A superfast algorithm for multi-dimensional Padé systems

Stan Cabay [1]

Department of Computing Science, University of Alberta, Edmonton, Alberta, Canada T6G 2H1

George Labahn [2]

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

Communicated by C. Brezinski
Received 12 September 1991; revised 27 December 1991

For a vector of $k+1$ matrix power series, a superfast algorithm is given for the computation of multi-dimensional Padé systems. The algorithm provides a method for obtaining matrix Padé, matrix Hermite Padé and matrix simultaneous Padé approximants. When the matrix power series is normal or perfect, the algorithm is shown to calculate multi-dimensional matrix Padé systems of type $(n_0, \ldots, n_k)$ in $O(\|n\| \cdot \log^2 \|n\|)$ block-matrix operations, where $\|n\| = n_0 + \cdots + n_k$. When $k=1$ and the power series is scalar, this is the same complexity as that of other superfast algorithms for computing Padé systems. When $k>1$, the fastest methods presently compute these matrix Padé approximants with a complexity of $O(\|n\|^2)$. The algorithm succeeds also in the non-normal and non-perfect case, but with a possibility of an increase in the cost complexity.

Keywords: Matrix Padé approximants, simultaneous Padé approximants, Hermite Padé approximants, rational approximation.

1. Introduction

Given a vector of $k+1$ power series

$$A_i(z) = \sum_{j=0}^{\infty} a_{i,j} z^j, \quad i = 0, 1, \ldots, k, \tag{1.1}$$

with coefficients from a field $\mathscr{F}$, a Hermite Padé approximant of type $n =$

[1] Supported in part by NSERC grant No. A8035.
[2] Partially supported by NSERC operating grant No. 6194.

$(n_0, \ldots, n_k)$ is a set of $k+1$ polynomials $P_i(z)$ having degrees bounded by the $n_i - 1$ and satisfying

$$A_0(z) \cdot P_0(z) + \cdots + A_k(z) \cdot P_k(z) = O(z^{\|n\|-1}), \tag{1.2}$$

where $\|n\| = n_0 + \cdots + n_k$. A simultaneous Padé approximant of type n is a set of $k+1$ polynomials $P_i(z)$ having degrees bounded by the $\|n\| - n_i$ and satisfying

$$P_0(z) \cdot A_i(z) - P_i(z) \cdot A_0(z) = O(z^{\|n\|+1}), \quad i = 1, \ldots, k. \tag{1.3}$$

(In the latter it is usually assumed that $A_0(z) = 1$.) These approximants are also known as type I and type II polynomial approximations, respectively. When the coefficients of the $A_i(z)$ and $P_i(z)$ come from the ring of $p \times p$ matrices over $\mathscr{F}$, rather than $\mathscr{F}$ itself, then we obtain matrix Hermite Padé and matrix simultaneous Padé approximants, respectively, of type n. When $k = 1$, these coincide with right and left matrix Padé approximants of a matrix of power series (cf., Labahn and Cabay [19]).

In the scalar case, both types of approximants originated with Hermite [15,16] and Padé [26]. Simultaneous Padé approximants were used extensively by Hermite when he proved the transcendence of e. Both types of approximants have been widely studied and include many classical approximation problems (e.g., Padé approximants [13], integral approximants [18], directed vector approximants [14] and G^3J approximants [1]). The general definition of both types of approximants, along with an extensive study of their properties is originally due to Mahler [23], who also noted strong relationships between the two types of approximants [24]. Additional properties and relationships can also be found in [7,10,11,17,20].

There exist a number of fast algorithms for computing these approximants. In the scalar case Della Dora and Dicrescenzo [12] and Paszkowski [29] present algorithms that compute a Hermite Padé approximant of type n in $O(\|n\|^2)$ operations. However, they require the input vector of power series to have the property of being *perfect* (also called *normal*), a strong restriction. The algorithm of Cabay et al. [10] also computes a Hermite Padé approximant of type n in $O(\|n\|^2)$ operations. Their method has the advantage that it also succeeds in the non-perfect case, but with a potential increase in complexity. There are no such problems with the algorithms of Beckermann [3] and Van Barel and Bultheel [2], which are of complexity of $O(\|n\|^2)$ even in the non-perfect case. In terms of cost complexity, their algorithms are at present best, since they have the additional advantage that there is no increase in complexity in the non-perfect case. In the case of simultaneous Padé approximants, a fast $O(\|n\|^2)$ algorithm that requires no restrictions on the input vector of power series has been given by De Bruin [6] and Beckermann [3].

In this paper, we present superfast algorithms for computing both approximants of a given type n. These algorithms can be applied to any vector of power series; the requirement of being perfect is not needed. In addition, the algorithms also can be applied when the coefficients of the power series are square matrices rather than just scalars. When fast polynomial arithmetic is possible, either approximant can be computed with $O(\|n\| \cdot \log^2 \|n\|)$ block-matrix operations in the case of perfect power series. The algorithms can also compute these matrix Padé approximants in the non-perfect case, with only a slight increase in complexity. We note that there are pathological cases for which the algorithm requires up to $O(\|n\|^3)$ block-matrix operations to compute the approximants. This is not the case with the scalar algorithms of Beckermann and Van Barel and Bultheel. However, when $k = 1$ and $p = 1$, (i.e. the scalar case of Padé approximation) the complexity is always $O(\|n\| \cdot \log^2 \|n\|)$ operations. This is the same complexity as that of other superfast scalar Padé algorithms such as those given by Brent et al. [5], Cabay and Choi [9] and Sugiyama [30].

Our methods actually provide superfast methods for computing matrix Hermite Padé and matrix simultaneous Padé systems, rather than just approximants. These Padé systems, introduced in [10] and [21], consist of the desired Padé approximants along with additional weaker type of Padé approximants. Many applications require the entire Padé system rather than simply the Padé approximant. Thus, for example, our results combined with [20] and [21] provide superfast algorithms for computing inverses of block Toeplitz and block Toeplitz-like matrices along with their block Hankel counterparts.

2. Matrix Hermite Padé systems

In this section we discuss the notion of a matrix Hermite Padé system for a vector of matrix power series. This is a natural generalization of the corresponding scalar definitions given in Cabay et al. [10].

Let

$$A(z) = [A_0(z) \mid A_1(z), \ldots, A_k(z)] = [B(z) \mid C(z)] \tag{2.1}$$

be a $1 \times (k+1)$ vector of $p \times p$ matrix power series with $\det(A_0(0)) \neq 0$. Let $n = (n_0, \ldots, n_k)$ be a vector of nonnegative integers and

$$S(z) = \left[\begin{array}{c|ccc} S_{0,0}(z) & S_{0,1}(z) & \cdots & S_{0,k}(z) \\ \hline S_{1,0}(z) & S_{1,1}(z) & \cdots & S_{1,k}(z) \\ \vdots & \vdots & & \vdots \\ S_{k,0}(z) & S_{k,1}(z) & \cdots & S_{k,k}(z) \end{array}\right] = \left[\begin{array}{c|c} z^2P(z) & U(z) \\ \hline z^2Q(z) & V(z) \end{array}\right], \tag{2.2}$$

a $p(k+1) \times p(k+1)$ polynomial matrix with each $S_{ij}(z)$ a $p \times p$ matrix polynomial. In (2.2) the constant and linear terms in $S_{i,0}(z)$, $i = 0, \ldots, k$, are zero and component-wise

$$\operatorname{degree}(S(z)) \leqslant \left[\begin{array}{c|ccc} n_0+1 & n_0 & \cdots & n_0 \\ \hline n_1+1 & n_1 & \cdots & n_1 \\ \vdots & \vdots & & \vdots \\ n_k+1 & n_k & \cdots & n_k \end{array}\right]. \tag{2.3}$$

DEFINITION 2.1

Let $n = (n_0, \ldots, n_k)$ be a vector of nonnegative integers with $n_i > 0$ for at least one n_i. The polynomial matrix $S(z)$ is called a *Matrix Hermite Padé System (MHPS)* of type n for $A(z)$ if

(I) $S(z)$ satisfies the degree bounds (2.3);

(II) $A(z) \cdot S(z) = z^{\|n\|+1} \cdot \hat{A}(z)$, where $\hat{A}(z)$ is a $1 \times (k+1)$ vector of $p \times p$ matrix power series;

(III) $V(0)$ and $\hat{B}(0)$ are nonsingular matrices, where $\hat{B}(z)$ is the $p \times p$ matrix determined by partitioning $\hat{A}(z)$ as

$$\hat{A}(z) = \left[\hat{B}(z) \mid \hat{C}(z)\right]. \tag{2.4}$$

A MHPS $S(z)$ is said to be *normalized* if in condition (III) we have $V(0) = I_{pk}$ and $\hat{B}(0) = I_p$. □

In the scalar (i.e. $p = 1$) case we refer to the above as simply a Hermite Padé System (HPS). For the sake of simplicity and without loss of generality, it is assumed in the remainder of this presentation that $B(0) = I_p$ and that the components of n are ordered so that $n_1 \geqslant \cdots \geqslant n_a > n_{a+1} = \cdots = n_k = 0$, where $0 \leqslant a \leqslant k$.

