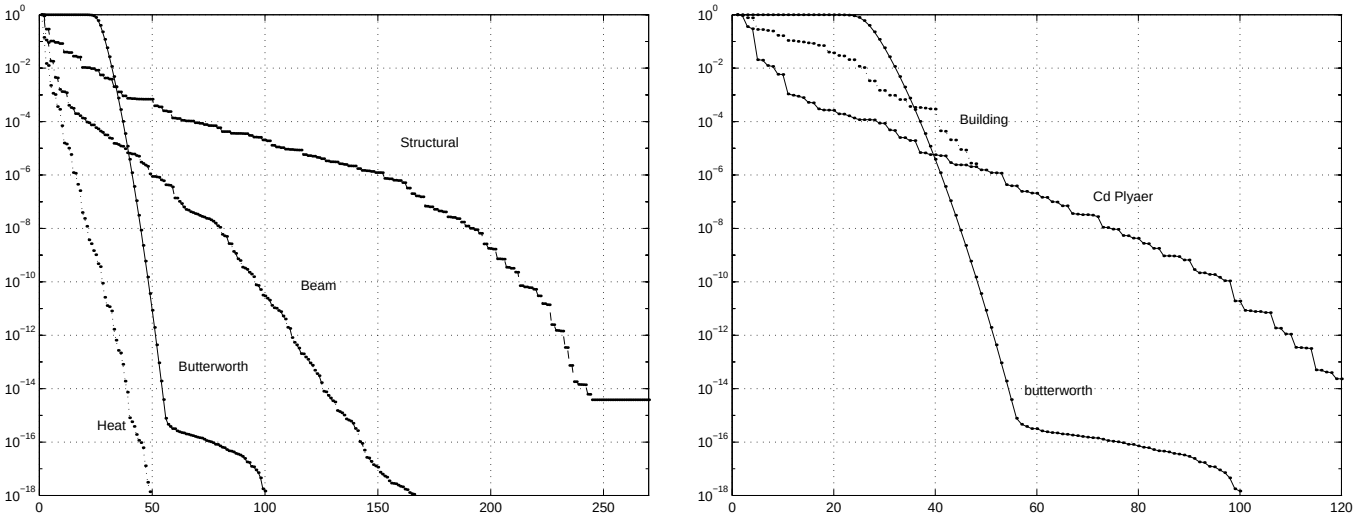


Figure 2: Normalized Hankel Singular Values of (a) Heat Model, Butterworth Filter , Clamped Beam Model and Structural Model; (b) Butterworth Filter, Building Example, CD Player.

[1]The *tolerance* corresponding to a k^{th} order reduced system is given by the ratio $\frac{\sigma_k}{\sigma_1}$ where σ_1 and σ_k are the largest and k^{th} singular value of the system respectively.

[2]For comparison, we normalize the highest Hankel Singular Value of each system to 1.

In each subsection below, we briefly describe the systems and then apply the algorithms. For each the largest singular value of the frequency response of full order, reduced order, and of the corresponding error systems; and the nyquist plots of the full order and reduced order systems are shown. Moreover, the relative $\mathcal{H}_\infty$ and $\mathcal{H}_2$ norms of the error systems are tabulated. Since balanced reduction and approximate balanced reduction approximants were *almost* the same for all the models except the heat model, we show and tabulate results just for the former for those cases.

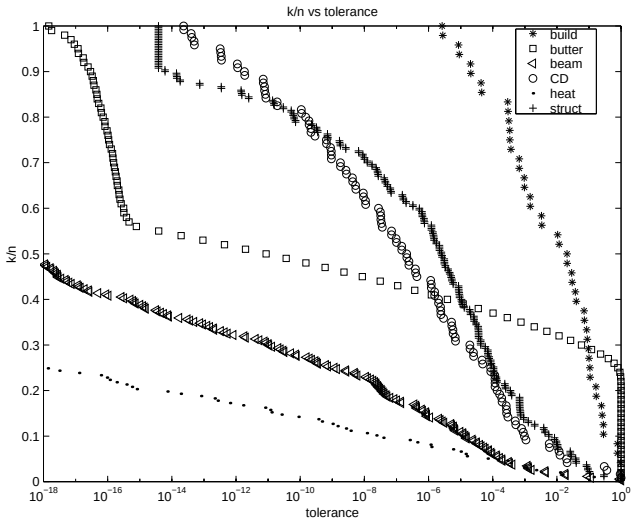


Figure 3: Relative degree reduction $\frac{k}{n}$ vs error tolerance $\frac{\sigma_k}{\sigma_1}$

