

On Some Recent Developments in Projection-based Model Reduction

E. Grimme

Intel Corporation

Santa Clara, CA 95051, USA

E-mail: egrimme1@td2cad.intel.com

K. A. Gallivan

*Computational Science and Engineering Program
and Computer Science Department*

Florida State University

Tallahassee, FL 32306-4530, USA

E-mail: gallivan@cs.fsu.edu

P. M. Van Dooren

Université Catholique

Louvain, Belgium

E-mail: vandooren@anma.ucl.ac.be

In this paper, we describe some recent developments in the use of projection methods to produce reduced-order models for linear time-invariant dynamic systems. Previous related efforts in model reduction problems from various applications are also discussed. An overview is given of the theory governing the definition of the family of Rational Krylov methods, the practical heuristics involved and the important future research directions.

1 Introduction

A surprisingly large variety of physical phenomena is modeled with linear, time-invariant (LTI) dynamic systems. The advantages of this approach include the relative ease by which both the initial model development and the eventual mathematical treatment can be achieved. Models can frequently be acquired through discretizations such as the common finite difference and finite element approaches. A range of techniques from the backward Euler method to multistep methods exists for solving the ordinary differential equations (ODE) that describe the system. Stable, well-understood numerical linear algebra algorithms, e.g., a reduction to Schur form by orthogonal transformations, dominate the low-level mathematical operations. When combined in various fashions, techniques such as the above enable the robust analysis, control or simulation of a large class of physical applications.

However, there is a pressing need for novel approaches for treating LTI

dynamic systems, particularly with respect to the linear algebra algorithms. Many physical models are becoming more complex due to either increased system size or an increased desire for detail. Discretizations of three-dimensional behavior are becoming common. Sources for such applications include the modeling of off-chip (and increasingly on-chip) interconnects in high-speed circuit designs, and modeling of the North American power grid arising from planning problems in an increasingly deregulated power industry. Although such models tend to accurately describe the behavior of the underlying physical system, their complexity leads to high analysis and simulation costs with traditional numerical techniques. Methods which exploit structure in the models such as sparsity are critical. This has resulted in a surge in interest in iterative methods for solving large sparse linear systems and large sparse eigenvalue problems. Such techniques are key computational kernels in most simulation algorithms for large dynamic systems.

In some cases, however, there is a need to go even further. Despite the use of efficient computational kernels, the model may still require an unacceptable amount of time to evaluate. It is then necessary to create a second model that is significantly lower in complexity while preserving important aspects of the original system. This is the model reduction problem for linear time-invariant dynamic systems.

It is assumed that the original system is described by the generalized state-space equations

$$\begin{cases} E\dot{x}(t) = Ax(t) + bu(t) \\ y(t) = c^T x(t) + du(t). \end{cases} \quad (1)$$

The vector $x(t) \in \mathbb{R}^{N \times 1}$ is the vector of state variables. For simplicity, it is assumed that the system is single-input single-output (SISO) so that the input $u(t)$ and output $y(t)$ are scalar functions of time. Extensions to multiple-input multiple-output (MIMO) systems exist but will not be treated here. Finally, and as is the case for nearly all large-scale problems, it is assumed that the system matrix $A \in \mathbb{R}^{N \times N}$ and descriptor matrix $E \in \mathbb{R}^{N \times N}$ are large and sparse or structured.

A reduced-order approximation to (1) takes the corresponding form

$$\begin{cases} \hat{E}\dot{\hat{x}}(t) = \hat{A}\hat{x}(t) + \hat{b}u(t) \\ \hat{y}(t) = \hat{c}^T \hat{x}(t) + du(t). \end{cases} \quad (2)$$

The dimension of the reduced-order model is designated as M . The output $\hat{y}(t)$ approximates the true output $y(t)$. However, in general, no simple relation exists between $\hat{x}(t)$ and the state vector $x(t)$. For instance, the tenth element of $\hat{x}(t)$ does not need to be directly related to the tenth element of $x(t)$.

It is also possible to represent the linear dynamic systems via the *transfer function* of the system that results from a Laplace transform. The transfer function of the original system is $\mathbf{h}(s) = \mathbf{c}^T(s\mathbf{E} - \mathbf{A})^{-1}\mathbf{b}$, where s represents complex frequency. Without a loss of generality, the feed-through term d of the original model can be assumed to be zero (see later). Laplace transform of the input $\mathbf{u}(s)$ to the Laplace transform of the output $\mathbf{y}(s)$. The transfer function of the reduced-order model can be defined analogously : $\hat{\mathbf{h}}(s) = \hat{\mathbf{c}}^T(s\hat{\mathbf{E}} - \hat{\mathbf{A}})^{-1}\hat{\mathbf{b}}$.

Several measures of the accuracy of the reduced-order model are possible. Formally, there tends to be an interest in the difference between the actual and low-order outputs, $y(t) - \hat{y}(t)$, given some set of inputs $u(t)$. This difference can be characterized via a system norm. The popular \mathcal{H}_∞ error norm, for example, measures, in the time domain, the worst ratio of output error energy to input energy. Equivalently, but in the frequency domain, it represents the largest magnitude of the frequency-response error.

