SIAM J. MATRIX ANAL. APPL.
Vol. 15, No. 3, pp. 804–823, July 1994

007

A UNIFORM APPROACH FOR THE FAST COMPUTATION OF MATRIX-TYPE PADÉ APPROXIMANTS *

BERNHARD BECKERMANN† AND GEORGE LABAHN‡

Abstract. Recently, a uniform approach was given by B. Beckermann and G. Labahn [*Numer. Algorithms*, 3 (1992), pp. 45–54] for different concepts of matrix-type Padé approximants, such as descriptions of vector and matrix Padé approximants along with generalizations of simultaneous and Hermite Padé approximants. The considerations in this paper are based on this generalized form of the classical scalar Hermite Padé approximation problem, *power Hermite Padé approximation.* In particular, this paper studies the problem of computing these new approximants.

A recurrence relation is presented for the computation of a basis for the corresponding linear solution space of these approximants. This recurrence also provides bases for particular subproblems. This generalizes previous work by Van Barel and Bultheel and, in a more general form, by Beckermann. The computation of the bases has complexity $\mathcal{O}(\sigma^2)$, where σ is the order of the desired approximant and requires no conditions on the input data. A second algorithm using the same recurrence relation along with divide-and-conquer methods is also presented. When the coefficient field allows for fast polynomial multiplication, this second algorithm computes a basis in the superfast complexity $\mathcal{O}(\sigma \log^2 \sigma)$. In both cases the algorithms are reliable in exact arithmetic. That is, they never break down, and the complexity depends neither on any normality assumptions nor on the singular structure of the corresponding solution table. As a further application, these methods result in fast (and superfast) reliable algorithms for the inversion of striped Hankel, layered Hankel, and (rectangular) block-Hankel matrices.

Key words. vector Padé approximant, Hermite Padé approximant, simultaneous Padé approximant, matrix Padé approximant, Hankel matrices

AMS subject classifications. 65D05, 41A21, CR: G.1.2

1. Introduction. Let $\mathbf{F} = (f_1, \ldots, f_m)^T$ (with $m \geq 2$) be an m-tuple of formal power series with coefficients from a field $\mathbb{K}$ (typically a subfield of either the real or complex numbers) and $\mathbf{n} = (n_1, \ldots, n_m)$ an m-tuple of integers, $n_l \geq -1$. A *Hermite Padé approximant* for $\mathbf{F}$ of type $\mathbf{n}$ is a nontrivial tuple $\mathbf{P} = (P_1, \ldots, P_m)$ of polynomials P_l over $\mathbb{K}$ having degrees bounded by the n_l such that

$$(1) \quad \mathbf{P}(z) \cdot \mathbf{F}(z) = P_1(z) f_1(z) + \cdots + P_m(z) f_m(z) = c_N z^N + c_{N+1} z^{N+1} + \cdots,$$

with $N = n_1 + \cdots + n_m + m - 1$.

The *Hermite Padé approximation problem* was introduced in 1873 by Hermite and has been studied widely by several authors (for a bibliography, see, e.g. [2]–[4] or [25]). Note that when $m = 2, \mathbf{F} = (f, -1)^T$, Eq. (1) is the same as

$$P_1(z) f(z) - P_2(z) = O\left(z^{n_1+n_2+1}\right),$$

and hence as a special case we have the classical Padé approximation problem for a power series f. Hermite Padé approximation also includes other classical approximation problems such as algebraic approximants ($\mathbf{F} = (1, f, f^2, \ldots, f^{m-1})^T$) (e.g. [23] for the special case $m = 2$) and $G^3 J$ approximants ($m = 3, \mathbf{F} = (f', f, 1)^T$). We refer the reader to [1, pp. 32–40] for additional examples. More generally, there is the

* Received by the editors April 16, 1992; accepted for publication (in revised form) February 25, 1993.

† Laboratoire d'Analyse Numérique et d'Optimisation, UFR IEEA-M3, Université des Sciences et Technologies de Lille, 59655 Villeneuve d'Ascq Cedex, France.

‡ Department of Computing Science, University of Waterloo, Waterloo, Ontario, Canada (glabahn@daisy.waterloo.edu).

M-Padé approximation problem that requires that $\mathbf{P} \cdot \mathbf{F}$ vanishes at a given set of knots $z_0, z_1, \ldots, z_{N-1}$, counting multiplicities ([2]–[4], [20], [21]). The case where all the z_i are equal to 0 is just the Hermite Padé problem.

Hermite also defined a second type of approximant to a vector of power series, the so-called *simultaneous Padé approximants* and used them in his proof of the transcendence of e. Close connections between these two approximation problems have been pointed out in [7], [14], [16], [17], [21].

In recent years, several vector and matrix generalizations of these approximation problems have been given (see §2). The aim of this paper is to study a uniform approach not only to Hermite Padé and simultaneous Padé approximants but also to their matrix-type generalizations. To this end, we consider the following generalized scalar Hermite Padé approximation problem [5].

DEFINITION 1.1. *Let $\sigma \geq 0, s > 0, n_1, \ldots, n_m$ be integers, $n_l \geq -1$ and $\mathbf{n} = (n_1, \ldots, n_m)$. Then a* power Hermite Padé approximant *(PHPA) $\mathbf{P} = (P_1, \ldots, P_m)$ of type $(\mathbf{n}, \sigma, s)$ consists of scalar polynomials P_l having degrees bounded by the n_l with*

$$(2)\quad R(z) = \mathbf{P}(z^s) \cdot \mathbf{F}(z) = P_1(z^s) f_1(z) + \cdots + P_m(z^s) f_m(z) = c_\sigma z^\sigma + c_{\sigma+1} z^{\sigma+1} + \cdots,$$

that is, has order σ. The power series R is referred to as the s-residual.