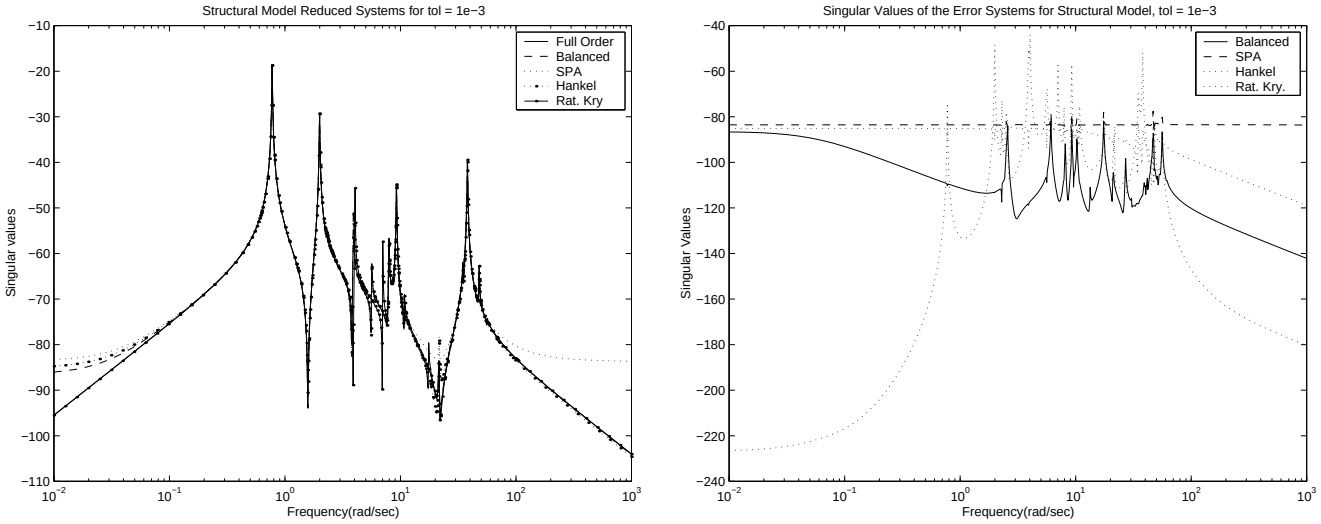
4.1 Structural Model

This is a model of component 1r (Russian service module) of the International Space Station. It has 270 states, 3 inputs and 3 outputs. The real part of the pole closest to imaginary axis is -3.11×10^{-3}. The normalized Hankel Singular Values of the system are shown in Figure 2-a. We approximate the system with reduced models of order 37. Since the system is MIMO, the Arnoldi and Lanczos Algorithms do not apply. The resultant reduced order systems are shown in Figure 4-a. As seen from the figure, all the models work quite well. The peaks, especially the ones at the lower frequencies, are well approximated. Figure 4-b shows the largest singular value $\sigma_{\max}$ of the frequency response of the error systems. Rational Krylov does a perfect job at the lower and higher frequencies. But for the moderate peak frequency levels, it has the highest error amplitude. This is because of the fact that the selection of interpolation point is not an automated process and relies on ad-hoc specification by the user. Singular perturbation approximation is the worst for low and higher frequencies. Table 3 lists the relative[3] $\mathcal{H}_\infty$ and $\mathcal{H}_2$ norms of the errors system. As seen from the figure, Rational Krylov has the highest error norms. Considering

	$\mathcal{H}_\infty$ norm	$\mathcal{H}_2$ norm
Balanced	6.93×10^{-4}	5.70×10^{-3}
Hankel	8.84×10^{-4}	1.98×10^{-2}
Sing. Pert	1.08×10^{-3}	3.66×10^{-2}
Rat. Kry	4.46×10^{-2}	1.33×10^{-1}

Table 3: Relative Error Norms for Structural Model

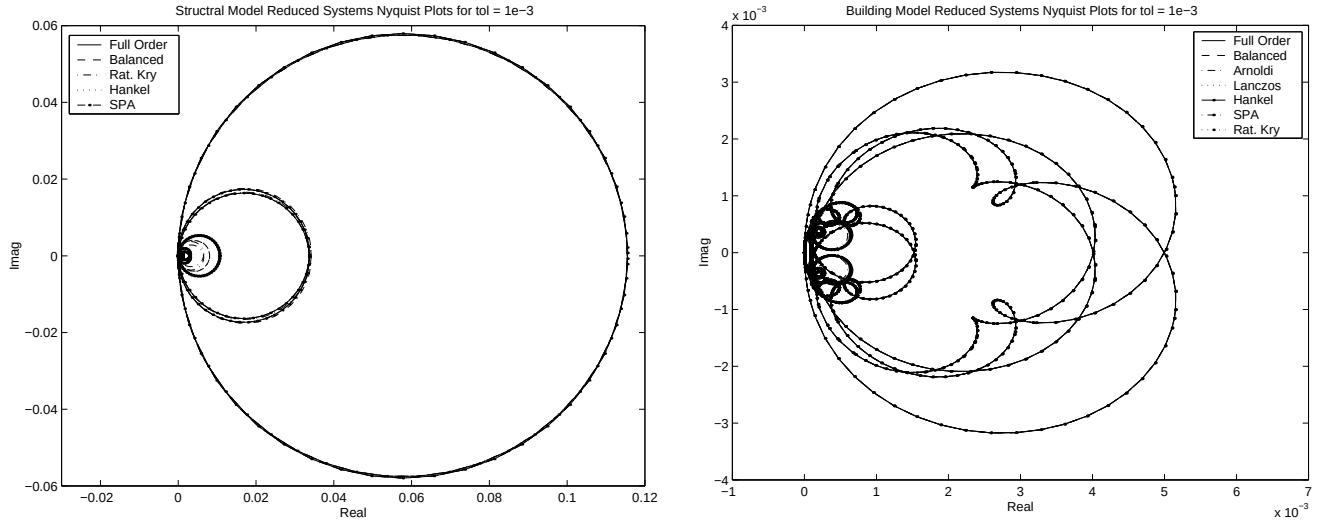
[3]To find the relative error, we divide the norm of the error system with the corresponding norm of the full order system



(a) (b)

Figure 4: $\sigma_{\max}$ of the frequency response of the (a) Reduced and (b) Error Systems of Structural Model

both the relative $\mathcal{H}_\infty$, $\mathcal{H}_2$ norms error norms and the whole frequency range, Balanced Reduction is the best. The nyquist plots of the full order and the reduced order systems are shown in Figure 5-a. Notice that all the approximants matches the full order model very well except the fact the rational Krylov deviates around origin.



(a) (b)

Figure 5: Nyquist plots of the full order and reduced order models for the (a) Structral (b) Building Model

4.2 Building Model

The full order model is a building (Los Angeles University Hospital) with 8 floors each of which has 3 degrees of freedom, namely displacements in x and y directions, and rotation. Hence we have 24 variables

with the following type of second order differential equation describing the dynamics of the system:

$$M\ddot{q}(t) + D\dot{q}(t) + Kq(t) = v\,u(t) \tag{4.1}$$

where $u(t)$ is the input. (4.1) can be put into state-space form by defining $x^* = [\ q^* \quad \dot{q}^*\ ]^*$:

$$\dot{x}(t) = \begin{bmatrix} 0 & I \\ -M^{-1}K & -M^{-1}D \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ M^{-1}v \end{bmatrix} u(t)$$

We are mostly interested in the motion in the first coordinate $q_1(t)$. Hence, we choose $v = [1 \ \ 0 \ \ \cdots \ \ 0]^*$ and the output $y(t) = \dot{q}_1(t) = x_{25}(t)$.