A second measure of the accuracy of the approximation is to assess which properties of the original model are retained in the reduced-order one. Those properties of interest are said to be invariant, that is, they are independent with respect to a similarity transform. By retaining certain original properties of the system in the reduced-order model, one hopes that the resulting approximation error is small. Of course, this error depends on the selection and pertinence of the retained invariant properties.

A common choice for these invariant properties are the so-called modal properties of the system. [1, 2, 3] The modal properties are based on the system's poles (eigenvalues) λ_n and residues ρ_n , which both arise in a partial fraction expansion of the frequency response, Hence, a reduced-order model that matches (or approximately matches) specific modal components of the original model retains certain time-dependent features of the original system in its response. Potentially, iterative eigenvalue techniques can be used to find these specific components so that this modal retention approach is feasible for large-scale problems. Unfortunately, it can be difficult to identify a priori which modes are the truly dominant modal components of the original system. [4]

Alternative invariant properties that may be retained in model reduction are the Hankel singular values. Hankel singular values are related to the controllability and observability properties of a system. [5] Constructing a reduced-order model to retain the largest Hankel singular values is known as balanced truncation.

The invariant properties of interest in this survey are the coefficients of some power series expansion of $\mathbf{h}(s)$. The solution techniques proposed determine a reduced-order model that accurately matches the leading coefficients μ_j arising in a chosen power series. In general, one can produce a reduced-order

model that interpolates the frequency response and its derivatives (the moments of the series) at multiple points. These K possible interpolation points $\{\sigma^{(1)}, \sigma^{(2)}, \dots, \sigma^{(K)}\}$ are differentiated by their superscripts. The first $2J_1$ moments are matched at $\sigma^{(1)}$, the next $2J_2$ moments are matched at $\sigma^{(2)}$, etc., where $J_1 + J_2 + \dots + J_K = M$. A model meeting these constraints is denoted a multipoint Padé approximation or a rational interpolant. [6, 7] By varying the location and number of interpolation points utilized with the underlying problem in mind, one can construct accurate reduced-order models in a variety of situations.

2 History

There are in fact several different avenues of implementation for moment matching-based model reduction algorithms. One is the explicit moment-matching approach that has been used extensively in circuit simulation. For large sparse systems, however, approaches based on projections and preconditioning result in a set of algorithms with superior numerical stability and the opportunity for iterative implementations in terms of basic efficient computational kernels for sparse matrices.

The methods forming the foundation for this work are relatively old. The history of Padé approximation, for example, spans more than one hundred years [8]; and the Lanczos algorithm, an important Krylov space projection-based iteration, is nearing its fiftieth anniversary. [9] Yet, as the tremendous amount of recent literature demonstrates, the understanding and application of these concepts is certainly not a closed topic.

A large number of the moment-matching methods, particularly the early ones, form a reduced-order model from an explicit knowledge of the desired moments of the original system (see, e.g., [10]). It is typically a two-step process. First, $2M$ selected moments μ_j of the original system are explicitly computed. In the second step, the reduced-order frequency response $\hat{\mathbf{h}}(s) = (\hat{\phi}_{M-1}s^{M-1} + \dots + \hat{\phi}_0)/(s^M + \hat{\psi}_{M-1}s^{M-1} + \dots + \hat{\psi}_0)$ is forced to correspond to the selected moments. That is, the numerator parameters $\hat{\phi}$ and denominator parameters $\hat{\psi}$ are chosen so that the moments of the reduced-order system $\hat{\mu}_j$ equal those of the original system μ_j for $j = 1, 2, \dots, 2M$. This parameter selection requires the solution of a linear systems of equations involving Hankel matrices for all of the standard approaches such as partial realization, Padé and shifted Padé. For the rational interpolation problem, equations involving the more general Loewner matrix must be solved. [6] In all cases, it is important to note that the system matrices and vectors only enter the modeling problem through the moments. For large dynamic sys-

tems, A and E contribute to the calculation of the moments through sparse matrix-vector multiplies and sparse linear system solves.

Explicit methods were utilized to construct Padé approximants in the area of control in the early 1970s. [11] Extensions of these techniques to multiple interpolation points followed. [12, 13, 14] Of more recent interest, circa 1990, is a class of explicit moment-matching methods known as asymptotic waveform evaluation (AWE). [15, 16] Although the AWE methods themselves vary little in basic concept from the earlier control implementations, the AWE techniques are applied to interconnect model reduction in the area of circuits. The methods received attention for their ability to reduce RC interconnect models involving tens of thousands of variables. A multipoint version of AWE, complex frequency hopping (CFH), is available as well. [16]

Unfortunately, all of these explicit moment-matching methods are known to exhibit numerical instabilities, particularly as the dimension of the reduced-order model M grows. The source of these difficulties was pointed out in [17] and in the independent work of [18]. The reader is referred to those papers and to [19] for a detailed discussion. The difficulties center around numerical problems when constructing the Hankel matrices involved and, even when accurate matrices are available, the ill-conditioning of the associated linear systems. Both [17] and [18] point out that moment-matching via the Lanczos method (and more generally (bi)orthogonalized Krylov-based projection) is a preferred numerical implementation.

