The Lanczos algorithm and Padé approximations

Paul Van Dooren
Dept. Mathematical Engineering
Université Catholique de Louvain
B-1348 Louvain-la-Neuve, BELGIUM

1 Introduction

In these two lectures we try to show the relations between the Lanczos algorithm and Padé approximations as used e.g. in identification and model reduction of dynamical systems. These notes are based on material in the papers [10, 17, 11, 12] for which a lot of credit ought to be given to the respective coauthors.

For simplicity we assume here that all systems are SISO, although some results do extend to the MIMO case. Let a n-th order dynamical system be described by

$$\dot{x} = Ax + bu \tag{1}$$
$$y = cx + du \tag{2}$$

where A is a square, b is a column vector, c is a row vector, and d is a scalar. It is well-known that the transfer function of this system :

$$h(s) = c(sI - A)^{-1}b + d$$

has a Taylor expansion around $s = \infty$ that looks like :

$$h(s) = d + cbs^{-1} + cAbs^{-2} + cA^2bs^{-3} + cA^3bs^{-4} + \dots$$

The coefficients m_{-i} of the powers of s^{-i} satisfy thus

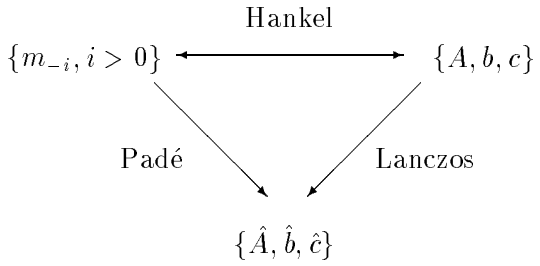
$$m_0 = d\ , \quad m_{-i} = cA^{i-1}b\ , \quad i \geq 1.$$

For $i \geq 1$ these are also called *moments* or *Markov parameters* of the system $\{A, b, c\}$. It follows already from the work of Hankel that the first $2n$ moments are sufficient to reconstruct the system triple $\{A, b, c\}$. This is done by solving a $n \times n$ *Hankel* system of equations :

$$\begin{bmatrix} h_1 & h_2 & h_3 & \dots & h_n \\ h_2 & h_3 & \iddots & \iddots & h_{n+1} \\ h_3 & \iddots & \iddots & \iddots & \vdots \\ \vdots & \iddots & \iddots & \iddots & \vdots \\ h_n & h_{n+1} & \dots & \dots & h_{2n-1} \end{bmatrix} \cdot \begin{bmatrix} p_0 \\ p_1 \\ \vdots \\ \vdots \\ p_{n-1} \end{bmatrix} = - \begin{bmatrix} h_n \\ h_{n+2} \\ \vdots \\ \vdots \\ h_{2n} \end{bmatrix}. \tag{3}$$

The Padé algorithm in fact produces the requested model $\{A, b, c\}$ from a LU factorization of this Hankel matrix. But this algorithm is recursive in nature and at each step k of the recurrence it

also produces an *approximate model* $\{\hat{A}, \hat{b}, \hat{c}\}$ of order k. The Lanczos algorithm is another recursive manner to compute the *same* approximate models $\{\hat{A}, \hat{b}, \hat{c}\}$ at each step k, but now starting from the system $\{A, b, c\}$ as input. We thus have the following triangle of relations :



Relations Hankel, Padé and Lanczos and model reduction

In the sequel we show in more detail the connections between these methods and stress the weaknesses and strengths of each approach. We also present variants of the Lanczos algorithm that have particularly attractive features for model reduction, especially for systems described by large sparse dynamical systems. Such models are typically found in circuit simulation of electronic devices, in power systems and in discretizations of distributed parameter systems.

Some history.

The relations of the Padé approximation problem to that of Hankel matrices was already observed in the late 1800's by people like H. Hankel (1862), G. F. Frobenius (1886) and T. J. Stieltjes (1894) (see [6] for more history on this). In the system theory literature the relation of partial realizations to Hankel matrices were rediscovered – and extended to the multi-input multi-output case – by Ho-Kalman, Youla-Tissi and Silverman, all three in 1966. The $O(n^2)$ algorithms for these partial realizations were rediscovered by Massey and Berlekamp in the context of convolutional codes, and later on by Rissanen in the context of identification. Several years later, de Jong showed that these algorithms all suffer from numerical instability. C. Lanczos was a a scientist working in relativity with also strong interests in numerical methods to approximate the spectrum of infinite dimensional operators. His papers on the tridiagonalization of the correspoding matrix discretizations appeared in the early 50's [22] and are still landmark papers in numerical linear algebra. These methods are now very popular for solving systems of equations as well as for computing part of the spectrum when the underlying matrix is large and sparse (see [14] for references). The first papers that made the link between these two problems are [15], [16] and [18].

2 Model reduction

Since the d term of the transfer function $h(s)$ does not play a role in what follows we may assume without loss of generality that it is zero. We are typically interested in a low-order approximation, defined by

$$\dot{\hat{x}} = \hat{A}\hat{x} + \hat{b}u \tag{4}$$

$$\hat{y} = \hat{c}\hat{x} \tag{5}$$

where the size of $\hat{A}$ is $k \ll n$. If the outputs are close (i.e., $\|y - \hat{y}\|$ is small) for some desired range of inputs, u, the low-order approximation is generally considered acceptable.

The zero-state (i.e., $x(0) = 0$) solution to (1) is

$$x(t) = \int_0^t e^{A(t-\tau)} b u(\tau)\, d\tau.$$

Thus determining a good low-order approximation (4,5) is intimately connected with finding a pair $\{\hat{A}, \hat{b}\}$ which yields a good approximation to the matrix exponential, $e^{At}b$. A method based on orthogonal Krylov projectors (the Arnoldi algorithm) is utilized in [3, 13, 29] for approximating $e^{At}b$ when A is sparse. But in fact, these concepts can be taken one step further by noting from (2) that one is ultimately only interested in that information in $e^{At}b$ which lies in the direction of c. For example, the impulse response (i.e., $u(t) = \delta(t)$) of the original system is $y(t) = ce^{At}b$. Taking a Taylor expansion of e^{At} around $t = 0$ yields

$$y(t) = ce^{At}b \approx \sum_{i=0}^{K-1} cA^i b \cdot \frac{t^i}{i!} = \sum_{i=1}^{K} m_{-i} \cdot \frac{t^{i-1}}{(i-1)!},$$

and this makes clear that matching momemts m_{-i} is a sensible thing to do. The Laplace transform of this impulse response is the transfer function and is also well approximated by a partial sum involving the first K moments :

$$h(s) = c(sI - A)^{-1}b \approx \sum_{i=1}^{K} cA^{i-1}b \cdot s^{-i} = \sum_{i=1}^{K} m_{-i} \cdot s^{-i}.$$

Model reduction via Padé approximation is well documented in the literature [30], and is known to be very fast. The complexity of computing *all* approximate models up to order k is in fact only $O(k^2)$. The reason for this is that each approximate model of order $i \leq k$ has the tridiagonal form

$$\left[\begin{array}{c|c} & \hat{c} \\ \hline \hat{b} & \hat{A} \end{array}\right] = \left[\begin{array}{c|ccc} & \gamma_1 & 0 & \dots & 0 \\ \hline \beta_1 & \alpha_1 & \gamma_2 & \ddots & \vdots \\ 0 & \beta_2 & \alpha_2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \gamma_i \\ 0 & \dots & 0 & \beta_i & \alpha_i \end{array}\right], \tag{6}$$

and hence only *three* additional coefficients α_{i+1}, β_{i+1}, γ_{i+1} have to be calculated to update this approximation to the one of order $i+1$. This is due to the underlying three term recurrence relation of orthogonal polynomials connected to the Padé approximation problem [18]. The complexity of such an update is $O(i)$ because only a few inner products between vectors (or polynomials) of length i are needed. This then explains why the accumulated complexity is $\sum_{i=1}^{k} O(i) = O(k^2)$.

The Padé algorithm is thus very appealing, even when the system is originally described by a system model $\{A, b, c\}$. One can then *evaluate* the moments m_{-i} up to a certain index K and construct from those a lower order system which matches these moments. This is particularly appealing when A is sparse since the moments are then cheap to construct (this involves merely a simulation of a discrete time system). This approach is e.g. followed in the AWE method for circuit simulation of electronic devices, which involves state space models of order up to 100,000 [1], [25], [26].

On the other hand, these AWE papers do not link Krylov projectors and Padé approximation together. Here we show that Padé approximation of the original system can be obtained *without* explicitly passing via the moments. Through the nonsymmetric Lanczos method [14, 22], one can

realize the reduced-order system $\{\hat{A}, \hat{b}, \hat{c}\}$ directly from the original system model. Approximating and eventually simulating the circuit through the Lanczos method requires approximately the same amount of effort as existing, explicit moment matching techniques. More importantly, the Lanczos method provides avenues for efficiently handling the shortcomings of Padé approximants. The oftentimes heuristic fixes developed for explicit moment matching can be replaced with more flexible and better conditioned Lanczos techniques.