The state-space model has order 48, and is single input and single output. For this example, the pole closest to imaginary axis has real part equal to -2.62×10^{-1}. We approximate the system with a model of order 31. The largest singular value $\sigma_{\max}$ of the frequency response of the reduced order and

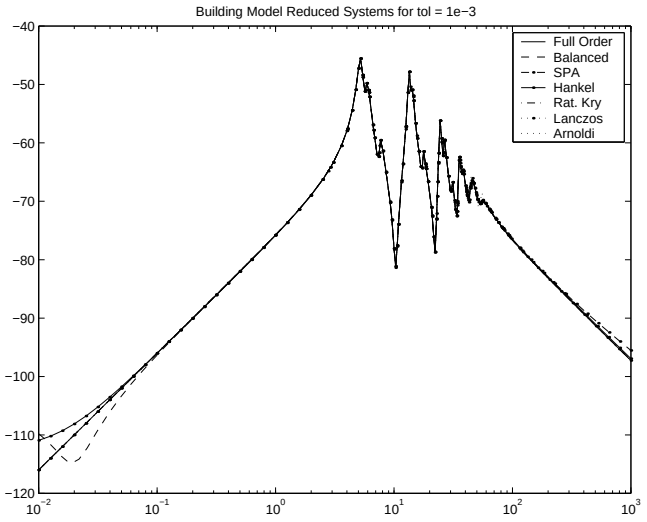
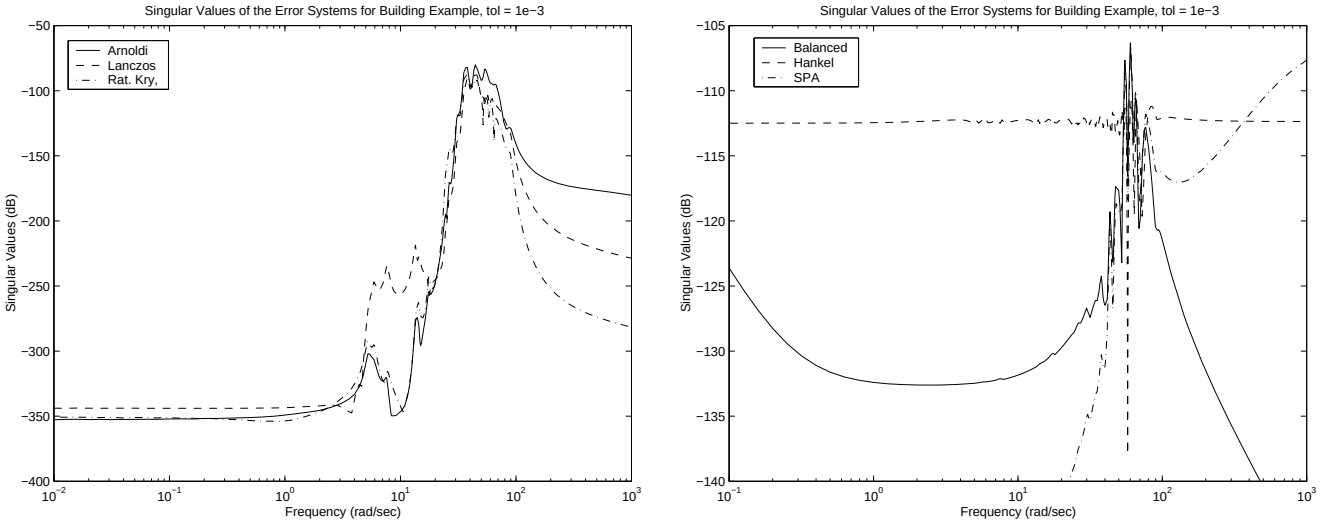


Figure 6: $\sigma_{\max}$ of the frequency response of the reduced systems of Building Model

of the error systems are shown in Figure 6 and 7 respectively. Since the expansion of transfer function $G(s)$ around $s_0 = \infty$ results in unstable reduced systems for Arnoldi and Lanczos procedures, we use the shifted version of these two methods with $s_0 = 1$. The effect of choosing s_0 as a low frequency point is very well observed in Figure 7-b that Arnoldi and Lanczos result in very good approximants for the low frequency range. The same is valid for rational Krylov methods as well, since $s_0 = 1$ was chosen as one of the interpolation points for this method. When compared to SVD based methods, the moments matching based methods are much better for low frequency range. Among the SVD based methods, Singular perturbation and balanced reduction methods are the best for the low frequency and high frequency range respectively. When we consider the whole frequency range, balancing and singular perturbation are closer to the original model. But in terms of relative $\mathcal{H}_\infty$ norm of error, Hankel Norm Approximation is the best. As expected rational Krylov, Arnoldi and Lanczos result in high relative errors due to being local in nature. Among them, rational Krylov is the best. Figure 5-b illustrates the nyquist plots of the full order and the reduced order systems. The figure shows that all the approximants matches the nyquist plots of the full order model quite well.

4.3 Heat diffusion model

The original system is a plate with two heat sources and two points of measurements. It is described by the heat equation. A model of order 197 is obtained by spatial discretization. The real part of the pole



(a) (b)

Figure 7: (a)-(b) $\sigma_{\max}$ of the frequency response of the error systems of Building Model