In projection methods, the M -th order reduced system is produced by applying two rectangular matrices Z and V to the matrices and vectors specifying the n -th order original system: $\hat{A} = Z^T A V$, $\hat{E} = Z^T E V$, $\hat{b} = Z^T b$, $\hat{c} = V^T c$, $\hat{d} = d$. The last equality explains also why d and \hat{d} can be chosen equal to zero without affecting the error between both models. The matrices are often taken to define projections onto Krylov spaces, $\mathcal{K}_j(G, g) = \text{span}\{g, Gg, G^2g, \dots, G^{j-1}g\}$ for specific choices of G and g . The first significant mathematical connection between the Lanczos algorithm, a Krylov-based technique, and model reduction occurred in the early 1980s. It was shown that partial realizations could be generated through the Lanczos algorithm. [20] Adaptations of Krylov subspaces were proposed in 1987 to generate Padé approximations and shifted Padé approximations. [21] Beyond the mathematical connections, the Lanczos method was utilized for model reduction in many application areas. The first of these areas chronologically was apparently structural dynamics. Even prior to the knowledge of the moment-matching connections, the Lanczos method was utilized in structural dynamics for model reduction based on eigenvalue analysis. [21, 22, 23] Later work in the field utilized the Lanczos method for Padé approximation including

MIMO systems. [24, 25, 26] The next wave of application work took place in the control literature . [27, 28, 29] A large amount of existing work was repeated, although new results did appear in the areas of error analysis and stability retention. [30, 31] Very recently, Lanczos-based model reduction has become a popular topic in the area of high-speed circuits. Existing Lanczos algorithms were applied to the standard [17, 18], MIMO [32] and symmetric problems. [33] New algorithms were proposed for stability retention. [34, 35] However, through all of these application areas, the approaches remained closely tied to the classical Lanczos algorithm. These approaches did not emphasize or exploit the fundamental structure in projection techniques for rational interpolation. A true multipoint rational Lanczos procedure was derived in [36] using a similar motivation, i.e., starting from the Lanczos procedure and modifying it to produce a reduced system that matched multiple moments at multiple frequency values.

3 General Projection Formulation

A large amount of the credit for connecting Krylov projection with Padè approximation belongs to Villemagne and Skelton. [37] However, recently in [19] and [38] the relationships between Krylov subspaces, the iterative algorithms for constructing these subspaces, and model reduction via rational interpolation has been explored in great detail. The results clearly demonstrate that the Lanczos-type methods are certainly not the only choice for iteratively constructing the rational interpolant. The family of Rational Krylov methods for model reduction contains many members and subsume previous approaches in the literature. The choice for a particular implementation is most often based on tradeoff between the desired level of numerical accuracy and the available computing resources. The remainder of this paper summarizes some of the important aspects of these recent developments in projection-based model reduction.

The basic relationship between the Krylov spaces and moment matching is summarized in the following theorem proven in [19].

Theorem 3.1 *If*

$$\bigcup_{k=1}^K \mathcal{K}_{J_{b_k}} \left((A - \sigma^{(k)} E)^{-1} (A - \zeta_k E), (A - \sigma^{(k)} E)^{-1} b \right) \subseteq \text{colsp} \{V\} \quad (3)$$

and

$$\bigcup_{k=1}^K \mathcal{K}_{J_{c_k}} \left((A - \sigma^{(k)} E)^{-T} (A - \zeta_k E)^T, (A - \sigma^{(k)} E)^{-T} c \right) \subseteq \text{colsp} \{Z\} \quad (4)$$

where ζ_k are arbitrary complex scalars, then the moments of (1) and (2) satisfy

$$\mu_k^{(j_k)} = -c^T \left\{ (A - \sigma^{(k)} E)^{-1} E \right\}^{j_k-1} (A - \sigma^{(k)} E)^{-1} b \quad (5)$$

$$= -\hat{c}^T \left\{ (\hat{A} - \sigma^{(k)} \hat{E})^{-1} \hat{E} \right\}^{j_k-1} (\hat{A} - \sigma^{(k)} \hat{E})^{-1} \hat{b} \quad (6)$$

$$= \hat{\mu}_k^{(j_k)} \quad (7)$$

for $j_k = 1, 2, \dots, J_{b_k} + J_{c_k}$ and $k = 1, 2, \dots, K$. \square

Assuming the non-singularity of the pencils $(\hat{A} - \sigma^{(k)} \hat{E})$ at these points, Theorem 3.1 guarantees that the desired rational interpolant is acquired. Any pair of projection bases satisfying (3) and (4) is sufficient to achieve the desired rational interpolant. Restrictions on V or Z , such as biorthogonality or orthogonality, are purely implementation specific choices. Various methods can be developed that create bases for the desired spaces. The cases where singularities are encountered in certain pencils can be handled by techniques described in [19].

The Rational Krylov Family of model reduction methods can be summarized using the RK algorithm presented in Table 1. Setting the parameters and selecting a particular point selection mechanism yields various members of the family. We point out here that $V = V_M = [v_1, \dots, v_M]$ and $Z = Z_M = [z_1, \dots, z_M]$ are constructed column by column by this algorithm. The parameters $\gamma_m^q, \beta_m^w, \gamma_m^v, \beta_m^z$ and vectors $\bar{v}_m, \bar{z}_m, \bar{q}_m, \bar{w}_m$, determine the different versions of the algorithm and are explained below.