3 Moment Matching

In previous AWE papers, the response of the original circuit is typically approximated via a two-step process. First, moments which correspond to frequency domain expansions of the circuit's impulse response are explicitly computed. Most commonly, the expansion is performed either about $s = 0$ to yield the low-frequency moments

$$m_i = cA^{-i-1}b, \quad i \geq 0$$

or about $s = \infty$ to yield the high-frequency moments (Markov parameters)

$$m_i = cA^{-i-1}b, \quad i < 0.$$

More generally, one can expand about an arbitrary $s = \sigma$ to obtain so-called shifted moments.

In the second step, the impulse response

$$\hat{h}(s) = \frac{n_{k-1}s^{k-1} + \ldots + n_1 s + n_0}{s_k + d_{k-1}s^{k-1} + \ldots + d_1 s + d_0} \tag{7}$$

of the approximate realization is forced to correspond to the first j low-frequency moments and $(2k-j)$ high-frequency moments of the original system. That is, given the Taylor series expansions

$$\hat{h}(s) = \sum_{i=1}^{\infty} \hat{h}_{-i} s^{-i} \quad \text{and} \quad \hat{h}(s) = \sum_{i=0}^{\infty} -\hat{h}_i s^i$$

one forces $\hat{h}_{-(2k-j)}$ through $\hat{h}_{j-1}$ to be $m_{-(2k-j)}$ through m_{j-1}. This feat is achieved by first solving

$$\begin{bmatrix} m_{-(2k-j-1)} & m_{-(2k-j-2)} & \cdots & m_{-(k-j)} \\ m_{-(2k-j-2)} & \ddots & \ddots & m_{-(k-j)} \\ \vdots & \ddots & \ddots & \vdots \\ m_{-(k-j)} & m_{-(k-j)} & \cdots & m_{(j-1)} \end{bmatrix} \begin{bmatrix} d_{k-1} \\ d_{k-2} \\ \vdots \\ d_0 \end{bmatrix} = - \begin{bmatrix} m_{-(2k-j)} \\ m_{-(2k-j-1)} \\ \vdots \\ m_{-(k-j+1)} \end{bmatrix} \tag{8}$$

for the coefficients in the denominator of (7). Then if $j \geq k$, one obtains the numerator coefficients from

$$\begin{bmatrix} n_0 \\ n_1 \\ \vdots \\ n_{k-1} \end{bmatrix} = - \begin{bmatrix} 0 & \cdots & 0 & m_0 \\ \vdots & \ddots & m_0 & m_1 \\ 0 & \ddots & \ddots & \vdots \\ m_0 & m_1 & \cdots & m_{k-1} \end{bmatrix} \begin{bmatrix} d_{k-1} \\ d_{k-2} \\ \vdots \\ d_0 \end{bmatrix};$$

otherwise, one uses

$$\begin{bmatrix} n_{k-1} \\ n_{k-2} \\ \vdots \\ n_0 \end{bmatrix} = \begin{bmatrix} 0 & \cdots & 0 & m_{-1} \\ \vdots & \ddots & m_{-1} & m_{-2} \\ 0 & \ddots & \ddots & \vdots \\ m_{-1} & m_{-2} & \cdots & m_{-k} \end{bmatrix} \begin{bmatrix} d_1 \\ \vdots \\ d_{k-1} \\ 1 \end{bmatrix}.$$

Via these expressions, the impulse response completely defines a reduced-order approximation (partial realization) which matches the desired moments of the original system. Thus (7) is a Padé approximation of the original circuit. Although it is not explicitly determined in existing AWE methods, a k^{th} order set of state space equations (i.e., $\{\hat{A}, \hat{b}, \hat{c}\}$) can be obtained which correspond exactly to $\hat{h}(s)$. Any k^{th} order set of state space equations satisfying

$$\hat{h}(s) = \hat{c}(sI - \hat{A})^{-1}\hat{b} \tag{9}$$

is in fact a valid description of the reduced-order system.

As an alternative to explicit moment matching, consider using the oblique Krylov projector $\pi_k = \pi_k^2 = V_k W_k^T$ to produce a k^{th} order model

$$\dot{\hat{x}} = (W_k^T A V_k)\hat{x} + (W_k^T b)u = \hat{A}\hat{x} + \hat{b}u \tag{10}$$

$$\hat{y} = (cV_k)\hat{x} = \hat{c}\hat{x} \tag{11}$$

for the original system in (1) and (2). The matrices $V_k \in \Re^{n\times k}$ and $W_k \in \Re^{n\times k}$ are biorthogonal, $W_k^T V_k = I$. Moreover, V_k and W_k are related to Krylov spaces, $\mathcal{K}_k$, in that

$$\text{COLSP}(V_k) = \mathcal{K}_k(A, b) = \text{span}\{b, Ab, \ldots, A^{k-1}b\} \tag{12}$$

$$\text{COLSP}(W_k) = \mathcal{K}_k(A^T, c^T) = \text{span}\{c^T, A^T c^T, \ldots, A^{k-1^T} c^T\}. \tag{13}$$

The utility of this Krylov projector comes from the fact that both V_k and W_k can be generated with only inner-products and matrix-vector multiplications. By taking advantage of the fact that the A matrix is sparse, one can compute the projector relatively cheaply.

But regardless of how quickly π_k can be computed, one is certainly also interested in the correspondence between the original system $\{A, b, c\}$ and the reduced-order system $\{\hat{A}, \hat{b}, \hat{c}\}$. An important insight into this relationship comes out of [16, 34].

Theorem 1 *Let the reduced-order system $\{\hat{A}, \hat{b}, \hat{c}\}$ be a restriction of the system $\{A, b, c\}$ by the projector π_k where V_k and W_k are defined as in (12) and (13) respectively. Then the first $2k$ Markov parameters of the original and reduced-order systems are identical, i.e.,*

$$cA^{i-1}b = \hat{c}\hat{A}^{i-1}\hat{b}$$

for $1 \le i \le 2k$.

Restating Theorem 1, the reduced-order model is a Padé approximation (partial realization) of the original system. And again, in a slightly different notation, the Krylov projector implicitly matches the first $2k$ high-frequency moments of the original system.

Through a projector corresponding to $\mathcal{K}_k(A, b)$ and $\mathcal{K}_k(A^T, c^T)$ one can obtain a state space realization which matches moments about $s = \infty$. In a completely analogous manner, a projector corresponding to $\mathcal{K}_k(A^{-1}, b)$ and $\mathcal{K}_k(A^{-T}, c^T)$ could be employed to generate a state space realization which matches moments about $s = 0$. And in fact, projectors can also be constructed (based on $\mathcal{K}_k(A, A^{-j/2}b)$ and $\mathcal{K}_k(A^T, (cA^{-j/2})^T)$) which are combination of both of these approaches.

4 The Nonsymmetric Lanczos Method

A popular technique for computing V_k and W_k in the Krylov projector is due to Lanczos [22]. This section serves as a brief review of the nonsymmetric Lanczos process. A standard implementation of the method is given below.

Algorithm 1 *A Standard Lanczos Algorithm*

1 Initiate the starting vectors as $v_1 = b/\beta_1$ *and* $w_1 = c^T/\gamma_1$ *so that* $\beta_1 = \pm\gamma_1$ *and* $w_1^T v_1 = 1$.

2 For $j = 1$ *to* k,

(a) set $\alpha_j = w_j^T A v_j$.

(b) set $r_j = Av_j - \alpha_j v_j - \gamma_j v_{j-1}$ *and* $q_j = A^T w_j - \alpha_j w_j - \beta_j w_{j-1}$.

(c) set $\beta_{j+1} = \sqrt{|r_j^T q_j|}$ *and* $\gamma_{j+1} = sign(r_j^T q_j) \cdot \beta_{j+1}$

(d) set $v_{j+1} = r_j/\beta_{j+1}$ *and* $w_{j+1} = q_j/\gamma_{j+1}$.