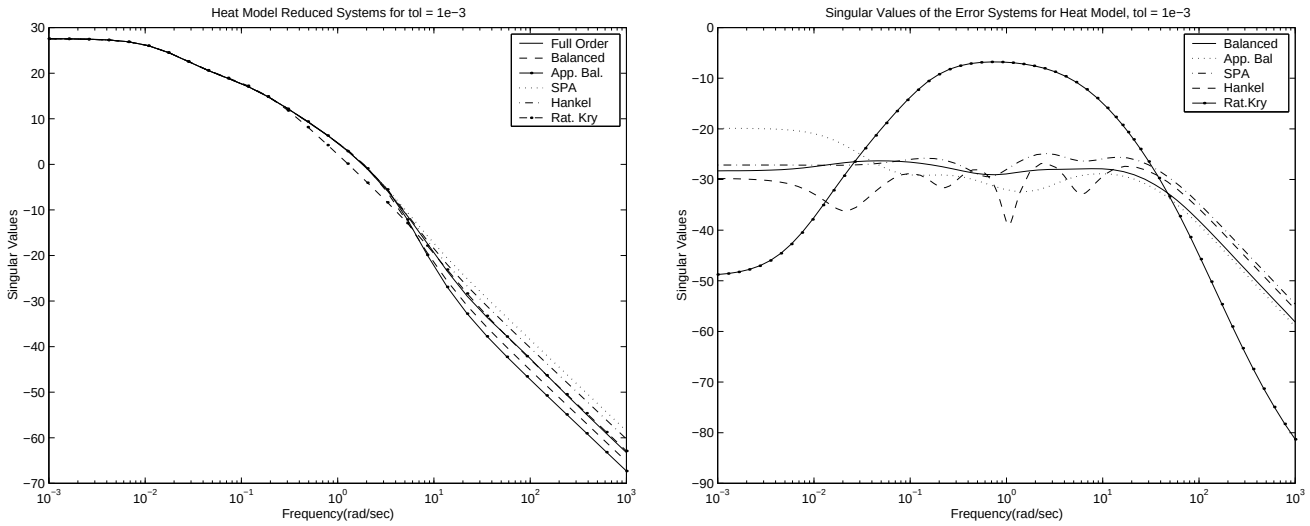
	$\mathcal{H}_\infty$ norm of error	$\mathcal{H}_2$ norm of error
Balanced	$9,64 \times 10^{-4}$	2.04×10^{-3}
Hankel	5.50×10^{-4}	6.25×10^{-3}
Sing. Pert	9.65×10^{-4}	$2,42 \times 10^{-2}$
Rat. Kry	7.51×10^{-3}	1.11×10^{-2}
Lanczos	7.86×10^{-3}	1.26×10^{-2}
Arnoldi	1.93×10^{-2}	3.33×10^{-2}

Table 4: Relative Error Norms Building Model

closest to imaginary axis is -1.52×10^{-2}. It is observed from Figure 2-a that this system is very easy to approximate since the Hankel singular values decay very rapidly. We approximate the model with a model of order 5. Since this is a MIMO system, Lanczos and Arnoldi do not apply. As expected due to the very low tolerance value, all the methods generate satisfactory approximants matching the full order model through the whole frequency range (see Figure 8). Only the Rational Krylov Method has some problems for moderate frequencies due to the unautomated choice of interpolation points. The nyquist plots of the full order and the reduced order systems are shown in Figure 9-a. The figure reveals that as in the structural model example, rational Krylov have problem matching the full order system around the origin. Except the rational Krylov approximant, all the methods very well approximate the nyquist plots of the full order model.

	$\mathcal{H}_\infty$ norm of error	$\mathcal{H}_2$ norm of error
Balanced	2.03×10^{-3}	5.26×10^{-2}
App. Balanced	4.25×10^{-3}	4.68×10^{-2}
Hankel	1.93×10^{-3}	6.16×10^{-2}
Sing. Pert	2.39×10^{-3}	7.39×10^{-2}
Rat. Kry	1.92×10^{-2}	2.01×10^{-1}

Table 5: Relative Error Norms of Heat Model



(a) (b)

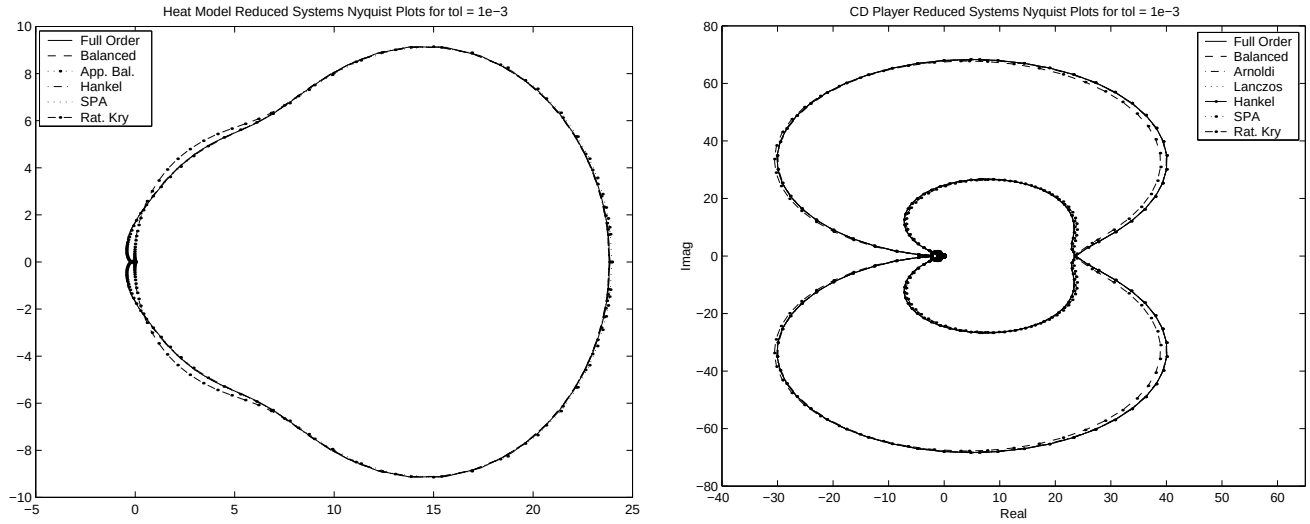
Figure 8: $\sigma_{\max}$ of the frequency response of the (a) Reduced and (b) Error Systems of Heat diffusion model

4.4 CD Player

This system describes the dynamics between the lens actuator and the radial arm position of a portable CD player. The model has 120 states with a single input and a single output. The pole closest to the imaginary axis has the real part equal to -2.43×10^{-2}. Approximants have order 12. The first moment of the system is zero. Hence, instead of expanding the transfer function around $s = \infty$, we expand it around $s_0 = 200$ rad/sec. This overcomes the breakdown in Lanczos procedure. We also use the shifted version of Arnoldi procedure with $s_0 = 200$ rad/sec. Figure 10-a illustrates the largest singular values of the frequency response of the reduced order models together with that of the full order model. One should notice that only the rational Krylov catch the peaks around the frequency range $10^4 - 10^5$ rad/sec. No SVD based method matches those peaks. Among the SVD based ones, Hankel Norm Approximation is the worst around $s = 0$, and also around $s = \infty$. The largest singular values of the frequency response of error systems in Figure 10-b reveal that the SVD based methods are better when we consider the whole frequency range. Despite doing a perfect job at $s = 0$ and $s = \infty$, Rational Krylov has the highest relative $\mathcal{H}_\infty$ and $\mathcal{H}_2$ error norms as listed in Table 6. But one should notice that the rational Krylov is superior to the Arnoldi and Lanczos procedures except the frequency range $10^2 - 10^3$ rad/sec. When we consider the whole frequency range, balanced reduction is again the best one. Figure 9-b illustrates the nyquist plots of the full order and the reduced order systems. Except rational Krylov's having some deviation from the full order model, all the methods result in satisfactory approximants.