Table 1: Rational Krylov Algorithm (General Version)

Initialize: $q_1 = (\gamma_1^q)^{-1} b$ and $w_1 = (\beta_1^w)^{-1} c$;
For $m = 1$ to M ,
(S1.1) Input: σ_m , the interpolation point for m^{th} iteration;
(S1.2) $\tilde{v}_m = (A - \sigma_m E)^{-1} q_{p_{m+1}}$ and $\tilde{z}_m = (A - \sigma_m E)^{-T} w_{p_{m+1}}$;
(S1.3) $\gamma_m^v v_m = \tilde{v}_m - V_{m-1} \bar{v}_m$ and $\beta_m^z z_m = \tilde{z}_m - Z_{m-1} \bar{z}_m$;
(S1.4) $\tilde{q}_{m+1} = E v_m$ and $\tilde{w}_{m+1} = E^T z_m$;
(S1.5) $\gamma_{m+1}^q q_{m+1} = \tilde{q}_{m+1} - Q_m \bar{q}_{m+1}$ and $\beta_{m+1}^w w_{m+1} = \tilde{w}_{m+1} - W_m \bar{w}_{m+1}$;
end

Three simplifying assumptions are made in going from Theorem 3.1 to the RK algorithm. First, the column spaces of V and Z are constructed to equal $(=)$ rather than contain (\supseteq) the union of Krylov subspaces on the left sides of (3) and (4). This assumption does not prevent rational interpolation; it is the

Krylov subspaces which contain the desired moment information. Second, it is assumed that the dimensions of the dual Krylov subspaces are consistent, i.e., $J_{b_k} = J_{c_k} = J_k$ for all k . These choices allow the matching of the maximum number of possible moments for a given model size M . Third, due to the shift invariant properties of Krylov spaces $A - \zeta_k E$ has been replaced by E , i.e., $\zeta_k = \infty$. The RK algorithm also uses the vectors q_m and w_m in addition to v_m and z_m . The associated matrices Q and W are related in direct fashions with V and Z via (S1.4). By initially incorporating all four matrices (rather than only V and Z) into the RK algorithm, more options become apparent.

Steps (S1.2) through (S1.5) of the RK algorithm generate the new columns of the projection matrices. Steps (S1.2) and (S1.4) introduces new information into the column spaces of V and Z , and Q and W respectively. The actual bases used to represent these columns spaces are determined in (S1.3) and (S1.5). The updates in these two steps correspond to the classical Gram-Schmidt procedure.

The choices for the vectors \bar{q}_m , \bar{v}_m , \bar{w}_m and \bar{z}_m in these updates determine what type of biorthogonality or orthogonality is produced among Q , V , W and Z . Three options have been investigated in [19] and subsume most of the ideas in the literature. If no orthogonality conditions are placed on V and Z and Q and W are removed by setting $\bar{q}_m = \bar{w}_m = \bar{v}_m = \bar{z}_m = 0$, the Rational Power Krylov method results. This is essentially multiple shifted inverse power methods run simultaneously. The Dual Rational Arnoldi method results from requiring V and Z to be orthogonal, i.e., $\bar{v}_m = V_{m-1}^T \tilde{v}_m$, $\bar{z}_m = Z_{m-1}^T \tilde{z}_m$, and scaling to unit vectors along with $\bar{q}_m = \bar{w}_m = 0$. The Rational Lanczos method results from requiring V and W to be biorthogonal, i.e., $\bar{v}_m = W_{m-1}^T \tilde{v}_m$, $\bar{w}_m = V_{m-1}^T \tilde{w}_m$, and $w_m^T v_m = 1$. This also results in an essentially sparse reduced order pencil (tridiagonal, banded or nearly so depending on the point selection process). The scalar subscript p_m that appears in (S1.2) locates the most recent iteration prior to the m^{th} iteration that employed the same interpolation point. It is shown in [19] that this can, in fact, be altered to other positions to, somewhat unexpectedly, improve the robustness of the algorithm in the presence of numerical loss of (bi)orthogonality.

Experiments comparing these methods can be found in [19]. In general, Dual Rational Arnoldi is the most robust numerically and Rational Krylov Power is the least robust. The Rational Lanczos algorithm tends to make a reasonable compromise between complexity per iteration and stability. Problems such as loss of biorthogonality and breakdown that are widely discussed in other contexts in the literature are also problems here. However, the use of multiple points makes breakdown avoidable in most useful circumstances and convergence, while slowed, is not destroyed by loss of biorthogonality. In

fact, with care and tuning, the Rational Lanczos algorithm can often be made the most efficient approach despite the increased iterations due to the reduced complexity of each iteration. When parallelism is introduced the preference begins to move back in the direction of methods more akin to the Rational Krylov Power algorithm. Leading evidence seems to indicate that hybrid approaches with characteristics between the Rational Krylov Power and Rational Lanczos algorithms are the most promising.

4 Efficiency Issues

The description above does not specify certain key heuristics that must be developed in order to make the production of the reduced-order model efficient. These are: error estimation, interpolation point selection/placement, and efficient approximate sparse system solving.

