On Padé-type model order reduction of J-Hermitian linear dynamical systems ⋆

Roland W. Freund

Department of Mathematics, University of California at Davis, One Shields Avenue, Davis, CA 95616, USA

Dedicated to Richard S. Varga

Abstract

A simple, yet powerful approach to model order reduction of large-scale linear dynamical systems is to employ projection onto block Krylov subspaces. The transfer functions of the resulting reduced-order models of such projection methods can be characterized as Padé-type approximants of the transfer function of the original large-scale system. If the original system exhibits certain symmetries, then the reduced-order models are considerably more accurate than the theory for general systems predicts. In this paper, the framework of J-Hermitian linear dynamical systems is used to establish a general result about this higher accuracy. In particular, it is shown that in the case of J-Hermitian linear dynamical systems, the reduced-order transfer functions match twice as many Taylor coefficients of the original transfer function as in the general case. An application to the SPRIM algorithm for order reduction of general RCL electrical networks is discussed.

Keywords: linear dynamical system, descriptor system, model order reduction, block Krylov subspace, projection, Padé-type approximation, J-Hermitian, sesquilinear form, SPRIM algorithm, RCL network

AMS classification: 65F30, 15A57, 15A63, 41A21, 65L80, 93A15

⋆ Research supported in part by NSF Grant DMS-0613032.
E-mail address: freund@math.ucdavis.edu
URL: http://www.math.ucdavis.edu/˜freund/

1. Introduction

Many fundamental methods in numerical analysis are closely related to Padé or Padé-type approximation [2]; see, e.g., [18,19,21,22,30]. Examples include implicit discretization schemes of parabolic differential equations, which can be viewed as certain Padé approximations to the exponential function [30], the 'Lanczos-Padé connection' [19,21,22] of the classical Lanczos process [23,24] for large-scale matrix computations, and Padé-based methods [3,7] for model order reduction of large-scale linear dynamical systems.

The Lanczos-Padé connection can be employed to devise efficient and numerically well behaved algorithms [4,17,5,6,20] for model order reduction of large-scale linear dynamical systems. For the case of single-input single-output systems, one such method is the Padé Via Lanczos (PVL) algorithm [4,5], and the extension to the case of general multi-input multi-output systems is the Matrix-Padé Via Lanczos (MPVL) algorithm [6]. These methods are optimal in the sense of Padé approximation of the underlying transfer functions. More precisely, the transfer function H of the original large-scale system is a $p \times m$-matrix-valued rational function of the form

$$H(s) = C^H (sE - A)^{-1} B, \quad s \in \mathbb{C},$$

where A, $E \in \mathbb{C}^{N \times N}$, $B \in \mathbb{C}^{N \times m}$, $C \in \mathbb{C}^{N \times p}$, N is the state-space dimension, m is the number of inputs, and p is the number of outputs. The transfer function H_n of any reduced-order model of state-space dimension n $(< N)$ of the original system is a $p \times m$-matrix-valued rational function of the form

$$H_n(s) = C_n^H (sE_n - A_n)^{-1} B_n, \quad s \in \mathbb{C},$$

where A_n, $E_n \in \mathbb{C}^{n \times n}$, $B_n \in \mathbb{C}^{n \times m}$, and $C_n \in \mathbb{C}^{n \times p}$. The reduced-order transfer function H_n is called an *n-th Padé approximant* of H (with respect to the expansion point $s_0 \in \mathbb{C}$) if

$$H_n(s) = H(s) + \mathcal{O}\Big((s - s_0)^{q(n)}\Big), \tag{1}$$

where $q(n)$ is as large as possible. It turns out that

$$q(n) \geq \left\lfloor \frac{n}{m} \right\rfloor + \left\lfloor \frac{n}{p} \right\rfloor,$$

with equality in the generic case; see, e.g., [6]. For any fixed expansion point s_0, the MPVL algorithm generates reduced-order models of state-space dimension n that are optimal in the sense that the corresponding transfer functions H_n are n-th Padé approximants of the original transfer function H.

Unfortunately, Padé-based order reduction methods such as MPVL do not preserve all the important properties and structures of the original large-scale system. For example, any meaningful original system will at least be stable, but Padé-based reduced-order models are not stable in general. One remedy is to relax (1) and to require instead that H_n is only an *n-th Padé-type approximant* of H, i.e.,

$$H_n(s) = H(s) + \mathcal{O}\Big((s - s_0)^{\tilde{q}(n)}\Big), \tag{2}$$

where $\tilde{q}(n) < q(n)$. A simple approach to constructing reduced-order models characterized by such a Padé-type approximation property is to employ projection onto suitable block Krylov subspaces [29,26–28]. For example, if block Krylov subspaces with the input matrix B as initial block are used, the transfer functions H_n of the resulting reduced-order models are n-th Padé-type approximants of the original transfer function H. Moreover, in (2), we have

$$\tilde{q}(n) \geq \left\lfloor \frac{n}{m} \right\rfloor, \tag{3}$$

with equality in the generic case; see, e.g., [29,28,8,9]. Algorithms characterized by such a Padé-type approximation property include PRIMA [26–28] and SPRIM [10,13,12]. While the approximation order (3) is best possible in the general case, in the case of SPRIM, there are certain special cases [10,11] where the Padé-type approximation property (2) even holds true with $\tilde{q}(n)$ replaced by $2\tilde{q}(n)$. The reason is that the original system exhibits certain symmetries and that these symmetries are preserved in the reduced-order models.

