Inversion Components of Block Hankel-like Matrices

George Labahn*
Department of Computer Science
University of Waterloo
Waterloo, Ontario, Canada N2L 3G1

Submitted by Peter Lancaster

ABSTRACT

The inversion problem for square matrices having the structure of a block Hankel-like matrix is studied. Examples of such matrices include Hankel striped, Hankel layered, and vector Hankel matrices. It is shown that the components that both determine nonsingularity and construct the inverse of such matrices are closely related to certain matrix polynomials. These matrix polynomials are multidimensional generalizations of Padé-Hermite and simultaneous Padé approximants. The notions of matrix Padé-Hermite and matrix simultaneous Padé systems are also introduced. These are shown to provide a second set of inverse components for block Hankel-like matrices. A recurrence relation is presented that allows for efficient computation of matrix Padé-Hermite and matrix simultaneous Padé systems. As a result it is shown that the inverse components can be computed via either the matrix Euclidean algorithm or a matrix Berlekamp-Massey algorithm applied to an associated matrix power series. An alternative algorithm based on this recurrence relation is also presented. For a block Hankel-like matrix of type $(n_0, n_1, \dots, n_k)$ this algorithm is shown to compute the inverse components with a complexity of $O(k \cdot (n_0 + \cdots + n_k)^2)$ block matrix operations, although this can be higher in some pathological cases. This is the same complexity as with existing algorithms. This algorithm has the significant advantage, however, that no extra conditions are required on the input matrix. Other block algorithms require that certain submatrices be nonsingular. Similar results hold in the case of block Toeplitz-like matrices.

*Partially supported by NSERC operating grant 6194.

655 Avenue of the Americas, New York, NY 10010 0024-3795/92/$5.00

1. INTRODUCTION

For any vector of $k+1$ nonnegative integers $n=(n_0,\ldots,n_k)$, let

$$H_n = \left[\begin{array}{ccc|c|ccc} a_{n_0-n_1+1,1} & \cdots & a_{n_0,1} & & a_{n_0-n_k+1,k} & \cdots & a_{n_0,k} \\ a_{n_0-n_1+2,1} & \cdots & a_{n_0+1,1} & & a_{n_0-n_k+2,k} & & a_{n_0+1,k} \\ \vdots & & \vdots & \cdots & \vdots & & \vdots \\ a_{\|n\|-n_1-1,1} & \cdots & a_{\|n\|-2,1} & & a_{\|n\|-n_k-1,k} & & a_{\|n\|-2,k} \\ a_{\|n\|-n_1,1} & \cdots & a_{\|n\|-1,1} & & a_{\|n\|-n_k,k} & \cdots & a_{\|n\|-1,k} \end{array}\right] \tag{1.1}$$

with $\|n\| = n_0 + \cdots + n_k$. The entries $a_{i,j}$ in (1.1) are assumed to be $p \times p$ matrices over a field F, with $a_{i,j} = 0$ for $i < 0$. If we let $m = \|n\| - n_0$, then (1.1) represents an $mp \times mp$ square matrix.

In the special case when $k = 1$, (1.1) represents the classical notion of a block Hankel matrix. For example, setting $a_i = a_{i,1}$ gives

$$H_{(s,s)} = \begin{bmatrix} a_1 & a_2 & \cdots & a_{s-1} & a_s \\ a_2 & a_3 & \cdots & a_s & a_{s+1} \\ \vdots & \vdots & & \vdots & \vdots \\ a_{s-1} & a_s & \cdots & a_{2s-3} & a_{2s-2} \\ a_s & a_{s+1} & \cdots & a_{2s-2} & a_{2s-1} \end{bmatrix}. \tag{1.2}$$

These have applications in such diverse branches of mathematics as systems theory (cf. Kalman [22]), partial realizations (cf. Gragg and Lindquist [15]), and rational function (i.e. Padé) approximation (cf. Gragg [14]). When $k = 2$ the matrix (1.1) is called a paired Hankel matrix (cf. Semencul [35]). When $n_0 = 0$, a paired Hankel matrix is a generalization of the classical Sylvester matrix. Indeed, if $p(z)$ and $q(z)$ are polynomials of degree r and s, respectively, then setting $p_i = a_{i,1}$ and $q_i = a_{i,2}$ gives

$$H_{(0,s,r)} = \left[\begin{array}{cccc|ccc} & & & p_0 & & & q_0 \\ & & \cdot^{\cdot^{\cdot}} & p_1 & & \cdot^{\cdot^{\cdot}} & q_1 \\ & p_0 & \cdot^{\cdot^{\cdot}} & \vdots & q_0 & \cdot^{\cdot^{\cdot}} & \vdots \\ p_0 & p_1 & & \vdots & q_1 & & \\ p_1 & \vdots & & p_r & \vdots & & q_s \\ \vdots & \vdots & \cdot^{\cdot^{\cdot}} & & & \cdot^{\cdot^{\cdot}} & \\ \vdots & p_r & & & \vdots & & \\ p_r & & & & q_s & & \end{array}\right], \tag{1.3}$$

which is precisely the Sylvester matrix of $p(z)$ and $q(z)$. These are of central importance in the study of polynomial GCDs (e.g., Sylvester's criterion) and theory of equations (via resultants) (cf. Geddes, Czapor, and Labahn [13]). In addition, vector Hankel matrices such as

$$\begin{bmatrix} v_1 & \cdots & v_s \\ \vdots & & \vdots \\ v_{s\cdot k} & \cdots & v_{s(k+1)-1} \end{bmatrix} \tag{1.4}$$

(where each v_i is a $1 \times k$ vector) can easily be transformed via column permutations into the matrix $H_{(s,\dots,s)}$, with $a_{i,j}$ being the jth component of v_i. Thus vector Hankel matrices can also be viewed as matrices having the structure of (1.1). For general k, matrices of the form (1.1) or its transpose are called block Hankel striped and block Hankel layered matrices, respectively, of type n.

We consider the problem of deciding when H_n is nonsingular and, when this is the case, of computing its inverse. This can always be solved by Gaussian elimination at a cost of $O(m^3p^3)$ operations. However, because of the added structure of matrices of the form (1.1), we expect more efficient methods to be available for solving the inversion problem. This is certainly true in the scalar case. Indeed, when $k = 1$ and $p = 1$, algorithms such as those given by Rissanen [34] or Cabay and Kao [9] requires only $O(m^2)$ operations to solve the inversion problem for H_n. For $k > 1$ and $p = 1$, the algorithm of Heinig [19] gives a fast inversion algorithm requiring less operations than Gaussian elimination. In addition, when fast polynomial arithmetic is available there are superfast methods by Brent, Gustavson, and Yun [7] or Sugiyama [36] in the $k = 1$ scalar case which compute the inverses with $O(m \log^2 m)$ operations. These methods all compute the inverse without any added conditions required by the original matrix. Other fast methods such as that of Levinson [28] and Trench [37] or the superfast methods of Ammar and Gragg [2] or de Hoog [20] all require that certain submatrices of H_n also be nonsingular.

When $k = 1$ and $p > 1$, fast $O(p^3m^2)$ inversion methods include those of Akaike [1], Rissanen [33], and Watson [38]. Bitmead and Anderson [6] give an example of a superfast $O(p^3m \log^2 m)$ method. However, in the block case all these algorithms require added conditions to be satisfied by the input block matrix. A fast algorithm that works in the block case without any added restrictions is given by Labahn, Choi, and Cabay [25]. For block matrices with $k > 1$ arbitrary, the algorithm of Kailath et al. [21] uses the notion of displacement rank to obtain the inverse of (1.1) in $O(kp^3m^2)$ operations. However, their method does not work in all cases.

In this paper we obtain fast methods for the inversion problem for arbitrary k and p. There is no restriction on the input matrix. Our work depends on some fundamental results of Lerer and Tismenetsky [27]. They show that the nonsingularity of H_n can be determined from the solution of a specific set of $2(k + 1)$ block equations having either H_n or its transpose as their matrix of coefficients. In addition, when such solutions exist, Lerer and Tismenetsky provide a formula to compute the inverse in terms of these solutions. We show that these solutions ("inverse components of the first kind") can be described in terms of matrix Padé-Hermite and matrix simultaneous Padé approximants of a related matrix of power series. These are generalizations of scalar concepts originally introduced by Mahler [29]. The inversion criterion for H_n is shown to be equivalent to certain normalizations of these associated approximants. When $k = 1$ these components of the first kind represent right and left matrix Padé forms, as defined in Labahn and Cabay [24]. When $p = 1$ they can be computed in $O(m \log^2 m)$ operations using the algorithm of Brent, Gustavson, and Yun [7].

The work of Lerer and Tismenetsky also results in an alternative set of inverse components for H_n ("inverse components of the second kind"). As before, these components are solutions of a (different) set of $2(k + 1)$ block equations having either H_n or its transpose as their matrix of coefficients. We show that these alternative inverse components can be described in terms of matrix Padé-Hermite and matrix Simultaneous Padé systems. These are generalizations of scalar concepts introduced by Cabay, Labahn, and Beckermann [11].

When $k = 1$ and $p = 1$, these Padé systems, and hence the inverse components, can be computed in $O(m \log^2 m)$ operations using the algorithm of Cabay and Choi [10]. When $k = 1$ and $p > 1$, both sets of inverse components are represented by right and left matrix Padé forms which can be computed via the algorithm of Labahn and Cabay [24]. This method computes the inverse components with a complexity of $O(p^3 m^2)$ operations, although there are pathological cases where the complexity becomes $O(p^3 m^3)$. The method of Labahn [23] computes the inverse components in $O(p^3 m \log^2 m)$ operations. However, the block Hankel matrix must satisfy additional properties (such as positive-definiteness for example) for this method to be applicable.

For arbitrary k and p a generalization of the recurrence relation of [11] gives a fast algorithm to compute the matrix Padé-Hermite and matrix simultaneous Padé systems. This in turn gives a fast algorithm for computing inverse components. This recurrence relation is shown to be a special case of the recurrence relation of Antoulas [3] used for the computation of minimal realizations of matrix power series. The work of Antoulas then implies that the inverse components for H_n can be computed via such algorithms as the matrix Euclidean algorithm (cf. Bultheel and Van Barel [8]) or a matrix version of the Berlekamp-Massey algorithm. These algorithms have the significant advantage that they require no extra conditions on the input block Hankel-like matrix to compute the inverse components.

A simplified algorithm based on the recurrence relation for Padé systems combined with Gaussian elimination is also presented. This algorithm also requires no extra conditions other than that H_n itself be nonsingular. The algorithm is iterative and computes, for no added cost, all the inverse components for certain nonsingular block Hankel-like submatrices of H_n. The complexity of this algorithm is normally given as $O(kp^3m^2)$ operations, although there are isolated examples where this cost can increase to $O(p^3m^3)$. The operation count for matrix inversion, however, is the former for those matrices that satisfy the conditions required by Kailath et al.

For a vector $m = (m_0, \dots, m_k)$ of integers let

$$T_m = \left[\begin{array}{cccc|c|cccc} b_{0,1} & b_{-1,1} & \cdots & b_{-m_1,1} & & b_{0,k} & \cdots & b_{-m_k,k} \\ b_{1,1} & b_{0,1} & \cdots & b_{-m_1+1,1} & & b_{1,k} & & b_{-m_k+1,k} \\ \vdots & \vdots & & \vdots & \cdots & \vdots & & \vdots \\ b_{M-2,1} & b_{M-1,1} & \cdots & b_{M-m_1-2,1} & & b_{M-2,k} & \cdots & b_{M-m_k-2,k} \\ b_{M-1,1} & b_{M-2,1} & \cdots & b_{M-m_1-1,1} & & b_{M-1,k} & \cdots & b_{M-m_k-1,k} \end{array}\right], \tag{1.5}$$

where $M = m_1 + \cdots + m_k + k$. A matrix of the form (1.4) is called a block Toeplitz striped matrix, while its (block) transpose is called a block Toeplitz layered matrix. If for each i we set $n_i = m_i + 1$ and $a_{\|n\|-n_i-i,j} = b_{i,j}$, then $T_m = JH_n$, where J is an $Mp \times Mp$ matrix with I_p on the block antidiagonal. Therefore any algorithm that computes the inverse of a nonsingular block Hankel-like matrix can also be used to compute the inverse of a nonsingular block Toeplitz-like matrix.

2. CHARACTERIZATION OF NONSINGULARITY

In this section we give necessary and sufficient conditions for a block Hankel-like matrix to be nonsingular, along with a formula to compute the inverse. Let

$$H_n \begin{bmatrix} q_{m-1} \\ \vdots \\ q_0 \end{bmatrix} = \begin{bmatrix} I_p \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{2.1}$$

$$H_n \begin{bmatrix} x_{m-1}^{(i)} \\ \vdots \\ x_0^{(i)} \end{bmatrix} = \begin{bmatrix} a_{n_0, i-1} \\ \vdots \\ a_{\|n\|-1, i-1} \end{bmatrix} - \begin{bmatrix} a_{n_0 - n_i, i} \\ \vdots \\ a_{\|n\|-n_i-1, i} \end{bmatrix} \quad \text{for} \quad 1 < i \leqslant k, \tag{2.2}$$

and

$$H_n \begin{bmatrix} x_{m-1}^{(1)} \\ \vdots \\ x_0^{(1)} \end{bmatrix} = - \begin{bmatrix} a_{n_0 - n_1, 1} \\ \vdots \\ a_{\|n\|-n_1-1, 1} \end{bmatrix}. \tag{2.3}$$

Clearly the $k + 1$ block equations (2.1)–(2.3) have solutions in the case when H_n is nonsingular.