	$\mathcal{H}_\infty$ norm of error	$\mathcal{H}_2$ norm of error
Balanced	9.74×10^{-4}	3.92×10^{-3}
Approx. Balanc.	9.74×10^{-4}	3.92×10^{-3}
Hankel	9.01×10^{-4}	4.55×10^{-3}
Sing. Pert	1.22×10^{-3}	4.16×10^{-3}
Rat. Kry	5.60×10^{-2}	4.06×10^{-2}
Arnoldi	1.81×10^{-2}	1.84×10^{-2}
Lanczos	1.28×10^{-2}	1.28×10^{-2}

Table 6: Relative Error Norms of CD Player



(a) (b)

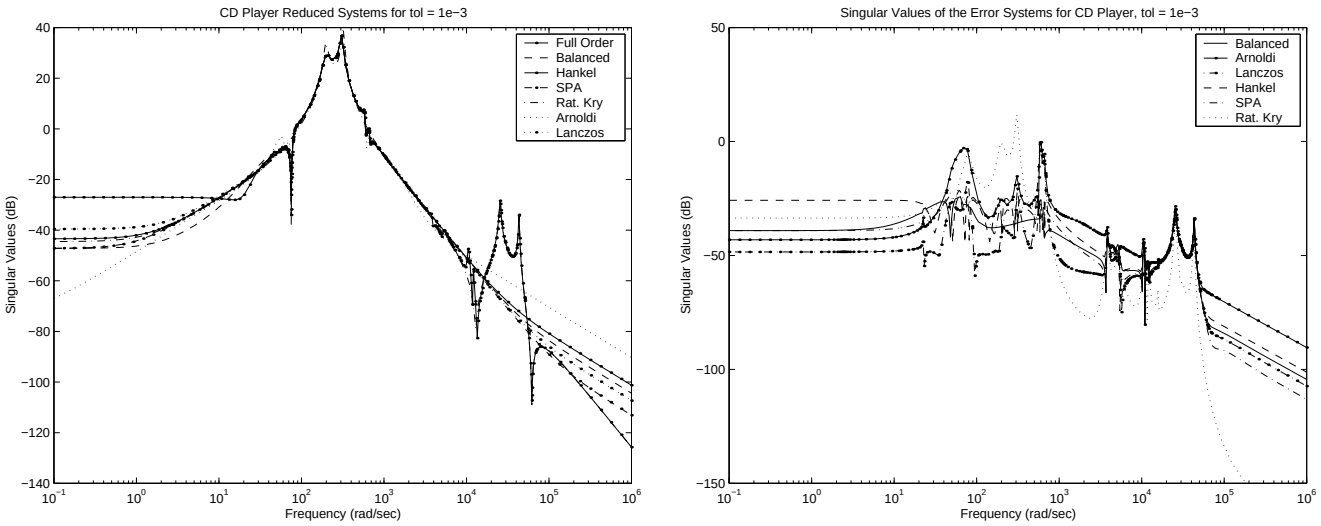
Figure 9: Nyquist plots of the full order and reduced order models for the (a) Heat Model (b) CD Player

4.5 Clamped Beam Model

The clamped beam model has 348 states and is SISO. It is again obtained by spatial discretization of an appropriate partial differential equation. The input represents the force applied to the structure, and the output is the displacement. For this example, the real part of the pole closest to imaginary axis is -5.05×10^{-3}. We approximate the system with a model of order 13. The plots of the largest singular value of the frequency response of the approximants and error systems are shown in Figure 11-a and 11-b respectively. Since $CB = 0$, we expand the transfer function $G(s)$ of the original system around $s_0 = 0.1$ instead of $s = \infty$ to prevent the breakdown of Lanczos. Moreover, to obtain better result, we use the shifted Arnoldi with $s_0 = 0.1$ rad/sec. Rational Krylov is again the best one for both $s = 0$ and $s = \infty$. Indeed except for the frequency range between 0.6 and 30 rad/sec, this method gives the best approximant among all the methdos. Lanczos and Arnoldi procedures also lead to a very good approximant especially for the frequency range $0 - 1$ rad/sec. This is due to the choice of s_0 as a low frequency point. Balanced model reduction is the best one among the SVD methods after $s = 1$ rad/sec. In terms of error norms, SVD based methods are better than moment matching based methods, but the difference are not as high as the previous examples. Again, the rational Krylov is the best among moment matching based methods. The nyquist plots of the full order and the reduced order systems are shown in Figure 12-a. The figure shows that all the approximants match the the nyquist plots of the full order model very well. Indeed, this is the best match of the nyquist plots among all the six examples.

