Pergamon

Appl. Math. Lett. Vol. 7, No. 5, pp. 75–80, 1994

0893-9659/94 $7.00 + 0.00

0893-9659(94)00075-1

Asymptotic Waveform Evaluation via a Lanczos Method

K. GALLIVAN, E. GRIMME AND P. VAN DOOREN
Coordinated Science Laboratory, University of Illinois at Urbana-Champaign
307 N. Mathews, Urbana, IL 61801, U.S.A.

(Received January 1994; accepted February 1994)

Abstract—In this paper we show that the two-sided Lanczos procedure combined with implicit restarts, offers significant advantages over Padé approximations used typically for model reduction in circuit simulation.

Keywords—Dynamical systems, Model reduction, Numerical methods, Padé approximation, Lanczos algorithm.

1. INTRODUCTION

Asymptotic Waveform Evaluation (AWE) is a recently developed technique for acquiring low-order approximations for large, linearized circuits [1–3]. Such a circuit is generally described by the set of state space equations

$$\dot{x} = Ax + bu \quad \text{and} \quad y = cx + du, \tag{1}$$

where A is a square, sparse matrix of size $n \gg 1$, b is a column vector, and c is a row vector. For simplicity, it will be assumed that the direct-coupling term d is zero. A corresponding, low-order approximation can then be defined by

$$\dot{\hat{x}} = \hat{A}\hat{x} + \hat{b}u \quad \text{and} \quad \hat{y} = \hat{c}x, \tag{2}$$

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 ($x(0) = 0$) solution to the first expression in (1) is $x(t) = \int_0^t e^{A(t-\tau)}bu(\tau)\,d\tau$. Thus, determining a good low-order approximation (2) 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 [4,5] for approximating $e^{At}b$. But in fact, these concepts can be taken one step further by noting 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 $y(t) = ce^{At}b$. Numerous papers [6–9] are beginning to explore this last fact in the context of control (but not circuit simulation). An oblique Krylov projector (i.e., the Lanczos algorithm) is employed to generate the reduced-order system $\{\hat{A}, \hat{b}, \hat{c}\}$.

Intertwined with the Krylov projections performed in all of these papers is the Padé approximation of the transfer function, $h(s) = c(sI - A)^{-1}b$. Model reduction via Padé approximation is well

This work was supported in part by the National Science Foundation under Grants CCR-9120105 and CCR-9209349.

documented in the literature [10], and Padé approximants actually form the backbone of AWE. But unfortunately, existing AWE papers do not link Krylov projectors and Padé approximation together.

In this paper, it will be demonstrated that a Padé approximation of the original circuit can be obtained without explicitly passing via the moments. Through the nonsymmetric Lanczos method [11,12], one can realize the reduced-order system $\{\hat{A}, \hat{b}, \hat{c}\}$ directly from the original circuit. Approximating 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 (Section 3).

2. 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, $m_i = cA^{-i-1}b$, $i \geq 0$, or about $s = \infty$ to yield the high-frequency moments (Markov parameters) $m_i = cA^{-i-1}b$, $i < 0$.

In the second step, the impulse response

$$\hat{h}(s) = \frac{n_{k-1}s^{k-1} + \cdots + n_1 s + n_0}{s_k + d_{k-1}s^{k-1} + \cdots + d_1 s + d_0} \tag{3}$$

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}. As a result, the impulse response completely defines a Padé approximation (partial realization) which matches the desired moments of the original system. Although it is not explicitly determined in existing AWE methods, note that state space equations (i.e., $\{\hat{A}, \hat{b}, \hat{c}\}$) can be obtained which correspond exactly to $\hat{h}(s)$.