The power s appearing in Definition 1.1 provides a method of converting a vector problem into a scalar problem (see §2). By defining these approximants in a similar way to Hermite Padé approximants we can borrow from the (successful) computational techniques for the Hermite Padé problem used in [2], [4], [25]. Of course the classical Hermite Padé approximation problem is included by setting $s = 1$ and $\sigma = \|\mathbf{n}\| - 1$, where the norm of multiindices $\mathbf{n} = (n_1, \ldots, n_m) \in (\mathbb{N}_0 \cup \{-1\})^m$ is defined by $\|\mathbf{n}\| := (n_1 + 1) + \cdots + (n_m + 1)$. Note that, by equating coefficients, (2) results in a system of homogeneous linear equations. By comparing the number of unknowns to equations, one can conclude that there exists at least $\|\mathbf{n}\| - \sigma$ PHPAs of type $(\mathbf{n}, \sigma, s)$ that are linearly independent over $\mathbb{K}$.

Section 2 gives examples of matrix-type generalizations of existing approximation problems. These are shown to be special cases of the PHPA problem for various values of s and σ. In §3 we provide a recursive algorithm to efficiently and reliably solve the PHPA problem in exact arithmetic. Some interesting properties of our algorithm along with a cost analysis are given in §4. It is shown that the algorithm is at least as fast or faster than existing methods for special cases. Thus, our results provide a uniform method of computing matrix-type generalizations of Padé approximation problems. Section 5 gives an example of the use of this algorithm in the context of square-matrix Padé approximants. Section 6 considers a modification of our algorithm that combines divide-and-conquer techniques along with the recurrence relation of §3. When the field $\mathbb{K}$ allows fast polynomial multiplication, the resulting new algorithm solves the PHPA problem with superfast complexity. Finally, the paper closes with a discussion of a number of research directions that follow from our work.

For purposes of presentation, we adopt the following notations. Let $\mathcal{S}$ be a space with scalars from $\mathbb{K}$, for instance, $\mathcal{S} = \mathbb{K}^{(p \times q)}$, the space of $p \times q$ matrices over $\mathbb{K}$. Then $\mathcal{S}[z]$ denotes the set of polynomials in z with coefficients from $\mathcal{S}$, whereas $\mathcal{S}[[z]]$ represents the set of formal power series in z with coefficients from $\mathcal{S}$. Multiindices and PHPAs are denoted in boldface letters; they are both $(1 \times m)$ row vectors. Also, throughout this paper the parameter s and the multiindex $\mathbf{n}$ are fixed. The algorithm of §3 follows along an m-dimensional "diagonal" path $(\mathbf{n}(\delta))_{\delta \in \mathbb{Z}}$ induced by $\mathbf{n}$, which is defined as follows:

$$(3) \quad \delta \in \mathbb{Z}, \mathbf{n} = (n_1, \ldots, n_m) : \mathbf{n}(\delta) = \left(n_1', \ldots, n_m'\right) \quad \text{with } n_l' = \max\{-1, n_l + \delta\}.$$

This notion allows us to discuss not only one approximation problem corresponding to $\mathbf{n} = \mathbf{n}(0)$ but also simultaneously all subproblems associated with $\mathbf{n}(\delta), \delta < 0$ (cf. Table 3). Finally, the set of all PHPAs of type $(\mathbf{n}(\delta), \sigma, s)$ is denoted as $\mathcal{L}_\delta^\sigma$; it is a finite-dimensional space over $\mathbb{K}$.

Parallel to and independent of [5] and our present work, another uniforming approach was proposed in [26] by Van Barel and Bultheel based on the concept of vector M-Padé approximation. Their approach does not reduce to a simple scalar concept as does the notion of our PHPAs. However, their approach does have the advantage of handling matrix rational interpolation and is seen as complementary to this paper.

2. Matrix-type Padé approximants as special PHPAs.

In this section we give examples of a number of matrix-type generalizations of classical Padé approximation problems. Let A be a $p \times q$ matrix of power series over $\mathbb{K}$ and suppose that $r \in \mathbb{N}$ and $M, N \in \mathbb{N}_0$.

Example 2.1 (Right-hand square and rectangular Matrix-Padé forms). Find $P \in \mathbb{K}^{(p\times r)}[z], Q \in \mathbb{K}^{(q\times r)}[z]$, with deg $P \le M$, deg $Q \le N$, and the columns of Q being linearly independent over $\mathbb{K}$ such that

$$A(z) \cdot Q(z) - P(z) = z^{M+N+1} \cdot R(z),$$

with $R \in \mathbb{K}^{(p\times r)}[[z]]$.

Example 2.2 (Left-hand square and rectangular Matrix-Padé forms). Find $P \in \mathbb{K}^{(r\times q)}[z], Q \in \mathbb{K}^{(r\times p)}[z]$, with deg $P \le M$, deg $Q \le N$, and the rows of Q being linearly independent over $\mathbb{K}$ such that

$$Q(z) \cdot A(z) - P(z) = z^{M+N+1} \cdot R(z),$$

with $R \in \mathbb{K}^{(r\times q)}[[z]]$.

When $p = q = r = 1$ this is the classical scalar Padé approximation problem. When $p = q = r > 1$ these are square right-hand or left-hand matrix Padé approximants [19]. In the rectangular ($p \ne q$) case, two natural matrix Padé approximations occur when either $p = r$ or $q = r$. Both of these rectangular-matrix types of Padé forms are used, for example, to compute the inverse of matrices partitioned into a rectangular-block Hankel structure [18].