Let

$$\left[q_{m-1}^{*(i)}, \dots, q_0^{*(i)}\right] H_n = E_{n_1 + \cdots + n_{i-1} + 1} \quad \text{for} \quad 1 \leqslant i \leqslant k, \tag{2.4}$$

where E_j denotes the jth block row of the $mp \times mp$ identity matrix, and

$$\left[v_{m-1}^{*}, \dots, v_0^{*}\right] H_n = -\left[a_{n_0 - n_1, 1}, \dots, a_{n_0 - 1, 1}, \dots, a_{n_0 - 1, k}\right]. \tag{2.5}$$

Again, when H_n is nonsingular one can solve the $k + 1$ block equations given by (2.4) and (2.5). Central to our work is:

THEOREM 2.1 (Lerer and Tismenetsky [27]). *Let n be a vector of nonnegative integers with $n_i \geqslant 1$ for at least one i. Then*

(a) *H_n is nonsingular if and only if there are solutions to Equations* (2.1)–(2.3);

(b) *H_n is nonsingular if and only if there are solutions to Equations* (2.4)–(2.5).

Furthermore, when there are solutions to Equations (2.1)–(2.5), *the inverse is given by*

$$\sum_{j=1}^{k} \begin{bmatrix} x_m^{(j)} & & \\ \vdots & \ddots & \\ x_1^{(j)} & \cdots & x_m^{(j)} \end{bmatrix} \begin{bmatrix} q_{m-1}^{*(j)} & \cdots & q_0^{*(j)} \\ \vdots & \cdot^{\cdot^{\cdot}} & \\ q_0^{*(j)} & & \end{bmatrix} - \begin{bmatrix} 0 & & & \\ q_{m-1} & & & \\ \vdots & \ddots & & \\ q_1 & \cdots & q_{m-1} & 0 \end{bmatrix} \begin{bmatrix} v_{m-1}^{*} & \cdots & v_0^{*} \\ \vdots & \cdot^{\cdot^{\cdot}} & \\ v_0^{*} & & \end{bmatrix}, \tag{2.6}$$

where $x_m^{(i)} = I_p \delta_{i,1}$.

Proof. Theorem 2.1 is simply a application of Theorem 4.2 of [27] to matrices of the form (1.1) given the natural rank decomposition. ∎

REMARK 1. In the $p = 1$ scalar case, Theorem 2.1 was first given by Heinig and Rost [17].

REMARK 2. By arranging the solutions of (2.1)–(2.5) into $p \times kp$ matrices

$$x_i = \left[x_i^{(1)}, \dots, x_i^{(k)}\right] \tag{2.7}$$

and $kp \times p$ matrices

$$q_i^* = \left[q_i^{*(1)}, \dots, q_i^{*(k)}\right]^T, \tag{2.8}$$

we can write (2.6) in the shortened form

$$H_n^{-1} = \begin{bmatrix} x_m & & \\ \vdots & \ddots & \\ x_1 & \cdots & x_m \end{bmatrix} \begin{bmatrix} q_{m-1}^* & \cdots & q_0^* \\ \vdots & \cdot^{\cdot^{\cdot}} & \\ q_0^* & & \end{bmatrix}$$

$$- \begin{bmatrix} 0 & & & \\ q_{m-1} & & & \\ \vdots & \ddots & & \\ q_1 & \cdots & q_{m-1} & 0 \end{bmatrix} \begin{bmatrix} v_{m-1}^* & \cdots & v_0^* \\ \vdots & \cdot^{\cdot^{\cdot}} & \\ v_0^* & & \end{bmatrix}. \tag{2.9}$$

COROLLARY 2.2. *H_n is nonsingular if and only if there are solutions to* (2.1) *and*

$$H_n \begin{bmatrix} v_{m-1}^{(i)} \\ \vdots \\ v_0^{(i)} \end{bmatrix} = - \begin{bmatrix} a_{n_0 - n_i, i} \\ \vdots \\ a_{\|n\| - n_i - 1, i} \end{bmatrix} \quad \textit{for} \quad 1 \leqslant i \leqslant k. \tag{2.10}$$

Proof. Since

$$H_n\{E_{n_1+\cdots+n_{i-1}}\}^T = \begin{bmatrix} a_{n_0,\, i-1} \\ \vdots \\ a_{\|n\|-1,\, i-1} \end{bmatrix} \qquad \text{for} \quad 1 < i \leqslant k, \tag{2.11}$$

it is clear that the existence of solutions to Equations (2.2) and (2.3) is equivalent to the existence of solutions to (2.10). ■

By Corollary 2.2 (and hence Theorem 2.1), the existence of solutions to the block linear equations (2.1), (2.4), (2.5), and (2.10) provide necessary and sufficient conditions for H_n to be nonsingular. In addition, when the inverse exists it can be determined from these solutions. For this reason we refer to the solutions of these four equations as the set of *inverse components of the first kind.*

A second set of inverse components also is represented as solutions of linear block equations. Let

$$H_n \begin{bmatrix} q_{m-1} \\ \vdots \\ q_0 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ I_p \end{bmatrix}, \tag{2.12}$$

$$H_n \begin{bmatrix} x_m^{(i)} \\ \vdots \\ x_1^{(i)} \end{bmatrix} = -\begin{bmatrix} a_{n_0+1,\, i} \\ \vdots \\ a_{\|n\|,\, i} \end{bmatrix} + \begin{bmatrix} a_{n_0-n_{i+1}+1,\, i+1} \\ \vdots \\ a_{\|n\|-n_{i+1},\, i+1} \end{bmatrix} \qquad \text{for} \quad 1 \leqslant i < k, \tag{2.13}$$

and

$$H_n \begin{bmatrix} x_m^{(k)} \\ \vdots \\ x_1^{(k)} \end{bmatrix} = -\begin{bmatrix} a_{n_0+1,\, k} \\ \vdots \\ a_{\|n\|,\, k} \end{bmatrix}. \tag{2.14}$$

Also let

$$\left[q_{m-1}^{*(i)}, \dots, q_0^{*(i)}\right] H_n = E_{n_1+\cdots+n_i} \qquad \text{for} \quad 1 \leqslant i \leqslant k \tag{2.15}$$

and

$$[v_m^*, \dots, v_1^*] H_n = -\left[a_{\|n\|-n_1+1,\, 1}, \dots, a_{\|n\|,\, 1}, \dots, a_{\|n\|,\, k}\right]. \tag{2.16}$$

THEOREM 2.3. *Let n be a vector of nonnegative integers with $n_i \geqslant 1$ for at least one i. Then*

(a) *H_n is nonsingular if and only if there are solutions to Equations* (2.12)–(2.14);

(b) *H_n is nonsingular if and only if there are solutions to Equations* (2.15)–(2.16).

Furthermore, when there are solutions to Equations (2.12)–(2.16), *the inverse is given by*

$$\sum_{j=1}^{k} \begin{bmatrix} x_{m-1}^{(j)} & \cdots & x_0^{(j)} \\ \vdots & \cdot^{\cdot^{\cdot}} & \\ x_0^{(j)} & & \end{bmatrix} \begin{bmatrix} q_{m-1}^{*(j)} & \cdots & q_0^{*(j)} \\ & \ddots & \vdots \\ & & q_{m-1}^{*(j)} \end{bmatrix}$$

$$- \begin{bmatrix} q_{m-2} & \cdots & q_0 & 0 \\ \vdots & \cdot^{\cdot^{\cdot}} & & \\ q_0 & & & \\ 0 & & & \end{bmatrix} \begin{bmatrix} v_m^* & \cdots & v_1^* \\ & \ddots & \vdots \\ & & v_m^* \end{bmatrix}, \tag{2.17}$$

where $x_0^{(i)} = I_p \delta_{i,k}$.

Proof. Theorem 2.3 is simply another form of Lerer and Tismenetsky's formula. However, for our purposes it is useful to show exactly how (2.17) can be derived from (2.6).

For every i, j set $\hat{a}_{i,j} = a_{\|n\|+n_0-n_{k-j+1}-i,\,k-j+1}$ and $\hat{n} = (n_0, n_k, \ldots, n_1)$. Let $\hat{H}_{\hat{n}}$ be the corresponding block Hankel-like matrix. Then

$$H_n = J\hat{H}_{\hat{n}}J, \tag{2.18}$$

where J is a matrix with I_p along the antidiagonal and zeros everywhere else. Equation (2.12) is therefore equivalent to

$$\hat{H}_{\hat{n}} \begin{bmatrix} q_0 \\ \vdots \\ q_{m-1} \end{bmatrix} = \begin{bmatrix} I_p \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{2.19}$$

which in turn corresponds to (2.1). Similarly, Equations (2.13)–(2.16) correspond to (2.2)–(2.5). Statements (a) and (b) of Theorem 2.3 therefore follow directly from Theorem 2.1. The inverse formula (2.17) is a direct consequence of (2.6) and the identity

$$H_n^{-1} = J\hat{H}_{\hat{n}}^{-1}J. \qquad \blacksquare \qquad (2.20)$$

REMARK 3. As before, by arranging the solutions of (2.13)–(2.15) into $p \times kp$ matrices

$$x_i = \left[x_i^{(1)}, \dots, x_i^{(k)}\right] \qquad (2.21)$$

and $kp \times p$ matrices

$$q_i^* = \left[q_i^{*(1)}, \dots, q_i^{*(k)}\right]^T \qquad (2.22)$$

we can write (2.17) in the shortened form

$$\begin{bmatrix} x_{m-1} & \cdots & x_0 \\ \vdots & \iddots & \\ x_0 & & \end{bmatrix} \begin{bmatrix} q_{m-1}^* & \cdots & q_0^* \\ & \ddots & \vdots \\ & & q_{m-1}^* \end{bmatrix}$$

$$- \begin{bmatrix} q_{m-2} & \cdots & q_0 & 0 \\ \vdots & \iddots & & \\ q_0 & & & \\ 0 & & & \end{bmatrix} \begin{bmatrix} v_m^* & \cdots & v_1^* \\ & \ddots & \vdots \\ & & v_m^* \end{bmatrix}. \qquad (2.23)$$

COROLLARY 2.4. *H_n is nonsingular if and only if there are solutions to* (2.12) *and*

$$H_n \begin{bmatrix} v_m^{(i)} \\ \vdots \\ v_1^{(i)} \end{bmatrix} = - \begin{bmatrix} a_{n_0+1,i} \\ \vdots \\ a_{\|n\|,i} \end{bmatrix} \qquad \textit{for} \quad 1 \leqslant i \leqslant k. \qquad (2.24)$$

We will refer to the solutions of Equations (2.12), (2.15), (2.16), and (2.24) as the *inverse components of the second kind.* The corresponding notions of inverse components of the first and second kinds for block Toeplitz-like matrices can also be given.

3. MATRIX PADÉ-HERMITE AND SIMULTANEOUS PADÉ APPROXIMANTS

In this section we introduce the notion of matrix Padé-Hermite and matrix simultaneous Padé forms for a matrix power series. These are multidimensional generalizations of scalar forms originally introduced by Mahler [29]. A fundamental tool used by Mahler (defined as a specific pair of matrix polynomials) is also generalized to the matrix case. The inversion characterization given in Theorem 2.1 of the previous section is then conveniently described in terms of a specific normalization of these fundamental tools of Mahler.

For each $0 \leqslant i \leqslant k$, let

$$A_i(z) = \sum_{i=0}^{\infty} a_i z^i \in \mathbf{D}[[z]] \tag{3.1}$$

be a formal power series with coefficients from the ring $\mathbf{D}$ of $p \times p$ matrices over a field $\mathbf{F}$. Let $n = (n_0, \dots, n_k)$ be a vector of nonnegative integers with $n_i \geqslant 1$ for at least one i, and $(P_0(z), \dots, P_k(z))$ a vector of $k+1$ matrix polynomials, each of size $p \times p$.

Definition 3.1. The $k+1$-vector of matrix polynomials $P = (P_0, \dots, P_k)^T$ is a *matrix Padé-Hermite form* (MPHFo) of type n for the $k+1$-vector of power series $A = (A_0, \dots, A_k)$ *if*

(I) the columns of P are linearly independent over F,
(II) $\deg_{(i)}(P) = \deg(P_i(z)) \leqslant n_i - 1$ for all $0 \leqslant i \leqslant k$, and
(III) one has

$$A_0(z) P_0(z) + \cdots + A_k(z) P_k(z) = z^{\|n\|-1} R(z), \tag{3.2}$$

where $R(z) \in \mathbf{D}[[z]]$.