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

$$\dot{\hat{x}} = \left(W_k^\top A V_k\right)\hat{x} + \left(W_k^\top b\right)u = \hat{A}\hat{x} + \hat{b}u \quad \text{and} \quad \hat{y} = (cV_k)\,\hat{x} = \hat{c}\hat{x} \tag{4}$$

for the original system (1). The matrices V_k and W_k are related to Krylov spaces, $\mathcal{K}_k$, in that

$$\text{COLSP}(V_k) = \mathcal{K}_k(A, b) = \text{span}\left\{b, Ab, \ldots, A^{k-1}b\right\} \tag{5}$$

$$\text{COLSP}(W_k) = \mathcal{K}_k(A^\top, c^\top) = \text{span}\left\{c^\top, A^\top c^\top, \ldots, A^{k-1^\top} c^\top\right\}. \tag{6}$$

The utility of selecting V_k and W_k in Krylov spaces comes from the fact that they can be generated with only inner-products and matrix-vector multiplications. By taking advantage of the sparsity of A, 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 from [13,14].

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 (5) and (6), respectively. Then the first $2k$ Markov parameters of the original and reduced-order systems are identical.*

Restating Theorem 1, the reduced-order model is a Padé approximation (partial realization) which matches the first $2k$ high-frequency moments of the original system.

Through a projector corresponding to $\mathcal{K}_k(A, b)$ and $\mathcal{K}_k(A^\top, c^\top)$, 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 $\mathcal{K}_k(A^{-\top}, c^\top)$ can be employed to generate a realization which matches moments about $s = 0$. And in fact, projectors can be constructed which are combinations of both approaches.

A popular technique for computing V_k and W_k in the Krylov projector is due to Lanczos [12]. Given the starting vectors v_1 and w_1, the Lanczos algorithm produces the rectangular matrices $V_k = [v_1, \ldots, v_k] \in \mathbb{R}^{n \times k}$ and $W_k = [w_1, \ldots, w_k] \in \mathbb{R}^{n \times k}$ which satisfy the recursive identities

$$AV_k = V_k T_k + \beta_{k+1}\, v_{k+1}\, e_k^\top \tag{7}$$

$$A^\top W_k = W_k T_k^\top + \gamma_{k+1}\, w_{k+1}\, e_k^\top. \tag{8}$$

The vector e_k is the k^{th} standard basis vector while T_k is a truncated reduction of A that is in tridiagonal form. Generally, the elements β_i and γ_i are chosen so that $V_{k+1}^\top W_{k+1} = I$. When V_{k+1} and W_{k+1} are biorthogonal, multiplying (7) on the left by $W_k^\top$ yields the relationship $W_k^\top AV_k = T_k$.

To choose the starting vectors, v_1 and w_1, it is important to note from (7) and (8) that $v_{k+1} \in \mathcal{K}_{k+1}(A, v_1)$ and $w_{k+1} \in \mathcal{K}_{k+1}(A^\top, w_1)$. Then if $v_1 = b/\beta_1$ and $w_1 = c^\top/\gamma_1$, the matrices V_k and W_k correspond to the Krylov spaces $\mathcal{K}_k(A, b)$ and $\mathcal{K}_k(A^\top, c^\top)$, respectively. And more importantly from (4), $\hat{A} = W_k^\top AV_k = T_k$, $\hat{b} = W^\top b = e_1\, \beta_1$ and $\hat{c} = cV_k = e_1^\top\, \gamma_1$, which is our reduced-order model!

3. ADVANTAGES OF LANCZOS-BASED MODEL REDUCTION

Compared to explicit moment matching, the Lanczos method provides superior results and/or greater flexibility in several areas: the sensitivity of the reduced-order realization, the scaling of the moments, the stability of the reduced-order approximation, and singularities in the Padé table. This section explores these issues in greater detail.