	$\mathcal{H}_\infty$ norm of error	$\mathcal{H}_2$ norm of error
Balanced	2.14×10^{-4}	7.69×10^{-3}
Hankel	2.97×10^{-4}	8.10×10^{-3}
Sing. Pert	3.28×10^{-4}	4.88×10^{-2}
Rat. Kry	5.45×10^{-4}	8.88×10^{-3}
Arnoldi	3.72×10^{-3}	1.68×10^{-2}
Lanczos	9.43×10^{-4}	1.67×10^{-2}

Table 7: Relative Error Norms of Clamped Beam Model



(a) (b)

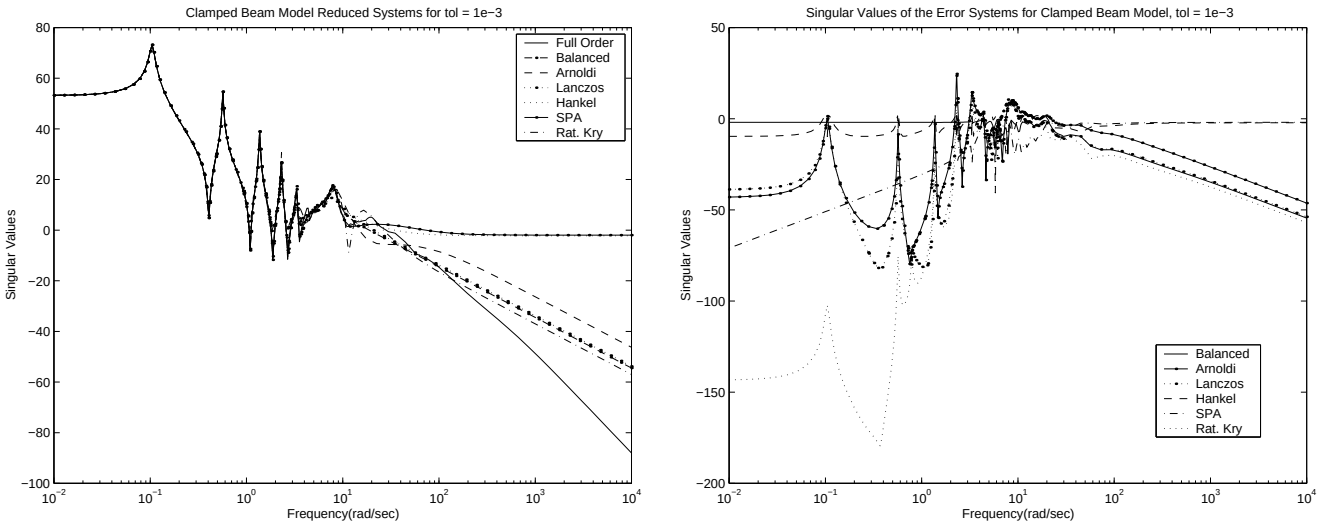
Figure 10: $\sigma_{\max}$ of the frequency response of the (a) Reduced and (b) Error Systems of CD Player

4.6 Low-Pass Butterworth Filter

The full order model is a Low-Pass Butterworth filter of order 100 with the cutoff frequency being 1 rad/sec. The normalized Hankel Singular Values corresponding to this system are shown in Figure 2-a and Figure 2-b. It should be noticed that unlike the other systems, Hankel Singular Values stay constant at the beginning, and then start to decay. Therefore, we cannot reduce the model to order less than 25. We approximate the system with a model of order 35. One should notice that the transfer function of this example has no zeros. Thus Arnoldi and Lanczos procedures do not work if we expand the transfer function $G(s)$ around $s = \infty$. Instead, we expand $G(s)$ around $s_0 = 0.1$. As Figure 13-a illustrates, all the moment matching based methods have difficulty especially around the cutoff frequency. Among them, Lanczos and Arnoldi show very similar results and are better than Rational Krylov Method. On the other hand, SVD based methods work without any problem producing quite good approximants for the whole frequency range. Although the Hankel norm approximation is the best in terms $\mathcal{H}_\infty$ norm, it is the worst in terms of $\mathcal{H}_2$ norm among the SVD based methods. Singular perturbation methods and balanced reduction shows very close behaviors for the frequencies less than 1 rad/sec. But after that, balanced reduction is better. Figure 12-b depicts the nyquist plots of the full order and the reduced order systems. As seen from the figure, moment matching methods are far from matching the full order model as in matching the frequency response. SVD based methods do not yield very good approximants, but compared to former, they are much better.

	$\mathcal{H}_\infty$ norm of error	$\mathcal{H}_2$ norm of error
Balanced	6.29×10^{-4}	5.19×10^{-4}
Approx. Balanc.	6.29×10^{-4}	5.19×10^{-4}
Hankel	5.68×10^{-4}	1.65×10^{-3}
Sing. Pert	6.33×10^{-4}	5.21×10^{-4}
Rat. Kry	1.02×10^{0}	4.44×10^{-1}
Arnoldi	1.02×10^{0}	5.38×10^{-1}
Lanczos	1.04×10^{0}	3.68×10^{-1}

Table 8: Relative Error Norms of Butterworth Filter



(a) (b)

Figure 11: $\sigma_{\max}$ of the frequency response of the (a) Reduced and (b) Error Systems of Clamped Beam

5 Projectors and computational complexity

The unifying feature of all model reduction methods presented above is that they are obtained by means of *projections*. Let $\pi = VW^*$ be a projection, i.e. $\pi^2 = \pi$. The corresponding reduced order model $\hat{\mathcal{S}}$ in (1.3) is obtained as follows:

$$\begin{aligned} \sigma\hat{x} &= (W^*AV)\hat{x} + (W^*B)u \\ \hat{y} &= (CV)\hat{x} \end{aligned} \tag{5.1}$$

The quality of the approximant is measured in terms of the frequency response $G(j\omega) = C(j\omega I - A)^{-1}B$. Optimal Hankel norm and Balancing emphasize energy of Grammians

$$\mathcal{P} = \frac{1}{2\pi}\int_{-\infty}^{+\infty}(j\omega I - A)^{-1}BB^*(j\omega I - A^*)^{-1}d\omega,\ \mathcal{Q} = \frac{1}{2\pi}\int_{-\infty}^{+\infty}(j\omega I - A^*)^{-1}C^*C(j\omega I - A)^{-1}d\omega$$

Krylov methods adapt to frequency response and emphasize relative contributions of $C(j\omega I - A)^{-1}B$. The new method emphasizes the energy of the cross grammian

$$X = \frac{1}{2\pi}\int_{-\infty}^{+\infty}(j\omega I - A)^{-1}BC(-j\omega I - A)^{-1}d\omega$$