In this paper, we present a general framework that explains this higher accuracy of Padé-type reduced-order models. In particular, we use the notion of J-Hermitian linear dynamical systems to show that in the case of J-Hermitian linear dynamical systems, the reduced-order transfer functions match twice as many Taylor coefficients of the original transfer function as in the general case. An application to the SPRIM algorithm for order reduction of general RCL electrical networks is also discussed.

The remainder of this article is organized as follows. In Section 2, we recall some basic facts about linear dynamical systems, Krylov subspace-based model order reduction, and the resulting Padé-type reduced-order models. In Section 3, we introduce the notion of J-Hermitian time-invariant linear dynamical systems and establish the main result of this paper on Padé-type reduced-order models in the J-Hermitian case. In Section 4, we apply this main result to obtain the Padé-type approximation property of the SPRIM algorithm. Finally, in Section 5, we make some concluding remarks.

Throughout this paper the following notation is used. The set of real and complex numbers is denoted by $\mathbb{R}$ and $\mathbb{C}$, respectively. Unless stated otherwise, all vectors and matrices are allowed to have real or complex entries. For (real or complex) matrices $M = (\, m_{jk} \,)$, we denote by $M^T = (\, m_{kj} \,)$ the *transpose* of M, and by $M^H := (\, \overline{m_{kj}} \,)$ the *Hermitian* (or *complex conjugate transpose*) of M. The $n \times n$ identity matrix is denoted by I_n. If the dimension of I_n is apparent from the context, we drop the index and simply use I. The zero matrix is denoted by 0. The actual dimension of 0 will always be apparent from the context. For any matrix V, we denote by range(V) the range of V, which is defined as the subspace spanned by the columns of V.

2. Padé-type model order reduction of linear dynamical systems

In this section, we recall some basic facts about linear dynamical systems, Krylov subspace-based model order reduction, and the resulting Padé-type reduced-order models.

2.1. Time-invariant linear dynamical systems

We consider m-input p-output time-invariant linear dynamical systems given by a *state-space description* of the form

$$\begin{aligned} E\frac{\mathrm{d}x}{\mathrm{d}t} &= Ax + Bu(t), \\ y(t) &= C^H x(t), \end{aligned} \tag{4}$$

together with suitable initial conditions. Here, $A,\ E \in \mathbb{C}^{N\times N}$, $B \in \mathbb{C}^{N\times m}$, and $C \in \mathbb{C}^{N\times p}$ are given matrices, $x(t) \in \mathbb{C}^N$ is the vector of state variables, $u(t) \in \mathbb{C}^m$ is the vector of inputs, $y(t) \in \mathbb{C}^p$ is the vector of outputs, N is the state-space dimension, and m and p are the number of inputs and outputs, respectively. We remark that the second relation in (4) can be replaced by the more general equation

$$y(t) = C^H x(t) + Du(t),$$

where $D \in \mathbb{C}^{p\times m}$ is an additional given matrix. The resulting transfer function differs from the transfer function (see (6) below) of the system (4) only in the additive constant D, which has no effect on Padé-type model order reduction. Therefore, for simplicity, we restrict ourselves to the case (4).

We stress that the matrix E in (4) is allowed to be singular; in this case, the first equation in (4) is a system of differential-algebraic equations. However, we always assume that the matrix pencil

$$sE - A, \quad s \in \mathbb{C}, \tag{5}$$

is *regular*, i.e., the characteristic polynomial $\phi(s) = \det(sE - A)$ of (5) is not the zero polynomial. Since ϕ is a polynomial of degree at most N, the matrix $sE - A$ is then nonsingular for all except at most N values of $s \in \mathbb{C}$. Furthermore, to check regularity of the pencil (5) it is sufficient to show that the matrix $s_0 E - A$ is nonsingular for one $s_0 \in \mathbb{C}$.

The regularity of the matrix pencil (5) guarantees that

$$H(s) := C^H (sE - A)^{-1} B, \quad s \in \mathbb{C}, \tag{6}$$

is a well-defined rational $m \times p$-matrix-valued function with possible poles at the finitely many values $s \in \mathbb{C}$ for which $sE - A$ is singular. We remark that (6) is called the *transfer function* of the linear dynamical system (4).

2.2. Model order reduction via projection

A *reduced-order model* of the linear dynamical system (4) is a system of the same form as (4), but with smaller state-space dimension n $(< N)$. More precisely, a reduced-order model of (4) with state-space dimension n is a system of the form

$$\begin{aligned} E_n \frac{\mathrm{d}x_n}{\mathrm{d}t} &= A_n x_n + B_n u(t), \\ y_n(t) &= C_n^H x_n(t), \end{aligned} \tag{7}$$

where A_n, $E_n \in \mathbb{C}^{n \times n}$, $B_n \in \mathbb{C}^{n \times m}$, and $C_n \in \mathbb{C}^{n \times p}$. The problem of model order reduction of linear dynamical systems (4) is to determine an appropriate reduced state-space dimension n and to construct matrices A_n, E_n, B_n, and C_n such that the reduced-order model (7) is a sufficiently accurate approximation to the original system (4). In analogy to (6), the transfer function of (7) is defined by

$$H_n(s) := C_n^H (sE_n - A_n)^{-1} B_n, \quad s \in \mathbb{C}. \tag{8}$$

Again, H_n is a well-defined rational $m \times p$-matrix-valued function, provided that the reduced-order matrix pencil

$$sE_n - A_n, \quad s \in \mathbb{C}, \tag{9}$$

is assumed to be regular.