The first and most important is the characterization of the error $\epsilon(s) = \hat{\mathbf{h}}(s) - \mathbf{h}(s)$. This is necessary to control the termination of the process so as to keep the reduced-order model degree as low as possible. A simple approach for estimating the frequency-response error, $\epsilon(s) = \hat{\mathbf{h}}(s) - \mathbf{h}(s)$, between the original and reduced-order models is to compute the difference between two reduced-order models, $\hat{\epsilon}(s) = \hat{\mathbf{h}}(s) - \hat{\mathbf{h}}_{err}(s)$. The transfer function of $\hat{\mathbf{h}}_{err}(s)$ corresponds to some second and completely different low-order approximation of the original system. Both of these approximations can be generated by any (and not necessarily the same) Krylov-based projection algorithm. Hence, $\hat{\epsilon}(s)$ is a suitable and achievable error estimate for any of the previously discussed modeling techniques. The two low-order approximations used in $\epsilon(s)$ should contrast in their approximations of the original system because this difference estimates the modeling error. That is, two points of view of the original system are sought, which are designed to be complementary. The use of drastically different viewpoints typically suggests that $\hat{\mathbf{h}}(s)$ and $\hat{\mathbf{h}}_{err}(s)$ agree consistently only at those frequencies where both approximations are accurate.

The generation of two distinct reduced-order models requires the construction of two different projection pairs of dimension M , the previously seen V, Z and the second pair V_{err}, Z_{err} . The flexibility in forming these two pairs of projection matrices resides in the choice of interpolation points. In [19] the use of two sets of interleaved interpolation points is investigated. In addition to the cost of producing a second model, various possibilities and associated costs exist for evaluating $\hat{\epsilon}(s)$. A practical alternative approach to computing $\hat{\epsilon}$ is to simply evaluate it at multiple carefully selected points. In practice, approximately $8m$ well-placed points in the fashion of the frequency-response algorithms of [39] give reasonable results for our purposes. The cost can be

further reduced by performing some dense state space transformations on the (assumed small) reduced-order models before evaluation.

The error estimate can often be improved through a simple modification of that above. Recall that we are ultimately interested in the acceptable, low-order model represented by $\hat{\mathbf{h}}(s)$. The second low-order function $\hat{\mathbf{h}}_{err}(s)$ serves only to estimate the true frequency response $\mathbf{h}(s)$. A better approximation for $\mathbf{h}(s)$ can be found by combining the information in the two sets of projection matrices V, Z and V_{err}, Z_{err} to obtain a $2M^{th}$ order model, $\hat{\mathbf{h}}_{\cup}(s)$. Unlike $\hat{\mathbf{h}}_{err}(s)$, $\hat{\mathbf{h}}_{\cup}(s)$ includes the original V and Z directions and, thus, tends to at least converge wherever $\hat{\mathbf{h}}$ does. This does however increase the cost of each iteration and care should be taken to insure that the improvement warrants the extra expense.

If $\hat{\mathbf{x}}_b$ and $\hat{\mathbf{x}}_c$ are defined to be the solutions of the dual reduced-order, shifted systems of equations

$$\begin{aligned} (s\hat{E} - \hat{A})\hat{\mathbf{x}}_b &= \hat{b} \\ (s\hat{E} - \hat{A})^T\hat{\mathbf{x}}_c &= \hat{c}, \end{aligned} \tag{8}$$

the transfer function of the reduced system can be written $\hat{\mathbf{h}}(s) = \hat{\mathbf{x}}_c^T Z^T (sE - A)V\hat{\mathbf{x}}_b$. By rewriting the original transfer function in a similar way one sees that the two differ only in that the reduced model approximates $\mathbf{x}_b = (sE - A)^{-1}b$ and $\mathbf{x}_c = (sE - A)^{-T}c$ with $V\hat{\mathbf{x}}_b$ and $Z\hat{\mathbf{x}}_c$. These approximations $V\hat{\mathbf{x}}_b$ and $Z\hat{\mathbf{x}}_c$ satisfy the Petrov-Galerkin conditions for any chosen V and Z .

In fact, in [19] it is shown that the difference between the frequency responses of the original and reduced-order systems is $\mathbf{r}_c^T(A - sE)^{-1}\mathbf{r}_b$, where

$$\begin{aligned} \mathbf{r}_b(s) &= b - (sE - A)V\hat{\mathbf{x}}_b \\ \mathbf{r}_c(s) &= c - (sE - A)^T Z\hat{\mathbf{x}}_c. \end{aligned}$$

Note that monitoring \mathbf{r}_b and/or \mathbf{r}_c does not directly lead to an estimate for the modeling error. Acquiring the modeling error requires an inverse of $(sE - A)$ which is not possessed. Rather, one must concentrate on the trends in the residual behavior as s and m vary. Attempting to gauge these trends demands the evaluation of \mathbf{r}_b and/or \mathbf{r}_c at numerous values of s . Fortuitously, the residual expressions of many of the RK family implementations can be simplified by exploiting the structure of the matrices involved. These simplifications and experiments comparing the two approaches can be found in [19]. The two approaches are often complementary and a robust code would probably need to use a combination.