We remark that, in the examples where $Q(z)$ is square, it is of special interest to determine those cases where we can form a Padé fraction $P(z)\cdot Q(z)^{-1}$ or $Q(z)^{-1}\cdot P(z)$ as an approximant to $A(z)$. In both cases we are therefore interested in necessary and sufficient conditions under which $Q(z)$ is nonsingular.

Motivated by the well-known connections between left-hand and right-hand square matrix Padé forms and by inversion formulas of block Hankel-like matrices, one of the authors [17] introduced for $p, \mu \in \mathbb{N}$ and $\rho_0, \ldots, \rho_\mu \ge 0, \rho = \rho_0 + \cdots + \rho_\mu, A_0, \ldots, A_\mu \in \mathbb{K}^{p\times p}[[z]]$.

Example 2.3 (Matrix Hermite Padé form). Find polynomials $P_0, \ldots, P_\mu \in \mathbb{K}^{p\times p}[z]$ with deg $P_l \le \rho_l - 1, 0 \le l \le \mu$, and

$$A_0(z)P_0(z) + \cdots + A_\mu(z)P_\mu(z) = z^{\rho-1} \cdot R(z),$$

$R \in \mathbb{K}^{p\times p}[[z]]$ such that the matrix $[P_0,\ldots,P_\mu] \in \mathbb{K}^{p\times(\mu+1)p}[z]$ has full rank over $\mathbb{K}$.

Example 2.4 (Matrix simultaneous Padé form). Find polynomials $Q_0,\ldots,Q_\mu \in \mathbb{K}^{p\times p}[z]$ with $\deg Q_l \le \rho - \rho_l, 0 \le 1 \le \mu$, and

$$Q_0(z)A_l(z) - Q_l(z)A_0(z) = z^{\rho+1}\cdot R_l(z),$$

$1 \le l \le \mu, R_l \in \mathbb{K}^{p\times p}[[z]]$ such that the matrix $[Q_0,\ldots,Q_\mu] \in \mathbb{K}^{p\times(\mu+1)p}[z]$ has full rank over $\mathbb{K}$.

Beside the classical scalar simultaneous Padé approximants ($p = 1, A_0(z) = 1$), Example 2.4 also includes the *simultaneous partial Padé approximation problem* where we have prescribed poles and zeros for the approximants [8]. Following [22], the question of irreducible Hermite Padé forms is of special interest, i.e., we also require that $[P_0(0),\ldots,P_\mu(0)] \in \mathbb{K}^{p\times(\mu+1)p}$ is different from zero (or moreover has full rank over $\mathbb{K}$). Similarly, in Example 2.4 we are interested in approximants where $Q_0(0)$ is a nonsingular matrix.

We remark that the order conditions in Examples 2.1–2.4 are all such that at least one solution exists for each approximation problem. In addition, the so-called *weak matrix Hermite Padé* and *weak matrix simultaneous Padé forms* are connected to Examples 2.3 and 2.4 (see [17]). In this case the order conditions are weakened to allow for more linearly independent solutions. Other examples of matrix-type generalizations of Padé approximants include Hermite Padé [11] and simultaneous Padé systems [12], [17]. These, however, only exist in certain cases.

Note that the matrix simultaneous Padé form is closely connected to a rectangular matrix Hermite Padé form if the interpolation conditions are written as follows:

$$\begin{bmatrix} A_1^T(z) \\ A_2^T(z) \\ A_3^T(z) \\ \vdots \\ A_\mu^T(z) \end{bmatrix} \cdot Q_0^T(z) + \begin{bmatrix} -A_0^T(z) \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \cdot Q_1^T(z) + \cdots + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ -A_0^T(z) \end{bmatrix} \cdot Q_\mu^T(z) = z^{\rho+1}\cdot \begin{bmatrix} R_1^T(z) \\ R_2^T(z) \\ R_3^T(z) \\ \vdots \\ R_\mu^T(z) \end{bmatrix}.$$

All examples given here are special cases of so-called vector Hermite Padé approximants.

Example 2.5 (vector Hermite Padé approximant). Let $m,\ s,\ \tau \in \mathbb{N}_0,\ m,\ s \ge 2, G_1,\ldots,G_m \in \mathbb{K}^{s\times 1}[[z]]$ and let $\mathbf{n}$ be a multiindex. Find linearly independent polynomial tuples $(P_1,\ldots,P_m), P_l \in \mathbb{K}[z]$ with $\deg P_l \le n_l, 1 \le l \le m$ such that

$$G_1(z)P_1(z) + \cdots + G_m(z)P_m(z) = z^\tau \cdot R(z),$$

with $R \in \mathbb{K}^{s\times 1}[[z]]$.

By setting

$$\text{(4)} \qquad \text{for } 1 \le l \le m\text{:} \quad f_l(z) = (1, z, z^2, \ldots, z^{s-1}) \cdot G_l(z^s),$$

we see that computing vector Hermite Padé approximants of type $(\mathbf{n},\tau)$ and dimension s is equivalent to the determination of PHPAs of type $(\mathbf{n},\tau s, s)$, i.e. of the solution set $\mathcal{L}_0^{\tau s}$. Indeed, the above technique of converting a vector problem to a scalar problem via the raising of z to the sth power provides the motivation for Definition 1.1.

In Table 1, we list the particular choices of $m, \mathbf{n}, s, \sigma, \mathbf{F}$ with respect to Examples 2.1, 2.2, and 2.3. Instead of Example 2.4, we consider the special case of scalar simultaneous Padé approximation.