A simple approach to model order reduction is to use projection. Let n be an appropriate reduced state-space dimension, and let

$$V_n \in \mathbb{C}^{N \times n}, \quad \operatorname{rank}(V_n) = n, \tag{10}$$

be any given matrix with full column rank. Then, by setting

$$A_n := V_n^H A V_n, \quad E_n := V_n^H E V_n, \quad B_n := V_n^H B, \quad C_n := V_n^H C, \tag{11}$$

one obtains reduced data matrices that define a reduced-order model (7). We remark that the rank condition in (10) is necessary for the regularity of the reduced-order matrix pencil (9).

The simple projection approach (11) yields powerful model-order reduction techniques when the subspace spanned by the columns of the matrix (10), V_n, contains certain block Krylov subspaces.

2.3. Block Krylov subspaces

Let $s_0 \in \mathbb{C}$ be arbitrary, but fixed, such that the matrix $s_0 E - A$ is nonsingular. Note that, in view of the regularity of the matrix pencil (5), this assumption only excludes at most N values of $s_0 \in \mathbb{C}$.

We can rewrite the transfer function (6), H, as follows:

$$H(s) = C^H \Big(s_0 E - A + (s - s_0)E\Big)^{-1} B = C^H \Big(I + (s - s_0)M\Big)^{-1} R, \tag{12}$$

$$\text{where} \quad M := (s_0 E - A)^{-1} E, \quad R := (s_0 E - A)^{-1} B. \tag{13}$$

We will use block Krylov subspaces induced by the matrices M and R in (13) to generate the projected data matrices (11).

Next, we briefly review the notion of block Krylov subspaces; see [1] for a more detailed discussion. The matrix sequence $R, MR, M^2R, \dots, M^{j-1}R, \dots$ is called a *block Krylov sequence*. The columns of the matrices in this sequence are vectors of length N, and thus at most N of these columns are linearly independent. By scanning the columns of the matrices in the block Krylov sequence from left to right and deleting each column that is linearly dependent on earlier columns, we obtain the *deflated* block Krylov sequence

$$R^{(1)}, MR^{(2)}, M^2R^{(3)}, \ldots, M^{j-1}R^{(j)}, \ldots, M^{j_{\max}-1}R^{(j_{\max})}. \tag{14}$$

This process of deleting linearly dependent vectors is called *deflation*. In (14), each $R^{(j)}$ is a submatrix of $R^{(j-1)}$, and $R^{(1)}$ is a submatrix of R. Denoting by m_j the number of columns of $R^{(j)}$, we thus have

$$m \geq m_1 \geq m_2 \geq \cdots \geq m_j \geq \cdots \geq m_{j_{\max}} \geq 1. \tag{15}$$

By construction, the columns of the matrices (14) are linearly independent, and for each $\hat{n}$, the subspace spanned by the first $\hat{n}$ of these columns is called the *$\hat{n}$-th block Krylov subspace* (induced by M and R) and denoted by $\mathcal{K}_{\hat{n}}(M, R)$ in the sequel. Note that, by construction, we have

$$\dim \mathcal{K}_{\hat{n}}(M, R) = \hat{n}. \tag{16}$$

For $j = 1, 2, \ldots, j_{\max}$, we set

$$\hat{n}(j) := m_1 + m_2 + \cdots + m_j. \tag{17}$$

For $\hat{n} = \hat{n}(j)$, the $\hat{n}$-th block Krylov subspace is given by

$$\mathcal{K}_{\hat{n}}(M, R) = \operatorname{range}\Big(R^{(1)} \quad MR^{(2)} \quad M^2R^{(3)} \quad \cdots \quad M^{j-1}R^{(j)} \Big). \tag{18}$$

Note that, by (17), $\hat{n}(j) \leq m \cdot j$ with $\hat{n}(j) = m \cdot j$ if no deflation has occurred.

Remark 1. *The deflation process described in this subsection assumes exact arithmetic, and is sometimes referred to as* exact *deflation. In actual implementations of block Krylov subspace methods in finite-precision arithmetic, one also needs to deflate columns that are 'almost' linearly independent on earlier columns. The need for this so-called* inexact *deflation arises in any block Krylov subspace method. We refer the reader to* [14] *for a discussion of inexact deflation in the case of the block QMR method for systems of linear equations with multiple right-hand sides.*

2.4. A Padé-type approximation property

We now employ the projection approach (11) with matrices (10), V_n, that satisfy

$$\mathcal{K}_{\hat{n}}(M, R) \subseteq \operatorname{range}(V_n). \tag{19}$$

We remark that, in view of (10) and (16), the condition (19) implies that $n \geq \hat{n}$. Moreover, $n = \hat{n}$ if, and only if, the two subspaces in (19) are equal; in this case, the matrix $V_{\hat{n}}$ is said to be a *basis matrix* of the $\hat{n}$-th block Krylov subspace $\mathcal{K}_{\hat{n}}(M, R)$.

Let (7) be the associated reduced-order model defined by the reduced data matrices (11). We assume that the matrix $s_0 E_n - A_n$ is nonsingular. Note that this assumption guarantees the regularity of the reduced-order matrix pencil (9).

In analogy to (12) and (13), the reduced-order transfer function (8), H_n, can be rewritten as follows:

$$H_n(s) = C_n^H \Big(I_n + (s - s_0) M_n \Big)^{-1} R_n, \tag{20}$$

$$\text{where} \quad M_n := (s_0 E_n - A_n)^{-1} E_n, \quad R_n := (s_0 E_n - A_n)^{-1} B_n. \tag{21}$$

It turns out that the Taylor expansions of the reduced-order transfer function (8), H_n, and of the original transfer function (6), H, agree in a number of leading terms. This means that H_n is a Padé-type approximant of H. More precisely, we have the following result, which is well known; see, e.g., [8] for the general case, [3] for the special case of linear dynamical systems (4) with $E = I$, and [20] for the special case that no deflation occurs in the underlying block Krylov subspaces.