When, in addition, we have $R(0) = I_p$, then P is said to be a *normed* MPHFo (of type n).

Definition 3.2. The $k+1$-vector of matrix polynomials $P^* = (P_0^*, \dots, P_k^*)$ [with each $P_i^*(z)$ a $p \times p$ matrix polynomial] is defined to be a *matrix simultaneous Padé form* (MSPFo) of type n for the $k+1$-vector of power series $A = (A_0, \dots, A_k)$ if

(I) the rows of P^* are linearly independent over F,

(II) $\deg_{[i]}(P^*) = \deg(P_i^*(z)) \leqslant \|n\| - n_i$ for all $0 \leqslant i \leqslant k$,
(III) one has

$$P_0^*(z)A_i(z) - P_i^*(z)A_0(z) = z^{\|n\|+1}R_i^*(z), \qquad 1 \leqslant i \leqslant k, \quad (3.3)$$

where $R_i^*(z) \in \mathbf{D}[[z]]$.

A MSPFo is called a *matrix simultaneous Padé fraction* (MSPFr) if $P_0^*(0)$ is also nonsingular; it is called a *normed* MSPFr if, in addition, $P_0^*(0) = I_p$.

REMARK 1. When $k = 1$, Definition 3.1 is the definition of a right matrix Padé form of type $(n_0 - 1, n_1 - 1)$, while Definition 3.2 gives a left matrix Padé form of type (n_0, n_1) (cf. Labahn and Cabay [24]). When $p = 1$ (i.e., the scalar case) and $k \geqslant 1$, Definition 3.1 is the classical notion of a Padé-Hermite form of type n for the vector of power series $(A_0(z), \ldots, A_k(z))$, while Definition 3.2 becomes the standard definition of a simultaneous Padé form when $A_0(0) \neq 0$. These were both originally introduced by Mahler [29] in 1932 as generalizations of some polynomial approximants used by Hermite in the special case of vectors of exponential functions. Simultaneous Padé approximants are also the same as directed vector Padé approximants in the unit coordinate directions (cf. Graves-Morris [16]).

For historical reasons, one can also refer to Definition 3.1 as defining the (matrix) Latin polynomial approximation problem and Definition 3.2 as the (matrix) German polynomial approximation problem. This naming convention comes from the notation used in their original definitions.

Note that when $A_0(0)$ is nonsingular, then one can define MPHFo's and MSPFo's even in the case $n_i = 0$. In this case we set $P_i(z) = 0$ and $P_i^*(z) = P_0^*(z)A_i(z)A_0(z)^{-1} \bmod z^{\|n\|+1}$.

If, for each i, we write $P_i(z) = \sum_{j=0}^{n_i-1} p_{i,j}z^j$, then Equation (3.2) is equivalent to solving a system of $\|n\| - 1$ block liner equations having $\|n\|$ block unknowns $\{p_{i,j}\}$. Therefore one can always find p linearly independent solutions. When these are arranged into columns of a matrix, we obtain at least one block solution of the linear system. Similarly, if for each i we write $P_i^*(z) = \sum_{j=0}^{\|n\|-n_i} p_{i,j}^* z^j$, then Equation (3.3) is equivalent to solving a system of $k \cdot (\|n\| + 1)$ block linear equations having $k \cdot (\|n\| + 1) + 1$ block unknowns $\{p_{i,j}^*\}$. Therefore one can always find p linearly independent solutions. When these are arranged into rows of a matrix, we get at least one block solution of the linear system.

THEOREM 3.3 (Existence of MPHFo's and SMPFo's). *For a given vector of integers n and vector of power series A, there always exists at least one MPHFo and one MPHFo of type n for A.*

Since Definition 3.1 depends on the integer vector n, we follow Mahler [29] in using the notation $P_i(z) = P_i(z|n)$ and $P_i^*(z) = P_i^*(z|n)$ when we wish to emphasize the type n. For a given integer vector n define the $p(k+1) \times p(k+1)$ matrix polynomials $\Lambda(z) = \Lambda(z|n)$ and $\Lambda^*(z) = \Lambda(z|n)$ having as their i, j entries

$$\Lambda_{i,j}(z|n) = P_i(z|n + e_j) \qquad \text{for} \quad 0 \leqslant i, j \leqslant k \tag{3.4}$$

and

$$\Lambda^*_{i,j}(z|n) = P_j^*(z|n - e_i) \qquad \text{for} \quad 0 \leqslant i, j \leqslant k, \tag{3.5}$$

where e_l denotes the lth row of I_{k+1}. In the $p = 1$ scalar case, Λ and Λ^* were first used by Mahler. Indeed, Mahler based much of his work in this area on the properties of these matrix polynomials. By Theorem 3.3 matrix polynomials Λ and Λ^* exist for any integer vector n. Of course, for an arbitrary integer vector n it could happen that there are many choices for them. However, in the cases when H_n is nonsingular these matrices are unique up to normalization. Indeed, we have:

THEOREM 3.4. *Let $A_1, \ldots, A_k$ be matrix power series determined by H_n. Then H_n is nonsingular if and only if there exists a matrix polynomial Λ of type n for the vector of power series $(I_p, A_1(z), \ldots, A_k(z))$ normalized so that*

$$\operatorname{lcoeff}_n(\Lambda(z|n)) = I_{p(k+1)}. \tag{3.6}$$

In (3.6) lcoeff$_n$ is the $p(k+1) \times p(k+1)$ matrix whose i, j entry is the coefficient of degree n_i in column j for $0 \leqslant i, j \leqslant k$.

Proof. Let $q_0, \ldots, q_{m-1}$ be a vector of $p \times p$ matrices. Partition this vector by

$$[q_{m-1}, \ldots, q_0] = -\left[p_{1,n_1-1}, \ldots, p_{1,0} | \cdots | p_{k,n_k-1}, \ldots, p_{k,0}\right], \tag{3.7}$$

and define matrix polynomials $P_i(z)$ as

$$P_i(z) = \sum_{j=0}^{n_i - 1} p_{i,j} z^j \qquad \text{for} \quad 1 \leqslant i \leqslant k \tag{3.8}$$

and

$$P_0(z) = -\{A_1(z)P_1(z) + \cdots + A_k(z)P_k(z)\} \bmod z^{\|n\|}. \qquad (3.9)$$

By (3.9), the matrix polynomials satisfy Equation (3.2) [with $A_0(z) = I_p$], while by (3.8) the matrix polynomials $P_i(z)$ for $1 \leqslant i \leqslant k$ have degrees at most $n_i - 1$. Therefore the matrix polynomials $P_0(z), \dots, P_k(z)$ form a MPHFo for $(I_p, A_1(z), \dots, A_k(z))$ of type $n + e_0$ if and only if the degree of $P_0(z)$ is at most n_0. Since

$$\begin{aligned} p_{0,n_0} &= -a_{n_0-n_1+1,1}p_{1,n_1-1} - \cdots - a_{n_0,1}p_{1,0} - \cdots - a_{n_0,k}p_{k,0} \\ &= a_{n_0-n_1+1,1}q_{m-1} + \cdots + a_{n_0,k}q_0, \end{aligned} \qquad (3.10)$$

$P_0(z)$ will have degree n_0 with a leading coefficient I_p if and only if the q_i's are the components of a solution of (2.1).

Similarly, for each $1 \leqslant j \leqslant k$ we can write

$$\left[v_{m-1}^{(j)}, \dots, v_0^{(j)}\right] = \left[p_{1,n_1-1}, \dots, p_{1,0} | \cdots | p_{k,n_k-1}, \dots, p_{k,0}\right] \qquad (3.11)$$

with $p_{j,n_j} = I_p$ and define $P_i(z)$ as

$$P_i(z) = \sum_{l=0}^{n_i-1} p_{i,l}z^l \qquad \text{for} \quad 1 \leqslant i \leqslant k, \quad i \neq j, \qquad (3.12)$$

$$P_j(z) = \sum_{l=0}^{n_j} p_{j,l}z^l, \qquad (3.13)$$

and

$$P_0(z) = -\{A_1(z)P_1(z) + \cdots + A_k(z)P_k(z)\} \bmod z^{\|n\|}. \qquad (3.14)$$

Then, for a given j, the matrix polynomials $P_0(z), \dots, P_k(z)$ defined by (3.12)–(3.14) form a MPHFo of type $n + e_j$ if (2.10) is true.

Clearly the matrix Λ generated by (3.8), (3.9), and (3.12)–(3.14) satisfies (2.1) and (2.10) if and only if (3.6) is true. Thus Theorem 3.4 follows directly from Corollary 2.2. ■

THEOREM 3.5. *Let* $A_1, \ldots, A_k$ *be matrix power series determined by* H_n. *Then* H_n *is nonsingular if and only if there exists a matrix polynomial* Λ^* *of type n for the vector of power series* $(I_p, A_1(z), \ldots, A_k(z))$ *normalized so that*

$$\operatorname{lcoeff}_{(\|n\|-n_0, \ldots, \|n\|-n_k)}(\Lambda^*(z|n)) = I_{p(k+1)}. \qquad (3.15)$$

Proof. Let

$$P_0^*(z) = v_0^* + \cdots + v_{m-1}^* z^{m-1} + v_m^* z^m \qquad (3.16)$$

with $v_m^* = I_p$ and

$$P_i^*(z) = P_0^*(z) A_i(z) \bmod z^{\|n\|} \quad \text{for} \quad 1 \leqslant i \leqslant k. \qquad (3.17)$$

Then the matrix polynomials $P_0^*(z), \ldots, P_k^*(z)$ defined by (3.16)–(3.17) form a MSPFo of type $n - e_0$ if (2.5) is true.

Similarly, for a given j with $1 \leqslant j \leqslant k$, set

$$P_0^*(z) = q_0^{(j)*} + \cdots + q_{m-1}^{(j)*} z^{m-1}, \qquad (3.18)$$

and define $P_i^*(z)$ by

$$P_i^*(z) = P_0^*(z) A_i(z) \bmod z^{\|n\|} \quad \text{for} \quad 1 \leqslant i \leqslant k. \qquad (3.19)$$

Then, for such a j, the matrix polynomials $P_0^*(z), \ldots, P_k^*(z)$ defined by (3.18)–(3.19) form a MSPFo of type $n - e_j$ if (2.5) is true.

Clearly the matrix Λ^* generated by (3.16)–(3.19) satisfies (2.4) and (2.5) if and only if (3.15) is true. Thus Theorem 3.5 follows directly from part (b) of Theorem 2.1. ∎

REMARK 2. Theorems 3.4 and 3.5 are equally valid if the leading coefficient is simply a nonsingular matrix.

EXAMPLE 3.6. Let $H_{(2,2,1)}$ be determined from the power series

$$A_1(z) = 2z^2 - 2z^3 + z^4 - 2z^5 + O(z^6),$$

$$A_2(z) = z - z^3 - z^4 + O(z^6).$$

Then $H_{(2,2,1)}$ is nonsingular. Solving Equations (2.1) and (2.10) and partitioning the solutions as in (3.7) and (3.11) gives

$$\Lambda(z) = \begin{bmatrix} z^2 - \frac{1}{4}z & -z & \frac{1}{4}z \\ -\frac{3}{8}z - \frac{1}{2} & z^2 + \frac{1}{2}z & -\frac{5}{8}z - \frac{1}{2} \\ \frac{1}{4} & 1 & z - \frac{1}{4} \end{bmatrix}.$$

Similarly, solving (2.4)–(2.5) and using (3.16)–(3.19) gives

$$\Lambda^*(z) = \begin{bmatrix} z^3 + \frac{1}{4}z^2 + \frac{1}{2}z + \frac{1}{2} & z^2 & -\frac{1}{4}z^3 + \frac{1}{2}z^2 + \frac{1}{2}z \\ \frac{3}{8}z^2 + \frac{1}{4}z - \frac{1}{4} & z^3 - \frac{1}{2}z^2 & \frac{5}{8}z^3 + \frac{1}{4}z^2 - \frac{1}{4}z \\ -\frac{1}{4}z^2 - \frac{1}{2}z - \frac{1}{2} & -z^2 & z^4 + \frac{1}{4}z^3 - \frac{1}{2}z^2 - \frac{1}{2}z \end{bmatrix}.$$

4. MATRIX PADÉ-HERMITE AND SIMULTANEOUS PADÉ SYSTEMS

In the previous section we used Theorems 2.1 and Corollary 2.2 along with some generalizations of the work of Mahler to provide a matrix polynomial characterization of the invertibility of block Hankel-like matrices. However this formulation does not lead to a simple mechanism to compute our inverse components.

In the case of scalar Padé-Hermite forms, Cabay, Labahn, and Beckermann [11] introduce a matrix polynomial that is similar to the matrix polynomial Λ used by Mahler. This matrix polynomial, called a Padé-Hermite system, is a combination of a Padé-Hermite form and a weak Padé-Hermite form. The latter is similar to a Padé-Hermite form except for a weakening of the order condition allowing for "larger" solutions. Cabay, Labahn, and Beckermann then give an efficient algorithm to compute these Padé-Hermite systems. A similar approach is found in [12] in the case of scalar simultaneous Padé systems.