Remark 1

If $B(z) = I_p$ and $C(z) = [A_1(z), \ldots, A_k(z)]$, then

$$C(z) \cdot V(z) + U(z) = z^{\|n\|+1} \cdot \hat{C}(z) \tag{2.5}$$

and

$$C(z) \cdot Q(z) + P(z) = z^{\|n\|-1} \cdot \hat{B}(z). \tag{2.6}$$

When $k = 1$, the pair $(U(z), V(z))$ generates a right matrix Padé fraction for $C(z)$ of type (n_0, n_1) while the pair $(P(z), Q(z))$ generates a right matrix Padé form of type $(n_0 - 1, n_1 - 1)$ for $C(z)$ (cf., Labahn and Cabay [19]). □

Remark 2

The definition of a MHPS is the natural extension of the concept of a Hermite Padé system, introduced in Cabay et al. [10]. In general, the first block column of a MHPS defines a matrix Hermine Padé form for $A(z)$ of type n, while block columns 2 to $k+1$ define a "weak" matrix Hermite Padé fraction for $A(z)$ of type n (cf., Labahn [21]). □

Let

$$D(z) = B^{-1}(z) \cdot C(z) \tag{2.8}$$

be a $1 \times k$ vector of $p \times p$ matrix power series and define

$$H_n = \left[\begin{array}{ccc|c|ccc} d_{n_0-n_1+1,1} & \cdots & d_{n_0,1} & & d_{n_0-n_k+1,k} & \cdots & d_{n_0,k} \\ d_{n_0-n_1+2,1} & & d_{n_0+1,1} & & d_{n_0-n_k+2,k} & & d_{n_0+1,k} \\ \vdots & & \vdots & \cdots & \vdots & & \vdots \\ d_{\|n\|-n_1,1} & \cdots & d_{\|n\|-1,1} & & d_{\|n\|-n_k,k} & \cdots & d_{\|n\|-1,k} \end{array}\right], \tag{2.9}$$

where $d_{i,j}$ is the coefficient of z^i in the jth component $D_j(z)$. Then a MHPS of type n can be obtained by solving a set of linear equations with H_n as the coefficient matrix.

The component $Q(z)$ of $S(z)$ in (2.2) satisfying $\hat{B}(0) = I_p$ corresponds to the block solution X of

$$H_n \cdot X = E_n, \tag{2.10}$$

where E_n is the unit column vector of length $\|n\| - n_0$ with a single I_p in the last block row. That is, if X is partitioned as

$$X = \left[x_{n_1-1,1}, \ldots, x_{0,1} \mid \cdots \mid x_{n_k-1,k}, \ldots, x_{0,k}\right]^t, \tag{2.11}$$

where each component $x_{i,j}$ is a $p \times p$ matrix, then the jth component $Q_j(z)$ of $Q(z)$ is given by

$$Q_j(z) = \sum_{i=0}^{n_j-1} x_{i,j} \cdot z^i. \tag{2.12}$$

The remaining components of the first column of $S(z)$ are then given by

$$P(z) = -D(z) \cdot Q(z) \mod z^{\|n\|-1}. \tag{2.13}$$

Similarly, the components $U(z)$ and $V(z)$ (with $V(0) = I_{pk}$) of $S(z)$ in (2.2) can be obtained from the block solution Y of

$$H_n \cdot Y = -\begin{bmatrix} d_{n_0+1,1} & d_{n_0+1,2} & \cdots & d_{n_0+1,k} \\ d_{n_0+2,1} & d_{n_0+2,2} & \cdots & d_{n_0+2,k} \\ \vdots & \vdots & & \vdots \\ d_{\|n\|,1} & d_{\|n\|,2} & \cdots & d_{\|n\|,k} \end{bmatrix}. \tag{2.14}$$

Note that (2.14) is valid even in those cases where $n_i = 0$. (In the special case when $n_j = 0$ for $1 \leqslant j \leqslant k$, the matrix H_n is null and we simply set $V(z) = I_{pk}$.) The component $U(z)$ of $S(z)$ is then given by

$$U(z) = -D(z) \cdot V(z) \mod z^{\|n\|+1}. \tag{2.15}$$

Clearly, when H_n is nonsingular, solutions of (2.10) and (2.14) are possible. It is then easy to show that (2.10), (2.11), (2.14) and (2.15) provide for the existence of a normalized MHPS of type n. Theorem 2.2 states that this is both a necessary and sufficient condition for existence.

THEOREM 2.2

A MHPS of type n exists if and only if $\det(H_n) \neq 0$.

Proof

A proof of this result is a straightforward generalization of that given for the scalar case in Cabay et al. [10]. The result also follows in a natural way from the work of Lerer and Tismenetsky [22]. □

Note that theorem 2.2 includes the case of $n_i = 0$, $1 \leqslant i \leqslant k$, by setting $\det(H_n) = 1$ when H_n is the null matrix. From (2.10) and (2.15), the normalized MHPS of type $(n_0, 0, \ldots, 0)$ is determined here to be

$$S(z) = \left[\begin{array}{c|c} z^{n_0+1} I_p & U(z) \\ \hline 0 & I_{pk} \end{array}\right], \tag{2.16}$$

where $U(z) = -D(z) \bmod z^{n_0+1}$. In section 5, for algorithmic purposes, we adopt (2.16) even in the case when $n_0 = 0$, despite the fact that it does not meet all the requirements set forth in (2.2).

EXAMPLE 2.3

Let

$$A(z) = [1,\ -1 + z^2 + z^5 - z^6 + z^7 + z^8 + z^9 + \cdots,$$
$$-z - z^2 + 2z^6 - z^7 + z^8 + 2z^9 + \cdots],$$

with $n = (2, 3, 1)$. Then

$$D(z) = [-1 + z^2 + z^5 - z^6 + z^7 + z^8 + z^9 + \cdots,$$
$$-z - z^2 + 2z^6 - z^7 + z^8 + 2z^9 + \cdots],$$

and the corresponding matrix H_n is nonsingular. By theorem 2.2, a HPS of type n exists. Using eqs. (2.10)–(2.13), the normalized HPS of type n is given by

$$S(z) = \begin{bmatrix} z^2 + z^3 & 1 + z + z^2 & -z - z^2 \\ z^2 & 1 + z - z^3 & -2z \\ z^2 & 2z & 1 - 2z \end{bmatrix},$$

with the first few terms of the residual being

$$\hat{A}(z) = [1 + z + \cdots, 4 - z + 5z^2 + \cdots, -3 + z - 2z^2 + \cdots]. \quad \square$$

3. Matrix simultaneous Padé systems

In this section we give the dual concept of simultaneous Padé systems. These correspond to Hermite Padé systems except with alternate degrees restrictions and with matrix multiplication on the left rather than the right. We show that results parallel to those of the previous section can also be given for these systems.

Let

$$A(z)=\left[\begin{array}{ccc} A_{0,1}(z) & \cdots & A_{0,k}(z) \\ \hline A_{1,1}(z) & \cdots & A_{1,k}(z) \\ \vdots & & \vdots \\ A_{k,1}(z) & \cdots & A_{k,k}(z) \end{array}\right]=\left[\begin{array}{c} B(z) \\ \hline C(z) \end{array}\right] \tag{3.1}$$

be a $(k+1)\times k$ matrix of $p\times p$ matrix power series with $\det(C(0))\neq 0$. Let $n=(n_0,\ldots,n_k)$ be a vector of nonnegative integers and define

$$S(z)=\left[\begin{array}{c|ccc} S_{0,0}(z) & S_{0,1}(z) & \cdots & S_{0,k}(z) \\ \hline S_{1,0}(z) & S_{1,1}(z) & \cdots & S_{1,k}(z) \\ \vdots & \vdots & & \vdots \\ S_{k,0}(z) & S_{k,1}(z) & \cdots & S_{k,k}(z) \end{array}\right]=\left[\begin{array}{c|c} V(z) & U(z) \\ \hline z^2\cdot Q(z) & z^2\cdot P(z) \end{array}\right], \tag{3.2}$$

a $p(k+1)\times p(k+1)$ polynomial matrix with each $S_{i,j}(z)$ a $p\times p$ polynomial matrix. In (3.2) the constant and linear terms in $S_{i,j}(z)$, $i=1,\ldots,k$, $j=0,\ldots,k$ are zero, and component-wise

$$\operatorname{degree}(S(z))\leqslant \|n\|-\left[\begin{array}{c|ccc} n_0 & n_1 & \cdots & n_k \\ \hline n_0-1 & n_1-1 & \cdots & n_k-1 \\ \vdots & \vdots & & \vdots \\ n_0-1 & n_1-1 & \cdots & n_k-1 \end{array}\right]. \tag{3.3}$$