Theorem 2. *Let $V_n \in \mathbb{C}^{N \times n}$ and $s_0 \in \mathbb{C}$ be such that the matrices $s_0 E - A$ and $s_0 E_n - A_n$ are nonsingular and* (19) *holds true for an $\hat{n} = \hat{n}(j)$ of the form* (17) *for some $1 \leq j \leq j_{\max}$. Then the transfer function* (6)*, H, of the linear dynamical system* (4) *and the transfer function* (8)*, H_n, of the reduced-order model* (7) *defined by the projected data matrices* (11) *satisfy*

$$H_n(s) = H(s) + \mathcal{O}\Big((s - s_0)^j \Big). \tag{22}$$

Theorem 2 readily follows from Lemma 3 below, which we will also need to establish the main result of this paper in Section 3.

For the sake of completeness, we include this short proof of Theorem 2.

Proof of Theorem 2. Using the representations (12) and (20) to expand H and H_n about s_0, we obtain

$$\begin{aligned} H(s) &= \sum_{i=0}^{\infty}(-1)^i C^H M^i R (s-s_0)^i, \\ H_n(s) &= \sum_{i=0}^{\infty}(-1)^i C_n^H M_n^i R_n (s-s_0)^i. \end{aligned} \tag{23}$$

With these expansions, the claim (22) is equivalent to

$$C^H M^i R = C_n^H M_n^i R_n \quad \text{for all} \quad i = 0, 1, \ldots, j-1.$$

However, these relations follow directly from (24) below, by multiplying (24) from the left by C^H and using the definition of C_n in (11). $\square$

Lemma 3. *Under the assumptions of Theorem 2,*

$$M^i R = V_n M_n^i R_n \quad \textit{for all} \quad i = 0, 1, \ldots, j-1. \tag{24}$$

Here, M, R, M_n, and R_n are the matrices defined in (13) *and* (21).

The result of Lemma 3 is established as part of the proof of [11, Theorem 1]. In order to keep this paper self-contained, we include the proof of Lemma 3 in Appendix A.

3. Model order reduction of J-Hermitian systems

In this section, we introduce the notion of J-Hermitian time-invariant linear dynamical systems and establish the main result of this paper on Padé-type reduced-order models in the J-Hermitian case.

3.1. J-Hermitian time-invariant linear dynamical systems

Definition 4. *Let $J \in \mathbb{C}^{N\times N}$ be a nonsingular matrix. A matrix $M \in \mathbb{C}^{N\times N}$ is called J-*Hermitian *if*

$$JM = M^H J.$$

We stress that J is allowed to be any nonsingular matrix. In some applications, such as the one to the SPRIM algorithm in Section 4, the matrix J is Hermitian, and in Remark 7 below, we will briefly comment on the case

$$J = J^H. \tag{25}$$

Remark 5. *The matrix* $J \in \mathbb{C}^{N\times N}$ *induces the* sesquilinear form

$$[\cdot,\cdot] : \mathbb{C}^N \times \mathbb{C}^N \mapsto \mathbb{C} \quad \textit{defined by} \quad [x,y] := y^H J x, \quad x,y \in \mathbb{C}^N;$$

see, e.g., [25, Section XIII, §7]. *A matrix* $M \in \mathbb{C}^{N\times N}$ *is* J*-Hermitian if, and only if,* M *is self-adjoint with respect to the sesquilinear form* $[\cdot,\cdot]$*, i.e.,*

$$[Mx, y] = [x, My] \quad \textit{for all} \quad x,y \in \mathbb{C}^N.$$

Definition 6. *Let* $J \in \mathbb{C}^{N\times N}$ *be a nonsingular matrix. The time-invariant linear dynamical system* (4) *is said to be* J-Hermitian *if the following three conditions are satisfied:*

(i) *The matrices* A *and* E *are* J*-Hermitian;*
(ii) *The number of inputs and outputs are the same, i.e.,* $m = p$*;*
(iii) *The matrices* B *and* C *satisfy*

$$JB = CF \tag{26}$$

for some nonsingular matrix $F \in \mathbb{C}^{m\times m}$.

Note that the condition (26) is equivalent to $\mathrm{range}(JB) = \mathrm{range}(C)$.

Remark 7. *If the linear dynamical system* (4) *is* J*-Hermitian with* $F = I$ *in* (26)*, then its transfer function* (6)*,* H*, can be rewritten as follows:*

$$H(s) = B^H J^H J^{-1} (sE^H - A^H)^{-1} C, \quad s \in \mathbb{C}.$$

In particular, in the Hermitian case (25)*, we have*

$$H^H(s) := B^H (sE^H - A^H)^{-1} C = H(s), \quad s \in \mathbb{C}. \tag{27}$$

The property (27) *is called* Hamiltonian symmetry [15,16]*; transfer functions that satisfy this kind of symmetry are studied in detail in* [15,16].

3.2. Reduced-order models of J-Hermitian systems

We now assume that the linear dynamical system (4) is J-Hermitian, and we consider model order reduction of such systems via the projection approach (10), (11). The following result gives conditions that guarantee that the n-th projected reduced-order model is J_n-Hermitian.