In this section we extend the results of [11] and [12] to the case of matrix systems. Using Theorems 2.3 and Corollary 2.4 we show that the inversion characterization can be given in terms of certain normalizations of these systems. The strength of these results comes from the next sections where it is shown how to efficiently compute these normalized systems. Thus the inversion characterization of this section along with the computational method of later sections provides a fast method to compute the inversion components of the second kind, and hence also the inverse.

We begin by introducing the notion of weak matrix Padé-Hermite and weak matrix simultaneous Padé approximants.

DEFINITION 4.1. A $p(k+1) \times pk$ matrix polynomial $P = (P_0, \dots, P_k)^T$ is a *weak matrix Padé-Hermite form* (WMPHFo) of type n for the $k+1$-vector of power series $(A_0, \dots, A_k)$ if

(I) the columns of P are linearly independent over F,
(II) $\deg_{(i)}(P) = \deg(P_i(z)) \leqslant n_i$ for all $0 \leqslant i \leqslant k$, and
(III) one has

$$A_0(z)P_0(z) + \cdots + A_k(z)P_k(z) = z^{\|n\|+1}W(z), \tag{4.1}$$

where $W(z)$ is a $p \times pk$ matrix power series.

If $P_{i,j}(z)$ denotes the i,j block entry of P, then we set $U(z) = [P_{0,j}]_{j=1}^{k}$ and $V(z) = [P_{i,j}(z)]_{i,j=1}^{k}$. If, in addition, $V(0) = I_{pk}$, then P is called a *weak matrix Padé-Hermite fraction* (WMPHFr).

REMARK 1. When $k = 1$, Definition 4.1 is the definition of a right matrix Padé form of type (n_0, n_1) (cf. Labahn and Cabay [24]). When $p = 1$ (i.e., the scalar case) and $k \geqslant 1$, Definition 4.1 is the generalization of the weak Padé-Hermite form introduced by Cabay, Labahn, and Beckermann [11].

REMARK 2. When $A_0(z) = -I_p$, Equation (4.1) becomes

$$A(z)V(z) - U(z) = z^{\|n\|+1}W(z), \tag{4.2}$$

where $A = (A_1, \dots, A_k)$. This is the form of the usual order equation found in Padé approximation. Because of Equation (4.1), the matrix polynomials $U(z)$, $V(z)$, and $W(z)$ are called the weak matrix Padé-Hermite numerator, denominator, and residual, respectively, of type n.

DEFINITION 4.2. A $pk \times p(k+1)$ matrix polynomial $P^* = (P_0^*, \dots, P_k^*)$ is a *weak matrix simultaneous Padé form* (WMSPFo) of type n for the $k+1$-vector of power series $(A_0, \dots, A_k)$ if

(I) the rows of P^* are linearly independent over $\mathbf{F}$,
(II) $\deg_{[i]}(P^*(z)) = \deg(P_i^*(z)) \leqslant \|n\| - n_i - 1$ for all $0 \leqslant i \leqslant k$, and
(III) one has

$$P_0^*(z)A_i(z) - P_i^*(z)A_0(z) = z^{\|n\|-1}R_i^*(z), \tag{4.3}$$

where $R_i^*(z)$ is a $pk \times p$ matrix power series.

The matrix polynomials $V^*(z) = P_0^*(z)$, $U^*(z) = (P_1^*(z), \dots, P_k^*(z))$, and $W^*(z) = (R_1^*(z), \dots, R_k^*(z))$ are called the denominator, numerator, and residual, respectively. When $W^*(0) = I_{kp}$, then we have a *normed* WMSPFo.

REMARK 3. When $k = 1$, Definition 4.2 is the definition of a left matrix Padé form of type $(n_0 - 1, n_1 - 1)$ (cf. Labahn and Cabay [24]). When $A_0(z) = I_p$ and $A = (A_1, \dots, A_k)$, Equation (4.3) becomes

$$V^*(z)A(z) - U^*(z) = z^{\|n\|-1}W^*(z). \tag{4.4}$$

This accounts for the naming convention used for $U^*(z)$, $V^*(z)$, and $W^*(z)$.

If, for each i, we write $P_i(z) = \sum_{j=0}^{n_i} p_{i,j} z^j$, then Equation (4.1) is equivalent to a system of $\|n\| + 1$ block linear equations having $\|n\| + k + 1$ block unknowns $\{p_{i,j}\}$. Therefore one can always find pk linearly independent solutions. When these are arranged into columns of a matrix, we obtain at least one block solution of the linear system. Similarly, if for each i we write $P_i^*(z) = \sum_{j=0}^{\|n\|-n_i-1} p_{i,j}^* z^j$, then Equation (4.2) is equivalent to a system of $k \cdot (\|n\| - 1)$ block linear equations having $k\|n\|$ block unknowns $\{p_{i,j}^*\}$. Therefore one can always find pk linearly independent solutions. When these are arranged into rows of a matrix, we get at least one block solution of the linear system.

THEOREM 4.3 (Existence of WMPHFo's and WSMPFo's). *For a given vector n of integers and vector A of power series, there always exists at least one WMPHFo and one WMPHFo of type n for A.*

DEFINITION 4.4. For a given A and n define a $p(k + 1) \times p(k + 1)$ matrix polynomial $\Gamma(z)$ as follows:

(I) block column 0 of $\Gamma(z)$ is a MPHFo for A of type n;
(II) block columns $1, \dots, k$ of $\Gamma(z)$ form a WMPHFo for A of type n.

The matrix $\Gamma(z)$ will be referred to as a *matrix Padé-Hermite system* (MPHS) of type n for A. When the WMPHFo is a WMPHFr and the MPHFo is normed, then the MPHS is a *normed* MPHS.

DEFINITION 4.5. For a given A and n we also define a $p(k + 1) \times p(k + 1)$ matrix polynomial $\Gamma^*(z)$ as follows:

(I) block row 0 of $\Gamma^*(z)$ is a MPHFo for A of type n;
(II) block rows $1, \dots, k$ of $\Gamma^*(z)$ form a WMSPFo for A of type n.

The matrix $\Gamma^*(z)$ will be referred to as a *matrix simultaneous Padé system* (MSPS) of type n for A. When both the MSPFo and the WSPFo are normed, then the MSPS is a *normed* MSPS.

We will also use the notation $\Gamma(z|n)$ and $\Gamma^*(z|n)$ when we wish to emphasize the type of the systems.

In the $p = 1$ scalar case, PHSs are introduced in Cabay, Labahn, and Beckermann [11]. By Theorem 4.3 there exists a MPHS and a MSPS for every integer vector n. Of course, for an arbitrary integer vector n it could happen that there are many choices for these matrix polynomials. However, when H_n is nonsingular we have:

THEOREM 4.6. *Let $A_1, \dots, A_k$ be matrix power series determined by H_n. Then H_n is nonsingular if and only if there exists a normed MPHS of type n for the vector of power series $(I_p, A_1(z), \dots, A_k(z))$.*

Proof. Writing

$$[q_{m-1}, \dots, q_0] = [p_{1,n_1-1}, \dots, p_{1,0} | \cdots | p_{k,n_k-1}, \dots, p_{k,0}] \quad (4.5)$$

and defining $\Gamma_{i,0}(z)$ as

$$\Gamma_{i,0}(z) = \sum_{j=0}^{n_i-1} p_{i,j} z^j \qquad \text{for} \quad 1 \leqslant i \leqslant k \quad (4.6)$$

with

$$\Gamma_{0,0}(z) = -\{A_1(z)\Gamma_{1,0}(z) + \cdots + A_k(z)\Gamma_{k,0}(z)\} \bmod z^{\|n\|-1} \quad (4.7)$$

shows that Equation (2.12) is equivalent to the existence of a normed MPHFo of type n. Similarly, for each $1 \leqslant i \leqslant k$ we can write

$$[v_m^{(i)}, \dots, v_1^{(i)}] = [p_{1,n_1}^{(i)}, \dots, p_{1,1}^{(i)} | \cdots | p_{k,n_k}^{(i)}, \dots, p_{k,1}^{(i)}] \quad (4.8)$$

and

$$p_{j,0}^{(i)} = I_p \delta_{i,j}. \quad (4.9)$$

Defining $\Gamma_{i,j}(z)$ by

$$\Gamma_{i,j}(z) = \sum_{l=0}^{n_i} p_{i,j}^{(j)} z^l \qquad \text{for} \quad 1 \leqslant i, j \leqslant k \tag{4.10}$$

and $\Gamma_{0,j}(z)$ by

$$\Gamma_{0,j}(z) = -\{A_1(z)\Gamma_{1,j}(z) + \cdots + A_k(z)\Gamma_{k,j}(z)\} \bmod z^{\|n\|+1} \tag{4.11}$$

shows that Equation (2.24) is equivalent to the existence of a normed WMPHFo of type n. Theorem 4.6 then follows directly from Corollary 2.4. ∎

THEOREM 4.7. *Let $A_1, \ldots, A_k$ be matrix power series determined by H_n. Then H_n is nonsingular if and only if there exists a normed MSPS of type n for the vector of power series $(I_p, A_1(z), \ldots, A_k(z))$.*

Proof. Writing

$$\Gamma_{0,0}^*(z) = v_0^* + v_1^* z + \cdots + v_m^* z^m \tag{4.12}$$

with $v_0^* = I_p$ and

$$\Gamma_{0,j}^*(z) = \Gamma_{0,0}(z) A_j(z) \bmod z^{\|n\|+1} \qquad \text{for} \quad 1 \leqslant j \leqslant k \tag{4.13}$$

shows that Equation (2.16) is equivalent to the existence of a normed MSPFr of type n. Similarly, for each $1 \leqslant i \leqslant k$, we write

$$\Gamma_{i\ ,0}^*(z) = q_0^{(i)*} + \cdots + q_{m-1}^{(i)*} z^{m-1} \tag{4.14}$$

and define $\Gamma_{i\ ,j}^*(z)$ by

$$\Gamma_{i\ ,j}^*(z) = \Gamma_{i\ ,0}^*(z) A_j(z) \bmod z^{\|n\|+1}. \tag{4.15}$$

Then Equation (2.15) is equivalent to the existence of a normed WSPFo of type n. Thus Theorem 4.7 follows directly from part (b) of Theorem 2.3. ∎

EXAMPLE 4.6. Let $A_1(z)$ and $A_2(z)$ be as in Example 3.6. Then $H_{(2,2,1)}$ is nonsingular. Using Equations (2.12) and (2.24) along with (4.5)–(4.11) gives

$$\Gamma(z) = \begin{bmatrix} \frac{1}{2}z & -z^2 & \frac{3}{4}z^2 - z \\ -\frac{1}{4}z & z+1 & \frac{5}{8}z^2 + \frac{1}{2}z \\ -\frac{1}{2} & -z & -\frac{3}{4}z + 1 \end{bmatrix}.$$

Similarly, solving (2.15)–(2.16) along with (4.12)–(4.15) gives

$$\Gamma^*(z) = \begin{bmatrix} \frac{5}{8}z^3 - \frac{1}{4}z^2 + \frac{1}{4}z + 1 & -\frac{3}{2}z^3 + 2z^2 & -\frac{5}{8}z^4 - \frac{5}{4}z^3 + \frac{1}{4}z^2 + z \\ \frac{1}{2}z^2 & 0 & \frac{1}{2}z^3 \\ -\frac{1}{4}z^2 - \frac{1}{2}z - \frac{1}{2} & -z^2 & \frac{1}{4}z^3 - \frac{1}{2}z^2 - \frac{1}{2}z \end{bmatrix}.$$

5. A GENERALIZATION OF MAHLER'S THEOREM

One of the major results of the work of Mahler on Padé-Hermite and simultaneous Padé approximants was the observation that, at least in the scalar case, we have the identity

$$\Lambda^*(z)\Lambda(z) = Dz^{\|n\|}, \qquad (.)$$

where D is a diagonal matrix. In this section we generalize this result to obtain a similar relationship between matrix Padé-Hermite and matrix simultaneous Padé approximants. While the generalization from the scalar to the matrix case is straightforward, our use of and derivation of this result differs substantially from that of Mahler. In our case we work with the matrix polynomials Γ and Γ^* of the previous section, rather than the matrices Λ and Λ^* of Section 3. Our primary observation is that an identity similar to (5.1) provides a relation between the matrix polynomials of the previous section and the main tools used by Antoulas [3] in his study of recursiveness in linear systems. This in turn leads to the recurrence relation of the next section.