TABLE 1

Specification of the PHPA parameters used in (2) for some matrix-type Padé approximation problems.

Example	m	s	σ	$\mathbf{n}$, $\mathbf{F}$	Number solutions*
Classical Hermite Padé	m	1	$\Vert\mathbf{n}\Vert - 1$	$(n_1, \dots, n_m)$, $\mathbf{F}^T(z) = (f_1(z), \dots, f_m(z))$	1
2.1	$p+q$	p	$p(M+N+1)$	$(M, \dots, M, N, \dots, N)$, $\mathbf{F}^T(z) = (1, z, \dots, z^{p-1}) \cdot (\mathbf{I}, -A(z^p))$	r
2.2	$p+q$	q	$q(M+N+1)$	$(M, \dots, M, N, \dots, N)$, $\mathbf{F}(z) = \begin{pmatrix} \mathbf{I} \\ -A(z^q) \end{pmatrix} \cdot (1, z, \dots, z^{q-1})^T$	r
2.3	$p(\mu+1)$	p	$p(\rho-1)$	$(\rho_0 - 1, \dots, \rho_0 - 1, \dots, \rho_\mu - 1, \dots, \rho_\mu - 1)$, $\mathbf{F}^T(z) = (1, z, \dots, z^{p-1}) \cdot (A_0(z^p), \dots, A_\mu(z^p))$	p
2.4 with $p = 1, A_0 = 1$	$\mu+1$	μ	$\mu(\rho+1)$	$(\rho - \rho_0, \dots, \rho - \rho_\mu)$, $f_1(z) = -\sum_{1 \le j \le \mu} z^{j-1} A_j(z^\mu)$, $j \ge 1 \ : \ f_{j+1}(z) = z^{j-1}$	1

* Number of PHPA solutions required to construct the corresponding matrix-type Padé approximant.

3. Recursive computation of PHPA bases. In this section, we construct systems of m PHPAs by recurrence on σ. This allows us to describe all the PHPAs of type $(\mathbf{n}(\delta), \sigma, s), \delta \le 0$, when a fixed $s, \mathbf{F}$, and $\mathbf{n}$ are given. Therefore, we not only solve the Hermite Padé approximation problem of type $\mathbf{n}$ or the corresponding matrix-type Padé approximation problem (see §2) but also all subproblems of type $\mathbf{n}(\delta), \delta \le 0$ (cf. (3)) belonging to a "diagonal path" in the solution table. The recurrence formula and the resulting algorithm do not require any assumptions on the input data $\mathbf{F}$. Moreover, the algorithm is fast, i.e. it always has a complexity of $\mathcal{O}(\Vert\mathbf{n}\Vert^2)$ arithmetic operations, whereas the classical Gaussian algorithm, applied on the corresponding system of linear equations, has complexity $\mathcal{O}(\Vert\mathbf{n}\Vert^3)$ because it does not take into account the special structure of the matrix of coefficients. Finally, our method is also reliable, which in this context means that it also recognizes insoluble problems or gives representations if the solution sets of type $\mathbf{n}(\delta), \delta < 0$ are multidimensional (assuming that exact arithmetic is available). We remark that our algorithm does not consider the case of floating point arithmetic and hence does not consider the issue of numerical stability in the presence of roundoff errors.

Several fast algorithms for special cases of PHPAs are well known, but most of them require a normal or perfect solution table (i.e., PHPAs of different type are distinct). As far as we know, only the methods proposed in [19] for square matrix Padé approximation and the Jacobi–Perron continued fraction algorithms of [6] for simultaneous Padé approximation and [2], [4], [11], [12], [25] for scalar Hermite Padé approximation are also reliable. All of them still require slight assumptions on the input data ($A(0)$ regular, $\mathbf{F}(0)$ nontrivial); moreover, the algorithms of [11], [12], [19] might reach a complexity $\mathcal{O}(\Vert\mathbf{n}\Vert^3)$ if none of the subproblems of type $\mathbf{n}(\delta), \delta < 0$ has a unique solution.

For the special case $s = 1$ (i.e., scalar Hermite Padé approximation), the recurrence formula of our new algorithm is similar to that used in [2], [4], [25]. The *fast Gaussian algorithm* [2, §5] is motivated by the close connections to the factorization of the corresponding matrix of coefficients via the Gaussian algorithm with partial pivoting; a "special rule" reduces the complexity to $\mathcal{O}(\Vert\mathbf{n}\Vert^2)$. It provides solutions to all subproblems on the diagonal path $(\mathbf{n}(\delta))_{\delta\leq 0}$. The methods of [2], [4] both are developed for the more general M-Padé approximation problem (arbitrary interpolation knots); moreover, by the algorithm given in [4] we can compute solutions by recurrence on "arbitrary paths" or "staircases" $(\mathbf{n}_k)$ where the multiindex $\mathbf{n}_{k+1}$ differs from $\mathbf{n}_k$ by increasing one component (also the decreasing of a second component is allowed). Parallel to [2], [4], Van Barel and Bultheel proposed a fast, reliable method for computing Hermite Padé approximants on diagonal paths [25]. Their version is similar to [2] but notationally less complicated. The ideas developed in [26] for a recursive computation of vector M-Padé approximants have close connections to [4], [25]. The authors propose three alternative "basic steps" that include considerable freedom in solving certain subproblems.

There seems to be no connection between the methods described above and the reliable Jacobi–Perron continued fraction algorithm of [6] for simultaneous Padé approximation. For this approximation problem, using our formalism we obtain a more compact method with at most the same complexity; in addition, we get more information about singular cases.