The error estimation cost can also be amortized by addressing the second important heuristic needed for efficient model reduction: point selection

and placement. At each step of the algorithm it is necessary to choose the interpolation point at which the next moment will be added. This choice has two purposes: avoiding unnecessary work and improving the accuracy of the reduced model. Forcing a certain number of moments at all of the points $\{\sigma^{(1)}, \sigma^{(2)}, \dots, \sigma^{(K)}\}$ typically yields a model that can be reduced further and that required extra work, i.e., unnecessary linear system solves due to the fact that the grid of points was made too fine. In addition to the selection of points from a predetermined grid, there is also the issue of point placement for extending the grid or refining it when it is found to be too coarse. Accompanying this is the need to determine where the points should be located in the complex plane. Intuitively, using purely imaginary $\sigma^{(i)}$ makes sense since the model is approximating the frequency response of the system. Imaginary points only give local information and therefore the initial grid on the imaginary axis must be placed so as to include at least one point in all areas of significant activity. Unfortunately, these areas are not always known a priori. For dynamically stable systems, however, taking positive real $\sigma^{(i)}$ can also yield useful information (and reduce the need for complex arithmetic). The reduced order model generated using real points tends to track large scale behavior of the frequency response and therefore can be used to identify regions of activity that need to be probed using imaginary $\sigma^{(i)}$. A strategy that combines real and imaginary point along with the error estimation strategies above is discussed in [19].

Common to all RK methods that match moments is the solution of linear systems with coefficient matrix $\Phi_m = (A - \sigma^{(k_m)}E)^{-1}$ where $\sigma^{(k_m)} \in \{\sigma^{(1)}, \dots, \sigma^{(K)}\}$, i.e., it is the interpolation point chosen from the set of K possible points for use on the m -th iteration. When direct sparse methods are used Φ_m must be factored only when the $\sigma^{(i)}$ in step m is being used for the first time. Subsequent solves with the same shift are simply applications of forward and backward substitution. However, for many large scale problems direct methods are not cost effective and an approximate solution must be found. Forms of the RK algorithms that use approximate solutions can be developed based on a general form similar to that in Table 1 but with the matrix $(A - \zeta_k E)$ and its transpose from Theorem 3.1 kept in the bases definitions.

The approximate form of the algorithm allows the choice of Φ_m and ζ_m . The operator Φ_m approximates the action of $(A - \sigma^{(k_m)}E)^{-1}$ on a vector. Numerous possibilities exist for finding a preconditioner that approximates $(A - \sigma^{(k_m)}E)^{-1}$. Alternatively, and more generally, one can think of Φ_m as an operation that takes in the vectors q_{p_m+1} , w_{p_m+1} and outputs the vectors \tilde{v}_m , \tilde{z}_m . Hence, Φ_m can represent an iterative system solver that computes approximate solutions to the equations. However, in this case, Φ_m is no longer

associated with a fixed matrix P_k that is associated with an interpolation point $\sigma^{(k)}$. This is due to the fact that a single interpolation point $\sigma^{(k)}$ may be used in more than one step of the iteration with different right-hand side vectors and therefore there would be more than one approximation to $(A - \sigma^{(k)}E)^{-1}$. If the methods are always iterated to near working precision then this can be treated as roundoff and covered by the appropriate error bounds. If, however, to reduce the time required the iterations are stopped early then the theory must be generalized. As can be seen from the experiments in [40] and [19], in practice it is not an important distinction.

If (and only if) Φ_m is chosen so that moment matching results then the choice of ζ_m in the RK algorithm does not contribute to the specification of the V and Z column spaces and ζ_k can be replaced in Theorem 3.1 by a variable s and removed from the algorithm, i.e., $\zeta_m = \infty$. When $P_m \approx (A - \sigma^{(k_m)}E)^{-1}$ it is possible to tune ζ_m to improve the results. For example, $\zeta_m = \infty$ can still be used, or more commonly, the setting $\zeta_m = \sigma^{(k_m)}$ is used. This can be motivated by subspace and/or eigenvalue mapping considerations. [19] The choice of ζ_m can lead to significant, but often unpredictable differences in the convergence of the reduced-order model. In [19] it is seen that the $\zeta_m = \sigma^{(k_m)}$ case performs well when the model reduction preconditioner, Φ_m is poorer (fewer iterative method steps used), but is unacceptable when the Φ_m is more accurate. This is similar to behavior observed in the eigenvalue literature. The opposite behavior occurs when ζ_m is ∞ , i.e., $(A - \zeta_m E)$ is replaced with E . In practice, ζ_m can be tuned between ∞ and $\sigma^{(k_m)}$ by using available information on the preconditioner quality and/or the convergence behavior of previous solves. Various choices for ζ_m appear in related approaches to problems in the eigenvalue literature. [41, 42] Alternatively, more sophisticated approaches for implementing the approximate solvers can be used to reduce significantly the importance of ζ_m . [19] This was done by noting the fact that at step m a reduced order model of size $m - 1$ and its projectors V_{m-1} and Z_{m-1} are available. The projectors can be used to define an initial guess at the solution of the system that ultimately provides a vector that extends the column spaces to V_m and Z_m when appropriate. As a side note, it is claimed that the derivation above provides an alternative path for obtaining Davidson's method for the eigenvalue problem.