Let $A = (A_1, \ldots, A_k)$ be a vector of $p \times p$ power series and $n = (n_0, \ldots, n_k)$ a vector of nonnegative integers. Partition the $p \cdot (k+1)$ square matrix polynomials $\Gamma(z|n)$ and $\Gamma^*(z|n)$ into

$$\Gamma(z|n) = \begin{bmatrix} P(z) & U(z) \\ Q(z) & V(z) \end{bmatrix}, \qquad \Gamma^*(z|n) = \begin{bmatrix} V^*(z) & U^*(z) \\ Q^*(z) & P^*(z) \end{bmatrix}, \qquad (5.2)$$

where $P(z)$ and $V^*(z)$ are $p \times p$ matrix polynomials and $V(z)$ and $P^*(z)$ are pk-square matrix polynomials. It is easy to check that

$$\deg(P(z)) \leqslant n_0 - 1, \qquad \deg_{(i)}(Q(z)) \leqslant n_i - 1, \quad 1 \leqslant i \leqslant k, \tag{5.3}$$

and

$$\deg(U(z)) \leqslant n_0, \qquad \deg_{(i)}(V(z)) \leqslant n_i, \quad 1 \leqslant i \leqslant k. \tag{5.4}$$

where $\deg_{(i)}$ denotes the degree of the ith block row. Also

$$\deg(Q^*(z)) \leqslant \|n\| - n_0 - 1, \qquad \deg_{[i]}(P^*(z)) \leqslant \|n\| - n_i + 1 \tag{5.5}$$

for all $1 \leqslant i \leqslant k$ and

$$\deg(V^*(z)) \leqslant \|n\| - n_0, \qquad \deg_{[i]}(U^*(z)) \leqslant \|n\| - n_i \tag{5.6}$$

for all i, where $\deg_{[i]}$ denotes the degree of the ith block column.

The four pairs of matrix polynomials $(P(z), Q(z))$, $(U(z), V(z))$, $(P^*(z), Q^*(z))$, and $(U^*(z), V^*(z))$ satisfy

$$A(z)Q(z) + P(z) = z^{\|n\|-1}R(z), \tag{5.7}$$

$$A(z)V(z) + U(z) = z^{\|n\|+1}W(z), \tag{5.8}$$

$$Q^*(z)A(z) - P^*(z) = z^{\|n\|-1}R^*(z), \tag{5.9}$$

$$V^*(z)A(z) - U^*(z) = z^{\|n\|+1}W^*(z), \tag{5.10}$$

with $R(z)$, $W(z)$, $R^*(z)$, and $W^*(z)$ the appropriate residual power series. Theorems 4.6 and 4.7 imply that H_n is nonsingular if and only if the normalization conditions

$$R(0) = I_p, \qquad V(0) = I_{kp}, \qquad R^*(0) = I_{kp} \quad \text{and} \quad V^*(0) = I_p \tag{5.11}$$

also hold true.

THEOREM 5.1. *When H_n is nonsingular, we have*

$$\begin{bmatrix} V^*(z) & U^*(z) \\ -Q^*(z) & -P^*(z) \end{bmatrix}\begin{bmatrix} P(z) & U(z) \\ Q(z) & V(z) \end{bmatrix} = z^{\|n\|-1}I_{p(k+1)}, \quad (5.12)$$

$$\begin{bmatrix} P(z) & U(z) \\ Q(z) & V(z) \end{bmatrix}\begin{bmatrix} V^*(z) & U^*(z) \\ -Q^*(z) & -P^*(z) \end{bmatrix} = z^{\|n\|-1}I_{p(k+1)}, \quad (5.13)$$

$$\begin{bmatrix} V^*(z) & -zW^*(z) \\ -zQ^*(z) & R^*(z) \end{bmatrix}\begin{bmatrix} R(z) & zW(z) \\ zQ(z) & V(z) \end{bmatrix} = I_{p(k+1)}, \quad (5.14)$$

$$\begin{bmatrix} R(z) & zW(z) \\ zQ(z) & V(z) \end{bmatrix}\begin{bmatrix} V^*(z) & -zW^*(z) \\ -zQ^*(z) & R^*(z) \end{bmatrix} = I_{p(k+1)}. \quad (5.15)$$

Proof. Multiplying Equation (5.8) on the left by $Q^*(z)$ and Equation (5.9) on the right by $V(z)$ and subtracting the first from the second gives

$$-Q^*(z)U(z) - P^*(z)V(z) = z^{\|n\|-1}\{R^*(z)V(z) - z^2Q^*(z)W(z)\}. \quad (5.16)$$

The degree of the i,j entry of the $kp \times kp$ matrix polynomial $Q^*(z)U(z)$ is at most

$$\deg(Q^*(z)) + \deg(U(z)) \leqslant \|n\| - n_0 + n_0 - 1 = \|n\| - 1, \quad (5.17)$$

while the degree of an arbitrary block entry of $P^*(z)V(z)$ is bounded by

$$\max_{1\leqslant j\leqslant k}\left\{\deg_{[j]}(P^*(z)) + \deg_{(j)}(V(z))\right\} \leqslant \max_{1\leqslant j\leqslant k}\left\{\|n\| - n_j - 1 + n_j\right\}$$
$$= \|n\| - 1. \quad (5.18)$$

Therefore the left-hand side, and consequently also the right-hand side, of Equation (5.16) has degree at most $\|n\| - 1$. Thus

$$-Q^*(z)U(z) - P^*(z)V(z) = z^{\|n\|-1}r_0^*v_0 = z^{\|n\|-1}I_{pk}. \quad (5.19)$$

Multiplying Equation (5.8) on the left by $V^*(z)$ and Equation (5.10) on the right by $V(z)$ and subtracting the second from the first gives

$$V^*(z)U(z) + U^*(z)V(z) = z^{\|n\|+1}\{V^*(z)W(z) - W^*(z)V(z)\}. \quad (5.20)$$

The degree of the ith component of $V^*(z)U(z)$ is at most

$$\deg(V^*(z)) + \deg_{(i)}(U(z)) \leqslant \|n\| - n_0 + n_0 = \|n\|, \tag{5.21}$$

while the degree of the ith component of $U^*(z)v(z)$ is bounded by

$$\begin{aligned}\max_{1 \leqslant l \leqslant k} \left\{\deg_{[l]}(U^*(z)) + \deg_{(l)}(V(z))\right\} &\leqslant \max_{1 \leqslant l \leqslant k} \{\|n\| - n_l + n_l\} \\ &= \|n\|. \end{aligned} \tag{5.22}$$

Therefore the left-hand side, and consequently also the right-hand side, of Equation (5.20) is of degree at most $\|n\|$. Thus

$$V^*(z)U(z) + U^*(z)V(z) = 0. \tag{5.23}$$

In a similar fashion we can combine Equations (5.7) and (5.9) and obtain, after simplification,

$$Q^*(z)P(z) + P^*(z)Q(z) = 0, \tag{5.24}$$

and combine Equations (5.7) and (5.10) to get

$$V^*(z)P(z) + U^*(z)Q(z) = z^{\|n\|-1}I_p. \tag{5.25}$$

Equations (5.19), (5.23), (5.24), and (5.25) make up (5.12). Equation (5.13) follows from (5.12), since matrix inverses are two-sided.

Equations (5.16) and (5.19) also imply that

$$R^*(z)V(z) - z^2Q^*(z)W(z) = I_{pk}, \tag{5.26}$$

while Equations (5.17) and (5.20) imply

$$V^*(z)W(z) - W^*(z)V(z) = 0. \tag{5.27}$$

Similarly, we also have

$$R^*(z)Q(z) - Q^*(z)R(z) = 0 \tag{5.28}$$

and

$$V^*(z)R(z) - z^2W^*(z)Q(z) = I_p. \tag{5.29}$$

Equations (5.26)–(5.29) give (5.14). Equation (5.15) follows directly from (5.14). ■

REMARK 1. When $k = 1$, Theorem 5.1 is found in Labahn, Choi, and Cabay [25]. It is used there as a type of commutativity relationships between left and right matrix Padé approximants as defined in [24]. The method of proof and the partitioning of the matrices Γ and Γ^* as in (5.2) come from this paper.

Set

$$\mathbf{V}_n = \begin{bmatrix} \hat{P}(z) & \hat{U}(z) \\ \hat{Q}(z) & \hat{V}(z) \end{bmatrix}, \qquad \mathbf{W}_n = \begin{bmatrix} \tilde{V}^*(z) & \tilde{U}^*(z) \\ -\tilde{Q}^*(z) & -\tilde{P}^*(z) \end{bmatrix}, \tag{5.30}$$

where the $\hat{}$ denotes reversing the order of the coefficients of the matrix polynomials (on a per row basis) and the $\tilde{}$ denotes reversing the order of the coefficients of the matrix polynomials (on a per-column basis).

COROLLARY 5.2. *If H_n is nonsingular, then $\mathbf{V}_n$ and $\mathbf{W}_n$ are inverses of each other.*

Proof. Corollary 5.2 is simply a restatement of Equations (5.13) and (5.14) in Theorem 5.1. ■

REMARK 2. Corollary 5.2 implies that the matrices $\mathbf{V}_n$ and $\mathbf{W}_n$ are unimodular polynomial matrices. From Equations (5.8) and (5.10) we also have

$$A(z) = -U(z)V(z)^{-1} \bmod z^{\|n\|+1}, \tag{5.31}$$

$$A(z) = V^*(z)^{-1}U^*(z) \bmod z^{\|n\|+1}. \tag{5.32}$$

These properties are the central conditions needed by Antoulas [3] in his study of recursiveness in linear systems theory.

Let $A_1, \dots, A_k$ be matrix power series determined by H_n. Assume that the A_i are in fact matrix polynomials with $\deg(A_i) \leqslant \|n\| + n_0 - n_i$. Define

$$\hat{A}_i(z) = z^{\|n\|+n_0-n_i}A_i(z^{-1}), \tag{5.33}$$

that is, the matrix polynomials obtained by reversing the order of the coefficients of the A_i. Then H_n is nonsingular if and only if $\hat{H}_n$ is nonsingular (cf. Theorem 2.3).

THEOREM 5.3. *Let H_n be nonsingular with associated matrix polynomials $A_1(z), \dots, A_k(z)$. Let $(P(z), Q(z), R(z))$ and $(U(z), V(z), W(z))$ be* *determined from the MPHS of type n for the matrix polynomials* $(I_p, -\hat{A}_1, \ldots, -\hat{A}_k)$, *and let* $(P^*(z), Q^*(z), R^*(z))$ *and* $(U^*(z), V^*(z), W^*(z))$ *be determined from the corresponding MSPS of type n for the same matrix polynomials. Then*

$$\Lambda(z|n) = \begin{bmatrix} \hat{R}(z) & \hat{W}(z) \\ \hat{Q}(z) & \hat{V}(z) \end{bmatrix} \quad \text{and} \quad \Lambda^*(z|n) = \begin{bmatrix} \tilde{V}^*(z) & \tilde{W}^*(z) \\ \tilde{Q}^*(z) & \tilde{R}^*(z) \end{bmatrix}. \tag{5.34}$$

Proof. Let $P_{0,j}(s), \ldots, P_{k,j}(s)$ be the jth block column of $\Lambda(z|n)$. For any j this column represents a MPHFo of type $n + e_j$ with the leading coefficient of $P_{j,j}(z)$ being the identity. In particular we have that $\deg(P_{i,j}(z)) \leqslant n_i - 1 + \delta_{i,j}$ and

$$P_{0,j}(z) + A_1(z)P_{1,j}(z) + \cdots + A_k(z)P_{k,j}(z) = z^{\|n\|}R_j(z). \tag{5.35}$$

Since $\deg(A_i(z)) \leqslant \|n\| + n_0 - n_i$ for each i we also have $\deg(R_j(z)) \leqslant n_0 - \delta_{0,j}$. When $j = 0$ we can substitute $z = z^{-1}$ into (5.35) and multiply both sides of the result by $z^{|n|+n_0-1}$ to obtain

$$z^{\|n\|-1}\hat{P}_{0,0}(z) + \hat{A}_1(z)\hat{P}_{1,0}(z) + \cdots + \hat{A}_k(z)\hat{P}_{k,0}(z) = \hat{R}_0(z). \tag{5.36}$$

This shows that

$$\left(\hat{R}_0(z), \hat{P}_{1,0}(z), \ldots, \hat{P}_{k,0}(z)\right)^T \tag{5.37}$$

is a normed MPHFo of type n for $(I_p, -\hat{A}_1(z), \ldots, -\hat{A}_k(z))$ [having a residual $\hat{P}_{0,0}(z)$].