Before describing bases for PHPA solution sets let us introduce the following definition.

DEFINITION 3.1 (defect, order). *The defect of a* $\mathbf{P} = (P_1, \ldots, P_m) \in \mathbb{K}^m[z]$ (*with respect to the fixed multiindex* $\mathbf{n} = (n_1, \ldots, n_m)$) *is*

$$\text{dct } \mathbf{P} := \min_l \ \{n_l + 1 - \deg P_l\},$$

where the zero polynomial has degree $-\infty$. *The order of* $\mathbf{P}$ (*with respect to* $s \in \mathbb{N}$ *and* $\mathbf{F}$) *is defined by*

$$\text{ord } \mathbf{P} := \sup\{\sigma \in \mathbb{N}_0 \,:\, \mathbf{P}(z^s) \cdot \mathbf{F}(z) = z^\sigma \cdot R(z) \quad \textit{with } R \in \mathbb{K}[[z]]\}.$$

The definition of the defect is a natural extension of that found in the case of the M-Padé problem (cf. [3], [4]) and its special case of rational interpolation (some authors use a slightly different definition). The defect is also closely connected to the τ-degree of [25].

Using Definition 3.1, we get an equivalent characterization for PHPA solution sets:

$$\text{(5)} \quad \text{For } \sigma \in \mathbb{N}_0, \quad \delta \in \mathbb{Z} \cup \{+\infty\} : \mathcal{L}_\delta^\sigma = \{\mathbf{P} \in \mathbb{K}^m[z] \,:\, \text{dct } \mathbf{P} > -\delta, \text{ord } \mathbf{P} \geq \sigma\}.$$

Now we are able to describe so-called σ-bases of PHPAs.

DEFINITION 3.2 (σ-bases). *Let* $\sigma \in \mathbb{N}_0$. *The system* $\mathbf{P}_1, \ldots, \mathbf{P}_m \in \mathbb{K}^m[z]$ *is called a* σ*-basis if and only if:*

(a) $\mathbf{P}_1, \ldots, \mathbf{P}_m \in \mathcal{L}_{+\infty}^\sigma$, *i.e.*, ord $\mathbf{P}_l \geq \sigma$.

(b) *For each* $\delta \in \mathbb{Z}\cup\{+\infty\}$ *and for each* $\mathbf{Q} \in \mathcal{L}_\delta^\sigma$ *there exists one and only one tuple of polynomials* $(\alpha_1, \ldots, \alpha_m)$, deg $\alpha_l <$ dct $\mathbf{P}_l + \delta$ *such that* $\mathbf{Q} = \alpha_1 \cdot \mathbf{P}_1 + \cdots + \alpha_m \cdot \mathbf{P}_m$.

Note that, as a consequence of Definition 3.2, a σ-basis $\mathbf{P}_1, \ldots, \mathbf{P}_m$ must be linearly independent with respect to polynomial coefficients. Moreover, we have

$$\text{(6)} \qquad \mathcal{L}_\delta^\sigma = \text{span}\left\{z^j \cdot \mathbf{P}_l \,:\, 1 \leq l \leq m, 0 \leq j < \text{dct } \mathbf{P}_l + \delta\right\},$$

$$\text{(7)} \qquad \dim \mathcal{L}_\delta^\sigma = \max\ \{\text{dct } \mathbf{P}_1 + \delta, 0\} + \cdots + \max\ \{\text{dct } \mathbf{P}_m + \delta, 0\}.$$

The existence of σ-bases for the case $s = 1$ was given in [2]–[4], [25] and for the case $s > 1$ in [5]. Before giving an algorithm for their computation, let us state some simple rules for the defect and order of linear combinations of PHPAs.

LEMMA 3.3. *For* $\mathbf{P}, \mathbf{Q} \in \mathbb{K}^m[z], c \in \mathbb{K} \setminus \{0\}$:

$$\text{dct } (c \cdot \mathbf{P}) = \text{dct } \mathbf{P}, \quad \text{dct } (\mathbf{P} + \mathbf{Q}) \geq \min \{\text{dct } \mathbf{P}, \text{dct } \mathbf{Q}\} \quad \text{dct } (z \cdot \mathbf{P}) = \text{dct } \mathbf{P} - 1, \tag{8}$$

$$\text{ord } (c \cdot \mathbf{P}) = \text{ord } \mathbf{P}, \quad \text{ord } (\mathbf{P} + \mathbf{Q}) \geq \min \{\text{ord } \mathbf{P}, \text{ord } \mathbf{Q}\}, \quad \text{ord } (z \cdot \mathbf{P}) = \text{ord } \mathbf{P} + s. \tag{9}$$

Proof. The proof is left to the reader. □

From the characterization (5), it is clear that $\mathcal{L}_\delta^\sigma \subset \mathcal{L}_{\delta+1}^\sigma$ and $\mathcal{L}_\delta^{\sigma+1} \subset \mathcal{L}_\delta^\sigma$. In addition, if $\mathbf{P} \in \mathcal{L}_\delta^\sigma \setminus \mathcal{L}_\delta^{\sigma+1}$, i.e. ord $\mathbf{P} = \sigma$, then from (8), (9) it is easy to see that for each $\mathbf{Q} \in \mathcal{L}_\delta^\sigma$ there exists a $c \in \mathbb{K}$ such that $\mathbf{Q} - c \cdot \mathbf{P} \in \mathcal{L}_\delta^{\sigma+1}$. This proves the statement

$$\mathcal{L}_\delta^{\sigma+1} \subset \mathcal{L}_\delta^\sigma, \qquad \dim \mathcal{L}_\delta^{\sigma+1} \geq \dim \mathcal{L}_\delta^\sigma - 1 \tag{10}$$

and already gives an idea about the computation of σ-bases by recurrence on the order as proposed in the procedure FPHPS (*fast power Hermite Padé solver*) below. We show in Theorem 3.4 that this method is both correct and produces the desired σ-bases.