DEFINITION 3.1

Let $n=(n_0,\ldots,n_k)$ be a vector of nonnegative integers. The polynomial matrix S is called a *Matrix Simultaneous Padé System (MSPS)* of type n for $A(z)$ if

(I) $S(z)$ satisfies the degree bounds (3.3);

(II) $S(z)\cdot A(z)=z^{\|n\|+1}\cdot\hat{A}(z)$, where $\hat{A}(z)$ is a $(k+1)\times k$ matrix of $p\times p$ matrix power series; and

(III) $\det(S_{0,0}(0))\neq 0$ and $\det(\hat{C}(0))\neq 0$, where $\hat{C}(z)$ is the $k\times k$ matrix determined by partitioning $\hat{A}(z)$ as

$$\hat{A}(z)=\left[\begin{array}{c} \hat{B}(z) \\ \hline \hat{C}(z) \end{array}\right]. \tag{3.4}$$

A MSPS $S(z)$ is said to be *normalized* if in condition (III) we have $S_{0,0}(0) = I_p$ and $\hat{C}(0) = I_{pk}$. □

As in the previous section, a SPS is the notation used to denote the scalar $(p = 1)$ case. For the sake of simplicity and without loss of generality, it is assumed in the remainder of this presentation that $C(0) = I_{pk}$ and that the components of n are ordered so that $n_1 \geqslant \cdots \geqslant n_a > n_{a+1} = \cdots = n_k = 0$ where $0 \leqslant a \leqslant k$.

Remark 1

When $C(z) = I_k$, we obtain the equations

$$V(z) \cdot B(z) + U(z) = z^{\|n\|+1} \cdot \hat{B}(z) \tag{3.5}$$

and

$$Q(z) \cdot B(z) + P(z) = z^{\|n\|-1} \cdot \hat{C}(z). \tag{3.6}$$

When $k = 1$, the pair $(U(z), V(z))$ generates a left matrix Padé fraction for $B(z)$ of type (n_0, n_1) while the pair $(P(z), Q(z))$ generates a left matrix Padé form of type $(n_0 - 1, n_1 - 1)$ for $B(z)$ (cf., Labahn and Cabay [19]). □

Remark 2

The definition of a MSPS is the natural dual to MHPS of section 2. In this case we have that the first row of a MSPS defines a matrix simultaneous Padé fraction of type n for $B(z) \cdot C(z)^{-1}$, whereas block rows 2 to $k+1$ define a "weak" matrix simultaneous Padé form [21] for $B(z) \cdot C(z)^{-1}$ of type n. □

Let

$$D(z) = B(z) \cdot C^{-1}(z) \tag{3.7}$$

and define H_n via (2.9). Then a MSPS of type n can be obtained by solving a set of linear equations with H_n as the coefficient matrix.

The component $V(z)$ (with $V(0) = I_p$) of $S(z)$ in (3.2) corresponds to the solution X of

$$X \cdot H_n = -\left[d_{\|n\|-n_1+1,1}, \ldots, d_{\|n\|,1} \mid \cdots \mid d_{\|n\|-n_k+1,k}, \ldots, d_{\|n\|,k}\right]. \tag{3.8}$$

(In the special case when $n_j = 0$ for $1 \leqslant j \leqslant k$, the matrix H_n is null and we simply set $V(z) = I_p$.) The component $U(z)$ of $S(z)$ is then given by

$$U(z) = -V(z) \cdot D(z) \mod z^{\|n\|+1}. \tag{3.9}$$

The components $P(z)$ and $Q(z)$ of $S(z)$ in (3.2) can be obtained as follows. For $n_i > 0$, $1 \leqslant i \leqslant a$, the solution Y of

$$Y \cdot H_n = E_{n_1 + \cdots + n_i}, \tag{3.10}$$

where $E_{n_1+\cdots+n_i}$ is the unit block row vector of length $\|n\|-n_0$ with its single I_p in column $n_1+\cdots+n_i$, yields the component $Q_i(z)=z^{-2}*S_{i,0}(z)$. The remaining components of the ith row of $S(z)$ in (3.2) is then given by

$$S_{i,j}(z) = -S_{i,0}(z)\cdot D_j(z) \mod z^{\|n\|-n_j+2} \quad \text{for } j=1,\ldots,k. \tag{3.11}$$

For $n_i=0$, $a<i\leqslant k$, define

$$S_{i,j}(z)=\begin{cases} z^{\|n\|+1}I_p, & j=i,\\ 0, & j\neq i.\end{cases} \tag{3.12}$$

Clearly, when H_n is nonsingular, solutions of (3.8) and (3.10) are possible. It is then easy to show that (3.8)–(3.12) provide for the existence of a normalized MSPS of type n. Theorem 3.2 states that this is both a necessary and sufficient condition for existence.

THEOREM 3.2

A MSPS of type n exists if and only if $\det(H_n)\neq 0$.

Proof

We again refer the reader to the dual argument presented in the scalar case for Hermite Padé systems in Cabay et al. [10]. □

Note that theorem 3.2 includes the case of $n_i=0$, $1\leqslant i\leqslant k$, by setting $\det(H_n)=1$ when H_n is the null matrix. From (3.9) and (3.12), the normalized MSPS of type $(n_0, 0,\ldots,0)$ is determined here to be

$$S(z)=\left[\begin{array}{c|c} I_p & U(z)\\ \hline 0 & z^{n_0+1}I_{pk}\end{array}\right], \tag{3.13}$$

where $U(z)=-D(z) \mod z^{n_0+1}$. As for the MHPS case, we use (3.13) even in the case that $n_0=0$ despite the fact that it does not strictly satisfy the requirements of (3.2).

EXAMPLE 3.3

Let

$$A(z)=\begin{bmatrix} -1+z^2+z^5-z^6+z^7+z^8+z^9+\cdots & -z-z^2+2z^6-z^7+z^8+2z^9+\cdots\\ 1 & 0\\ 0 & 1\end{bmatrix}.$$

Then with $D(z)$ as in (3.7) and $n=(2, 3, 1)$, the matrix H_n is nonsingular, so a SPS of type n exists. Using eqs. (3.9)–(3.13), the normalized SPS of type n is given by

$$S(z)=\begin{bmatrix} 1-z+2z^2-z^3+2z^4 & 1-z+z^2 & z+z^3+z^4+z^5\\ z^2 & z^2-z^4 & z^3+z^4\\ z^2-z^3-z^5 & z^2-z^3-z^4 & z^3-z^5-z^6\end{bmatrix},$$

with the first few terms of the residual being

$$\hat{A}(z)=\begin{bmatrix} 4-3z+5z^2+\cdots & -3+6z-3z^2+\cdots \\ 1-z+z^2+\cdots & 2z-z^2+\cdots \\ -2z+2z^2+\cdots & 1+2z-3z^2+\cdots \end{bmatrix}. \quad \square$$

4. A recurrence relation for multi-dimensional Padé systems

Given a vector of matrix power series (2.1) and a vector of integers n, a corresponding MHPS can be determined via a method such as Gaussian elimination at a cost of

$$O(\|n\|^3+k\cdot\|n\|^2) \tag{4.1}$$

block operations. Similarly, given a matrix of matrix power series (3.1), a corresponding MSPS can be computed at a cost of

$$O(\|n\|^3+k^2\cdot\|n\|^2) \tag{4.2}$$

block operations. In both cases there is the advantage that there need be no restrictions on the input matrix of power series. However, in both cases such calculations do not take into account the special structure of the coefficient matrix of these systems. The goal of this section is to describe a recurrence relation that will lead to an efficient algorithm for the determination of a MHPS or MSPS of any type. The resulting algorithm takes advantage of the special structure of the coefficient matrix (2.9), without placing any additional restrictions on the input.

Consider first the case of computing a MPHS. Following [10] we let

$$N=1+\min\{n_0,\, n_1\} \tag{4.3}$$

and define integer vectors $n^{(i)}=(n_0^{(i)},\ldots,n_k^{(i)})$ for $1\leqslant i\leqslant N$ by

$$n_j^{(i)}=\max\{0,\, n_j-N+i\} \quad \text{for } j=0,\ldots,k. \tag{4.4}$$

Then the sequence $\{n^{(i)}\}_{i=1,\ldots,N}$ lies on a piecewise linear path with $n_j^{(i+1)}\geqslant n_j^{(i)}$ for each i, j with

$$n^{(1)}=\begin{cases}(0,\, n_1-n_0,\,\ldots), & n_1\geqslant n_0,\\ (n_0-n_1,\, 0,\ldots,0), & n_1<n_0,\end{cases} \tag{4.5}$$

and $n^{(N)}=n$. Also define $n^{(0)}=-e_0=-(1,0,\ldots,0)$.