Given the starting vectors v_1 and w_1, the Lanczos algorithm produces the rectangular matrices $V_k = [v_1, \ldots, v_k] \in I\!R^{n\times k}$ and $W_k = [w_1, \ldots, w_k] \in I\!R^{n\times k}$ which satisfy the recursive identities

$$AV_k = V_k T_k + \beta_{k+1} v_{k+1} e_k^T \tag{14}$$

$$A^T W_k = W_k T_k^T + \gamma_{k+1} w_{k+1} e_k^T. \tag{15}$$

The vector e_k is the k^{th} standard basis vector and

$$T_k = \begin{bmatrix} \alpha_1 & \gamma_2 & & \\ \beta_2 & \ddots & \ddots & \\ & \ddots & \ddots & \gamma_k \\ & & \beta_k & \alpha_k \end{bmatrix}$$

is a truncated reduction of A. Generally, the elements β_i and γ_i are chosen so that $V_{k+1}^T W_{k+1} = I$. When V_{k+1} and W_{k+1} are biorthogonal, multiplying (14) on the left by W_k^T yields the relationship $W_k^T A V_k = T_k$. It is also convenient to denote the residuals $\beta_{k+1} v_{k+1}$ and $\gamma_{k+1} w_{k+1}$ as the vectors r_k and q_k, respectively. Then the expressions $r_k \in \mathcal{K}_{k+1}(A, v_1)$ and $q_k \in \mathcal{K}_{k+1}(A^T, w_1)$ come from the Lanczos identities in (14) and (15).

The relationship between the residuals and the corresponding Krylov spaces indicates the clear choice for the starting vectors v_1 and w_1. Mainly, if $v_1 = b/\beta_1$ and $w_1 = c^T/\gamma_1$, the matrices V_k and W_k correspond to the Krylov spaces $\mathcal{K}_k(A, b)$ and $\mathcal{K}_k(A^T, c^T)$ respectively. Hence V_k and W_k meet the desired form. And more importantly from (10) and (11), $\hat{A} = W_k^T A V_k = T_k$, $\hat{b} = W^T b = e_1\beta_1$ and $\hat{c} = cV_k = e_1^T\gamma_1$. Note that the Lanczos algorithm generates directly a tridiagonal $\hat{A}$ and vectors $\hat{b}$ and $\hat{c}$ with all components zero except the first one. This shows again the relation with the Padé algorithm since it produces the same tridiagonal form (6).

5 Advantages of Lanczos-based Model Reduction

Compared to explicit moment matching, the Lanczos method provides superior results and/or greater flexibility in several areas, which are explored in some detail in the remainder of this section.

5.1 Sensitivity of the Realization

In past AWE papers [25], the reduced-order model is expressed via the partial fraction expansion (PFE) of (7)

$$\hat{h}(s) = \sum_{j=1}^{k} \frac{r_j}{s - p_j}. \tag{16}$$

The poles, p_j, are the roots of the denominator of (7), which for simplicity are assumed to be unique. The residuals, r_j, are computed as

$$r_j = (s - p_j)\hat{h}(s) \Big|_{s=p_j}.$$

Note that given (16), it is simple to obtain a state space representation in Jordan canonical form

$$\left[\begin{array}{c|c} & \hat{c}_J \\ \hline \hat{b}_J & \hat{A}_J \end{array}\right] = \left[\begin{array}{c|ccc} & 1 & \cdots & 1 \\ \hline r_1 & p_1 & & \\ \vdots & & \ddots & \\ r_k & & & p_k \end{array}\right]. \tag{17}$$

It is also a simple matter to show that $\{\hat{A}_J, \hat{b}_J, \hat{c}_J\}$ satisfies (9). Unfortunately, given an arbitrary representation $\{\hat{A}, \hat{b}, \hat{c}\}$, the transformation required to obtain the Jordan canonical form is oftentimes poorly conditioned [20, 35]. More specifically, computing the eigenvectors and eigenvalues of an arbitrary $\hat{A}$ (which is related to computing the residuals, r_j, and frequencies, p_j) can be extremely sensitive.

On the other hand, the realization produced by the Lanczos method takes the tridiagonal form

$$\left[\begin{array}{c|c} & \hat{c}_L \\ \hline \hat{b}_L & \hat{A}_L \end{array}\right] = \left[\begin{array}{c|ccc} & \gamma_1 & & \\ \hline \beta_1 & \alpha_1 & \gamma_2 & \\ & \beta_2 & \ddots & \ddots \\ & & \ddots & \gamma_k \\ & & \beta_k & \alpha_k \end{array}\right]. \tag{18}$$

The transformation to obtain such a realization is known to be better conditioned in general [14, 20].

As an example, consider the simple system defined by

$$\left[\begin{array}{c|c} & c \\ \hline b & A \end{array}\right] = \left[\begin{array}{c|ccc} & 1 & 4.5 \cdot 10^{-2} & 6.75 \cdot 10^{-4} & 3.3375 \cdot 10^{-6} \\ \hline 1 & -3.01 & -3.03 & -1.03 & -(0.01 + 10^{-11}) \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right]. \tag{19}$$

Via implicit state space transformations, one can also realize this system in Jordan form (moment matching) and tridiagonal form (Lanczos method).

One of the eigenvalues of A is at -0.01 while the remaining three lie clustered around -1, $\{-0.9997, -1.0001 \pm 0.0002j\}$. Although these three eigenvalues are close to each other, they are by no means identical relative to the machine precision. Yet the proximity of the three eigenvalues is sufficient to demonstrate the ill-conditioning of the Jordan realization. In Figure 1, relative errors are plotted which correspond to the step responses of the initial (19), Jordan, and tridiagonal realizations.

As one would expect, the response error (dashed line) between the initial and tridiagonal realizations,

$$\text{error}_{\text{dash}}(t) = \frac{|y_{\text{initial}}(t) - y_{\text{tridiag}}(t)|}{|y_{\text{initial}}(t)|},$$

is nearly on the order of the machine precision. Yet the error between the responses of the initial and Jordan realizations (solid line) demonstrates a significant loss of precision in the Jordan realization.

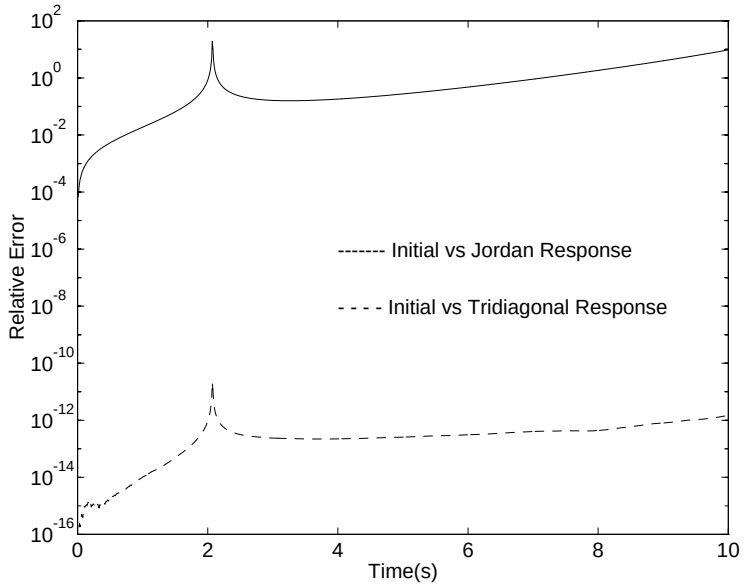


Figure 1: Relative error in the step responses of the Jordan and tridiagonal realizations vs. the step response of the original realization.

In fact, no digits are accurate during most of the response of the Jordan realization. Moreover, this difference is only for a fourth order system! In general, Jordan forms (17) are to be avoided. Better conditioned realizations (such as the one produced by the Lanczos method) must be employed when approximating the system.

5.2 Moment Scaling

Besides the sensitivity of the final realization, one must be concerned with the scaling of the moments (to simplify the discussion in this paragraph, consider only high-frequency ones for the time being.) If certain eigenvalues of A are extremely large, the size of the moments, cA^ib, will increase rapidly. As cA^ib becomes much larger than cb, the Hankel (moment) matrix in (8) will become singular to working precision, which results in a situation where realizations of size $\geq i$ cannot be computed. To overcome this difficulty, [25] proposes scaling the eigenvalues of A to keep the Hankel matrix nonsingular. Unfortunately, scaling does not address a second difficulty inherent to computing matrix-vector products containing increasing powers of A. Consider an A with one eigenvalue, λ_1, significantly larger than the others. Then as i increases, the product A^ib will converge to the eigenvector corresponding to λ_1. In finite precision, the information corresponding to the other eigenvectors will be lost in the higher moments. Regardless of how many moments are matched in this situation, the computed approximation never converges to the original circuit (see [1, example 1]). To contend with this difficulty, [1] suggests working around the problem by attempting to appropriately shift the moments. If forming a certain type of moment emphasizes an extreme eigenvalue of A, one must select other types of moments which emphasize information from other portions of the spectrum.

The Lanczos method, on the other hand, completely avoids both issues because it never computes the moments. Rather $\{\hat{A}, \hat{b}, \hat{c}\}$ are computed from the Krylov spaces $\mathcal{K}_k(A, b)$ and $\mathcal{K}_k(A^T, c^T)$.