Proposition 8. *Let $J \in \mathbb{C}^{N\times N}$ and $J_n \in \mathbb{C}^{n\times n}$ be nonsingular matrices, and let $V_n \in \mathbb{C}^{N\times n}$ be a matrix with full column rank n. Assume that the linear dynamical system* (4) *is J-Hermitian. Then the reduced-order model* (7) *defined by the projected data matrices* (11) *is J_n-Hermitian provided that the matrices A_n and E_n are J_n-Hermitian and the matrices J, J_n, and V_n satisfy the* compatibility condition

$$V_n^H J = J_n V_n^H. \tag{28}$$

Moreover, we have

$$J_n B_n = C_n F, \tag{29}$$

where F is the same matrix as in (26).

Proof. We only need to show (29). The remaining conditions for the reduced-order model (7) to be J_n-Hermitian are satisfied, in view of the assumptions of this proposition.

Multiplying (26) from the left by V_n^H and using (28) and the definitions of B_n and C_n in (11), we obtain

$$J_n B_n = J_n V_n^H B = V_n^H J B = V_n^H C F = C_n F.$$

Thus the proof is complete. $\square$

3.3. The Padé-type property in the J-Hermitian case

For the remainder of this section, let $J \in \mathbb{C}^{N\times N}$ and $J_n \in \mathbb{C}^{n\times n}$ be given nonsingular matrices and $V_n \in \mathbb{C}^{N\times n}$ be a given matrix with full column rank n. We assume that the time-invariant linear dynamical system (4) is J-Hermitian, and we consider the reduced-order model (7) defined by the projected data matrices (11). Moreover, we assume that the matrices A_n and E_n are J_n-Hermitian and that the matrices J, J_n, and V_n satisfy the compatibility condition (28). Note that, in view of Proposition 8, the reduced-order model (7) is J_n-Hermitian.

In this case of J_n-Hermitian reduced-order models of J-Hermitian systems, we have the following stronger Padé-type approximation property, instead of the corresponding property (22) of Theorem 2 for the general case.

Theorem 9. *Let $V_n \in \mathbb{C}^{N\times n}$ and $s_0 \in \mathbb{R}$ be such that the matrices $s_0E - A$ and $s_0E_n - A_n$ are nonsingular and (19) holds true for an $\hat{n} = \hat{n}(j)$ of the form (17) for some $1 \le j \le j_{\max}$. Then the transfer function (6), H, of the J-Hermitian linear dynamical system (4) and the transfer function (8), H_n, of the J_n-Hermitian reduced-order model (7) satisfy*

$$H_n(s) = H(s) + \mathcal{O}\Big((s-s_0)^{2j}\Big). \tag{30}$$

Proof. Recall the expansions (23) of H and H_n about s_0. By (23), the claim (30) is equivalent to

$$C^H M^i R = C_n^H M_n^i R_n \quad \text{for all} \quad i = 0, 1, \ldots, 2j-1. \tag{31}$$

In view of (34) below, we have

$$C^H M^{i_1} V_n = C_n^H M_n^{i_1}, \quad i_1 = 0, 1, \ldots, j, \tag{32}$$

and from Lemma 3, we have

$$M^{i_2} R = V_n M_n^{i_2} R_n, \quad i_2 = 0, 1, \ldots, j-1. \tag{33}$$

Using (33) and (32), it follows that

$$\begin{aligned} C^H M^i R = C^H M^{i_1+i_2} R = C^H M^{i_1} M^{i_2} R &= C^H M^{i_1} V_n M_n^{i_2} R_n \\ &= C_n^H M_n^{i_1} M_n^{i_2} R_n = C_n^H M_n^{i_1+i_2} R_n = C_n^H M_n^i R_n \end{aligned}$$

for all $i = i_1 + i_2 = 0, 1, \ldots, 2j-1$. Thus, the proof is complete. □

In the following proposition, we show that (34) indeed holds true.

Proposition 10. *Under the assumptions of Theorem 9,*

$$C^H M^i V_n = C_n^H M_n^i, \quad i = 0, 1, \ldots, j. \tag{34}$$

Proof. Instead of (34), we show the equivalent relation

$$V_n^H (M^H)^i C = (M_n^H)^i C_n, \quad i = 0, 1, \ldots, j. \tag{35}$$

For $i = 0$, (35) reduces to $V_n^H C = C_n$, which is just the definition of C_n in (11).
Now let $1 \le i \le j$. Since A and E are J-Hermitian and $s_0 \in \mathbb{R}$, we have

$$J(s_0E - A) = (s_0E - A)^H J,$$

or, equivalently,

$$(s_0E - A)^{-H} J = J(s_0E - A)^{-1}.$$

Recall that s_0 is assumed to be chosen such that the matrix $s_0E - A$ is nonsingular. Using the definition of M in (13) and the fact that E is assumed to be J-Hermitian, it follows that

$$M^H J = E^H (s_0E - A)^{-H} J = E^H J (s_0E - A)^{-1} = JE(s_0E - A)^{-1}.$$

This relation implies

$$\begin{aligned} (M^H)^i J &= JE\Big((s_0E - A)^{-1}E\Big)^{i-1}(s_0E - A)^{-1} \\ &= JEM^{i-1}(s_0E - A)^{-1}. \end{aligned} \tag{36}$$