3.1. Sensitivity of the Realization

In past AWE papers [2], the reduced-order model is expressed via the partial fraction expansion (PFE) of (3)

$$\hat{h}(s) = \sum_{j=1}^{k} \frac{r_j}{s - p_j}, \tag{9}$$

where the poles p_j are the (assumed unique) roots of the denominator of (3) and the coefficients r_j are denoted as the residuals. Note that given (9), it is simple to obtain a state space representation in Jordan canonical form

$$\left[\begin{array}{c|c} \hat{A}_j & \hat{b}_j \\ \hline \hat{c}_j & d \end{array}\right] = \left[\begin{array}{ccc|c} p_1 & & & r_1 \\ & \ddots & & \vdots \\ & & p_k & r_k \\ \hline 1 & \cdots & 1 & 0 \end{array}\right]. \tag{10}$$

Unfortunately, given an arbitrary representation $\{\hat{A}, \hat{b}, \hat{c}\}$, the transformation required to obtain the Jordan canonical form is oftentimes poorly conditioned [15,16]. More specifically, computing the eigenvectors and eigenvalues of an arbitrary $\hat{A}$ (which amounts 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

$$\left[\begin{array}{c|c} \hat{A}_l & \hat{b}_l \\ \hline \hat{c}_l & d \end{array}\right] = \left[\begin{array}{cccc|c} \alpha_1 & \gamma_2 & & & \beta_1 \\ \beta_2 & \ddots & \ddots & & \\ & \ddots & \ddots & \gamma_k & \\ & & \beta_k & \alpha_k & \\ \hline \gamma_1 & & & & 0 \end{array}\right]. \tag{11}$$

The transformation to obtain this realization is known to be better conditioned in general [11,15].

As an example, consider the simple system defined by

$$\left[\begin{array}{c|c} A & b \\ \hline c & d \end{array}\right] = \left[\begin{array}{cccc|c} -3.01 & -3.03 & -1.03 & -(0.01+10^{-11}) & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \hline 1 & 4.5\cdot 10^{-2} & 6.75\cdot 10^{-4} & 3.3375\cdot 10^{-6} & 0 \end{array}\right]. \tag{12}$$

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 (12), Jordan, and tridiagonal realizations.

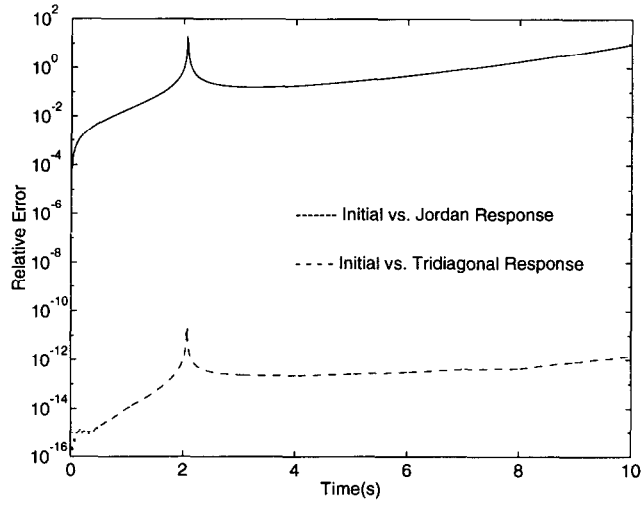

Figure 1. Relative error in the step responses of the Jordan and tridiagonal realizations vs. the step response of the original realization.

As one would expect, the response error (dashed line) between the initial and tridiagonal realizations, 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. 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, better conditioned realizations (such as a tridiagonal one) should be chosen over the Jordan form.