FPHPS ALGORITHM
INPUT: $m \geq 2, s \in \mathbb{N}, \mathbf{F} = (f_1, \dots f_m)^T$, multiindex $\mathbf{n} = (n_1, \dots, n_m)$
INITIALIZATION: Let for $\sigma = 0, l = 1, \dots, m$:
  $d_{l,0} = n_l, \mathbf{P}_{l,0} = (0, \dots, 0, 1, 0, \dots, 0)$($l$th unit vector)
RECURSIVE STEP: For $\sigma = 0, 1, 2, \dots$:
  Let for $l = 1, \dots, m$: $c_{l,\sigma} = z^{-\sigma} \cdot \mathbf{P}_{l,\sigma}(z^s) \cdot \mathbf{F}(z)|_{z=0}$ and $\Lambda_\sigma = \{l \ : \ c_{l,\sigma} \neq 0\}$
  CASE $\Lambda_\sigma = \{\}$, then for $l = 1, \dots, m$:
    $\mathbf{P}_{l,\sigma+1} = \mathbf{P}_{l,\sigma}, d_{l,\sigma+1} = d_{l,\sigma}$
  CASE $\Lambda_\sigma \neq \{\}$, then let $\pi = \pi_\sigma \in \Lambda_\sigma$ be defined by
    $d_{\pi,\sigma} = \max \{d_{l,\sigma} \ : \ l \in \Lambda_\sigma\}$
    and compute for $l = 1, \dots, m$:
    $l \in \Lambda_\sigma, l \neq \pi$: $\mathbf{P}_{l,\sigma+1} = \mathbf{P}_{l,\sigma} - \frac{c_{l,\sigma}}{c_{\pi,\sigma}} \cdot \mathbf{P}_{\pi,\sigma}, d_{l,\sigma+1} = d_{l,\sigma}$
    $l \notin \Lambda_\sigma$: $\mathbf{P}_{l,\sigma+1} = \mathbf{P}_{l,\sigma}, d_{l,\sigma+1} = d_{l,\sigma}$
    $l = \pi$: $\mathbf{P}_{\pi,\sigma+1} = z \cdot \mathbf{P}_{\pi,\sigma}, d_{\pi,\sigma+1} = d_{\pi,\sigma} - 1$
OUTPUT: For $\sigma = 0, 1, 2, \dots$:
  σ-bases $\mathbf{P}_{1,\sigma}, \dots, \mathbf{P}_{m,\sigma}$ with dct $\mathbf{P}_{l,\sigma} = d_{l,\sigma} + 1, l = 1, \dots, m$, i.e.
  for all δ: $\mathcal{L}_\delta^\sigma = \{\alpha_1 \cdot \mathbf{P}_{1,\sigma} + \dots + \alpha_m \cdot \mathbf{P}_{m,\sigma} \ : \ \deg \ \alpha_l \leq d_{l,\sigma} + \delta\}$.

THEOREM 3.4 (Feasibility of method FPHPS). *Method* FPHPS *is well defined and gives the specified results.*

Proof. We show the assertion by induction on σ for a fixed δ.

The case $\sigma = 0$ follows immediately from the definition of $\mathcal{L}_\delta^0$. Hence, suppose $\sigma \geq 0$ and that the algorithm is correct for σ. We show that the algorithm produces the correct output for $\sigma + 1$. Note that by assumption ord $\mathbf{P}_{l,\sigma} \leq \sigma$, i.e., its s-residual takes the form

$$\mathbf{P}_{l,\sigma}(z^s) \cdot \mathbf{F}(z) = z^\sigma \cdot R_l(z) \quad \text{with } R_l \in \mathbb{K}[[z]].$$

Hence, $c_{l,\sigma} = R_l(0)$ and the recurrence step is well defined. By construction we have

$$\text{ord } \mathbf{P}_{l,\sigma+1} \geq \sigma + 1 \quad \text{and} \quad \text{dct } \mathbf{P}_{l,\sigma+1} \geq d_{l,\sigma+1} + 1.$$

Moreover, it is easy to see that with $\mathbf{P}_{1,\sigma}, \ldots, \mathbf{P}_{m,\sigma}$ also $\mathbf{P}_{1,\sigma+1}, \ldots, \mathbf{P}_{m,\sigma+1}$ are linearly independent with respect to polynomial coefficients.

Consider first the case when $\mathcal{L}_\delta^\sigma = \mathcal{L}_\delta^{\sigma+1}$. By assumption, each $\mathbf{Q} \in \mathcal{L}_\delta^{\sigma+1}$ then has a representation

$$\mathbf{Q} = \alpha_1 \cdot \mathbf{P}_{1,\sigma} + \cdots + \alpha_m \cdot \mathbf{P}_{m,\sigma}, \quad \deg\ \alpha_l < \operatorname{dct} \mathbf{P}_{l,\sigma} + \delta.$$

This is already a suitable linear combination of $\mathbf{P}_{1,\sigma+1}, \ldots, \mathbf{P}_{m,\sigma+1}$. To see this, note that with $\alpha_l \neq 0$, we get $\operatorname{dct} \mathbf{P}_{l,\sigma} + \delta > 0$ and $\mathbf{P}_{l,\sigma} \in \mathcal{L}_\delta^\sigma = \mathcal{L}_\delta^{\sigma+1}$; hence, $c_{l,\sigma} = 0$ and $\mathbf{P}_{l,\sigma+1} = \mathbf{P}_{l,\sigma}$.