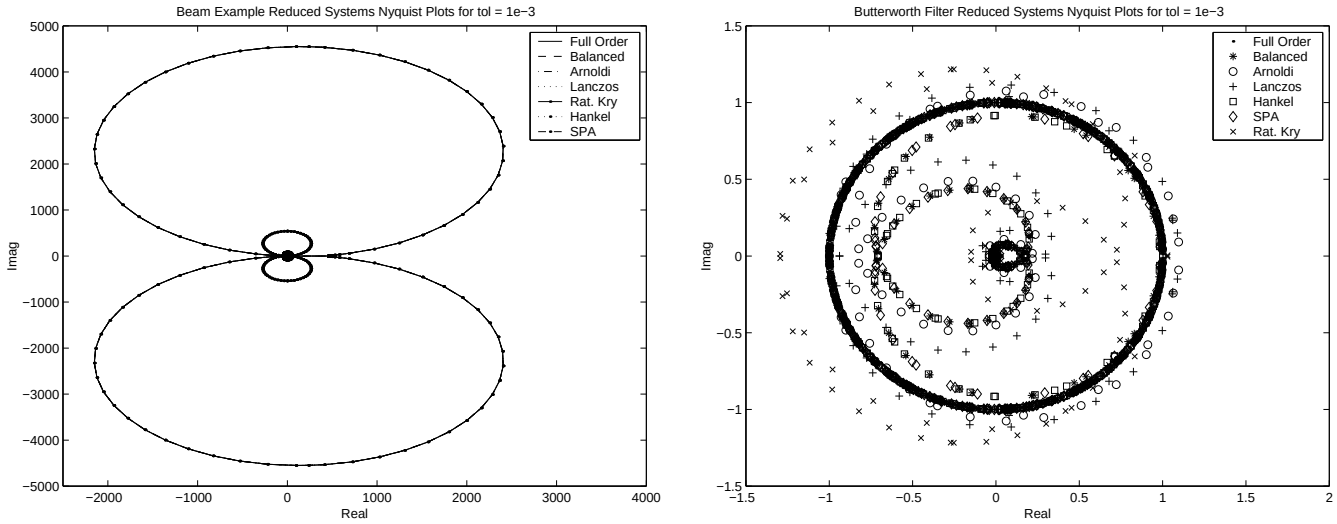
The *choices of projectors* for the different methods are as follows.

1. **Balanced truncation**. Solve: $A\mathcal{P} + \mathcal{P}A^* + BB^* = 0$, $A^*\mathcal{Q} + \mathcal{Q}A + C^*C = 0$, and *project* onto the dominant eigenspace of $\mathcal{P}\mathcal{Q}$.

2. **Optimal Hankel norm approximation**. Solve for the Grammians. Embed in a lossless transfer function and project onto its stable eigenspace.

3. **Krylov-based approximation**. Project onto controllability and/or observability spaces.

4. **New method**. Project onto the space spanned by the dominant right singular vectors or eigenvectors of the cross grammian.



(a) (b)

Figure 12: Nyquist plots of the full order and reduced order models for the (a) Clamped Beam Model (b) Butterworth Filter

The *complexity* of these methods using *dense* decompositions taking into account only dominant terms of the total cost, is as follows:

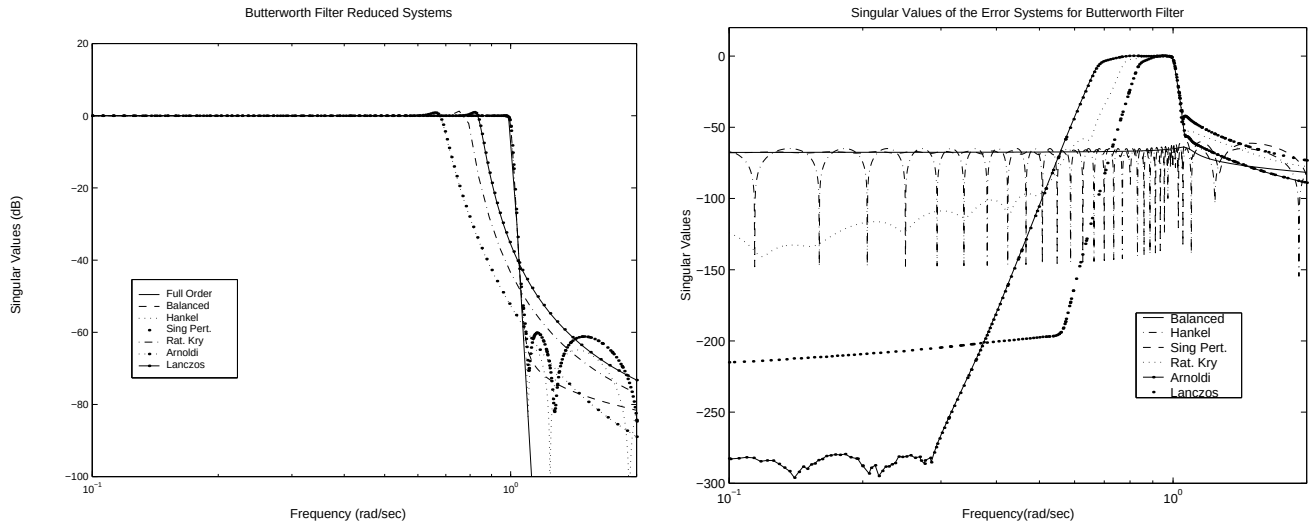
1. **Balanced truncation**. Compute Grammians $\approx 70N^3$ (QZ algorithm); perform balancing $\approx 30N^3$ (eigendecomposition).
2. **Optimal Hankel norm approximation**. Compute Grammians $\approx 70N^3$ (QZ algorithm); perform balancing and embedding $\approx 60N^3$.
3. **Krylov approximation**. $\approx kN^2$ operations.