Using (36), (26), and (13), one readily verifies that

$$(M^H)^i CF = (M^H)^i JB = JEM^{i-1}(s_0E - A)^{-1}B = JEM^{i-1}R. \tag{37}$$

Similarly, since the reduced-order model (7) is J_n-Hermitian and B_n and C_n satisfy (29), one shows that

$$(M_n^H)^i C_n F = J_n E_n M_n^{i-1} R_n. \tag{38}$$

Furthermore, note that, by (28) and the fact that E_n is J_n-Hermitian,

$$V_n^H JEV_n = J_n V_n^H EV_n = J_n E_n. \tag{39}$$

By multiplying (37) from the left by V_n^H and using (24) and subsequently (39), we obtain

$$\begin{aligned} V_n^H (M^H)^i C = V_n^H JE(M^{i-1}R) &= V_n^H JE(V_n M_n^{i-1} R_n) \\ &= J_n E_n M_n^{i-1} R_n. \end{aligned} \tag{40}$$

Since the right-hand sides of (38) and (40) are identical, the left-hand sides must agree as well, i.e.,

$$V_n^H (M^H)^i CF = (M_n^H)^i C_n F.$$

Recall that F is nonsingular, and by multiplying this last relation from the right by F^{-1}, we obtain (35). The proof of claim (35) is thus complete. □

4. Application to the SPRIM algorithm

In this section, we apply the result of Theorem 9 to establish a Padé-type approximation property of the SPRIM algorithm.

SPRIM is a reduction technique tailored to the problem of order reduction of the very large-scale RCL networks that arise in the simulation of electronic circuits. An RCL network is an electronic circuit that consists of only resistors, capacitors, and inductors, and that is powered by voltage and current sources. SPRIM was first proposed in [10] for the somewhat simpler case of RCL networks with only current sources. Recently [12,13], SPRIM was extended to the case of general RCL networks. Such networks can be described by m-input m-output linear dynamical systems of the form (4), where the matrices A, E, B, and C have additional structures; see, e.g. [13].

Here, we only recall from [13] the essential structures of the data matrices in (4) that are needed to verify that linear dynamical systems (4) describing general RCL networks are indeed J-Hermitian. The matrices A and E exhibit the following block structures:

$$A = \begin{pmatrix} A_1 & A_2 & A_3 \\ -A_2^H & 0 & 0 \\ -A_3^H & 0 & 0 \end{pmatrix}, \quad E = \begin{pmatrix} E_1 & 0 & 0 \\ 0 & E_2 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$
$$\text{where} \quad A_1 = A_1^H \in \mathbb{C}^{\nu_1 \times \nu_1}, \quad A_2 \in \mathbb{C}^{\nu_1 \times \nu_2}, \quad A_3 \in \mathbb{C}^{\nu_1 \times \nu_3},$$
$$E_1 = E_1^H \in \mathbb{C}^{\nu_1 \times \nu_1}, \quad E_2 = E_2^H \in \mathbb{C}^{\nu_2 \times \nu_2}. \tag{41}$$

Moreover, the matrices B and C are identical and of the form

$$B = C = \begin{pmatrix} B_1 & 0 \\ 0 & 0 \\ 0 & B_2 \end{pmatrix}, \quad \text{where} \quad B_1 \in \mathbb{C}^{\nu_1 \times \mu_1}, \quad B_2 \in \mathbb{C}^{\nu_3 \times \mu_2}. \tag{42}$$

Using (41) and (42), one readily verifies that the corresponding linear dynamical system (4) is J-Hermitian with J given by

$$J := \begin{pmatrix} I_{\nu_1} & 0 & 0 \\ 0 & -I_{\nu_2} & 0 \\ 0 & 0 & -I_{\nu_3} \end{pmatrix} \tag{43}$$

and the nonsingular matrix F in (26) given by

$$F := \begin{pmatrix} I_{\mu_1} & 0 \\ 0 & -I_{\mu_1} \end{pmatrix}. \tag{44}$$

Note that for the block sizes in (41)–(44), we have $\nu_1 + \nu_2 + \nu_3 = N$ and

$\mu_1 + \mu_2 = m$, and that the number of inputs and outputs in (4) is the same, i.e., $m = p$.

The SPRIM algorithm employs the projection approach (11) with a matrix (10), V_n, that satisfies the block-Krylov subspace inclusion (19). The main feature of SPRIM is that by carefully choosing the matrix V_n, it is possible to preserve the block structures (41), (42) of the data matrices of the original system (4).

More precisely, let $V_{\hat{n}}$ be a basis matrix of the $\hat{n}$-th block Krylov subspace $\mathcal{K}_{\hat{n}}(M, R)$, and let

$$V_{\hat{n}} = \begin{pmatrix} \hat{V}_1 \\ \hat{V}_2 \\ \hat{V}_3 \end{pmatrix} \tag{45}$$

be the partitioning of $V_{\hat{n}}$ corresponding to the block sizes of the matrices A and E in (41). For each $l = 1, 2, 3$, we first determine $n_l := \operatorname{rank}(\hat{V}_l)$ and set $V_l = \hat{V}_l$ if $n_l = n$. Furthermore, if $n_l < n$, we construct (e.g. via an LQ factorization) a matrix $V_l \in \mathbb{C}^{N_l \times n_l}$ with

$$\operatorname{range}(V_l) = \operatorname{range}(\hat{V}_l), \quad \operatorname{rank}(V_l) = n_l.$$

Finally, we set

$$n := n_1 + n_2 + n_3 \quad \text{and} \quad V_n := \begin{pmatrix} V_1 & 0 & 0 \\ 0 & V_2 & 0 \\ 0 & 0 & V_3 \end{pmatrix}. \tag{46}$$