The case where $\mathcal{L}_\delta^\sigma \neq \mathcal{L}_\delta^{\sigma+1}$ is also easy to handle. Let

$$\mathcal{L}_\delta := \{\alpha_1 \cdot \mathbf{P}_{1,\sigma+1} + \cdots + \alpha_m \cdot \mathbf{P}_{m,\sigma+1} \ : \ \deg\ \alpha_l < \operatorname{dct} \mathbf{P}_{l,\sigma+1} + \delta\},$$

so that in view of Lemma 3.3 we have $\mathcal{L}_\delta \subset \mathcal{L}_\delta^{\sigma+1}$. On the other hand, the dimension of $\mathcal{L}_\delta$ can be estimated as follows:

$$\begin{aligned} \dim\ L_\delta &= \max\ \{\operatorname{dct} \mathbf{P}_{1,\sigma+1} + \delta, 0\} + \cdots + \max\ \{\operatorname{dct} \mathbf{P}_{m,\sigma+1} + \delta, 0\}, \\ &\geq \max\ \{d_{1,\sigma+1} + 1 + \delta, 0\} + \cdots + \max\ \{d_{m,\sigma+1} + 1 + \delta, 0\}, \\ &\geq \max\ \{d_{1,\sigma} + 1 + \delta, 0\} + \cdots + \max\ \{d_{m,\sigma} + 1 + \delta, 0\} - 1, \\ &= \dim\ \mathcal{L}_\delta^\sigma - 1 = \dim\ \mathcal{L}_\delta^{\sigma+1}, \end{aligned}$$

where for the last two equalities we have applied (7) and (10). Consequently, $\mathcal{L}_\delta = \mathcal{L}_\delta^{\sigma+1}$, and we have equality in the estimation above. For all $\Delta \geq \delta$ we also have that $\mathcal{L}_\Delta^{\sigma+1} \neq \mathcal{L}_\Delta^\sigma$, since by definition $\emptyset \neq \mathcal{L}_\delta^\sigma \setminus \mathcal{L}_\delta^{\sigma+1} \subset \mathcal{L}_\Delta^\sigma \setminus \mathcal{L}_\Delta^{\sigma+1}$. Therefore, the above equations are also valid if we replace δ by $\Delta \geq \delta$. Choosing Δ sufficiently large, we can conclude that $\operatorname{dct} \mathbf{P}_{l,\sigma+1} = d_{l,\sigma+1} + 1$ for $l = 1, \ldots, m$ which proves the theorem. $\square$

4. Some properties of the FPHPS algorithm. In this section, we discuss some properties of the σ-bases obtained by the procedure FPHPS. In particular, we are interested in simple conditions describing whether some PHPAs are irreducible and whether given PHPAs $\mathbf{P}_1, \ldots, \mathbf{P}_\lambda$ (and its values at zero) are linearly independent with respect to polynomial coefficients (and constant coefficients, respectively)—questions that as explained in §2 naturally arise in the context of matrix Padé approximation. In addition, for multidimensional solution sets, we classify PHPAs having "best" approximation properties, i.e., maximal order and/or minimal degree. The complexity of method FPHPS is determined at the end of this section.

Let Λ_σ and π_σ (for a given σ) be defined as in the FPHPS algorithm. As given in all applications of Table 1, in the sequel we only discuss the case $s \leq m$ and $\Lambda_0 \neq \{\ \}, \ldots, \Lambda_{s-1} \neq \{\ \}$. This is equivalent to the fact that the matrix $(\mathbf{F}(0), \mathbf{F}'(0), \ldots, \mathbf{F}^{(s-1)}(0))$ has full rank. In Theorem 4.1, we summarize some facts about reducible PHPAs. These results are generalizations of ideas appearing in [25].

THEOREM 4.1. (a) *For all $\sigma \geq s$, we have* card $\Lambda_\sigma \geq 1$, *more precisely*

$$\begin{gathered} \pi_{\sigma-s} \in \Lambda_\sigma \subset L_\sigma \cup \{\pi_{\sigma-s}\}, \\ \textit{where } L_\sigma := \{1, \ldots, m\} \setminus \{\pi_{\sigma-s}, \pi_{\sigma-s+1}, \ldots, \pi_{\sigma-1}\}. \end{gathered} \tag{11}$$

(b) *Let U denote the $(m-s)$ dimensional subspace of vectors that are orthogonal to all $\mathbf{F}(0), \mathbf{F}'(0), \ldots, \mathbf{F}^{(s-1)}(0)$. Then for $\sigma \geq s$*

$$\operatorname{span}\ \{\mathbf{P}_{l,\sigma}(0) : l \in L_\sigma\} = U \quad \textit{and for all } l \notin L_\sigma : \ \mathbf{P}_{l,\sigma}(0) = 0. \tag{12}$$