The *complexity* using *approximate* and/or *sparse* decompositions, is as follows. Let α be the average number of non-zero elements per row in A, and let k be the number of expansion points. Then:

1. **Balanced truncation**. Grammians $\approx c_1 \alpha k N$; balancing $\mathcal{O}(n^3)$.
2. **Optimal Hankel norm approximation**. Grammians $\approx c_1 \alpha k N$; embedding $\mathcal{O}(n^3)$.
3. **Krylov approximation**. $\approx c_2 k \alpha N$ operations

6 Conclusions

In this note we presented a comparative study of seven algorithms for model reduction, namely: Balanced Model Reduction, Approximate Balanced Reduction, Singular Perturbation Method, Hankel Norm Approximation, Arnoldi Procedure, Lanczos Procedure, and Rational Krylov Method. These algorithms have been applied to six different dynamical systems. The first four make use of Hankel Singular Values and the latter three are based on matching the moments; i.e. the coefficients of the Laurent expansion of the transfer function around some point of the complex plane. The results show that Balanced Reduction and Approximate Balanced Reduction are the best when we consider the whole frequency range. Between



(a) (b)

Figure 13: $\sigma_{\max}$ of the frequency response of the (a) Reduced and (b) Error Systems of Butterworth Filter

these two, Approximate Balancing has the advantage that it computes an *almost* balanced reduced system *iteratively* without obtaining a balanced realization of the full order system first, and subsequently truncating, thus reducing the computational cost and storage requirements. Hankel Norm Approximation gives the worst approximation around $s = 0$ among the SVD based methods. Although it has the lowest $\mathcal{H}_\infty$ error norm in most of the cases, it leads to the highest $\mathcal{H}_2$ error norm. Being local in nature Moment Matching methods always lead a higher error norms than SVD based methods; but they reduce the computational cost and storage requirements remarkably when compared to the latter. Among them, the Rational Krylov Algorithm gives better results due to the flexibility of the selection of interpolation points. However, the selection of these points which determines the reduced model is not an automated process and has to be specified by the user, with little guidance from the theory on how to choose these points. In contrast, in other methods a given error tolerance value determines everything.

References

[1] A.C. Antoulas, "On recursiveness and related topics in linear systems," *IEEE Transactions on Automatic Control*, **AC-31**, pp. 1121-1135 (1986).

[2] A.C. Antoulas, J.A. Ball, J. Kang, and J.C. Willems, "On the solution of the minimal rational interpolation problem," *Linear Algebra and Its Applications*, Special Issue on Matrix Problems, **137/138**: 511-573 (1990).

[3] A.C. Antoulas and J.C. Willems, "A behavioral approach to linear exact modeling," *IEEE Transactions on Automatic Control*, **AC-38**: 1776-1802 (1993).

[4] A.C. Antoulas, "Recursive modeling of discrete-time time series," in *IMA volume on Linear Algebra for Control*, P. van Dooren and B.W. Wyman Editors, Springer Verlag, vol. **62**: 1-20 (1993).

[5] A.C. Antoulas, E.J. Grimme and D.C. Sorensen, "On behaviors, rational interpolation, and the Lanczos algorithm", Proc. 13th IFAC Triennial World Congress, San Francisco, Pergamon Press (1996).

[6] A.C. Antoulas, "Approximation of linear operators in the 2-norm", Special Issue of LAA (Linear Algebra and Applications) on *Challenges in Matrix Theory*, **278**: 309-316, (1998).

[7] A.C. Antoulas, "Approximation of linear dynamical systems", in the Wiley Encyclopedia of Electrical and Electronics Engineering, edited by J.G. Webster, volume **11**: 403-422 (1999).

[8] A.C. Antoulas and D.C. Sorensen, *Projection methods for balanced model reduction*, Technical Report ECE-CAAM Depts, Rice University, September 1999.

[9] D.L. Boley, *Krylov space methods on state-space control models*, Circuits, Systems, and Signal Processing, **13**: 733-758 (1994).

[10] P. Feldman and R.W. Freund, *Efficient linear circuit analysis by Padé approximation via a Lanczos method*, IEEE Trans. Computer-Aided Design, **14**, 639-649, (1995).

[11] W.B. Gragg and A. Lindquist, *On the partial realization problem*, Linear Algebra and Its Applications, Special Issue on Linear Systems and Control, **50**: 277-319 (1983).

[12] E.J. Grimme,*Krylov Projection Methods for Model Reduction*, Ph.D. Thesis, ECE Dept., U. of Illinois, Urbana-Champaign,(1997).

[13] K. Gallivan, E.J. Grimme, and P. Van Dooren, *Asymptotic waveform evaluation via a restarted Lanczos method*, Applied Math. Letters, **7**: 75-80 (1994).

[14] E.J. Grimme, D.C. Sorensen, and P. Van Dooren, *Model reduction of state space systems via an implicitly restarted Lanczos method*, Numerical Algorithms, **12**: 1-31 (1995).

[15] I.M. Jaimoukha, E.M. Kasenally, *Implicitly restarted Krylov subspace methods for stable partial realizations*, SIAM J. Matrix Anal. Appl., **18**: 633-652 (1997).

[16] P. V. Kokotovic, R. E. O'Malley, P. Sannuti, Singualr Perturbations and Order Reduction in Control Theory - an Overview, *Automatica*, **12**: 123-132, 1976.

[17] J. Li, F. Wang, J. White, *An efficient Lyapunov equation-based approach for generating reduced-order models of interconnect*, Proc. 36th IEEE/ACM Design Automation Conference, New Orleans, LA, (1999).

[18] Y. Liu and B. D. O. Anderson, Singular Perturbation Approximation of Balanced Systems, *Int. J. Control*, **50**: 1379-1405, 1989.

[19] B. C. Moore, *Principal Component Analysis in Linear System: Controllability, Observability and Model Reduction*, IEEE Transactions on Automatic Control, AC-26:17-32, 1981.

[20] K. Glover, All Optimal Hankel-norm Approximations of Linear Mutilvariable Systems and their L^∞-error Bounds, *Int. J. Control*, **39**: 1115-1193, 1984.

[21] T. Penzl, *Eigenvalue decay bounds for solutions of Lyapunov equations: The symmetric case*, Systems and Control Letters, to appear (2000).

[22] A. Ruhe, *Rational Krylov algorithms for nonsymmetric eigenvalue problems II: matrix pairs*, Linear Alg. Appl., **197**:283-295, (1984).

[23] D.C. Sorensen, *Implicit application of polynomial filters in a k-step Arnoldi method*, SIAM J. Matrix Anal. Applic., **13**: 357-385 (1992).

[24] P. Van Dooren, *The Lanczos algorithm and Padé approximations*, Short Course, Benelux Meeting on Systems and Control, (1995).