3.2. 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. Compared to cA^ib, the value of cb will be zero in finite precision, resulting in a situation where realizations of size $\geq i$ cannot be computed. To overcome this difficulty, [2] proposes scaling the matrix A. 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 $\mathcal{K}_k(A^\top, c^\top)$. As an example, consider the state space equations arising from a small, stiff RC ladder circuit

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

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 $\mathcal{K}_k(A^{-\top}, c^\top)$) 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

3.3. Stability of the Approximation

When moment matching is employed, the reduced-order model for a stable circuit may be unstable [17]. 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 sufficiently large 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, [18] 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 $\tilde{\pi}_k = \tilde{W}^\top \tilde{V}$ which corresponds to the new starting vectors

$$\tilde{v}_1 = (A - \mu_p I) \dots (A - \mu_1 I)\, v_1 \qquad \text{and} \qquad \tilde{w}_1 = \left(A^\top - \mu_p I\right) \dots \left(A^\top - \mu_1 I\right) w_1. \tag{13}$$

These implicit restarts (which correspond to LR-steps [16] with the tridiagonal matrix T_k) incorporate information from higher moments into the reduced-order model. Strategies for choosing the parameters μ_i in (13) can be employed to insure that this extra information stabilizes the partial realization. Examples in [18] demonstrate that when properly employed, implicit restarts can stabilize a realization with negligible computational effort.

3.4. Singularities in the Padé Table

Singularities can occur in the Padé table. Where these singularities exist, partial realizations (2) will not be possible. Additionally, poorly conditioned entries in the Padé table must 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, where it is termed a "serious" breakdown. By employing "look-ahead" into the Lanczos method [19], one possesses a powerful tool for detecting and avoiding ill-conditioned table entries.

4. CONCLUDING REMARKS

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

Lanczos methods are already being applied to model reduction problems in the areas of control [6,7,9] and dynamical structures [8]. Similar methods appear promising for partial circuit realizations in AWE.

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. L.T. Pillage and R.A. Rohrer, Asymptotic waveform evaluation for timing analysis, *IEEE Trans. Computer-Aided Design* **9**, 352–366 (1990).
3. 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).
4. E. Gallopoulos and Y. Saad, Efficient solution of parabolic equations by Krylov approximation methods, *SIAM J. Sci. Stat. Comput.* **13**, 1236–1264 (1992).
5. Y. Saad, Analysis of some Krylov subspace approximations to the matrix exponential operator, *SIAM J. Numer. Anal.* **29**, 209–228 (1992).
6. D.L. Boley, Krylov space methods on state-space control models, Technical Report, Univ. of Minnesota, Minneapolis, MN, (1992).
7. I.M. Jaimoukha and E.M. Kasenally, Oblique projection methods for large scale model reduction, *SIAM J. Matrix Anal. Appl.* (1994) (to appear).
8. 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).
9. 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).
10. Y. Shamash, Model reduction using the Routh stability criterion and the Padé approximation technique, *Int. J. Control* **21**, 475–484 (1975).
11. G.H. Golub and C. Van Loan, *Matrix Computations*, 2nd edition, Johns Hopkins University Press, Baltimore, MD, (1989).
12. C. Lanczos, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, *J. Res. Nat. Bur. Standards* **45**, 255–282 (1950).
13. W.B. Gragg and A. Lindquist, On the partial realization problem, *Lincar Algeb. Its Appl.* **50**, 277–319 (1983).
14. C.D. Villemagne and R.E. Skelton, Model reduction using a projection formulation, *Int. J. Control* **46**, 2141–2169 (1987).
15. B. Kågström and P. Van Dooren, A generalized state-space approach for the additive decomposition of a transfer matrix, *J. Num. Linear Alg. with Applns.* **1**, 165–181 (1992).
16. J.H. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford, England, Clarendon Press, (1965).
17. C.I. Byrnes and A.Lindquist, The stability and instability of partial realizations, *Sys. Control Lett.* **2**, 99–105 (1982).
18. E.J. Grimme, D.C. Sorensen and P. Van Dooren, Stable partial realizations via an implicitly restarted Lanczos method, *Proc. American Control Conf.* (1994) (to appear).
19. R.W. Freund, M.H. Gutknecht and N.M. Nachtigal, An implementation of the look-ahead Lanczos algorithm for non-Hermitian matrices, Technical Reports 90.45 and 90.46, RIACS, NASA Ames Research Center, (1990).