When $1 \leqslant j \leqslant k$, we can substitute $z = z^{-1}$ into (5.35) and multiply both sides of the result by $z^{\|n\|+n_0}$ to obtain

$$z^{\|n\|+1}\hat{P}_{0,j}(z) + \hat{A}_1(z)z\hat{P}_{1,j}(z) + \cdots + \hat{A}_j(z)\hat{P}_{j,j}(z) + \cdots + \hat{A}_k(z)z\hat{P}_{k,j}(z \\ = \hat{R}_j(z). \tag{5.38}$$

Therefore,

$$\begin{bmatrix} \hat{R}_1(z) & \hat{R}_2(z) & \cdots & \hat{R}_k(z) \\ \hat{P}_{1,1}(z) & z\hat{P}_{1,2}(z) & \cdots & z\hat{P}_{1,k}(z) \\ z\hat{P}_{2,1}(z) & \hat{P}_{2,2}(z) & \cdots & z\hat{P}_{2,k}(z) \\ \vdots & \vdots & & \vdots \\ z\hat{P}_{k,1}(z) & z\hat{P}_{k,2}(z) & \cdots & \hat{P}_{k,k}(z) \end{bmatrix} \tag{5.39}$$

is a normed WMPHFr of type n for $(I_p, -\hat{A}_1(z), \ldots, -\hat{A}_k(z))$ [with residual $(\hat{P}_{0,1}, \ldots, \hat{P}_{0,k})$]. Together (5.37) and (5.39) define a normed MPHS of type n for $(I_p, -\hat{A}_1(z), \ldots, -\hat{A}_k(z))$. Since $\hat{H}_n$ is nonsingular, these are unique by Theorem 4.6. This gives the first part of (5.34). The second part of (5.34) follows from a similar argument using the MSPS and Theorem 4.7. ∎

Theorem 5.3 provides a simple mechanism for computing the inverse components of one kind given that the components of the other kind are known.

As a result of Theorem 5.1 and Theorem 5.3 we have the following generalization of Mahler's theorem:

COROLLARY 5.4. *If H_n is nonsingular, then*

$$\Lambda^*(z|n)\Lambda(z|n) = z^{\|n\|} I_{p(k+1)} \tag{5.40}$$

and

$$\Lambda(z|n)\Lambda^*(z|n) = z^{\|n\|} I_{p(k+1)}. \tag{5.41}$$

Proof. From (5.34) we obtain (letting $s = z^{-1}$)

$$\begin{aligned}
\Lambda^*(z|n)\Lambda(z|n) &= \begin{bmatrix} \hat{R}(z) & \hat{W}(z) \\ \hat{Q}(z) & \hat{V}(z) \end{bmatrix} \begin{bmatrix} \tilde{V}^*(z) & \tilde{W}^*(z) \\ \tilde{Q}^*(z) & \tilde{R}^*(z) \end{bmatrix} \\
&= \begin{bmatrix} R(s) & sW(s) \\ sQ(s) & V(s) \end{bmatrix} \begin{bmatrix} z^{n_0} & & \\ & \ddots & \\ & & z^{n_k} \end{bmatrix} \\
&\quad \times \begin{bmatrix} z^{\|n\|-n_0} & & \\ & \ddots & \\ & & z^{\|n\|-n_k} \end{bmatrix} \begin{bmatrix} V^*(s) & -sW^*(s) \\ -sQ^*(s) & R^*(s) \end{bmatrix} \\
&= z^{\|n\|} \begin{bmatrix} R(s) & sW(s) \\ sQ(s) & V(s) \end{bmatrix} \begin{bmatrix} V^*(s) & -sW^*(s) \\ -sQ^*(s) & R^*(s) \end{bmatrix}.
\end{aligned} \tag{5.42}$$

Equations (5.40) and (5.41) therefore follow directly from (5.14) and (5.42). ∎

6. A RECURRENCE RELATION FOR INVERSION COMPONENTS

The previous section showed a strong relationship between normed MPHS and normed MSPS introduced in Section 4 and the main tools used by Antoulas. In this section we give a recurrence relation to construct these Padé systems that is a special case of the recurrence relation given by Antoulas. This will show how one can construct the inverse components of a nonsingular block Hankel-like matrix by finding the inverse components of two smaller block Hankel-like matrices.

By renumbering the $a_{i,j}$ if necessary we can assume

$$n_0 \geqslant n_1 \geqslant \cdots \geqslant n_k \geqslant 0. \tag{6.1}$$

Let r and s be integers such that

$$n_0 \geqslant n_1 \geqslant \cdots \geqslant n_r \geqslant s \geqslant n_{r+1} \geqslant \cdots \geqslant n_k, \tag{6.2}$$

and set

$$m = (n_0 - s, \dots, n_r - s, 0, \dots, 0). \tag{6.3}$$

Consider the problem of efficiently deciding when H_n is nonsingular given that H_m is known to be nonsingular. If this is the case, then is it possible to use the inverse components from the smaller matrix to build the inverse components of H_n.

From Section 4, the nonsingularity of H_m is equivalent to the existence of a matrix polynomial Γ satisfying conditions of Theorem 4.6. When we decompose the matrix Γ according to (5.2), we obtain pairs of matrix polynomials $(P(z), Q(z))$ and $(U(z), V(z))$ along with residual matrix power series $R(z)$ and $W(z)$. These matrix polynomials satisfy the degree constraints (5.2) and (5.3). In addition, by Equation (5.11), $R(0) = I_p$; hence the matrix power series $R(z)$ has an inverse. Set

$$A^{\#}(z) = zR(z)^{-1}W(z), \tag{6.4}$$

a $1 \times k$ vector of matrix power series, and let $H^{\#}_{n-m}$ be the block Hankel-like matrix of type $n - m = (s, s, \dots, s, n_{r+1}, \dots, n_k)$ associated to the vector of matrix power series $A^{\#}(z)$. The primary result of this section is:

THEOREM 6.1. *Suppose H_m is nonsingular. Then H_n is nonsingular if and only if $H^{\#}_{n-m}$ is nonsingular. In addition, the associated Γ and Γ^* matrices satisfy*

$$\Gamma(z|n) = \Gamma(z|m)S\Gamma^{\#}(z|n-m) \tag{6.5}$$

and

$$\Gamma^*(z|n) = \Gamma^{*\#}(z|n-m)T\Gamma^*(z|m), \tag{6.6}$$

where S and T are the matrices

$$S = \begin{bmatrix} zI_p & 0 \\ 0 & I_{pk} \end{bmatrix} \quad \text{and} \quad T = \begin{bmatrix} I_p & 0 \\ 0 & -zI_{pk} \end{bmatrix}. \tag{6.7}$$

Proof. Define a $p(k+1)$-matrix polynomial Γ by

$$\Gamma(z) = \Gamma(z|m)S\Gamma^{\#}(z|n-m), \tag{6.8}$$

where $\Gamma(z|m)$ and $\Gamma^{\#}(z|n-m)$ are arbitrary MPHSs for (I_p, A) and $(I_p, A^{\#})$, respectively.

By their definition, the degrees of the polynomial matrices are bounded on an element-by-element basis by

$$\deg(\Gamma(z|m)S) \leqslant \begin{bmatrix} n_0 - s & \cdots & n_0 - s & n_0 - s & \cdots & n_0 - s \\ n_1 - s & \cdots & n_1 - s & n_1 - s & \cdots & n_1 - s \\ \vdots & & \vdots & \vdots & & \vdots \\ n_r - s & \cdots & n_r - s & n_r - s & \cdots & n_r - s \\ -1 & \cdots & -1 & 0 & \cdots & -1 \\ \vdots & & \vdots & \vdots & & \vdots \\ -1 & \cdots & -1 & -1 & \cdots & 0 \end{bmatrix}, \tag{6.9}$$

$$\deg(\Gamma^{\#}(z|n-m)) \leqslant \begin{bmatrix} s-1 & s & \cdots & s & s & \cdots & s \\ s-1 & s & \cdots & s & s & \cdots & s \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ s-1 & s & \cdots & s & s & \cdots & s \\ n_{r+1}-1 & n_{r+1} & \cdots & n_{r+1} & n_{r+1} & \cdots & n_{r+1} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ n_k - 1 & n_k & \cdots & n_k & n_k & \cdots & n_k \end{bmatrix}; \tag{6.10}$$

hence (since $n_u - s < 0$ for all $r + 1 \leqslant u \leqslant k$) the degrees of the elements of Γ are bounded by

$$\deg(\Gamma) \leqslant \begin{bmatrix} n_0 - 1 & n_0 & \cdots & n_0 \\ n_1 - 1 & n_1 & \cdots & n_1 \\ \vdots & \vdots & & \vdots \\ n_k - 1 & n_k & \cdots & n_k \end{bmatrix}. \tag{6.11}$$

Therefore Γ satisfies the degree constraints for a MPHS of type n.

Also, we have

$$\begin{aligned} [I_p, A]\Gamma(z) &= [I_p, A]\Gamma(z|m)S\Gamma^{\#}(z|n-m) \\ &= z^{\|m\|-1}[R(z), z^2W(z)]S\Gamma^{\#}(z|n-m) \\ &= z^{\|m\|}R(z)[I_p, A^{\#}(z)]\Gamma^{\#}(z|n-m) \\ &= z^{\|m\|+\|n-m\|-1}R(z)[R^{\#}(z), z^2W^{\#}(z)], \\ &= z^{\|n\|-1}[R(z)R^{\#}(z), z^2R(z)W^{\#}(z)] \end{aligned} \tag{6.12}$$

so Γ also satisfies the order condition to be a MPHS of type n.

By (6.12) the residuals are given by

$$R(z)R^{\#}(z) \quad \text{and} \quad R(z)W^{\#}(z). \tag{6.13}$$

If we partition $\Gamma(z|m)$, $\Gamma(z|n-m)$, and $\Gamma(z)$ as in (5.2), then

$$\Gamma(z) = \begin{bmatrix} U(z)Q^{\#}(z) - zP(z)P^{\#}(z) & U(z)V^{\#}(z) - zP(z)U^{\#}(z) \\ V(z)Q^{\#}(z) - zU(z)P^{\#}(z) & V(z)V^{\#}(z) - zU(z)U^{\#}(z) \end{bmatrix}. \tag{6.14}$$

Since H_m is nonsingular, Theorem 4.6 implies that $\Gamma(z|m)$ can be chosen so that $V(0) = I_{pk}$ and $R(0) = I_p$. By (6.13) and (6.14) it is clear that Γ is a normalized MPHS of type n for A if and only if $\Gamma^{\#}(z|n-m)$ is a normalized MPHS of type $n - m$ for $A^{\#}$. This proves Theorem 6.1. ■

REMARK 1. Theorem 6.1 is a special case of the recurrence relation of Antoulas [3]. Indeed, by reversing the order of the polynomial coefficients, Equations (6.5) and (6.6) are the same as

$$\mathbf{V}_n = \mathbf{V}_m \mathbf{V}_{n-m} \quad \text{and} \quad \mathbf{W}_n = \mathbf{W}_{n-m} \mathbf{W}_m \tag{6.15}$$

(where the matrices $\mathbf{V}$ and $\mathbf{W}$ are defined in the previous section). These equations are used by Antoulas to obtain a minimal partial realization by interconnecting minimal partial realizations of two subsystems.

REMARK 2. Theorem 6.1 states that one can solve the inversion problem for H_m by solving the inverse problem for two smaller problems of the same type. The cost of such an approach is the cost of generating the matrix power series $A^{\#}(z)$, solving for the inverse components of the subproblems, and then expanding (6.5) and (6.6). This provides for an algorithm that builds the inverse components by solving a series of smaller problems, all of the same type.

EXAMPLE 6.2. Let

$$A_1(z) = 2z^2 - 2z^3 + z^4 - 2z^5 + 2z^6 - z^7 + z^8 + O(z^{12})$$

and

$$A_2(z) = z - z^3 - z^4 + z^6 + z^7 - z^9 + z^{11} + O(z^{12}).$$

We are interested in the inverse components of the second kind (if they exist) of $H_{(3,3,2)}$ and $H_{(4,4,3)}$. By Example 3.6 or 4.6, we have that $H_{(2,2,1)}$ is nonsingular with inverse components of the second kind given in Example 4.6. The residuals in this case are given by

$$R(z) = 1 - \tfrac{1}{4}z - z^3 + \tfrac{1}{4}z^4 + \tfrac{1}{4}z^5 + O(z^7)$$

and

$$W(z) = \left[-z^2 + z^3 + z^4 + O(z^7), \tfrac{5}{8} - \tfrac{9}{8}z^3 + \tfrac{11}{8}z^4 + z^5 - \tfrac{3}{4}z^6 + O(z^7)\right].$$

Therefore,

$$A_1^{\#}(z) = -z^3 + \tfrac{3}{4}z^4 + \tfrac{19}{16}z^5 - \tfrac{45}{64}z^6 + O(z^7)$$

and

$$A_2^{\#}(z) = \tfrac{5}{8}z + \tfrac{5}{32}z^2 + \tfrac{5}{128}z^3 - \tfrac{251}{512}z^4 + \tfrac{2565}{2048}z^5 + \tfrac{9477}{8192}z^6 + O(z^7).$$

Computing the matrices $H^{\#}_{(1,1,1)}$ and $H^{\#}_{(2,2,2)}$ shows that the first is singular, while the second is nonsingular with inverse components of the second kind given by

$$\Gamma^{\#}(z) = \begin{bmatrix} -\frac{5}{8}z & \frac{3}{4}z^2 & \frac{5}{8}z^2 - \frac{5}{8}z \\ -\frac{1}{2}z & \frac{13}{5}z^2 + z + 1 & \frac{3}{2}z^2 - \frac{1}{2}z \\ -\frac{1}{4}z + 1 & \frac{19}{10}z^2 - \frac{6}{5}z & \frac{1}{4}z^2 - \frac{5}{4}z + 1 \end{bmatrix}$$

and

$$\Gamma^{\#*}(z) = \begin{bmatrix} -\frac{11}{5}z^4 - \frac{1}{4}z^3 + z^2 - \frac{1}{4}z + 1 & z^4 - z^3 & -\frac{1}{2}z^4 + \frac{5}{8}z^3 + \frac{5}{8}z \\ \frac{1}{4}z^3 - z^2 & 0 & -\frac{5}{8}z^3 \\ \frac{3}{10}z^3 + \frac{7}{4}z^2 + \frac{3}{4}z + 1 & -z^3 & \frac{5}{4}z^3 + \frac{5}{8}z^2 + \frac{5}{8}z \end{bmatrix}.$$