As an example, consider the state space equations arising from a small, stiff RC ladder circuit

$$\left[\begin{array}{c|c} & c \\ \hline b & A \end{array}\right] = \left[\begin{array}{c|ccc} & 1 & -1 & 0 \\ \hline C_1^{-1} & -2C_1^{-1} & C_1^{-1} & 0 \\ 0 & C_2^{-1} & -2C_2^{-1} & C_2^{-1} \\ 0 & 0 & C_3^{-1} & -C_3^{-1} \end{array}\right]$$

where $C_1 = 10^{-3}$, $C_2 = 10^{-6}$, and $C_3 = 10^{-9}$. Allowing $k = 3$, the eigenvalues of the realization obtained with both explicit moment matching (about $s = 0$) and the Lanczos method (corresponding to $\mathcal{K}_k(A^{-1}, b)$ and $\mathcal{K}_k(A^{-T}, c^T)$) are presented in Table 1. Due to the poor scaling of the moments, explicit moment matching is unable to accurately determine the fastest pole. The Lanczos method, on the other hand, is able to capture all of the eigenvalues of A.

Table 1: Computed Eigenvalues of A

	eig 1	eig 2	eig 3
Exact	-9.98999000e2	-1.00000100e6	-1.00100100e9
Moment Match	-9.98999000e2	-1.00000078e6	-5.45486876e6
Lanczos	-9.98999000e2	-1.00000100e6	-1.00100100e9

5.3 Stability of the Approximation

When moment matching is employed, the reduced-order model for a stable circuits may be unstable [7]. To handle this problem, existing AWE papers prescribe searching the Padé table until a stable realization is located. Although such a technique must eventually succeed for a large enough k, it is both heuristic and potentially expensive. One cannot know a priori how many realizations must be generated before a stable one is acquired. Moreover, when a stable realization is determined, its size may exceed some desired value.

As an alternative to searching the Padé table, [17] stabilizes a realization of specified size k by incorporating implicit restarts into the Lanczos algorithm. With implicit restarts, the projector π_k is modified to $\tilde{\pi}_k = \tilde{W}^T\tilde{V}$ which corresponds to the new starting vectors

$$\begin{aligned} \tilde{v}_1 &= (A - \mu_p I)(A - \mu_{p-1} I) \ldots (A - \mu_1 I) v_1 \\ \tilde{w}_1 &= (A^T - \mu_p I)(A^T - \mu_{p-1} I) \ldots (A^T - \mu_1 I) w_1. \end{aligned} \tag{20}$$

These implicit restarts (which correspond to LR-steps [35] with the tridiagonal matrix, T_k) incorporate information from higher moments into the reduced-order model. Strategies for choosing the parameters μ_i in (20) can be employed to insure that this extra information stabilizes the partial realization. In a later section we demonstrate that when properly employed, implicit restarts can stabilize a realization with negligible computational effort.

5.4 Singularities in the Padé Table

Singularities can occur in the Padé table. Where these singularities exist in the table, partial realizations of the form (4) will not be possible. As pointed out by de Jong, the loss of numerical stability occurs in fact when "near singularities" are encountered in the Padé table. Such poorly conditioned entries in the Padé table should be avoided.

To the best of the authors' knowledge, this issue is not addressed in previous AWE papers. Yet the occurrence of ill-conditioned table entries is well-studied in the Lanczos algorithm [24], where it

is termed a "serious" breakdown. By employing "look-ahead" into the Lanczos method, [18, 8, 23], one possesses a powerful tool for detecting and avoiding ill-conditioned table entries. In the sequel we show also different techniques to avoid near sigularities in the Padé table, based on implicit shifts and multipoint expansions.

5.5 Reduced-order Simulation

One of the greatest advantages of the partial fraction expansion (16) is its low complexity in the evaluation of the discretized response of the system. After multiplying the summation, (16), by $u(s)$, one can determine the time response due to each individual frequency mode (eigenvalue) by analytically determining the inverse Laplace transform of the product. Then when calculating $y(t)$, the complexity per time step for the PFE representation is linear in the number of terms, k, in the expansion. At most $2k$ operations per time step are required if second order terms are included.

When using a tridiagonal representation (18) for $\{\hat{A}, \hat{b}, \hat{c}\}$, a simple integration scheme requires either multiplying with a tridiagonal matrix at each step or solving a system of equations involving such a tridiagonal matrix. In both cases, the complexity is again linear in k, requiring at most $3k$ operations per time step. If $\hat{A}$ is appropriately normalized, the number of operations can be reduced to $2k$ for the tridiagonal representation as well.

In addition, the implicitly restarted Lanczos method [17] provides the flexibility of "filtering" certain frequencies of the original circuit out of the realization. By removing undesired frequencies, one obtains an $\hat{A}_L$ whose eigenvalues are not too different from each other in scale. Since $\hat{A}_L$ is then not stiff, the step size needed for integrating $\dot{\hat{x}} = \hat{A}_L\hat{x} + \hat{b}_L u$ will not be too different from the step size used in the PFE approach (which is based on the individual eigenvalues of $\hat{A}$).

In a later section we show how to perform moment matching about multiple points to insure an accurate approximation of both the transient and steady state poles.

6 Implicitly Restarted Lanczos

The degree of success achieved in applying a Lanczos-type method is dependent upon the choice of starting vectors, v_1 and w_1. In some cases, such as the model reduction problem, one can make an educated initial guess for these starting vectors ($v_1 = b/\beta_1$ and $w_1 = c^T/\gamma_1$). But the stable plant, unstable reduced model issue demonstrates that what may appear as a good choice for the starting vectors can yield disastrous results. To overcome the results of a poor starting vector, one could repeatedly and explicitly recompute Krylov spaces with a modified pair of initial vectors. For lack of better data, one should use information from past results to refine these new starting vectors. Yet such an approach becomes computationally expensive when several such restarts are required. Each restart costs $O(k^2 n)$ flops when implemented in a numerically reliable manner (this requires re-biorthogonalization of the modified bases at each step of the recursion).

In this section, an implicit approach (an analogue to implicitly restarted Arnoldi [31]) is developed for generating the modified projector corresponding to the starting vectors in (12,13). It will be shown that given V_k and W_k, one can generate $\bar{V}_k$ and $\bar{W}_k$ more efficiently with implicit restarts. Also, experiments indicate a higher precision in $\bar{\pi}_k$ for the implicit method.

As a simple step between the standard Lanczos method and the new factorization corresponding to (12,13), we will first derive a technique for implicitly obtaining a $\tilde{V}_k$ and $\tilde{W}_k$ which correspond to the starting vectors $\tilde{v}_1 = \rho_v(A-\mu I)v_1$ and $\tilde{w}_1 = \rho_w(A^T-\mu I)w_1$. For the time being, the parameter, μ, is assumed to be real.

The first step in performing an implicit restart is obtaining the two LR-decompositions $L_v R_v = (T_k - \mu I)$ and $L_w R_w = (L_v^{-1} T_k L_v - \mu I)^T$ where in each case L is unit lower-triangular and R is upper-triangular. Due to the band preserving properties of the LR-decomposition [35], L and R are bidiagonal. With these LR-decompositions defined, (10) and (11) can be updated to

$$\begin{aligned} AV_k L_v L_w^{-T} &= V_k L_v L_w^{-T} (L_w^T L_v^{-1} T_k L_v L_w^{-T}) \\ &\quad + r_k e_k^T L_v L_w^{-T} \qquad (21) \\ A^T W_k L_v^{-T} L_w &= W_k L_v^{-T} L_w (L_w^{-1} L_v^T T_k^T L_v^{-T} L_w) \\ &\quad + q_k e_k^T L_v^{-T} L_w. \qquad (22) \end{aligned}$$

If one defines $\tilde{V}_k = V_k L_v L_w^{-T}$, $\tilde{W}_k = W_k L_v^{-T} L_w$ and $\tilde{T}_k = L_w^T L_v^{-1} T_k L_v L_w^{-T}$, then (21) and (22) become

$$\begin{aligned} A\tilde{V}_k &= \tilde{V}_k \tilde{T}_k + r_k e_k^T L_v L_w^{-T} \qquad (23) \\ A^T \tilde{W}_k &= \tilde{W}_k \tilde{T}_k^T + q_k e_k^T L_v^{-T} L_w. \qquad (24) \end{aligned}$$

To see the relationship between the new and old starting vectors (i.e., v_1 and w_1 versus $\tilde{v}_1$ and $\tilde{w}_1$), rewrite (10) as

$$\begin{aligned} (A - \mu I) V_k &= V_k (T_k - \mu I) + r_k e_k^T \\ &= V_k L_v R_v + r_k e_k^T. \qquad (25) \end{aligned}$$

Multiplying (25) on the right by $e_1 = L_w^{-T} e_1$ gives the relation

$$(A - \mu I) V_k e_1 = V_k L_v e_1 \rho_v^{-1} = V_k L_v L_w^{-T} e_1 \rho_v^{-1}$$

where $\rho_v^{-1} = e_1^T R_v e_1$. A similar derivation may be applied to (11) to yield that the two new starting vectors are in fact $\tilde{v}_1 = \rho_v (A - \mu I) v_1$ and $\tilde{w}_1 = \rho_w (A^T - \mu I) w_1$ where $\rho_w^{-1} = e_1^T R_w e_1$.