By construction,

$$\mathcal{K}_{\hat{n}}(M, R) = \operatorname{range}(V_{\hat{n}}) \subseteq \operatorname{range}(V_n),$$

and thus the inclusion (19) is indeed satisfied. Furthermore, in view of (46), the projected data matrices (11) of the n-th SPRIM reduced-order model now exhibit block structures corresponding to those of the original data matrices (41), (42). It follows that both A_n and E_n are J_n-Hermitian, with the matrix J_n given by

$$J_n := \begin{pmatrix} I_{n_1} & 0 & 0 \\ 0 & -I_{n_2} & 0 \\ 0 & 0 & -I_{n_3} \end{pmatrix}. \tag{47}$$

Using (46) and (47), one immediately verifies that

$$V_n^H J = \begin{pmatrix} V_1^H & 0 & 0 \\ 0 & -V_2^H & 0 \\ 0 & 0 & -V_3^H \end{pmatrix} = J_n V_n^H,$$

which is just the compatibility condition (28). In view of Proposition 8, the n-th SPRIM reduced-oder models are thus J_n-Hermitian. We can thus apply Theorem 9 and obtain the following result about the Padé-type approximation property of the SPRIM algorithm.

Corollary 11. *Assume that the data matrices of the linear dynamical system* (4) *are of the form* (41), (42). *Let* (7) *be the n-th SPRIM reduced-order model of* (4), *where $s_0 \in \mathbb{R}$ and the dimension $\hat{n} = \hat{n}(j)$ of the underlying block Krylov subspace $\mathcal{K}_{\hat{n}}(M, R)$ is of the form* (17) *for some $1 \leq j \leq j_{\max}$. Assume that the matrices $s_0 E - A$ and $s_0 E_n - A_n$ are nonsingular. Then the transfer function* (6), *H, of the linear dynamical system* (4) *and the transfer function* (8), *H_n, of the n-th SPRIM reduced-order model* (7) *satisfy*

$$H_n(s) = H(s) + \mathcal{O}\Big((s - s_0)^{2j}\Big).$$

We stress that for the PRIMA order-reduction method, a result analogous to Corollary 11 is not true. The reason is that the PRIMA reduced–order models of J-Hermitian linear dynamical systems do not exhibit any J_n-Hermitian structure.

5. Concluding remarks

Projection onto block Krylov subspaces is a simple, yet powerful approach to construct reduced-order models of large-scale linear dynamical systems. The resulting models can be characterized by a Padé-type approximation property of the reduced-order transfer functions. In this paper, we have used the notion of J-Hermitian linear dynamical systems to provide a general framework that explains the stronger Padé-type approximation property in the case of systems with certain additional structures. While the creation of this general framework was motivated by the SPRIM algorithm for model order reduction of RCL networks in circuit simulation, there are other classes of order reduction problems that are covered by the proposed framework. For example, the Padé-type approximation properties for Hermitian higher-order linear dynamical systems discussed in [11] can also be obtained as a corollary to the main result of this present paper.

Appendix A. A proof of Lemma 3

In this appendix, we give a proof of Lemma 3.

First, note that by construction of the block Krylov subspace (18), $\mathcal{K}_{\hat{n}}(M, R)$,

$$\mathrm{range}(M^i R) \subseteq \mathcal{K}_{\hat{n}}(M, R) \quad \text{for all} \quad i = 0, 1, \ldots, j-1.$$

Thus, the assumption (19) guarantees the existence of matrices $X_i \in \mathbb{C}^{N \times m}$ such that

$$M^i R = V_n X_i, \quad i = 0, 1, \ldots, j-1. \tag{A1}$$

Moreover, since V_n has full column rank n, each matrix X_i is unique. In fact, we will show that

$$X_i = M_n^i R_n, \quad i = 0, 1, \ldots, j-1. \tag{A2}$$

By inserting $(A2)$ into $(A1)$, we then obtain the claim (24) of Lemma 3.

It thus remains to prove $(A2)$. To this end, we use induction on i. For $i = 0$, we note that, by $(A1)$ and the definition of R in (13),

$$V_n X_0 = R = (s_0 E - A)^{-1} B.$$

Multiplying this relation from the left by the matrix

$$(s_0 E_n - A_n)^{-1} V_n^H (s_0 E - A) \tag{A3}$$

and using the definitions (11) of A_n and E_n and the definition (21) of R_n, it follows that $X_0 = R_n$. This is just the relation $(A2)$ for $i = 0$.

Now assume that $(A2)$ holds true for $i-1$, i.e., $X_{i-1} = M_n^{i-1} R_n$, for some $1 \leq i \leq j-1$. Together with $(A1)$, it follows that

$$V_n X_i = M^i R = M(M^{i-1} R) = M V_n X_{i-1} = M V_n M_n^{i-1} R_n. \tag{A4}$$

In view of the definitions of M, E_n, and M_n in (13), (11), and (22), we have

$$(s_0 E_n - A_n)^{-1} V_n^H (s_0 E - A) M V_n = (s_0 E_n - A_n)^{-1} V_n^H E V_n = M_n. \tag{A5}$$

Multiplying $(A4)$ from the left by the matrix $(A3)$ and using $(A5)$, we obtain

$$X_i = M_n(M_n^{i-1}R_n) = M_n^i R_n.$$

This completes the proof of Lemma 3.

Acknowledgments

The author would like to thank Peter Benner for bringing the references [15,16] to his attention. The author is grateful to the two anonymous referees for their careful reading of the original manuscript and their constructive comments that helped to improve the presentation of the paper.

References

[1] J. I. Aliaga, D. L. Boley, R. W. Freund, and V. Hernández, A Lanczos-type method for multiple starting vectors, Math. Comp. 69 (2000) 1577–1601.