Therefore $H_{(3,3,2)}$ is a singular matrix, while $H_{(4,4,3)}$ is nonsingular. Using Equations (6.5) and (6.6) gives the inverse components of the second kind for $H_{(4,4,3)}$ as

$$\Gamma(z) = \begin{bmatrix} z^2 - z & -\frac{4}{5}z^4 - \frac{19}{5}z^3 + \frac{1}{5}z^2 & -z^4 - z^3 + 2z^2 - z \\ 0 & z^4 + \frac{14}{5}z^3 + 3z^2 + 2z + 1 & z^3 + z^2 \\ z^2 - z + 1 & -\frac{22}{5}z^3 + \frac{9}{5}z^2 - \frac{11}{5}z & -2z^3 + 2z^2 - 2z + 1 \end{bmatrix}$$

with $\Gamma^*(z)$ given by

$$\begin{bmatrix} -2z^7 + \frac{4}{5}z^6 + \frac{1}{5}z^5 - \frac{11}{5}z^4 + z^3 + z^2 + 1 & \frac{14}{5}z^7 - \frac{17}{5}z^6 - 2z^5 + 3z^4 - 2z^3 + 2z^2 & z^8 + 3z^7 - \frac{4}{5}z^6 - \frac{16}{5}z^5 + z \\ z^6 + \frac{9}{5}z^5 + \frac{6}{5}z^4 + \frac{9}{5}z^3 + 2z^2 + z + 1 & \frac{4}{5}z^6 - \frac{7}{5}z^5 + 3z^4 + 2z^2 & -z^6 - \frac{9}{5}z^5 - \frac{1}{5}z^4 + z^3 + z^2 + z \end{bmatrix}.$$

Using Theorem 2.3 gives the inverse of $H_{(4,4,3)}$ as

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & -\frac{6}{5} & -\frac{4}{5} & \frac{4}{5} & -1 & -1 & 0 \\ 0 & -2 & -1 & 0 & -2 & -1 & 0 \\ 0 & -1 & 0 & 0 & -1 & 0 & 0 \\ 0 & \frac{13}{5} & \frac{7}{5} & \frac{3}{5} & 3 & 2 & 1 \\ 0 & -\frac{6}{5} & -\frac{4}{5} & -\frac{1}{5} & -1 & -1 & -1 \\ 1 & \frac{9}{5} & \frac{6}{5} & \frac{9}{5} & 2 & 1 & 1 \end{bmatrix}.$$

REMARK 3. Using Theorems 5.3 and 6.1, one can also obtain a recurrence relation for computing the matrices Λ and Λ^*. This will provide an algorithm for computing the inverse components of the first kind.

REMARK 4. For each i set $\hat{a}_{i,j} = a_{\|n\|+n_0-n_j-i,j}$. Then $\hat{H}_n$ is generated by the matrix power series (5.33). It is easy to see that H_n is nonsingular if and only if $\hat{H}_n$ is nonsingular, since they are the same matrices except for reordering of the rows and columns. Similarly, by writing out the entries of

$$H_{(n_0+rs,\, n_1-s,\, \dots,\, n_r-s,\, n_{r+1},\, \dots,\, n_k)}$$

one sees that this matrix is nonsingular if and only if $\hat{H}_{n-m}$ is nonsingular. Therefore the recurrence relation (6.5) applied to nonsingular matrices $\hat{H}_m$ and $\hat{H}_n$ results in a recurrence relation relating the inverse components of H_n to the components of

$$H_{(n_0+rs,\, n_1-s,\, \dots,\, n_r-s,\, n_{r+1},\, \dots,\, n_k)}$$

and the components of a second block Hankel-like matrix.

When $k = p = 1$, the inverse components (of either kind) are Padé forms. The recurrence relation (6.5) then gives an algorithm that computes these Padé forms along an off-diagonal path in the Padé table. By reversing the order of the coefficients as in (5.33), the resulting recurrence relation of the last paragraph computes the Padé forms along an antidiagonal path. In the first case the recurrence relation is the same as one used in the Berlekamp-Massey algorithm (cf. [5] and [30]), while the second is the same as derived from the Euclidean algorithm (cf. McEliece and Shearer [31]).

EXAMPLE 6.3. Let $A_1(z)$ and $A_2(z)$ be as in Example 6.2. Then

$$\hat{A}_1(z) = z^3 - z^4 + 2z^5 - 2z^6 + z^7 - 2z^8 + 2z^9 + O(z^{12})$$

and

$$\hat{A}_2(z) = z - z^3 + z^5 + z^6 - z^8 - z^9 + z^{10} + O(z^{12}).$$

The corresponding matrix $\hat{H}_{(4,4,3)}$ is the same, up to reordering of rows and columns, as $H_{(4,4,3)}$ from Example 6.2. In particular, the inverse components of the second kind for $\hat{H}_{(4,4,3)}$ are (up to reordering) the same as the inverse components of the first kind for $H_{(4,4,3)}$.

In order to compute the inverse components of the second kind for $\hat{H}_{(4,4,3)}$ one computes the inverse components of the second kind (where possible) for the sequence of submatrices

$$\hat{H}_{(1,1,0)}, \qquad \hat{H}_{(2,2,1)}, \quad \text{and} \quad \hat{H}_{(3,3,2)}.$$

One then uses Theorem 6.1 whenever one of these submatrices is nonsingular (in the above case this occurs only at $\hat{H}_{(3,3,2)}$). This corresponds to constructing the inverse components of the first kind (where possible) for the sequence of submatrices

$$H_{(10,1,0)}, \qquad H_{(8,2,1)}, \quad \text{and} \quad H_{(6,3,2)}.$$

(which in the above case only exist for $H_{(6,3,2)}$) in order to compute the inverse components of the first kind for $H_{(4,4,3)}$.

7. AN INVERSION ALGORITHM

Theorem 6.1 gives a recursive method to compute the inverse components of a block Hankel-like matrix by constructing the inverse components of a series of principal submatrices. This recursion is a special case of the method of Antoulas [3] for the study of recursiveness in linear systems theory. By the work of Antoulas the inversion components can therefore be determined by a matrix version of the Berlekamp-Massey algorithm or a matrix version of the Euclidean algorithm (cf. Bultheel and Van Barel [8]).

In this section we use Theorem 6.1 to develop an alternative algorithm for computing the inverse components of a block Hankel-like matrix, when such components exist. This simple algorithm stays entirely in the context of matrix theory, using Gaussian elimination to solve the individual subproblems. It is iterative, rather than recursive, and builds the inverse components by computing the inverse components of all nonsingular submatrices H_m for m of the form (6.3). The algorithm allows for a complete complexity analysis given in the next section.

Given a vector of nonnegative integers $n = (n_0, \ldots, n_k)$ ordered as in (6.1), the algorithm Inverse–Components below makes use of Theorem 6.1 to compute the polynomial matrices $\Gamma(z|n)$ and $\Gamma^*(z|n)$ for a given vector of matrix power series $A = (A_1, \ldots, A_k)$. Intermediate results available from this algorithm include those polynomial matrices $\Gamma(z|m$ and $\Gamma^*(z|m)$ at vectors m such that the submatrix H_m is nonsingular. These intermediate results allow for the construct the inverses of the corresponding "principal" submatrices.

The algorithm is presented in two parts. The first, First–Components, takes as its input H_n and n and returns the inverse components of the first principal submatrix H_m with m of the form (6.3). The main algorithm calls this routine to iteratively construct the inverse components for the block Hankel-like matrix associated to the residual matrix power series.

First–Components(H_n, n).

(FC-1) $d \leftarrow 0$, $s \leftarrow 0$
(FC-2) Do while $s < n_1$ and $d = 0$
(FC-3) $\quad s \leftarrow s + 1$
(FC-4) $\quad m_j \leftarrow \min(n_j, s)$, $j = 0, \ldots, k$
(FC-5) $\quad$ Compute $d \leftarrow \det(H_m)$, using Gaussian elimination
$\qquad\quad$ End While
(FC-6) If $d \neq 0$ then solve for $\Gamma(z|m)$ and $\Gamma^*(z|m)$
$\qquad\quad$ else Return("Matrix Singular");
(FC-7) Return $(s, \Gamma(z|m), \Gamma^*(z|m))$

The main algorithm, Inverse–Components, takes as its input a block Hankel-like matrix and an ordered vector of integers.

Inverse–Components(H_n, n).

(IC-1) $(s_0, \Gamma, \Gamma^*) \leftarrow$ First–Components(H_n, n); $s \leftarrow s_0$; $i \leftarrow 1$;
(IC-2) Do while $s \leqslant n_1$
(IC-3) $\quad m_j^{(i)} \leftarrow \max(n_j - s, 0)$, $j = 0, \ldots, k$
(IC-4) $\quad$ Determine R and W via Equations (5.7) and (5.8); form $A^\#$
(IC-5) $\quad (s_i, \Gamma^\#, \Gamma^{\#*}) \leftarrow$ First–Components(H_{n-m}, $n - m$)
(IC-6) $\quad \Gamma \leftarrow \Gamma \mathbf{S} \Gamma^\#$, $\Gamma^* \leftarrow \Gamma^{\#*} \mathbf{T} \Gamma^*$ [cf. (6.5) and (6.6)]
(IC-7) $\quad s \leftarrow s + s_{i-1}$,
$\qquad\quad$ End While
(IC-8) Return(s, Γ, Γ^*)

EXAMPLE 7.1. Let the 13×13 matrix $H_{(7,7,6)}$ be generated by

$$A_1(z) = 2z^2 - 2z^3 + z^4 - 2z^5 + 2z^6 - z^7 + z^8 + O(z^{14})$$

and

$$A_2(z) = z - z^3 - z^4 + z^6 + z^7 - z^9 + z^{11} + O(z^{14}).$$

then the inverse components of the second kind for $H_{(7,7,6)}$ are determined by computing the inverse components of the second kind for the sequence of submatrices

$$H_{(2,2,1)}, \quad H_{(4,4,3)}, \quad H_{(5,5,4)}, \quad \text{and} \quad H_{(6,6,5)}.$$

The submatrices $H_{(1,1,0)}$ and $H_{(3,3,2)}$ need to be "jumped," since they are singular. This is done via Gaussian elimination. Note that the inverse components of the first kind could be computed by this algorithm (using Remark 4 of the previous section) applied to the "reversed" polynomials. In this case the inverse components are computed by searching through

$$H_{(19,1,0)}, \quad H_{(17,2,1)}, \quad H_{(15,3,2)}, \quad H_{(13,4,3)}, \quad H_{(11,5,4)} \quad \text{and} \quad H_{(9,6,5)}$$

for nonsingular submatrices. In this case all the submatrices in the sequence are singular, so the inverse components of the first kind would just be determined by Gaussian elimination.

8. COMPLEXITY OF THE INVERSION ALGORITHM

In assessing the cost of our inversion algorithm we count the number of multiplications required by most of the steps of the algorithm, excluding from consideration the more trivial ones. The complexity computation closely follows that of Cabay, Labahn, and Beckermann [11] from the scalar case.

Let $s_1, \dots, s_l$ be the step sizes returned from the individual calls to First–Components in steps (IC-1) and (IC-5), and let $m^{[i]}$ be the corresponding integer vectors determined upon exit from the While loop (IC-3). Then the inverse components for H_n are computed iteratively by computing the inverse components for the submatrices $H_{m^{[1]}}, \dots, H_{m^{[l]}}$. The matrix H_n is then nonsingular if and only if $m^{[l]} = n$.