Let $\sigma_0=0$ and define

$$\sigma_i=\min\{\sigma>\sigma_{i-1}\colon \det(H_{n^{(\sigma)}})\neq 0\},\quad \text{for } i\geqslant 1. \tag{4.6}$$

Then the sequence $\{n^{(i)}\}$ determines a subsequence $\{m^{(i)}\}$ of *nonsingular points* $m^{(i)}=n^{(\sigma_i)}$.

For $i = 0$, let $S^{(0)}(z) = I_{p(k+1)}$, and for $i \geqslant 1$ let $S^{(i)}(z)$ be the uniquely determined normalized MHPS of type $m^{(i)}$ for $A(z)$ with $A^{(i)}(z)$ the corresponding residual. This gives

$$A(z) \cdot S^{(i)}(z) = z^{\|m^{(i)}\|+1} \cdot A^{(i)}(z), \tag{4.7}$$

with

$$B^{(i)}(0) = I_p \quad \text{and} \quad V^{(i)}(0) = I_{pk}, \tag{4.8}$$

where $A^{(i)}(z)$ is partitioned into $B^{(i)}(z)$ and $C^{(i)}(z)$ as in (2.4). Note that (4.6) and (4.7) also hold when $i = 0$.

The algorithm described in the next section for constructing a MHPS of type n requires the successive computation of $S^{(i+1)}(z)$ given $S^{(i)}(z)$. Theorem 4.1 gives a mechanism for doing this efficiently.

THEOREM 4.1

For $i \geqslant 0$, $\sigma > \sigma_i$, let $\nu^{(\sigma)} = n^{(\sigma)} - m^{(i)} - e_0$. Then $n^{(\sigma)}$ is a nonsingular point for $A(z)$ if and only if $\nu^{(\sigma)}$ is a nonsingular point for $A^{(i)}(z)$. Furthermore, we have the recurrence relation

$$S^{(i+1)}(z) = S^{(i)}(z) \cdot \hat{S}(z) \quad \text{and} \quad A^{(i+1)}(z) = \hat{A}(z), \tag{4.9}$$

where $\hat{S}(z)$ is the MHPS of type $m^{(i+1)} - m^{(i)} - e_0$ for $A^{(i)}(z)$, and $\hat{A}(z)$ is its residual.

Proof

The proof of theorem 4.1 follows naturally from the scalar version given in [10] and so will not be given here. □

The construction from eqs. (4.3)–(4.8) can also be accomplished for the computation of MSPS. The only difference is the matrix multiplication must appear on the left, rather than the right side.

THEOREM 4.2

For $i \geqslant 0$, $\sigma > \sigma_i$, let $\nu^{(\sigma)} = n^{(\sigma)} - m^{(i)} - e_0$. Then $n^{(\sigma)}$ is a nonsingular point for $A(z)$ if and only if $\nu^{(\sigma)}$ is a nonsingular point for $A^{(i)}(z)$. Furthermore, we have the recurrence relation

$$S^{(i+1)}(z) = \hat{S}(z) \cdot S^{(i)}(z) \quad \text{and} \quad A^{(i+1)}(z) = \hat{A}(z), \tag{4.11}$$

where $\hat{S}(z)$ is the MSPS of type $m^{(i+1)} - m^{(i)} - e_0$ for $A^{(i)}(z)$, and $\hat{A}(z)$ is its residual.

Proof

The proof of theorem 4.2 also follows closely from the arguments used in the scalar Hermite Padé case. Because of the differences in degree definitions and because a proof has not been given elsewhere, we give a more complete proof for the MSPS case.

Let a be such that

$$m_1^{(i)} \geqslant \cdots \geqslant m_a^{(i)} > m_{a+1}^{(i)} = \cdots = m_k^{(i)} = 0. \tag{4.12}$$

Then, element by element, the degrees of $S^{(i)}(z)$ are bounded by

$$\| m^{(i)} \| - \left[\begin{array}{ccc|cccc} m_0^{(i)} & \cdots & m_a^{(i)} & 0 & \cdots & 0 & \\ \hline m_0^{(i)}-1 & \cdots & m_a^{(i)}-1 & -1 & \cdots & -1 & \\ \vdots & & \vdots & \vdots & & \vdots & \\ m_0^{(i)}-1 & \cdots & m_a^{(i)}-1 & -1 & \cdots & -1 & \\ \hline \infty & \cdots & \infty & -1 & \infty & \cdots & \infty \\ \vdots & & \vdots & \infty & -1 & & \vdots \\ & & & \vdots & & \ddots & \infty \\ \infty & \cdots & \infty & \infty & \cdots & \infty & -1 \end{array}\right], \tag{4.13}$$

where the degrees of the zero polynomials in rows $a+1, \ldots, k$ are denoted by ∞. In addition, in rows $a+1, \ldots, k$ of $S^{(i)}(z)$, the only nonzero entries occur in position (j, j), where according to (3.12), $S_{j,j}(z) = z^{\| m^{(i)} \| + 1} I_p$ for $a \leqslant j \leqslant k$.

Now let $\nu^{(\sigma)} = n^{(\sigma)} - m^{(i)} - e_0$ be a nonsingular point for $A^{(i)}(z)$ and let $\hat{S}(z)$ be the MSPS of type $\nu^{(\sigma)}$ or $A^{(i)}(z)$ with residual $\hat{A}(z)$. Since $\sigma > \sigma_i$, then from (4.4) there is an integer $b \geqslant a$ such that

$$n_1^{(\sigma)} \geqslant \cdots \geqslant n_b^{(\sigma)} > n_{b+1}^{(\sigma)} = \cdots = n_k^{(\sigma)} = 0. \tag{4.14}$$

Consequently,

$$\nu^{(\sigma)} = \left(s-1,\, s, \ldots, s,\, \nu_{a+1}^{(\sigma)}, \ldots, \nu_b^{(\sigma)},\, 0, \ldots, 0\right), \tag{4.15}$$

where $s = n_j^{(\sigma)} - m_j^{(i)}$ for $0 \leqslant j \leqslant a$, and $1 \leqslant \nu_j^{(\sigma)} = n_j^{(\sigma)} \leqslant s$ for $a < j \leqslant b$. Thus, the degrees of the elements of $\hat{S}(z)$ are bounded by

$$\| n^{(\sigma)} \| - \| m^{(i)} \| - 1 - \left[\begin{array}{c|ccc|ccc|ccccc} s-1 & s & \cdots & s & n_{a+1}^{(\sigma)} & \cdots & n_b^{(\sigma)} & 0 & \cdots & & & 0 \\ \hline s-2 & s-1 & \cdots & s-1 & n_{a+1}^{(\sigma)}-1 & \cdots & n_b^{(\sigma)}-1 & -1 & \cdots & & & -1 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & & & \vdots \\ s-2 & s-1 & \cdots & s-1 & n_{a+1}^{(\sigma)}-1 & \cdots & n_b^{(\sigma)}-1 & -1 & \cdots & & & -1 \\ \hline \infty & & & & \cdots & & \infty & -1 & \infty & & \cdots & \infty \\ \vdots & & & & & & \vdots & \infty & -1 & & & \vdots \\ & & & & & & & & & \ddots & & \\ \infty & & & & \cdots & & \infty & \vdots & & & -1 & \infty \\ & & & & & & & \infty & \cdots & & \infty & -1 \end{array}\right], \tag{4.16}$$

where ∞ appears only in rows $b+1, \ldots, k$.

Now, set

$$S^*(z) = \hat{S}(z) \cdot S^{(i)}(z). \tag{4.17}$$

Then from (4.13), (4.15) and (4.16), the degrees of the elements of $S^*(z)$ are bounded by

$$\| n^{(\sigma)} \| - \left[\begin{array}{ccc|cccc} n_0^{(\sigma)} & \cdots & n_b^{(\sigma)} & 0 & \cdots & & 0 \\ \hline n_0^{(\sigma)}-1 & \cdots & n_b^{(\sigma)}-1 & -1 & \cdots & & -1 \\ \vdots & & \vdots & \vdots & & & \vdots \\ n_0^{(\sigma)}-1 & \cdots & n_b^{(\sigma)}-1 & -1 & \cdots & & -1 \\ \hline \infty & \cdots & \infty & -1 & \infty & \cdots & \infty \\ & & & \infty & -1 & & \vdots \\ \vdots & & \vdots & & & \ddots & \\ \infty & \cdots & \infty & \vdots & & -1 & \infty \\ & & & \infty & \cdots & \infty & -1 \end{array} \right], \tag{4.18}$$

where ∞ appears only in rows $b+1, \ldots, k$. Thus, $S^*(z)$ satisfies condition (I) for an MSPS of type $n^{(\sigma)}$.