Clearly we are nearing the desired result; new starting vectors have been obtained which fit the desired form. Unfortunately, (23) and (24) are not valid Lanczos identities. Define l_v and l_w to be elements of the products $L_v L_w^{-T}$ and $L_v^{-T} L_w$ respectively. Then in (23) and (24), the residuals are multiplied by

$$\begin{aligned} e_k^T L_v L_w^{-T} &= \left(0, 0, \ldots 0, l_v^{(k,k-1)}, l_v^{(k,k)}\right) \text{ or} \\ e_k^T L_v^{-T} L_w &= \left(0, 0, \ldots 0, l_w^{(k,k-1)}, l_w^{(k,k)}\right) \end{aligned}$$

rather than just e_k^T. However, one can obtain a valid Lanczos factorization by simply truncating off a portion of (23) and (24). Rewrite (23) as

$$A\tilde{V}_k = (\tilde{V}_{k-1}, \tilde{v}_k, r_k) \left(\begin{array}{c|c} \tilde{T}_{k-1} & \tilde{\gamma}_k e_{k-1} \\ \tilde{\beta}_k e_{k-1}^T & \tilde{\alpha}_k \\ l_v^{(k,k-1)} e_{k-1}^T & l_v^{(k,k)} \end{array} \right)$$

and (24) as

$$A^T \tilde{W}_k = (\tilde{W}_{k-1}, \tilde{w}_k, q_k) \left(\begin{array}{c|c} \tilde{T}_{k-1}^T & \tilde{\beta}_k e_{k-1} \\ \tilde{\gamma}_k e_{k-1}^T & \tilde{\alpha}_k \\ l_w^{(k,k-1)} e_{k-1}^T & l_w^{(k,k)} \end{array} \right).$$

Equating the first $k-1$ columns of these two expressions yields the new Lanczos identities

$$A\tilde{V}_{k-1} = \tilde{V}_{k-1}\tilde{T}_{k-1} + \tilde{r}_{k-1}e_{k-1}^T \quad (26)$$
$$A^T\tilde{W}_{k-1} = \tilde{W}_{k-1}\tilde{T}_{k-1}^T + \tilde{q}_{k-1}e_{k-1}^T. \quad (27)$$

The new starting vectors are still defined as above while the new residual vectors are

$$\tilde{r}_{k-1} = \tilde{\beta}_k\tilde{v}_k + l_v^{(k,k-1)}r_k$$
$$\tilde{q}_{k-1} = \tilde{\gamma}_k\tilde{w}_k + l_w^{(k,k-1)}q_k.$$

One can also show that $\tilde{V}_{k-1}$, $\tilde{W}_{k-1}$, $\tilde{r}_{k-1}$, and $\tilde{q}_{k-1}$ meet the biorthogonality condition. It is further claimed that one can only insure $\tilde{r}_{k-1}^T\tilde{q}_{k-1} \neq 0$ because both LR-decompositions are included in the above development. For example, if *only* L_v was incorporated into the above expressions, μ's would exist (mainly the eigenvalues of T_k) for which the new residual vectors would be orthogonal.

¿From the above work, an extension to the general case is straightforward. One is now interested in a series of LR-decompositions. Define

$$L_{v_i}R_{v_i} = (\bar{L}_{i-1}T_k\bar{L}_{i-1}^{-1} - \mu_i I)$$
$$L_{w_i}R_{w_i} = (L_{v_i}^T\bar{L}_{i-1}^{-T}T_k^T\bar{L}_{i-1}^T L_{v_i}^{-T} - \mu_i I)$$

where

$$\bar{L}_{i-1} = L_{w_{i-1}}^T L_{v_{i-1}}^{-1}\cdots L_{w_1}^T L_{v_1}^{-1}.$$

Note that in practice, one should determine L_{v_i} and L_{w_i} via an implicit LR approach (see §5). Pairs of complex conjugate shifts would be handled via double LR shifts [35].

Corresponding to p implicit restarts are the new identities

$$A\bar{V}_{k-p} = \bar{V}_{k-p}\bar{T}_{k-p} + \bar{r}_{k-p}e_{k-p}^T$$
$$A^T\bar{W}_{k-p} = \bar{W}_{k-p}\bar{T}_{k-p}^T + \bar{q}_{k-p}e_{k-p}^T$$

where $\bar{T}_{k-p}$, $\bar{V}_{k-p}$ and $\bar{W}_{k-p}$ are the appropriate submatrices of $\bar{T}_k = \bar{L}_p T_k \bar{L}_p^{-1}$, $\bar{V}_k = V_k\bar{L}_p^{-1}$ and $\bar{W}_k = W_k\bar{L}_p^T$. The new residuals are

$$\bar{r}_{k-p} = \bar{\beta}_{k-p+1}\bar{v}_{k-p+1} + \bar{l}_v^{(k,k-p)}r_k$$
$$\bar{q}_{k-p} = \bar{\gamma}_{k-p+1}\bar{w}_{k-p+1} + \bar{l}_w^{(k,k-p)}q_k \quad (28)$$

where the $\bar{l}$'s are elements of $\bar{L}_v = L_{v_1}\cdots L_{v_p}$ and $\bar{L}_w = L_{w_1}\cdots L_{w_p}$. Most importantly, the starting vectors do indeed satisfy (12) and (13). Note that only p additional standard Lanczos iterations are required to obtain an order-k Lanczos factorization corresponding to $\bar{v}_1$ and $\bar{w}_1$. Note that in [17] we use HR-steps instead of double LR-steps because these are better conditioned in general. Moreover, the existence of the HR-decomposition is tied to the lack of serious Lanczos breakdowns. Avoiding breakdowns is thus linked to finding shifts such that the HR-steps exist (see [17]).

7 Example: The Portable CD Player

The Compact Disc player is a well-known mechanism for reproducing sound from a disc. At the heart of the CD player is an optical unit (consisting of a laser diode, lenses, and photodectors) which is mounted on the end of a radial arm [5]. In particular, we will be interested in the relationship

between the voltage applied to the magnetic lens actuator and the resulting lens position. Traditionally, the behavior of the lens position is represented by a third-order set of equations. However, controllers designed from these simple, low-order systems experience difficulties when employed in newer, portable CD players [5].

To obtain a higher-order controller for the CD player, a better model of its behavior is required. Via finite element approximation, various portions of the CD player were modeled and combined to yield a system of equations of order $n = 120$. It is unfortunate that the size of A is relatively small. But, this example is very adequate in demonstrating both the severity of the unstable partial realization problem and the power of implicit restarts in solving this problem.

A very valid concern is the total number of Lanczos realizations (T_k, $W_k^T b$, cV_k, $1 \leq k \leq 120$) which are actually unstable. If there are only a few values of k for which T_k is unstable, then incorporating implicit restarts into the standard Lanczos method is unnecessary work. But Figure 2 demonstrates that T_k stable is the exception, not the rule, for this example. In general, one cannot count on stumbling upon stability at the appropriate recursion step k.

However, employing implicit restarts with appropriate choices for the parameters, μ_i, (see §5) quickly stabilizes the reduced-order model. The number of restarts needed to obtain a stable $\bar{T}$ given various T_k's is indicated in Table 2.

Table 2: Restarts Needed to Stabilize an Order-k Model

	$k = 20$	$k = 30$	$k = 40$	$k = 50$	k=60
Restarts	5	0	2	3	1

It is also important to note that in this example, implicit restarts do not have a detrimental effect on the accuracy of the final, stabilized model (and, in fact, they are extremely beneficial when the original model is unstable). For example, Figure 3 displays the impulse responses for both an initially stable Lanczos model (T_{47}) and a restarted (stabilized) Lanczos model ($\bar{T}_{50}$). Even with a modified projector, $\bar{\pi}_k$, the restarted model's response is closer to that of the actual system.

8 Implementation Remarks

Until now, several important implementation details have been glossed over. This section will quickly address some of these issues.

Paramount in arriving at a stabilizing projector from an initial projector is proper selection of the parameters (shifts), μ_i. Although there is certainly an endless number of possibilities for the shifts, the following theorem (an analogue to one in [31]) indicates a practical policy for choosing the restart parameters.