When $P_m \approx (A - \sigma^{(k_m)}E)^{-1}$, the V and Z generated by any version of the approximate RK algorithm no longer form bases for the unions of Krylov subspaces required for rational interpolation. However, the residuals of the reduced order model still satisfies the Petrov-Galerkin conditions, $Z^T \mathbf{r}_b(s) = 0$ and $W^T \mathbf{r}_c(s) = 0$ for all values of s . As long as reasonable approximations $V \hat{\mathbf{x}}_b$ and $Z \hat{\mathbf{x}}_c$ to \mathbf{x}_b and \mathbf{x}_c are acquired, a good reduced-order model is achievable.

It is interesting to note that even the exactly preconditioned V and Z do not necessarily lead to optimal approximations to \mathbf{x}_b and \mathbf{x}_c at all s .

In addition to loss of moment-matching the relationships between V , Z , W , and Q are also affected. The Rational Lanczos is the most dependent on these relationships holding in order to guarantee the sparsity of the reduced-order model. There are two ways to respond to this problem. The first is to ignore it and simply use the reduced-order model that is implicitly produced by the RK iteration. The second is to ignore the implicitly produced model and retain the projectors V and Z and apply them explicitly at the end of the iteration to produce the reduced order model. The second does have distinct accuracy advantages but destroys the efficiency of the Rational Lanczos iteration. In [19] and [43] it is shown that the approximate version of the Rational Lanczos that uses the implicit approach can still produce acceptable reduced order models if the systems are solved on each iteration to the level of the accuracy desired in the frequency response, i.e., errors introduced by approximations are not damped by later iterations of the Rational Lanczos algorithm. The implicit approach allows looser tolerance on solving the systems but at the cost of a larger reduced model order.

5 Generalizations

Generalizations of the basic RK algorithms and theory to the MIMO case have been derived. [19] Work remains, however, on the efficient implementation of such methods. For some applications, the linear systems are dependent on two parameters, e.g., frequency and azimuth. A transfer function can be derived and used to create a reduced order model. Recent work has developed analogs of Theorem 3.1 for such problems. [44] Finally, some applications yield transfer functions that involve linear systems whose matrix elements are not linear functions of frequency as assumed above. Addressing these problems often requires exploitation of structure arising out of the particular application. Such an approach has recently been investigated for use in analyzing frequency selective surfaces. [45]

Acknowledgments

This work was supported by a Department of Energy Fellowship, IBM Corporation, and the National Science Foundation under grants CCR-9120105 and CCR-9796315. Parts of this paper presents research results of the Belgian Programme on Inter-university Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture. The scientific responsibility rests with its authors.

References

1. L. A. Aguirre. Quantitative measure of modal dominance for continuous systems. In *32nd IEEE Conf. Decision Contr.*, San Antonio, TX, 1993.
2. D. Bonvin and D. A. Mellichamp. A unified derivation and critical review of modal approaches to model reduction. *Int. J. Contr.*, 35:829–848, 1982.
3. R. E. Skelton. *Dynamic System Control*. John Wiley & Sons, New York, NY, 1988.
4. Y. Shamash. Viability of methods for computing stable reduced-order models. *IEEE Trans. Automat. Contr.*, 26:1285–1286, 1981.
5. B. C. Moore. Principal component analysis in linear systems: controllability, observability and model reduction. *IEEE Trans. Automat. Contr.*, 26:17–32, 1981.
6. B. D. O. Anderson and A. C. Antoulas. Rational interpolation and state space variable realizations. *Lin. Algeb. Its Appl.*, 138:479–509, 1990.
7. G. A. Baker Jr. *Essentials of Padé approximation*. Academic Press, New York, NY, 1975.
8. C. Brezinski. *Padé-Type Approximation and General Orthogonal Polynomials*. Birkhauser, Basel, 1980.
9. C. Lanczos. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *J. Res. Nat. Bureau Stan.*, 45:255–282, 1950.
10. A. Bultheel and M. Van Barel. Padé techniques for model reduction in linear system theory: a survey. *J. Comp. Appl. Math.*, 14:401–438, 1986.
11. Y. Shamash. Linear system reduction using Padé approximation to allow retention of dominant modes. *Int. J. Contr.*, 21:257–272, 1975.
12. C. Hwang and M. Y. Chen. A multi-point continued fraction expansion for linear system reduction. *IEEE Trans. Automat. Contr.*, 31:648–651, 1986.
13. C. Kenney, A. J. Laub, and S. Stubberud. Frequency response computation via rational interpolation. *IEEE Trans. Automat. Contr.*, 38:1203–1213, 1993.
14. H. Xiheng. FF-Padé method of model reduction in frequency domain. *IEEE Trans. Automat. Contr.*, 32:243–246, 1987.
15. L. T. Pillage and R. A. Rohrer. Asymptotic waveform evaluation for timing analysis. *IEEE Trans. Computer-Aided Design*, 9:352–366, 1990.
16. E. Chiprout and M. S. Nakhla. *Asymptotic Waveform Evaluation and Moment Matching for Interconnect Analysis*. Kluwer Academic Publish-