In addition,

$$\begin{aligned} S^*(z) \cdot A(z) &= \hat{S}(z) \cdot S^{(i)}(z) \cdot A(z) = z^{\| m^{(i)} \| + 1} \cdot \hat{S}(z) \cdot A^{(i)}(z) \\ &= z^{\| n^{(\sigma)} \| + 1} \cdot \hat{A}(z) \end{aligned} \tag{4.19}$$

and

$$S_{0,0}^*(0) = \hat{S}_{0,0}(0) \cdot S_{0,0}^{(i)}(0) + \sum_{j=1}^{k} \hat{S}_{0,j}(0) \cdot S_{j,0}^{(i)}(0) = \hat{S}_{0,0}(0) \cdot S_{0,0}^{(i)}(0) = 1. \tag{4.20}$$

Hence, $S^*(z)$ is a MSPS of type $n^{(\sigma)}$ for $A(z)$. We have also shown that if $\nu^{(\sigma)}$ is a nonsingular point for $A^{(i)}(z)$, then $n^{(\sigma)}$ is a nonsingular point for $A(z)$.

The proof of the converse, that $\nu^{(\sigma)}$ is a nonsingular point for $A^{(i)}(z)$ if $n^{(\sigma)}$ is a nonsingular point for $A(z)$, follows a similar argument (cf. [10] for a parallel argument in the case of Padé–Hermite systems).

Thus, the smallest σ for which $\nu^{(\sigma)} = n^{(\sigma)} - m^{(i)} - e_0$ is a nonsingular point for $A^{(i)}(z)$ is $\sigma = \sigma_{i+1}$. This yields the MSPS $S^*(z)$ of type $m^{(i+1)} = n^{(\sigma_{i+1})}$ for $A(z)$. Thus $S^{(i+1)}(z) = S^*(z)$ and by eq. (4.19) $A^{(i+1)}(z) = \hat{A}(z)$. □

Remark 1

Theorems 4.1 and 4.2 reduce the problem of determining a MHPS and MSPS of type $m^{(i+1)}$ to two "smaller" problems: determine a MHPS or MSPS of type $m^{(i)}$ and then determine a MHPS or MSPS of type $m^{(i+1)} - m^{(i)} - e_0$. □

Remark 2

A MSPS is often computed for the purpose of determining a simultaneous Padé form of type n for $A(z)$. Let $S^{(i)}(z)$ be an MSPS of type $m^{(i)}$ for $A(z)$ with residual $A^{(i)}(z)$. Then, from the proof of theorem 4.2, it is clear that $[\hat{V}(z), \hat{U}(z)]$ is a simultaneous Padé form of type $\nu^{(\sigma)}$ for $A^{(i)}(z)$ if and only if

$$[V(z), U(z)] = \left[\hat{V}(z), \hat{U}(z)\right] \cdot S^{(i)}(z) \tag{4.21}$$

is a simultaneous Padé form of type $n^{(\sigma)}$ for $A(z)$, where $\sigma \geqslant \sigma_i$. Thus, we can compute a simultaneous Padé form of type n for $A(z)$ by computing all the SPSs at the nonsingular points and then computing a simultaneous Padé form of a much smaller type for the final residual. In addition, the problem of characterizing the family of simultaneous Padé forms of type n for $A(z)$ is reduced to the simpler problem of characterizing the family of simultaneous Padé forms for the final residuals. For further work in this area see Van Barel and Bultheel [8] and Beckermann [4].

Similarly, a MHPS is often computed for the purpose of determining a Hermite Padé form of type n for $A(z)$. Let $S^{(i)}(z)$ be a MHPS of type $m^{(i)}$ for $A(z)$ with residual $A^{(i)}(z)$. Then, it is easy to show that $[\hat{V}(z), \hat{U}(z)]^{\mathrm{T}}$ is a matrix Hermite Padé form of type $\nu^{(\sigma)}$ for $A^{(i)}(z)$ if and only if

$$[U(z), V(z)]^{\mathrm{T}} = S^{(i)}(z) \cdot \left[\hat{U}(z), \hat{V}(z)\right]^{\mathrm{T}} \tag{4.22}$$

is a matrix Hermite Padé form of type $n^{(\sigma)}$ for $A(z)$, where $\sigma \geqslant \sigma_i$. Thus, we can compute a matrix Hermite Padé form of type n for $A(z)$ by computing all the MHPSs at the nonsingular points and then computing a matrix Hermite Padé form of a much smaller type for the final residual. □

EXAMPLE 4.3

Let $A(z)$ be the matrix power series from example 3.3. Then (2, 3, 1) is a nonsingular point. In addition the residual matrix power series is nonsingular at (0, 1, 1). Computing the SPS of type (0, 1, 1) for this matrix power series gives

$$\hat{S}(z) = \begin{bmatrix} 1-6z-10z^2 & -4+29z & 3-22z \\ 3z^2+4z^3 & -12z^2 & 9z^2 \\ 4z^2+5z^3 & -16z^2 & 12z^2 \end{bmatrix}.$$

Therefore (3, 4, 2) is also a nonsingular point and theorem 3.1 gives the corresponding normalized SPS as

$$\begin{bmatrix} 1-7z-3z^2+z^3+10z^4-5z^5+2z^6 & 1-7z-4z^2+8z^3+13z^4-7z^5 & z-6z^2-10z^3-2z^4+11z^5+3z^6+12z^7 \\ 3z^2+z^3-z^4-4z^5+2z^6-z^7 & 3z^2+z^3-4z^4-5z^5+3z^6 & 3z^3+4z^4-5z^6-2x^7-5z^8 \\ 4z^2+z^3-z^4-6z^5+3z^6-2z^7 & 4z^2+z^3-5z^4-7z^5+4z^6 & 4z^3+5z^4-7z^6-3z^7-7z^8 \end{bmatrix},$$

with the first few terms of the residual given by

$$\begin{bmatrix} -24+15z+4z^2+\cdots & 2-25z+21z^2+\cdots \\ 1+9z-7z^2+\cdots & 9z^2+\cdots \\ 13z-11z^2+\cdots & 1-2z+13z^2+\cdots \end{bmatrix}. \quad \square$$

5. A superfast multi-dimensional Padé algorithm

Using the results of the previous section and following the approach of Cabay et al. [10], we can construct an algorithm that efficiently computes a MHPS and a MSPS of a given type n. The complexity of such an algorithm is generically $O(\|n\|^2)$ block operations, although there are cases where the complexity can be higher. The desired systems are computed by determining the systems from one nonsingular point to the next. Theorems 4.1 and 4.2 shows that such a process can be accomplished simply by providing an efficient method for computing an initial system along a given path and working with residuals.

In this section we present an algorithm that, when fast polynomial arithmetic is available, will lower the complexities in both cases from $\|n\|^2$ to $\|n\| \cdot \log^2(\|n\|)$. Instead of computing from nonsingular point to nonsingular point, we proceed iteratively doubling the step-size at each step. When we are at a nonsingular point we can use theorems 4.1 or 4.2 and work with the residual series. When we are not at a nonsingular point we cannot use the recurrence from these theorems. Instead we must continue until we get to the next nonsingular location and use the recurrence relation at this point. We will give the details of algorithm for the computation of a MHPS only – the case of a MSPS is similar.

Given a vector of nonnegative integers $n = (n_0, \ldots, n_k)$, the algorithm FAST_MHPS below makes use of theorem 4.1 to compute a subsequence of the MHPS $\{S^{(i)}(z)\}$ for a given block vector of matrix power series $A(z)$. The output gives results associated with the final point $m^{(l)}$. If this final point is a nonsingular point, then the output $S^{(l)}(z)$ is a MHPS of type n. If n is a singular point, then the output is the MHPS at the last nonsingular point in the path generated by the sequence $\{n^{(i)}\}$.

As in [10] we present the algorithm in two parts. The first, INITIAL_MHPS, takes as its input a vector of matrix power series, $A(z)$, with $\det(A_0(0)) \neq 0$, an integer vector n with $n_1 \geqslant \cdots \geqslant n_k$, and a MHPS $S(z)$ for $A(z)$ of type m, where m is one of the nonsingular points defined in (4.6). The procedure returns the MHPS for the residual $\hat{A}(z) = A(z) \cdot S(z)/z^{\|m\|+1}$ at the first nonsingular point, if such a point exists, along the piecewise path determined by $n - m - e_0$. The main algorithm, FAST_MHPS calls INITIAL_MHPS to iteratively construct MHPSs for the residuals, $A^{(i)}(z)$. The MHPSs $S^{(i)}(z)$ for $A(z)$ are computed using the results of theorem 4.1. In the case where INITIAL_MHPS does not return a MHPS, then FAST_MHPS returns the last computed MHPS.