Theorem 2 *Let $\{\theta_1, \ldots, \theta_k\} \cup \{\mu_1, \ldots, \mu_p\}$ be a disjoint partition of the spectrum of T_{k+p} and define $\bar{T}_k$ to be the the tridiagonal matrix resulting from p implicit restarts with shifts μ_1 through μ_p. The eigenvalues of $\bar{T}_k$ are $\{\theta_1, \ldots, \theta_k\}$.*

Restarting with exactly p eigenvalues of T_{k+p} as the shifts "tosses out" these p eigenvalues from $\bar{T}_k$. For our application, given that T_k is unstable, one needs to proceed until a T_{k+p} is determined with less than p unstable poles. Then via implicit restarts and Theorem 2, one can remove the unstable poles to yield a stable $\bar{T}_{k+q}$, $0 \leq q \leq p$. Note that the condition "find T_{k+p} with less than p unstable poles" is much less restrictive than finding a stable T_{k+p}.

Although the ultimate goal in choosing the shifts, μ_i, is to obtain a stable realization, there is also a more basic concern, the sensitivity of LR-decompositions. Because the LR-decomposition of

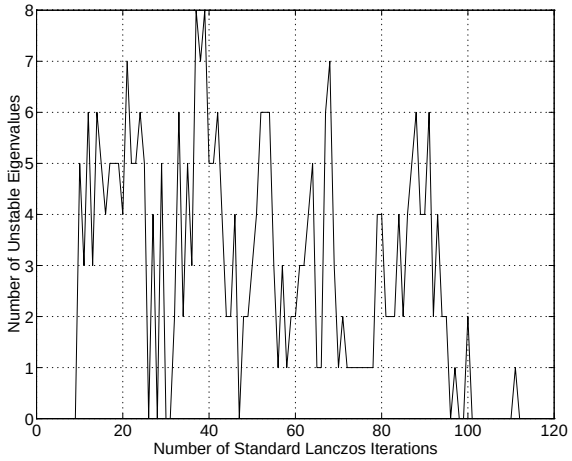


Figure 2: The number of unstable eigenvalues in T_k, where k is the number of Lanczos iterations.

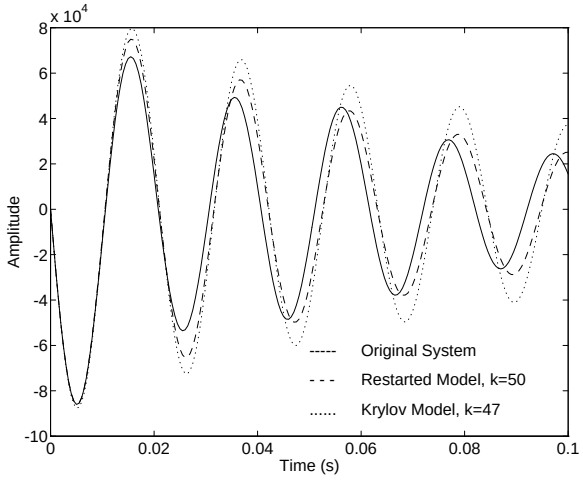


Figure 3: Impulse responses for CD player models.

T_{k+p} exists only if its $k+p-1$ leading principal minors are nonzero, a shift cannot be an eigenvalue of any of these $k+p-1$ minors of T_{k+p}. But the existence of L is not sufficient. L must be kept well-scaled to maintain the biorthogonality of $\bar{V}$ and $\bar{W}$. In general, selecting the shifts to be the unstable eigenvalues of T_{k+p} seems to generate well-scaled L's. However, one should be aware that a slight perturbation on a shift may be required in some cases.

In practice, these LR-decompositions should be performed implicitly [35]. That is, a series of elementary transformations should be used to chase a bulge down the tridiagonal of T_k. For the single-shift case, generating $\bar{T}_{k-1}$ in this manner costs only $O(k^2)$ flops while $\bar{V}_{k-1}$ and $\bar{W}_{k-1}$ can each be generated with $O(kn)$ flops. An additional $O(kn)$ flops is needed for the single Lanczos iteration (full reorthogonalization) yielding $\bar{T}_k$. Note that an explicit restart (with full reorthogonalization), on the other hand, requires $O(k^2 n)$ operations.

9 Multi-point Padé approximation with Lanczos

In this section, we extend the results and benefits of existing Lanczos methods to the case of generalized state space systems :

$$\left\{ \begin{array}{rcl} E\dot{x} & = & Ax + bu \\ y & = & cx \end{array} \right. \tag{29}$$

This new technique possesses the additional advantage of allowing for *multi-point* Padé approximation. That is, the resulting reduced-order model of dimension $k = \overline{\imath j}$ satisfies

$$m_j(s_i) = \hat{m}_j(s_i), \quad j = 1, 2, \ldots, 2\bar{\jmath}, \quad i = 1, 2, \ldots, \bar{\imath}, \tag{30}$$

where

$$m_j(s_i) = c\left\{(A - s_i E)^{-1} E\right\}^{j-1} (A - s_i E)^{-1} b,$$

is the j^{th} moment of (1) about the expansion frequency s_i and

$$\hat{m}_j(s_i) = \hat{c}\left\{(\hat{A} - s_i \hat{E})^{-1} \hat{E}\right\}^{j-1} (\hat{A} - s_i \hat{E})^{-1} \hat{b}$$

is the j^{th} moment of reduced-order model about s_i. It will be assumed for notational simplicity that $2\bar{\jmath}$ moments are to be matched about each of the frequencies s_1 through $s_{\bar{\imath}}$. In general, the number of moments matched may vary from expansion frequency to expansion frequency. One can easily modify the following results to handle this case.

The variant of the Lanczos method employed to generate a reduced-order model $\{\hat{E}, \hat{A}, \hat{b}, \hat{c}\}$ satisfying (30) will be denoted the *rational* Lanczos algorithm as it is an adaptation of the *rational* Arnoldi method of [27, 28]. The most glaring difference between the two rational methods is that rational Lanczos computes a biorthogonal V_k and W_k rather than an orthogonal V_k. There are, however, smaller dissimilarities between the two methods which are necessary to insure that the oblique projector, $\pi = V_k W_k$, of rational Lanczos yields multi-point Padé approximants.

Algorithm 2 *Rational Lanczos algorithm*

Initialize $r_0 = (A - s_1 E)^{-1} b$ *and* $q_0 = c^T$;

For $i = 1$ *to* $\bar{\imath}$,

For $j = 1$ *to* $\bar{\jmath}$,

(1) $k = (i-1)\bar{\jmath} + j$;

(2) $h_{k,k-1} = \sqrt{|r_{k-1}^T q_{k-1}|}$

(3) $v_k = (r_{k-1}/h_{k,k-1})$ *and*
$w_k = sign(r_{k-1}^T q_{k-1}) \cdot (q_{k-1}/h_{k,k-1})$;

(4) *if* $j < \bar{\jmath}$ *and* $i < \bar{\imath}$,

(4.1) $r_k = (A - s_i E)^{-1} E v_k$ *and*
$q_k = E^T (A - s_i E)^{-T} w_k$;

else if $j = \bar{\jmath}$ *and* $i < \bar{\imath}$,

(4.2) $r_k = (A - s_{i+1} E)^{-1} b / h_{1,0}$ *and*
$q_k = E^T (A - s_{i+1} E)^{-T} c^T$;

else

(4.3) $r_k = (A - s_1 E)^{-1} E v_{\bar{\jmath}}$ *and*

$$q_k = E^T(A - s_1E)^{-T} w_{\bar{\jmath}};$$

end

(5) $h_{1\ldots k,k} = W_k^T r_k$ *and* $g_{1\ldots k,k} = V_k^T q_k$;

(6) $r_k = r_k - V_k h_{1\ldots k,k}$ *and* $q_k = q_k - W_k g_{1\ldots k,k}$;

end

end

$v_{\overline{ij}+1} = (r_{\overline{ij}}/h_{\overline{ij}+1,\overline{ij}})$ *where* $h_{\overline{ij}+1,\overline{ij}} = \sqrt{|r_{\overline{ij}}^T q_{\overline{ij}}|}$.

Strong similarities exist between Algorithm 2 and the traditional Lanczos algorithm (Algorithm 1). The key difference between the standard and rational Lanczos algorithms lies in step (4) of Algorithm 1. In rational Lanczos the matrix, $(A - sE)^{-1}E$, multiplying the previous v vector changes with the expansion frequency. By making this matrix a function of s, the following sequences of Krylov spaces are computed (see [12] for a proof).