- ers, Boston, MA, 1994.
17. K. Gallivan, E. Grimme, and P. Van Dooren. Asymptotic waveform evaluation via a Lanczos method. *Appl. Math. Lett.*, 7:75–80, 1994.
 18. 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.
 19. E. Grimme. *Krylov Projection Methods for Model Reduction*. PhD thesis, University of Illinois at Urbana-Champaign, 1997.
 20. W. B. Gragg and A. Lindquist. On the partial realization problem. *Lin. Algeb. Its Appl.*, 50:277–319, 1983.
 21. B. Nour-Omid and R. W. Clough. Dynamic analysis of structures using Lanczos coordinates. *Earthquake Eng. and Struc. Dyn.*, 12:565–577, 1984.
 22. I. U. Ojalvo and M. Newman. Vibration modes of large structures by an automatic matrix-reduction method. *AIAA J.*, 8:1234–1239, 1970.
 23. E. L. Wilson, M. W. Yuan, and J. M. Dickens. Dynamic analysis by superposition of Ritz vectors. *Earthquake Eng. and Struc. Dyn.*, 10:813–821, 1982.
 24. H. M. Kim and R. R. Craig Jr. Structural dynamics analysis using an unsymmetric block Lanczos algorithm. *Int. J. Numer. Methods Eng.*, 26:2305–2318, 1988.
 25. H. M. Kim and R. R. Craig Jr. Computational enhancement of an unsymmetric block Lanczos algorithm. *Int. J. Numer. Methods Eng.*, 30:1083–1089, 1990.
 26. T. J. Su and R. R. Craig Jr. Krylov vector methods for model reduction and control of flexible structures. *Advances Contr. Dynamic Sys.*, 54:449–481, 1992.
 27. M. M. M. Al-Husari, B. Hendel, I. M. Jaimoukha, E. M. Kasenally, D. J. N. Limebeer, and A. Portone. Vertical stabilisation of tokamak plasmas. In *30th IEEE Conf. Decision Contr.*, Brighton, UK, 1991.
 28. D. L. Boley. Krylov space methods on state-space control models. *Circ. Sys. Signal Proc.*, 13:733–758, 1994.
 29. P. M. Van Dooren. Numerical linear algebra techniques for large scale matrix problems in systems and control. In *31st IEEE Conf. Decision Contr.*, Tucson, AZ, 1992.
 30. E. J. Grimme, D. C. Sorensen, and P. M. Van Dooren. Model reduction of state space systems via an implicitly restarted Lanczos method. *Numer. Algorithms*, 12:1–31, 1996.
 31. I. M. Jaimoukha and E. M. Kasenally. Oblique projection methods for large scale model reduction. *SIAM J. Matrix Anal. Appl.*, 16:602–627,

- 1995.
32. P. Feldman and R. W. Freund. Reduced-order modeling of large linear subcircuits via a block Lanczos algorithm. In *32nd ACM/IEEE Design Automation Conf.*, San Fransisco, CA, 1995.
 33. P. Feldman and R. W. Freund. Reduced-order modeling of large passive linear circuits by means of the SyPVL algorithm. In *IEEE/ACM Int. Conf. Computer-Aided Design*, San Jose, CA, 1996.
 34. K. J. Kerns, I. L. Wemple, and A. T. Yang. Stable and efficient reduction of substrate model networks using congruence transforms. In *IEEE/ACM Int. Conf. Computer-Aided Design*, San Jose, CA, 1995.
 35. L. M. Silveira, M. Kamon, I. Elfadel, and J. White. A coordinate-transformed Arnoldi algorithm for generating guaranteed stable reduced-order models of RLC circuits. In *IEEE/ACM Int. Conf. Computer-Aided Design*, San Jose, CA, 1996.
 36. K. Gallivan, E. Grimme, and P. Van Dooren. A rational Lanczos method for model reduction. *Numer. Algorithms*, 12:33–63, 1996.
 37. C. D. Villemagne and R. E. Skelton. Model reduction using a projection formulation. *Int. J. Contr.*, 46:2141–2169, 1987.
 38. E. Grimme, K. Gallivan, and P. Van Dooren. Krylov projection methods for rational interpolation. submitted for publication.
 39. A. Grace, A. J. Laub, J. N. Little, and C. Thompson. *Control System Toolbox*. The MathWorks, Natick, MA, 1990.
 40. V. de Clippel. Modèles réduits de grands systèmes dynamiques. Technical report, Université Catholique de Louvain, Belgium, 1997.
 41. G. L. G. Sleijpen and H. van der Vorst. A Jacobi-Davidson iteration method for linear eigenvalue problems. *SIAM J. Matrix Anal. Appl.*, 17:401–425, 1996.
 42. A. Stathopoulos, Y. Saad, and C. F. Fischer. Robust preconditioning of large sparse symmetric eigenvalue problems. *J. Comp. Appl. Math.*, 64:197–216, 1995.
 43. E. Grimme, K. A. Gallivan, Valerie de Clippel, and Paul Van Dooren. Approximate solves in krylov-based modeling methods. In *Proceedings of 1997 IEEE Conference on Decision and Control*, 1997.
 44. D. S. Weile, E. Michielssen, E. Grimme, and K. Gallivan. A method for generating rational interpolant reduced order models of two parameter systems. submitted for publication.
 45. D. S. Weile, E. Michielssen, and K. Gallivan. Rational krylov reduced order modeling of multiscreen frequency selective surfaces. In *Proceedings of 14th Annual Review of Progress in Applied Computational Electromagnetics*, 1998.