INITIAL_MHPS($A(z)$, n, $S(z)$, m)

I – 1) $\nu \leftarrow n - m - e_0$; $M \leftarrow \min\{\nu_0, \nu_1\} + 1$; $s \leftarrow 0$; $d \leftarrow 0$

I – 2) While $s < M$ and $d = 0$ do

I – 3) $s \leftarrow s+1$
I – 4) $\nu^{(s)} \leftarrow \max\{0, \nu_j - M + s\}, \quad j=0,\ldots,k$
I – 5) $\hat{A}(z) \leftarrow (A(z) \cdot S(z)/z^{\|m\|+1}) \bmod z^{\|\nu^{(s)}\|+1}$
I – 6) Compute $d \leftarrow \det(H_{\nu^{(s)}})$, using Gaussian elimination
End While
I – 7) If $d \neq 0$ then
solve eqs. (2.10) and (2.14) for $\hat{S}(z)$, the MHPS of type $\nu^{(s)}$ for $\hat{A}(z)$
else
$\hat{S}(z) \leftarrow I_{p(k+1)}$; $s \leftarrow 0$
I – 8) Return($\hat{S}(z)$, s) □

The main algorithm, FAST_MHPS takes as its input a vector of matrix power series and a vector of integers, each having $k+1$ components. The vector of integers must have non-negative entries (otherwise one calls FAST_MHPS with a smaller value of k).

FAST_MHPS($A(z)$, n)
F-1) $N \leftarrow \min\{n_0, n_1\}+1$; $i \leftarrow 0$; $\sigma \leftarrow 0$; $\alpha_0 \leftarrow 0$
$m \leftarrow -e_0$; $s \leftarrow 1$; $S(z) \leftarrow I_{p(k+1)}$
F-2) While ($\sigma < N$ and $s > 0$) do
At this stage $S(z)$ is a MHPS of type $m = n^{(\sigma)}$. We then compute
$\hat{S}(z)$ at the first nonsingular point $n^{(\sigma+s)} - n^{(\sigma)} - e_0$ for the
residual $A(z) \cdot S(z)/z^{\|m\|+1}$.
F-3) $(\hat{S}(z), s) \leftarrow$ INITIAL_MHPS($A(z)$, n, $S(z)$, m)
F-4) $\sigma \leftarrow \sigma + s$
Nonsingular point is of type $(m_0, \ldots, m_k)$, where
F-5) $m_j \leftarrow \max\{0, n_j - N + \sigma\}$, $j=0,\ldots,k$
$S(z)$ is the corresponding MHPS of type m
F-6) $S(z) \leftarrow S(z) \cdot \hat{S}(z)$
F-7) If $s > 0$ then
F-8) $\alpha_{i+1} \leftarrow 1 + \lfloor \log(\sigma) \rfloor$
F-9) $t \leftarrow \min\{N, 2^{\alpha_{i+1}} - 1\}$; $i \leftarrow i+1$
Next steps obtain solution at largest nonsingular point before
$n^{(t)}$
F-10) $M_j \leftarrow \max\{0, n_j - N + t\}$, $j=0,\ldots,k$
F-11) $\nu \leftarrow M - m - e_0$
F-12) $\bar{A}(z) \leftarrow (A(z) \cdot S(z)/z^{\|m\|+1}) \bmod z^{\|\nu\|+1}$
F-13) $(\bar{S}(z), \bar{\sigma}) \leftarrow$ FAST_MHPS($\bar{A}(z)$, ν)
F-14) $\sigma \leftarrow \sigma + \bar{\sigma}$
F-15) $m_j \leftarrow \max\{0, n_j - N + \sigma\}$, $j=0,\ldots,k$

F-16) $\quad S(z) \leftarrow S(z) \cdot \bar{S}(z)$
End If
End While
F-17) Return($S(z)$, σ). □

EXAMPLE 5.1
Let $A(z) = [1, A_1(z), A_2(z), A_3(z)]$ be a 4-tuple of power series with

$$A_1(z) = 1 + z + z^3 - z^5 + z^6 + z^7 - z^8 + z^{11} + z^{12} - 2z^{13} + z^{24} - z^{25} + z^{37} - z^{48} + O(z^{63}),$$

$$A_2(z) = 1 + z^3 + z^4 + z^{10} - z^{12} - z^{13} + z^{14} - 2z^{25} + z^{26} + 2z^{39} - z^{40} + O(z^{63}),$$

$$A_3(z) = z + z^2 + z^3 - 2z^4 - z^5 + z^7 + z^8 + z^9 + z^{14} + z^{21} + 2z^{22} + z^{33} + O(z^{63}),$$

and suppose $n = (15, 16, 16, 15)$. We assume that the coefficients of the power series come from $\mathbb{Q}$, the field of rational numbers. In this case, the path determined by n has one singular point, namely $n^{(6)} = (5, 6, 6, ,5)$. The algorithm then proceeds as follows:

Node at start of while loop	Node at start of if loop
$n^{(0)} = (-1, 0, 0, 0)$	$n^{(1)} = (0, 1, 1, 0)$
$n^{(1)} = (0, 1, 1, 0)$	$n^{(2)} = (1, 2, 2, 1)$
$n^{(3)} = (2, 3, 3, 2)$	$n^{(4)} = (3, 4, 4, 3)$
$n^{(7)} = (6, 7, 7, 6)$	$n^{(8)} = (7, 8, 8, 7)$
$n^{(15)} = (14, 15, 15, 14)$	$n^{(16)} = (15, 16, 16, 15)$

Thus we would call FAST_MHPS recursively to compute in step sizes of 1, 2, 4, and 8 respectively.

Suppose now that we consider the power series in $A(z)$ as having coefficients from the finite field $\mathbb{Z}_3$. In this case the path determined by n has singular nodes at locations $n^{(2)}$, $n^{(3)}$, $n^{(6)}$, $n^{(8)}$, $n^{(9)}$, $n^{(13)}$ and $n^{(16)} = n$. Therefore the algorithm proceeds by

Node at start of while loop	Node at start of if loop
$n^{(0)} = (-1, 0, 0, 0)$	$n^{(1)} = (0, 1, 1, 0)$
$n^{(1)} = (0, 1, 1, 0)$	$n^{(4)} = (3, 4, 4, 3)$
$n^{(7)} = (6, 7, 7, 6)$	$n^{(10)} = (9, 10, 10, 9)$
$n^{(15)} = (14, 15, 15, 14)$	

In this case, the algorithm returns the Hermite Padé system of type (14, 15, 15, 14) since this is the last nonsingular point along the path determined by n. $\square$

THEOREM 5.2
(*Correctness*)
For $A(z)$ such that $\det(A_0(0)) \neq 0$ and $n_1 \geqslant \cdots \geqslant n_k$, FAST_MHPS computes the normalized MHPS of type $n^{(\sigma)}$, where

$$\sigma = \max\{t\colon t \leqslant N,\ \det(H_{n^{(t)}}) \neq 0\}. \tag{5.1}$$

Proof
Assume inductively that prior to pass $i+1$, $i = 0, 1, \ldots,$ of the WHILE loop F-2 the following holds:
(I) $\alpha_i \geqslant i$,
(II) $\sigma = \max\{t\colon t \leqslant \{N, 2^{\alpha_i} - 1\}, \det(H_{n^{(t)}}) \neq 0\}$,
(III) $m = n^{(\sigma)}$,
(IV) $S(z)$ is the normalized MHPS of type m for which

$$A(z) \cdot S(z) = z^{\|m\|+1} \hat{A}(z). \tag{5.2}$$

Initially, for $i = 0$, $\alpha_i = 0$ and therefore $\sigma = 0$ (because $\det(H_{n^{(0)}}) = 1$ with $n^{(0)} = -e_0$). Also, for $i = 0$, $S(z) = I_{p(k+1)}$.
In step F-3, if $\det(H_{n^{(\sigma+t)}}) = 0$ for all t such that $1 \leqslant t \leqslant N - \sigma$, then INITIAL_MHPS returns $(\hat{S}(z), s) \leftarrow (I_{p(k+1)}, 0)$. Here $m = n^{(\sigma)}$ is already the last nonsingular point along the path from $n^{(0)}$ to n and the algorithm terminates at the earliest opportunity. Otherwise, let

$$s = \min\{t\colon t \leqslant N - \sigma,\ \det(H_{n^{(\sigma+t)}}) \neq 0\}, \tag{5.3}$$

which defines the next nonsingular point $n^{(\sigma+s)}$. The subroutine INITIAL_MHPS computes $\hat{S}(z)$ of type $\nu^{(s)}$ for the residual $\hat{A}(z)$ such that

$$\hat{A}(z) \cdot \hat{S}(z) = z^{\|\nu^{(s)}\|+1} \cdot \bar{A}(z), \tag{5.4}$$

where $\nu^{(s)} = n^{(\sigma+s)} - n^{(\sigma)} - e_0$. Using (5.2) and (5.3), it is easy to show that the computation in step F-6 yields $S(z) \cdot \hat{S}(z)$ satisfying

$$A(z) \cdot S(z) \cdot \hat{S}(z) = z^{\|n^{(\sigma+s)}\|+1} \cdot \bar{A}(z), \tag{5.5}$$

and all the other conditions of the normalized MHPS of type $n^{(\sigma+s)}$ for $A(z)$. Just prior to step F-7, the values σ, m and $S(z)$ have been redefined so that $S(z)$ is now the MHPS of type $m = n^{(\sigma)}$, where (for the case $s > 1$ in (5.3))

$$\sigma = \min\{t\colon 2^{\alpha_i} \leqslant t \leqslant N,\ \det(H_{n^{(t)}}) \neq 0\}. \tag{5.6}$$