Theorem 3 *If V_k and W_k are the results of the first k steps of the rational Lanczos algorithm with $1 \le k \le \overline{ij}$ then*

$$colsp(V_k) = \Big\{ \mathcal{K}_{k-\bar{\jmath}(i-1)}\left((A - s_iE)^{-1}E, (A - s_iE)^{-1}b\right) \bigcup_{l=1}^{i-1} \mathcal{K}_{\bar{\jmath}}\left((A - s_lE)^{-1}E, (A - s_lE)^{-1}b\right) \Big\}$$

where $i - 1$ is the quotient of $k/\bar{\jmath}$. Correspondingly,

$$colsp(W_k) = \Big\{ \mathcal{K}_{\bar{\jmath}}\left(E^T(A - s_1E)^{-T}, c^T\right) \bigcup_{l=2}^{i-1} \mathcal{K}_{\bar{\jmath}}\left(E^T(A - s_lE)^{-T}, E^T(A - s_lE)^{-T}c^T\right) \bigcup \mathcal{K}_{k-\bar{\jmath}(i-1)}\left(E^T(A - s_iE)^{-T}, E^T(A - s_iE)^{-T}c^T\right) \Big\}$$

if $i > 1$; otherwise

$$colsp(W_k) = \mathcal{K}_k\left(E^T(A - s_1E)^{-T}, c^T\right).$$

Recall that the V_k and W_k matrices resulting from the standard Lanczos method each corresponded to a single Krylov space. In the rational Lanczos method, multiple Krylov spaces are computed. Each space corresponds to an expansion frequency s_i. The cost of combining multiple Krylov spaces into V_k and W_k is the loss of a three-term recurrence in step (6) of the algorithm. Thus one should expect to see upper-Hessenberg rather than tridiagonal matrices appearing out of the rational Krylov projection.

For the remainder of this section, it will be assumed that Algorithm 1 is executed to completion and the value of k will be fixed as $k = \bar{k} \equiv \overline{ij}$. Then given the results of the rational Lanczos method, we will define the reduced-order model so that

$$\begin{aligned} \hat{A} &= K_{k,k} + s_1 H_{k,k}, & \hat{E} &= H_{k,k}, \\ \hat{b} &= W_k^T (A - s_1E)^{-1} b, & \hat{c} &= cV_k K_{k,k}. \end{aligned} \tag{31}$$

To begin to motivate the choices in (31), rewrite the definition (1) of the original system as

$$\begin{cases} (A - s_1E)^{-1}E\dot{x} = (A - s_1E)^{-1}(A - s_1E + s_1E)x + (A - s_1E)^{-1}bu \\ y = cx. \end{cases}$$

The restriction of the original system by the projector π is formed by replacing the state vector, x, with $V_kW_k^Tx$ and multiplying on the left by W_k^T to yield

$$\begin{cases} W_k^T(A - s_1E)^{-1}E\pi\dot{x} = W_k^Tx + s_1W_k^T(A - s_1E)^{-1}E\pi x + W_k^T(A - s_1E)^{-1}bu \\ \hat{y} = cV_kW_k^Tx. \end{cases} \tag{32}$$

If one *temporarily* assumes that $K_{k,k}$ is invertible, one can rewrite (32) as

$$\begin{cases} H_{k,k}K_{k,k}^{-1}W_k^T\dot{x} = s_1H_{k,k}K_{k,k}^{-1}W_k^Tx + W_k^Tx + W_k^T(A - s_1E)^{-1}bu \\ \hat{y} = cV_kW_k^Tx. \end{cases}$$

which in turn becomes

$$\begin{cases} H_{k,k}\dot{\hat{x}} = (K_{k,k} + s_1H_{k,k})\hat{x} + W_k^T(A - s_1E)^{-1}bu \\ \hat{y} = cV_kK_{k,k}\hat{x}. \end{cases} \tag{33}$$

by defining $\hat{x}(t)$ to be $K_{k,k}^{-1}W_k^Tx(t)$.

Comparing (29) and (33) indicates that the prescribed choices for $\hat{A}$, $\hat{E}$, $\hat{b}$ and $\hat{c}$ are quite logical. However, (33) was obtained assuming $K_{k,k}$ to be invertible. This assumption is in fact not necessary for our purposes. The following result (see [12] for a proof) states that the reduced-order model corresponding to (31) matches the desired moments of the original system without placing any restrictions on the invertibility of $K_{k,k}$ or E.

Theorem 4 *Let the j^{th} moments of the original and reduced order systems about the expansion frequency s_i be $m_j(s_i) = c\{(A - s_iE)^{-1}E\}^{j-1}(A - s_iE)^{-1}b$ and $\hat{m}_j(s_i) = \hat{c}\{(\hat{A} - s_i\hat{E})^{-1}\hat{E}\}^{j-1}(\hat{A} - s_i\hat{E})^{-1}\hat{b}$ respectively. If $\hat{A} = K_{k,k} + s_1H_{k,k}$, $\hat{E} = H_{k,k}$, $\hat{b} = W_k^T(A - s_1E)^{-1}b$ and $\hat{c} = cV_kK_{k,k}$ where $H_{k+1,k}$, $K_{k+1,k}$, V_{k+1} and W_{k+1} are the results of Algorithm 1 with $k = \overline{ij}$, then $m_j(s_i) = \hat{m}_j(s_i)$ for $i = 1, 2, \ldots, \bar{\imath}$ and $j = 1, 2, \ldots, 2\bar{\jmath}$.*

As pointed out earlier, the matrices $K_{k,k}$ and $H_{k,k}$ are no longer tridiagonal but are still guaranteed to be upper Hessenberg. In fact one shows that they are *essentially* tridiagonal, except for two nonzero columns above the diagonal at each transition to another expansion point. For a 2-point, 8-dimensional model, the matrices $K_{k,k}$ and $H_{k,k}$ have thus the form :

$$\begin{bmatrix} x & x & & & x & x & & \\ x & x & x & & x & x & & \\ & x & x & x & x & x & & \\ & & x & x & x & x & & \\ & & & x & x & x & & \\ & & & & x & x & x & \\ & & & & & x & x & x \\ & & & & & & x & x \end{bmatrix}$$

The projection resulting from the rational Lanczos method satisfies the multi-point condition of (30). It appears that the benefits of the standard Lanczos method can be easily extended to the

rational Lanczos approach. These are : avoiding computing the moments, avoiding break-downs and look-ahead methods, and the possibility of fast simulation methods.

We finally point out that just as Padé approximation comes up in system theory as the partial realization problem, the multi-point Padé approximation problem has been studied in system theory under the name rational interpolation. References include [9] and [2].

10 Example

As a brief example of the utility of multi-point Padé approximations, we take now the 120^{th} order system describing the effects of a magnetic actuator on the radial tracking arm of a portable compact disc player, see [17]. Figure 1 plots the frequency responses of the original system (solid line), a 24^{th} order Padé approximation about $s_0 = \infty$ (dashed line), a 12^{th} order Padé approximation about $s_0 = 0$ (dotted line) and a 6^{th} order multi-point Padé approximation (dashed-dotted line) for the CD player. The multi-point approximation matches six moments expanded about $s_1 = 0$, four moments about $s_2 = 10^5$, and two moments about $s_3 = 10^4$. To keep this example short, we do not discuss the algorithm used to choose these expansion points here.

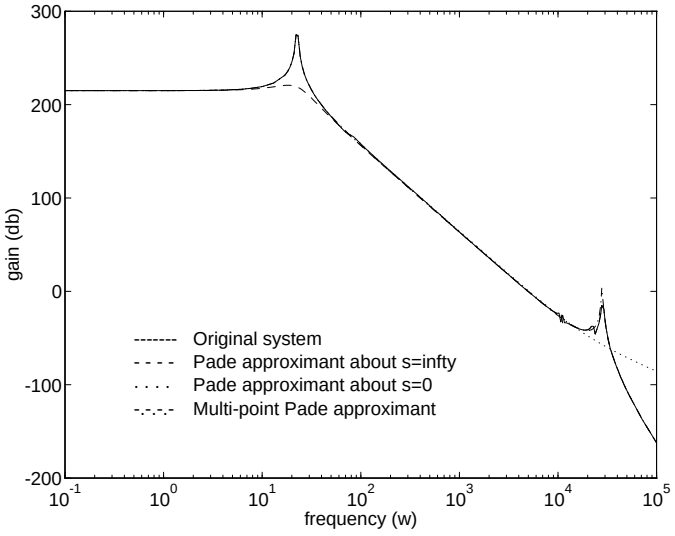


Figure 4: Frequency responses for the example.

Note that the frequency response of the original system displays two sharp peaks at $w \approx 30$ and $w \approx 30^4$. The frequency response of the multi-point Padé approximation captures both of these peaks and is almost indistinguishable with the response of the original system. As one should expect, the Padé approximation about $s_0 = 0$ displays the first peak but demonstrates significant error at high frequencies. The approximation about infinity, on the other hand, captures the second peak but smoothes over the peak at $w \approx 30$.