Consider first the cost of invoking the initialization algorithm. Gaussian elimination is used in step (FC-5) to obtain a triangular factorization of H_m. Assuming that the elimination is performed by applying bordering techniques (as s increases), step (FC-5) requires approximately $(\|m\|)^3 p^3/3$ multiplications in F, where m is the integer vector attained upon exit from the while loop (FC-2). In the case where $d \neq 0$, the solution of the equations resulting in the polynomial matrices Γ and Γ^* of type m can then be obtained by forward and backward substitution, requiring approximately $2(k+1)(\|m\|)^2 p^3$ multiplications in total. Therefore, if we set $m^{[0]} = 0$, then the ith invocation of First–Components costs approximately

$$\{\eta_i^3 + 2(k+1)\eta_i^2\}p^3 \tag{8.1}$$

operations, where $\eta_i = \|m^{[i]}\| - \|m^{[i-1]}\|$. If we set $\tau_i = \|m^{[i]}\|$, $i = 0, \ldots, l$, then it is easy to see that the total cost for the computation of residuals and $A^{\#}$ in step (IC-4) and the matrices Γ and Γ^* in step (IC-6) is approximately

$$2(k+1)\tau_i\eta_i p^3. \tag{8.2}$$

THEOREM 8.1. *The algorithm* Inverse–Components *requires approximately*

$$O\big((k+1)\|n\|^2 p^3 + (k+1)^2 s^2 \|n\| p^3\big) \tag{8.3}$$

multiplications in F, where $s = \max(s_1, s_2, \ldots, s_l)$ *and s_i is the ith step size.*

Proof. Equations (8.1) and (8.2) imply that the number of $p \times p$ matrix multiplications in Inverse–Components is asymptotically given by

$$\sum_{i=1}^{l} \big[\eta_i^3 + (k+1)\eta_i^2 + \eta_i\tau_i(k+1)\big], \tag{8.4}$$

where $\eta_i \leqslant s_i \cdot (k+1) \leqslant s \cdot (k+1)$, and

$$\sum_{i=1}^{l} \eta_i \leqslant \tau_l \leqslant \|n\|. \tag{8.5}$$

Therefore the number of matrix operations of the algorithm is bounded approximately by

$$(k+1)^2 s^2 \sum_{i=1}^{l} \eta_i + (k+1)\|n\| \sum_{i=1}^{l} \eta_i + (k+1) \sum_{i=1}^{l} \eta_i \tau_i$$

$$\leqslant (k+1)^2 s^2 \|n\| + (k+1)\|n\|^2 + (k+1) \sum_{i=1}^{l} (\tau_i - \tau_{i-1})\tau_i$$

$$\leqslant (k+1)^2 s^2 \|n\| + (k+1)\|n\|^2 + (k+1) \sum_{j=0}^{\|n\|} (j+1-j)j$$

$$\leqslant O\big((k+1)^2 s^2 \|n\| + (k+1)\|n\|^2\big). \qquad \blacksquare \qquad (8.6)$$

Note that the second term in the cost complexity expression (8.3) accounts for the costs arising from all invocations of First–Components, whereas the first term accounts for all the other costs. Generally speaking, if a large step s_i is required by Inverse–Components, then s is large and the second term in (8.3) dominates; whereas if all step sizes s_i are small, then the first term dominates.

EXAMPLE 8.2. When $s_i = 1$ for all i, the second term in (8.3) becomes $O((k+1)^2\|n\| p^3)$ and so the complexity of the algorithm becomes $O((k+1)\|n\|^2 p^3)$. At the other extreme, when all points with the possible exception of the last along the computational path are singular, that is, $s = s_0 = \max(n_j + 1)$ and $(k+1)s \geqslant \|n\|$, then the second term in (8.3) becomes $O(\|n\|^3 p^3)$, which corresponds to the cost of Gaussian elimination of H_n. The first term in (8.3) becomes irrelevant here; indeed, the solution returned by the algorithm is exactly that obtained by the first invocation of First–Components in step (IC-1).

EXAMPLE 8.3. When $k = 1$ and $p = 1$, it can be shown that the matrices appearing in the First–Components algorithm are always triangular; hence the cost of steps (IC-1) and (IC-5) is reduced to $2\eta_0^2$ and $2\eta_i^2$, respectively. The corresponding total cost of determining the inverse components of H_n is then bounded by $O(\|n\|^2)$. This is the case regardless of any assumptions on the size of the steps from one nonsingular node to the next.

When $n_0 = \cdots = n_k$, Example 8.2 shows that the complexity of Inverse–Components when the matrix is regular (i.e., $s_i = 1$ for all i) is

$O((k+1)pN^2)$, where $N = (k+1)n_0 p$ is the size of the block Hankel-like matrix. This agrees with the results of Kailath et al. [21] under the same assumptions. In the nonregular case, however, their algorithm breaks down, and so a method such as Gaussian elimination, with a cost of $O(N^3)$ operations, is required. With the use of Inverse–Components, even the existence of only one nonsingular principal submatrix can result in significant speedup.

9. FURTHER RESEARCH DIRECTIONS

There are still a number of directions for research in this area. Our formulae are subsequent algorithm are based on the use of exact arithmetic. However, no consideration is given to any coefficient growth in the numerical domain, such as would occur if the block matrices had their coefficients from the field of rational numbers. There needs to be a study of fraction-free methods such as that of Bareiss [4] that prevents inefficiency due to unnecessary coefficient growth.

In addition, our work also does not consider the case of floating-point arithmetic. In this case a major problem of any algorithm is to be efficient while at the same time returning numerically stable results. As suggested by Antoulas, a possible direction could be to avoid those subproblems that are based on unstable subsystems. In our case this would involve not necessarily taking the first nonsingular submatrix $H_{(s,\dots,s,n_{r+1},\dots,n_k)}$, but rather the first nonsingular submatrix that is also well conditioned. Note that by Theorem 6.1, such an approach will still lead to a recurrence relation. We refer the reader to the work of Meleshko [32] for details of the case when $k = p = 1$.

The complexity of the inverse-components algorithm given in Section 7 is $O(p^3km^2)$, with the possibility that it could reach as high as $O(p^3km^3)$ in pathological cases. However, in the case when $k = p = 1$, the algorithm requires at most $O(m^2)$ operations regardless of the singularities encountered in intermediate steps. The scalar algorithm of Heinig [19] has the added advantage in the $p = 1$ case that it is always $O(km^2)$, even for $k > 1$. It would be of interest to alter our algorithm so that this can also be true in the block case.

In the $k = p = 1$ case there are a number of algorithms (cf. Brent, Gustavson, and Yun [7] or Cabay and Choi [10]) which compute the inverse components in $O(m \log^2 m)$ steps, at least in those cases where fast polynomial arithmetic is possible. We conjecture that it is possible to convert the inverse component algorithm to one having complexity $O(p^3k^2m \log^2 m)$ operations. We conjecture that such an algorithm is possible, based on computing the components using quadratic steps as done in [10].

Lerer and Tismenetsky also give inverse characterizations and formulae for a wider class of matrices than (1.1), namely those matrices having a block displacement rank of k. It would be interesting to apply the ideas from this paper to the inverse problem for these general matrices. We also conjecture that it would be straightforward to apply our techniques to matrices of the form (1.1) whose components are allowed to be rectangular $p \times q$ matrices. This has already been done in the $k = 1$ case of rectangular-block Hankel matrices (cf. Labahn [26]). More generally, Heinig and Amdebrhan [18] provide inversion characterization and formulae in the case of mosaic Hankel and Toeplitz matrices. Again it would be interesting to apply the ideas from this paper to the inverse problem for mosaic matrices. Such a generalization would apply to both matrices of the form (1.1) along with rectangular-block Hankel matrices.

REFERENCES

1 H. Akaike, Block Toeplitz matrix inversion, *SIAM J. Appl. Math.* 24:234–241 (1973).

2 G. S. Ammar and W. B. Gragg, Superfast solution of real positive definite Toeplitz systems, *SIAM J. Matrix Anal. Appl.* 9(1):61–76 (1988).

3 A. C. Antoulas, On recursiveness and related topics in linear systems, *IEEE Trans. Automat. Control* AC-31(12):1121–1135 (1986).

4 E. H. Bareiss, Sylvester's identity and multistep integer-preserving Gaussian elimination, *Math. Comp.* 22(103):565–578 (1968).

5 E. R. Berlekamp, *Algebraic Coding Theory*, McGraw-Hill, New York, 1968.

6 R. R. Bitmead and B. D. O. Anderson, Asymptotically fast solutions of Toeplitz and related systems of linear equations, *Linear Algebra Appl.* 34:103–116 (1980).

7 R. Brent, F. G. Gustavson, and D. Y. Y. Yun, Fast solution of Toeplitz systems of equations and computation of Padé approximants, *J. Algorithms* 1:259–295 (1980).

8 A. Bultheel and M. Van Barel, A Matrix Euclidean Algorithm and Matrix Padé Approximation, Tech. Report, Dept. of Computer Science, Katholieke Univ. Leuven, Jan. 1988.

9 S. Cabay and T. T. C. Kao, The Diagonal Padé Table and the Triangular Decomposition of Hankel Matrices, Tech. Report TR83-10, Dept. of Computing Science, Univ. of Alberta, 1983.

10 S. Cabay and D. K. Choi, Algebraic computations of scaled Padé fractions, *SIAM J. Comput.* Feb. 1986, pp. 243–270.

11 S. Cabay, G. Labahn, and B. Beckermann, On the theory and computation of non-perfect Padé-Hermite approximants, *J. Comput. Appl. Math.*, to appear.

12 S. Cabay and G. Labahn, Simultaneous Padé systems, *J. Numer. Algorithms*, submitted for publication.

13 K. O. Geddes, S. R. Czapor, and G. Labahn, *Algorithms for Computer Algebra*, Kluwer, Boston, 1991.

14 W. B. Gragg, The Padé table and its relation to certain algorithms of numerical analysis, *SIAM Rev.* 14:1–61 (1972).
15 W. B. Gragg and A. Lindquist, On the partial realization problem, *Linear Algebra Appl.* 50:277–319 (1983).
16 P. Graves-Morris, Vector-valued rational interpolants II, *IMA J. Numer. Anal.* 4:209–224 (1984).
17 G. Heinig and K. Rost, *Algebraic Methods for Toeplitz-like Matrices and Operators*, Birkhäuser, Basel, 1984.
18 G. Heinig and T. Amdebrhan, On the inverses of Hankel and Toeplitz mosaic matrices, in *Seminar on the Analysis of Operator Equations and Numerical Analysis 1987 / 1988*, Karl-Weierstrass-Institut für Mathematik, 1988, pp. 53–65.
19 G. Heinig, Structure theory and fast inversion of Hankel striped matrices, *Integral Equations Operator Theory* 11:205–229 (1988).
20 F. de Hoog, A new algorithm for solving Toeplitz systems of equations, *Linear Algebra Appl.* 88:123–138 (1987).
21 T. Kailath, S.-Y. Kung, and M. Morf, Displacement ranks of matrices and linear equations, *J. Math. Anal. Appl.* 68:395–407 (1979).
22 R. E. Kalman, On minimal partial realizations of a linear input-output map, in *Aspects of Network and Systems Theory* (R. E. Kalman and N. de Clairs, Eds.), Holt, Reinhart and Winston, 1971, pp. 385–407.
23 G. Labahn, *Matrix Padé Approximants*, M.Sc. Thesis, Dept. of Computing Science, Univ. of Alberta, Edmonton, Canada, 1986.
24 G. Labahn and S. Cabay, Matrix Padé fractions and their computation, *SIAM J. Comput.* 18(4):639–657 (1989).
25 G. Labahn, D. K. Choi, and S. Cabay, Inverses of block Hankel and block Toeplitz matrices, *SIAM J. Comput.* 19(1):98–123 (1990).
26 G. Labahn, Inversion Algorithms for Rectangular-Block Hankel Matrices, Tech. Report CS-90-52, Dept. of Computing Science, Univ. of Waterloo, 1990.
27 L. Lerer and M. Tismenetsky, Generalized Bezoutians and the inversion problem for block matrices, *Integral Equations and Operator Theory* 9:790–819 (1986).
28 N. Levinson, The Wiener RMS (root mean square) error in filter design, *J. Math. and Phys.* 25:261–278 (1947).
29 K. Mahler, Perfect systems, *Compositio Math.* 19:95–166 (1968).
30 J. L. Massey, Shift-register synthesis and BCH decoding, *IEEE Trans. Inform. Theory* IT-15:122–127 (1969).
31 R. J. McEliece and J. B. Shearer, A property of Euclid's algorithm and an application to Padé approximation, *SIAM J. Appl. Math.* 34:611–617 (1978).
32 R. J. Meleshko, A Stable Algorithm for the Computation of Padé Approximants, Ph.D. Thesis, Dept. of Computing Science, Univ. of Alberta, Edmonton, Canada, 1990.
33 J. Rissanen, Algorithms for triangular decomposition of block Hankel and Toeplitz matrices with application to factoring positive matrix polynomials, *Math. Comp.* 27:147–154 (1973).
34 J. Rissanen, Solution of linear equations with Hankel and Toeplitz matrices, *Numer. Math.* 22:361–366 (1974).

35 A. A. Semencul, Inversion of finite sections of paired and transposed paired operators (in Russian) *Mat. Issled* 3(1):100–107 (1968).
36 Y. Sugiyama, An algorithm for solving discrete-time Wiener-Hopf equations based on Euclid's algorithm, *IEEE Trans. Inform. Theory* 32:394–409 (1986).
37 W. F. Trench, An algorithm for the inversion of finite Hankel matrices, *SIAM J. Appl. Math.* 13:1102–1107 (1965).
38 G. A. Watson, An algorithm for the inversion of block matrices of Toeplitz form, *J. Assoc. Comput. Mach.* 20:409–415 (1973).

Received 29 January 1991; final manuscript accepted 14 October 1991