[2] G. A. Baker, Jr. and P. Graves-Morris, Padé Approximants, Second Edition, Cambridge University Press, New York, New York, 1996.

[3] C. de Villemagne and R. E. Skelton, Model reductions using a projection formulation, Internat. J. Control 46 (1987) 2141–2169.

[4] P. Feldmann and R. W. Freund, Efficient linear circuit analysis by Padé approximation via the Lanczos process, in: Proceedings of EURO-DAC '94 with EURO-VHDL '94, IEEE Computer Society Press, Los Alamitos, California, 1994, pp. 170–175.

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

[6] P. Feldmann and R. W. Freund, Reduced-order modeling of large linear subcircuits via a block Lanczos algorithm, in: Proc. 32nd ACM/IEEE Design Automation Conference, ACM, New York, New York, 1995, pp. 474–479.

[7] L. Fortuna, G. Nunnari, and A. Gallo, Model Order Reduction Techniques with Applications in Electrical Engineering, Springer-Verlag, London, United Kingdom, 1992.

[8] R. W. Freund, Krylov-subspace methods for reduced-order modeling in circuit simulation, J. Comput. Appl. Math. 123 (2000) 395–421.

[9] R. W. Freund, Model reduction methods based on Krylov subspaces, Acta Numerica 12 (2003) 267–319.

[10] R. W. Freund, SPRIM: structure-preserving reduced-order interconnect macromodeling, in: Tech. Dig. 2004 IEEE/ACM International Conference on Computer-Aided Design, IEEE Computer Society Press, Los Alamitos, California, 2004, pp. 80–87.

[11] R. W. Freund, Padé-type model reduction of second-order and higher-order linear dynamical systems, in: P. Benner, V. Mehrmann, and D. C. Sorensen (Eds.), Dimension Reduction of Large-Scale Systems, Lecture Notes in Computational Science and Engineering, Vol. 45, Springer-Verlag, Berlin/Heidelberg, 2005, pp. 191–223.

[12] R. W. Freund, The SPRIM algorithm for structure-preserving order reduction of general RCL circuits, Technical report, Department of Mathematics, University of California, Davis, California, in preparation, 2008.

[13] R. W. Freund, Structure-preserving model order reduction of RCL circuit equations, in: H. van der Vorst, W. Schilders, and J. Rommes (Eds.), Model Order Reduction: Theory, Research Aspects and Applications, Mathematics in Industry, Springer-Verlag, Berlin/Heidelberg, Germany, 2008, to appear. Available online at `http://www.math.ucdavis.edu/~freund/reprints.html`.

[14] R. W. Freund and M. Malhotra, A block QMR algorithm for non-Hermitian linear systems with multiple right-hand sides, Linear Algebra Appl. 254 (1997) 119–157.

[15] P. A. Fuhrmann, On symmetric rational matrix functions, Linear Algebra Appl. 50 (1983) 167–250.

[16] P. A. Fuhrmann, On Hamiltonian rational transfer functions, Linear Algebra Appl. 63 (1984) 1–93.

[17] K. Gallivan, E. J. Grimme, and P. Van Dooren, Asymptotic waveform evaluation via a Lanczos method, Appl. Math. Lett. 7 (1994) 75–80.

[18] W. B. Gragg, The Padé table and its relation to certain algorithms of numerical analysis, SIAM Rev. 14 (1972) 1–62.

[19] W. B. Gragg, Matrix interpretations and applications of the continued fraction algorithm, Rocky Mountain J. Math. 4 (1974) 213–225.

[20] E. J. Grimme, Krylov projection methods for model reduction, Ph.D. Thesis, Department of Electrical Engineering, University of Illinois at Urbana-Champaign, Urbana-Champaign, Illinois, 1997.

[21] M. H. Gutknecht, A completed theory of the unsymmetric Lanczos process and related algorithms, Part I, SIAM J. Matrix Anal. Appl. 13 (1992) 594–639.

[22] M. H. Gutknecht, A completed theory of the unsymmetric Lanczos process and related algorithms, Part II, SIAM J. Matrix Anal. Appl. 15 (1994) 15–58.

[23] C. Lanczos, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, J. Res. Nat. Bur. Standards 45 (1950) 255–282.

[24] C. Lanczos, Solution of systems of linear equations by minimized iterations, J. Res. Nat. Bur. Standards 49 (1952) 33–53.

[25] S. Lang, Algebra, Third Edition, Addison-Wesley, Reading, Massachusetts, 1993.

[26] A. Odabasioglu, Provably passive RLC circuit reduction, M.S. thesis, Department of Electrical and Computer Engineering, Carnegie Mellon University, 1996.

[27] A. Odabasioglu, M. Celik, and L. T. Pileggi, PRIMA: passive reduced-order interconnect macromodeling algorithm, in: Tech. Dig. 1997 IEEE/ACM International Conference on Computer-Aided Design, IEEE Computer Society Press, Los Alamitos, California, 1997, pp. 58–65.

[28] A. Odabasioglu, M. Celik, and L. T. Pileggi, PRIMA: passive reduced-order interconnect macromodeling algorithm, IEEE Trans. Computer-Aided Design 17 (1998) 645–654.

[29] 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: Tech. Dig. 1996 IEEE/ACM International Conference on Computer-Aided Design, IEEE Computer Society Press, Los Alamitos, California, 1996, pp. 288–294.

[30] R. S. Varga, On higher order stable implicit methods for solving parabolic partial differential equations, J. Math. and Phys. 40 (1961) 220–231.