The impulse response of the stable, original system is dominated by those modes corresponding to the frequency response peak at $w \approx 30$. The multi-point approximation is stable and its impulse response recreates that of the original system with great precision. The Padé approximation about $s_0 = 0$ identifies those modes corresponding to the low-frequency peak but its impulse response is unstable. However, the techniques of Section 6 can stabilize the $s_0 = 0$ approximation so that the

$s_0 = 0$ response to an impulse (which may or may not be the input of interest for a given application) follows that of the original system with great precision. Finally, the Padé approximation about infinity is not stable nor does it capture those modes corresponding to the low-frequency peak. As a result, even a stabilized model about $s_0 = \infty$ does a poor job of approximating the system's impulse response. Note that because the Krylov sequences corresponding to $s_0 = \infty$ do not invert A, the oftentimes desirable low-frequency information is lost.

11 Conclusion

Both explicit moment matching and the Lanczos algorithm are efficient techniques for generating partial realizations of large-scale systems. But by either avoiding a difficulty or providing well-defined techniques for fixing it, the Lanczos method is better suited for handling the problems inherent to Padé approximation.

Through Algorithm 2, the Lanczos method can be extended to treat multiple expansion frequencies. Multi-point approximation shows promise in several applications [9, 26, 36] as an approach for handling false instabilities and frequency response errors in the reduced-order model. However, the techniques for choosing the expansion frequencies are still rather heuristic; a more formal approach should be explored in future work.

The Lanczos method must also still be extended to the multiple input-multiple output (MIMO) case. Block versions of the Lanczos method are already in existence, see [21], [4] for example. Combining the multi-point method with implicit shifts in order to guarantee stability of the reduced order system is also a point of further research.

Finally, we note that the inversion of $(A - s_i E)$ is an area requiring additional work. Regardless of whether explicit moment matching or the Lanczos method is being employed, one must avoid explicit inversions and instead utilize sparse factorizations of the matrix or iterative techniques. One must further insure that the values for s_i are chosen so that $(A - s_i E)$ is well-conditioned.

References

[1] M. M. Alaybeyi, J. Y. Lee and R. A Rohrer, "Numerical integration and asymptotic waveform evaluation," *Tech. Digest IEEE International Conf. on Computer-Aided Design*, 1992.

[2] A.C. Antoulas and B.D.O. Anderson, Rational interpolation and state variable realizations, *Linear Algebra and Appl.*, vol. 137, pp. 479–509, 1990.

[3] C. Baldwin and E. Gallopoulos, "High accuracy solution of parabolic partial differential equations by rational approximation techniques on Cedar," Technical Report, Center for Supercomputing Research and Development, Univ. of Illinois, Urbana, IL, 61801.

[4] D. L. Boley, "Krylov space methods on state-space control models," Technical Report, Univ. of Minnesota, Minneapolis, MN 55455, 1992.

[5] O. H. Bosgra, G. Schoolstra and M. Steinbuch, "Robust control of a compact disc player," in *Proc. IEEE 31th Conf. on Decision and Control*, (Tucson, AZ), 1992.

[6] C. Brezinski, *Padé-Type Approximation and General Orthogonal Polynomials*, ISNM vol. 50, Birkhäuser, Basel, 1980.

[7] C. I. Byrnes and A. Lindquist, "The stability and instability of partial realizations," *Sys. Control Lett.*, vol. 2, pp. 99-105, 1982.

[8] R.W. Freund, M.H. Gutknecht and N.M. Nachtigal, An implementation of the look-ahead Lanczos algorithm for non-Hermitian matrices, *SIAM J. Sci. Comp.*, vol. 14, pp. 137–158, 1993.

[9] C. Kenney, A. J. Laub and S. Stubberud, "Frequency response computation via rational interpolation," *IEEE Trans. Automatic Control*, vol. AC-38, pp. 1203–1213, 1993.

[10] K. Gallivan, E. Grimme and P. Van Dooren, "Asymptotic waveform evaluation via a Lanczos method", *Appl. Math. Lett.*, Vol. 7, pp.75-80, 1994.

[11] K. Gallivan, E. Grimme and P. Van Dooren, "Reduction and simulation of large-scale dynamical systems with Lanczos methods", *IEEE Conf. Dec. Contr.*, Lake Buena Vista FL, pp. 443-448, 1994.

[12] K. Gallivan, E. Grimme and P. Van Dooren, "Multi-point Padé approximation of large-scale systems via a two-sided rational Krylov algorithm," Techical Report, Univ. of Illinois, Urbana, IL 61801, 1994.

[13] E. Gallopoulos and Y. Saad "Efficient solution of parabolic equations by Krylov approximation methods," *SIAM J. Sci. Stat. Comput.*, vol. 13, pp. 1236-1264, 1992.

[14] G. H. Golub and C. Van Loan, *Matrix Computations*, 2nd ed. Baltimore, MD: Johns Hopkins University Press, 1989.

[15] W. B. Gragg, "The Padé table and its relation to certain algorithms of numerical analysis", *SIAM Rev.*, vol. 14 pp. 1-62, 1972.

[16] W. B. Gragg and A. Lindquist, "On the partial realization problem," *Linear Alg. Appl.*, vol. 50, pp. 277-319, 1983.

[17] E. Grimme, D. Sorensen and P. Van Dooren, "Model reduction of state space systems via an implicitly restarted Lanczos method", *Numerical Algorithms*, 1995.

[18] M. H. Gutknecht, "A completed theory of the unsymmetric Lanczos process and related algorithms, Part I", *SIAM J. Matrix Anal. Appl.*, vol. 13, pp. 594-639, 1992, "Part II", idem, vol. 15, pp. 15-58, 1994.

[19] I. M. Jaimoukha and E. M. Kasenally, "Oblique projection methods for large scale model reduction," *SIAM J. Matrix Anal. Appl.*, 1995.

[20] B. Kågström and P. Van Dooren, "A generalized state-space approach for the additive decomposition of a transfer matrix", *J. Num. Linear Alg. Applns.*, vol. 1, pp. 165-181, 1992.

[21] H. M. Kim and R. R. Craig Jr., Structural dynamics analysis using an unsymmetric block Lanczos algorithm, *Inter. J. Numer. Methods in Engineering*, vol. 26, pp. 2305–2318, 1988.

[22] C. Lanczos "An iteration method for the solution of the eigenvalue problem of linear differential and integral operators," *J. Res. Nat. Bur. Standards*, vol. 45, pp. 255-282, 1950.

[23] B.N. Parlett, D.R. Taylor and Z.S. Liu, A look-ahead Lanczos algorithm for unsymmetric matrices, *Math. Comp.*, vol. 44, pp. 105–124, 1985.

[24] B. N. Parlett, "Reduction to tridiagonal form and minimal realizations," *SIAM J. Matrix Anal. Appl.*, vol. 13, pp. 567-593, 1992.

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

[26] V. Raghavan, R. A. Rohrer, L. T. Pillage, J. Y. Lee, J. E. Bracken and M. M. Alaybeyi, "AWE-inspired," *Proceedings of the IEEE Custom Integrated Circuits Conference*, 1993.

[27] A. Ruhe, "Rational Krylov algorithms for nonsymmetric eigenvalue problems, II: Matrix pairs," *Lin. Alg. Appl.*, vol. 194, 1993.

[28] A. Ruhe, "The rational Krylov algorithm for nonsymmetric eigenvalue problems, III: Complex shifts for real matrices," *BIT*, vol. 34, 1994.

[29] Y. Saad, "Analysis of some Krylov subspace approximations to the matrix exponential operator," *SIAM J. Numer. Anal.* vol. 29, pp. 209-228, 1992.

[30] Y. Shamash, "Model reduction using the Routh stability criterion and the Padé approximation technique," *Int. J. Control*, vol. 21, pp. 475-484, 1975.

[31] D. C. Sorensen, "Implicit application of polynomial filters in a K-step Arnoldi method," *SIAM J. Matrix Anal. Appl.*, vol. 13, pp. 357-385, 1992.

[32] T. J. Su and R. R. Craig Jr., "An unsymmetric Lanczos algorithm for damped structural dynamics systems," *Proc. 33rd Conf. on Structures, Structural Dynamics and Materials*, 1992.

[33] P. Van Dooren, "Numerical linear algebra techniques for large scale matrix problems in systems and control", *Proc. IEEE 31st Conf. on Decision and Control*, (Tucson, AZ), 1992.

[34] C. D. Villemagne and R. E. Skelton, "Model reduction using a projection formulation," *Int. J. Control*, vol. 46, pp. 2141-2169, 1987.

[35] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford, England: Claredon Press, 1965.

[36] H. Xiheng, "FF-Padé method of model reduction in frequency domain", *IEEE Trans. Autom. Control*, vol. AC-32, pp. 243-246, 1987.