With σ given by (5.6), in step F-8, α_{i+1} is determined so that

$$2^{\alpha_{i+1}-1} \leqslant \sigma \leqslant 2^{\alpha_{i+1}} - 1. \tag{5.7}$$

Clearly $i+1 \leqslant \alpha_i + 1 \leqslant \alpha_{i+1}$ and so the inductive hypothesis (I) holds. In steps F-9 through F-16, the algorithm computes the MHPS of type $n^{(\sigma+\bar{\sigma})}$ for the largest $\bar{\sigma}$ in the range

$$\sigma \leqslant \sigma + \bar{\sigma} \leqslant \min\{N, 2^{\alpha_{i+1}} - 1\} \tag{5.8}$$

for which $H_{n^{(\sigma+\bar{\sigma})}} \neq 0$. This is accomplished by recursively invoking FAST_MHPS for the residual $\bar{A}(z)$. Computed is the normalized MHPS $\bar{S}(z)$ of type ν for $\bar{A}(z)$ such that

$$\bar{A}(z) \cdot \bar{S}(z) = O(z^{\|\nu\|+1}), \tag{5.9}$$

where $\nu = n^{(\sigma+\bar{\sigma})} - n^{(\sigma)} - e_0$. It then follows that $S(z) \cdot \bar{S}(z)$ satisfies

$$A(z) \cdot S(z) \cdot \bar{S}(z) = O(z^{\|n^{(\sigma+\bar{\sigma})}\|+1}), \tag{5.10}$$

and all the other conditions of the normalized MHPS of type $n^{(\sigma+\bar{\sigma})}$ for $A(z)$. By accordingly redefining σ, m and $S(z)$ in F-14, F-15 and F-16, the inductive hypotheses II, III and IV are shown to hold for the next pass through the WHILE loop F-2. □

Remark 1

The changes to our algorithm for the computation of a MSPS are straightforward. In this case the input to the main algorithm would be a $(k+1) \times k$ block matrix of $p \times p$ power series, $A(z)$. The only other changes would be the multiplication in steps I-5, F-6, F-12 and F-16, which would now be on the left instead of the right. □

Remark 2

The algorithms given in this section compute either a MHPS or a MSPS of type n. They can also be used to compute matrix Hermite Padé and matrix simultaneous Padé approximants of type n as long as n is a nonsingular point. At present, the algorithm used to compute the first nonsingular point and its corresponding Padé system returns the identity if no nonsingular point can be found. If this is changed to return all possible Padé approximants and weak Padé approximants (determined by solving the linear system of equations) then, because of remark 2 from the previous section, the main algorithm will return all possible Padé forms of type n. Hence our approach can be used to find all possible Padé forms (either Hermite Padé or simultaneous Padé) of type n, regardless of whether n is a singular or nonsingular point. □

6. Complexity of the multi-dimensional Padé algorithm

THEOREM 6.1

For $A(z)$ such that $\det(A_0(0)) \neq 0$ and n such that $n_1 \geqslant \cdots \geqslant n_k$, the cost of FAST_MHPS is

$$O\big((k+1)^2 \cdot \|n\| \cdot \log^2(\|n\|) + (k+1)^2 \cdot s^2 \cdot \|n\|\big) \tag{6.1}$$

$p \times p$ matrix operations, where $s = \max\{s_i \mid s_i = m_0^{(i+1)} - m_0^{(i)}\}$ is the maximum step-size used.

Proof

We first estimate the cost of FAST_MHPS for those $n = (n_0, \ldots, n_k)$ which are restricted by

$$n_0 = n_1 - 1. \tag{6.2}$$

This restriction holds for all recursive calls of FAST_MHPS in step F-13. For n so restricted and for all $A(z)$ let $T(\beta)$ be an estimate of the cost of using FAST_MHPS to compute the normalized MHPS of type $n^{(\sigma)}$, where

$$\sigma = \max\{t \mid t \leqslant \beta,\ \det(H_{n^{(t)}}) \neq 0\}. \tag{6.3}$$

With $N = \min\{n_0, n_1\} + 1$, then $T(N)$ gives an estimate for the cost of computing the MHPS of type n (since $n^{(N)} = n$).

To obtain $T(N)$, we examine the cost of the ith pass through the WHILE loop F-2. Just prior to the ith pass, $S(z)$ is the normalized MHPS of type $n^{(\sigma_i)}$, where

$$\sigma_i = \max\{t \mid t \leqslant 2^{\alpha_i} - 1,\ \det(H_{n^{(t)}}) \neq 0\}. \tag{6.4}$$

Let s_i be the step-size computed in F-3. Then (except when $s = 0$ and the algorithm is about to terminate)

$$s_i = \min\{t \mid \sigma_i + t \geqslant 2^{\alpha_i},\ \det(H_{n^{(\sigma_i+t)}}) \neq 0\}. \tag{6.5}$$

Also let

$$\nu^{(i)} = n^{(\sigma_i + s_i)} - n^{(\sigma_i)} - e_0. \tag{6.6}$$

Table 1
Cost estimates for one iteration.

Step	Estimate of number of $p \times p$ matrix operations
F-3	$O((k+1) \cdot \|\nu^{(i)}\| \cdot \|n^{(2^{\alpha_i})}\| + \|\nu^{(i)}\|^3)$
F-6	$O((k+1) \cdot \|\nu^{(i)}\| \cdot \|n^{(2^{\alpha_i})}\|)$
F-12	$O((k+1)^2 \cdot \|n^{(2^{\alpha_i+1})}\| \cdot \log(\|n^{(2^{\alpha_i+1})}\|))$
F-13	$T(2^{\alpha_{i+1}-1})$
F-16	$O((k+1) \cdot \|n^{(2^{\alpha_i+1})}\| \cdot \log(\|n^{(2^{\alpha_i+1})}\|))$

Then estimates of the number of $p \times p$ matrix operations performed in the main steps of FAST_MHPS are given in table 1.

In table 1, for the step F-3 the term $(k+1) \cdot \| \nu^{(i)} \| \cdot \| n^{(2^{\alpha_i})} \|$ relates to the cost of computing the residual $\hat{A}(z)$ in step I-5 of INITIAL_MHPS, whereas $\| \nu^{(i)} \|^3$ relates to the cost of using the Gaussian elimination method to solve systems (2.10) and (2.14) in steps I-6 and I-7 in INITIAL_MHPS. In steps F-12 and F-16 it is assumed the last polynomial multiplication methods (based on fast Fourier transforms) are used. Note that the degrees of the polynomials computed in step F-12 are $O(k+1)$ larger than those in step F-16, which accounts for the difference in the complexity of the two steps.

The result (6.1) follows from table 1 by summing the costs of iteration i for $i = 1, \dots, \lfloor \log(N) \rfloor$. The term $(k+1)^2 \cdot s^2 \cdot \| n \|$ in (6.1) accounts for the cost $\| \nu^{(i)} \|^3$ of solving systems (2.10) and (2.14) in step F-3 whereas the term $(k+1) \cdot \| n \| \cdot \log^2(\| n \|)$ in (6.1) accounts for all the other costs identified in table 1. The cost $(k+1)^2 \cdot s^2 \cdot \| n \|$ is precisely that of solving systems (2.10) and (2.14) of type $\nu^{(j)} = m^{(j+1)} - m^{(j+1)} - e_0$ for all nonsingular points $m^{(j)}$, $j = 1, \dots, l$, where

$$l = \max\{ t \mid t \leqslant N, \det(H_{n^{(t)}}) \neq 0 \}. \tag{6.7}$$

Namely,

$$\sum_{j=1}^{l} \| \nu^{(j)} \|^3 \leqslant (k+1)^2 \cdot s^2 \cdot \| n \|. \tag{6.8}$$

These systems (2.10) and (2.14) (for different residual power series) arise as a consequence of the accumulation of the recursive call of FAST_MHPS in step F-13.

To remove the restriction (6.2) we note that $S(z)$ at $m^{(i)}$ is obtained in step F-3 by solving (2.10) and (2.14) at the first nonsingular point $\nu^{(0)} = m^{(1)} - m^{(0)} - e^0 = m^{(1)}$. We then add to this the cost of (6.8) to obtain

$$\sum_{j=0}^{l} \| \nu^{(j)} \|^3 \leqslant (k+1)^2 \cdot s^2 \cdot \| n \|. \quad \square \tag{6.9}$$

As a follow-up to remark 1 of the previous section, we have a similar result for computing a MSPS.

COROLLARY 6.2

An algorithm FAST_MSPS for computing a MSPS of type $(n_0, \dots, n_k)$ can be given which requires

$$O\big((k+1)^3 \cdot \| n \| \cdot \log^2(\| n \|) + (k+1)^3 \cdot s^3 \big) \tag{6.10}$$

$p \times p$ matrix operations, where $s = \max\{ s_i \mid s_i = m_0^{(i+1)} - m_0^{(i)} \}$ is the maximum step-size used. $\square$

EXAMPLE 6.3

In the perfect case, $s_i = 1$ for all i. In this case the second term in (6.1) becomes $O((k+1)^3 \cdot \| n \|)$ and so the complexity of the algorithm becomes $O(\| n \| \cdot \log^2(\| n \|))$. 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 $(k+1) \cdot s \geqslant \| n \|$, then the second term in (6.1) becomes $O(\| n \|^3)$ which corresponds to the cost of Gaussian elimination of the full Hankel-like system (2.9). Indeed the solution is exactly that obtained by a single invocation of INITIAL_MHPS. □

EXAMPLE 6.4

When $k = 1$ and $p = 1$ the FAST_MHPS algorithm coincides with that given for Padé approximants by Cabay and Choi [9]. In the scalar case a call to INITIAL_MHPS always results in solving a triangular system of linear equations. Thus the cubic terms in (6.1) resulting from Gaussian elimination are in fact not present in the scalar case. As such the algorithm computes a Padé approximant of type (m, n) with the superfast complexity $O((m+n) \cdot \log^2(m+n))$. This is the case regardless of any assumptions on the size of the steps from one nonsingular node to the next. Gaussian elimination would require $O((m+n)^3)$ operations in this case. □

7. Conclusions

We have given a new reliable algorithm for the computation of matrix Hermite Padé and matrix simultaneous Padé systems. This in turn provides a reliable algorithm for the computation of the corresponding matrix Padé approximants. The algorithm is superfast, that is, when fast polynomial multiplication is possible the algorithm in most cases computes a system of type n in $O(\| n \| \cdot \log^2 \| n \|)$ block matrix operations.

There are a number of possible research directions that follow from our work. Our algorithm depends on the distribution of the nonsingular points along a diagonal path in k-dimensional space. When most of the points are nonsingular the algorithm is superfast, and hence faster than existing algorithms such as proposed by Beckermann [11] or Van Barel and Bultheel [2]. However, when there are very few nonsingular points and the distance between such nonsingular points is large then the algorithm has a potential complexity of $O(\| n \|^3)$. The algorithms of both Beckermann and Van Barel and Bultheel, on the other hand, do not depend on the singular structure of their computational path. It is of interest to generalize their algorithms (in both the scalar and matrix cases), using similar divide-and-conquer methods such as found in section 5. The hope would be for algorithms that compute in superfast complexity, regardless of any singularities in the path of computation. At present our algorithm has this important property, but only in the special case of scalar Padé approximants (i.e. $k=1$, $p=1$).

Our approach has similarities with look-ahead [28] and block-pivoting [27] methods for solving associated block Hankel-like systems. Thus, as in the scalar algorithm [25] the algorithm promises to be a numerically stable one. This can be done by recursing at a stable point rather than a nonsingular point. Stable points are those having an "acceptable" condition number for the corresponding matrix of the system of equations. Further work is under way in this direction.

References

[1] G.A. Baker and P. Graves-Morris, *Padé Approximants Part II: Extensions and Applications* (Addison-Wesley, Reading, 1981).

[2] M. Van Barel and A. Bultheel, The computation of non-perfect Padé–Hermite approximants, Preprint, to appear in Num. Algorithms (1991).

[3] B. Beckermann, A reliable method for computing M-Padé approximants on arbitrary staircases, J. Comp. Appl. Math. Preprint (1990).

[4] B. Beckermann, The structure of the singular solution table of the M-Padé approximation problem, J. Comp. Appl. Math. 32 (1990) 3–15.

[5] 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 (1980) 259–295.

[6] M.G. de Bruin, The interruption phenomenon for generalized continued fractions, Bull. Austral. Math. Soc. 19 (1978) 245–272.

[7] M.G. de Bruin, Some aspects of simultaneous rational approximation, in: *Proc. PAN* (1987).

[8] A. Bultheel and M. Van Barel, Minimal vector Padé approximation, J. Comp. Appl. Math. 32 (1990) 27–37.

[9] S. Cabay and D.K. Choi, Algebraic computations of scaled Padé fractions, SIAM J. Comput. (Feb. 1986) 243–270.

[10] S. Cabay, G. Labahn and B. Beckermann, On the theory and computation of non-perfect Padé–Hermite approximants, to appear in J. Comp. Appl. Math. (1991).

[11] J. Coates, On the algebraic approximation of functions, I–III, K. Nederl. Ak. Wetenschappen 69 (1966) 421–461.

[12] J. Della Dora and C. Dicrescenzo, Approximants de Padé–Hermite, Numer. Math. 43 (1984) 23–57.

[13] W.B. Gragg, The Padé table and its relation to certain algorithms of numerical analysis, SIAM Rev. 14 (1972) 1–61.

[14] P. Graves-Morris, Vector-valued rational interpolants II, IMA J. Numer. Anal. 4 (1984) 209–224.

[15] M.Ch. Hermite, Sur l'Expression $U\sin x + V\cos x + W$, Extrait d'une lettre a M. Paul Gordan, *Oeuvres Tome III* (1873) pp. 135–145.

[16] M.Ch. Hermite, Sur quelques approximations algebriques, Extrait d'une lettre a M. Borchardt, *Oeuvres Tome III* (1873) pp. 146–149.

[17] H. Jager, A multidimensional generalization of the Padé table I–VI, K. Nederl. Ak. Wetenschappen 67 (1964) 192–249.

[18] G.A. Baker Jr. and P.R. Graves-Morris, Definition and uniqueness of integral approximants, J. Comput. Appl. Math. 31 (1990) 357–372.

[19] G. Labahn and S. Cabay, Matrix Padé fractions and their computation, SIAM J. Comput. 18 (1989) 639–657.

[20] G. Labahn, D.K. Choi and S. Cabay, Inverses of block Hankel and block Toeplitz matrices, SIAM J. Comput. 19 (1990) 98–123.
[21] G. Labahn, Inversion components for block Hankel-like matrices, to appear in Lin. Algebra Appl. (1991).
[22] L. Lerer and M. Tismenetsky, Generalized Bezoutians and the inversion problem for block matrices, Integral Eq. Oper. Th. 9 (1986) 790–819.
[23] K. Mahler, Zur Approximation der Exponentialfunktion und des Logarithmus, Teil I, Journ. f.d.r.u.a. 166 (1932) 118–136.
[24] K. Mahler, Perfect systems, Compositio Math. 19 (1968) 95–166.
[25] R.J. Meleshko, A stable algorithm for the computation of Padé approximants, Ph.D. Thesis, Dept. of Computing Science, University of Alberta, Edmonton, Canada (1990).
[26] H. Padé, Sur la représentation approchée d'une fonction par des fractions rationelles, Ann. Ecole Nor., supplement Thesis (1892) pp. 1–93.
[27] B.N. Parlett and J.R. Bunch, Direct methods for solving symmetric indefinite systems of linear equations, SIAM J. Numer. Anal. 8 (1971) 639–655.
[28] B.N. Parlett, D.R. Taylor and Z.A. Liu, A look-ahead Lanczos algorithm for unsymmetric matrices, Math. Comput. 44 (1985) 105–124.
[29] S. Paszkowski, Recurrence relations in Padé–Hermite approximation, J. Comp. Appl. Math. 19 (1987) 99–107.
[30] Y. Sugiyama, An algorithm for solving discrete-time Wiener–Hopf equations based on Euclid's algorithm, IEEE Trans. Inform. Theory IF-32 (1986) 394–409.