Proof. Note that in the FPHPS algorithm we always have ord $\mathbf{P}_{\pi_\sigma,\sigma+1} = \sigma + s$. Therefore, $\pi_\sigma \in \Lambda_{\sigma+s} \neq \{\quad\}$ but $\pi_\sigma \notin \Lambda_{\sigma+1}, \ldots, \pi_\sigma \notin \Lambda_{\sigma+s-1}$. This proves (a). The second part of (b) follows directly from the fact that $\mathbf{P}_{l,\sigma+1} = \mathbf{P}_{l,\sigma}$ for all $l \notin L_{\sigma+1} \cup L_\sigma$, and $\mathbf{P}_{l,\sigma+1}(0) = 0$ for $l = \pi_\sigma$. The first assertion of (b) can be shown by a simple recurrence argument on $\sigma \geq s$. Let

$$U_\sigma := \text{span}\,\{\mathbf{P}_{l,\sigma}(0) \,:\, l \in L_\sigma\}.$$

Then $U_\sigma \subset U$ since we have ord $\mathbf{P}_{l,\sigma} \geq s$. With $\mathbf{P}_{l,\sigma}(0), l \in L_\sigma$, also the vectors $\mathbf{P}_{l,\sigma+1}(0), l \in L_\sigma \bigcap L_{\sigma+1}$, together with $\mathbf{P}_{\pi_\sigma,\sigma}(0)$ are linearly independent. From the recurrence relations we know that $\mathbf{P}_{\pi_{\sigma-s},\sigma}(0) = 0$ and $\mathbf{P}_{\pi_{\sigma-s},\sigma+1}(0) = c \cdot \mathbf{P}_{\pi_\sigma,\sigma}(0)$ with $c \neq 0$. Consequently, $\mathbf{P}_{l,\sigma+1}(0), l \in L_{\sigma+1}$, are linearly independent that proves part (b). □

Supposing that the vector $\mathbf{F}$ contains only polynomial entries, we expect that the solution set $\mathcal{L}_\delta^\sigma$ becomes stationary for sufficiently large σ. In contrast, due to Theorem 4.1(a) the σ-bases will always change if σ is increased. In fact, we observe that for large σ, the nonconstant part of the σ-basis described by the sets Λ_σ consists only of approximants with defect smaller than $-\delta$ and that, for sufficiently large σ, for the representation (6) of the solution set $\mathcal{L}_\delta^\sigma$ we need at most $(m - s)$ elements of the σ-basis.

Theorem 4.1(b) yields a simple criterion determining whether the solution set $\mathcal{L}_\delta^\sigma$ contains an irreducible element. By definition, the components of an element of the σ-basis can only have a common factor that vanishes at zero. Hence, there exists an element $\mathbf{P}$ of $\mathcal{L}_\delta^\sigma$ being irreducible, i.e., $\mathbf{P}(0) \neq 0$, if and only if there is an $l \in L_\sigma$ with $d_{l,\sigma} \geq -\delta$. Moreover, we immediately get the following corollary.

COROLLARY 4.2. (a) *$\mathcal{L}_\delta^\sigma$ contains $\lambda \leq m$ elements $\mathbf{P}_1, \ldots, \mathbf{P}_\lambda$ being linearly independent over $\mathbb{K}[z]$ if and only if there are distinct $l_1, \ldots, l_\lambda \in \{1, \ldots, m\}$ with $d_{l_j,\sigma} \geq -\delta$.*

(b) *$\mathcal{L}_\delta^\sigma$ contains $\lambda \leq m - s$ elements $\mathbf{P}_1, \ldots, \mathbf{P}_\lambda$ such that $\mathbf{P}_1(0), \ldots, \mathbf{P}_\lambda(0)$ are linearly independent over $\mathbb{K}$ if and only if there are distinct $l_1, \ldots, l_\lambda \in L_\sigma$ with $d_{l_j,\sigma} \geq -\delta$.*

(c) *In both cases linearly independent approximants from $\mathcal{L}_\delta^\sigma$ are given by $\mathbf{P}_j = \mathbf{P}_{l_j,\sigma}, j = 1, \ldots, \lambda$.*

In most applications, the first s components of $\mathbf{F}$ take the simple form $f_j(z) = z^{j-1}$. Here we consider the first s components $\mathbf{p}$ and the last $(m - s)$ components $\mathbf{q}$ of a PHPA $\mathbf{P} = (\mathbf{p}, \mathbf{q})$ separately and ask for approximants $\mathbf{P}_1, \ldots, \mathbf{P}_\lambda \in \mathcal{L}_\delta^\sigma$ with $\mathbf{q}_1, \ldots, \mathbf{q}_\lambda$ (or $\mathbf{q}_1(0), \ldots, \mathbf{q}_\lambda(0)$) being linearly independent. Here also the criteria given in Corollary 4.2.(a) and (b) can be applied as long as we can guarantee there is no $\mathbf{P} = (\mathbf{p}, \mathbf{q}) \in \mathcal{L}_\delta^\sigma$ with $\mathbf{p} \neq 0$ and $\mathbf{q} = 0$ ($\mathbf{p}(0) \neq 0$ and $\mathbf{q}(0) = 0$, respectively). But due to the simple form of $\mathbf{F}$ it can be easily verified that $\mathbf{P} = (\mathbf{p}, \mathbf{q}) \in \mathcal{L}_\delta^\sigma$ with $\mathbf{q}(0) = 0$ and $\sigma \geq s$ also implies that $\mathbf{p}(0) = 0$. Similarly, if $s \cdot (n_j + \delta) + j \leq \sigma$ for $j = 1, \ldots, s$ (which for the most interesting PHPA cases of §2 is true) and $\mathbf{P} = (P_1, \ldots, P_m) = (\mathbf{p}, 0) \in \mathcal{L}_\delta^\sigma$, then $\mathbf{p}$ must also be identical zero since ord $\mathbf{P} \leq \max\,\{s \cdot \deg P_j + j - 1 : j = 1, \ldots, s\}$.

If the solution set is multidimensional, we are interested in classifying particular solutions that have certain uniqueness properties. The concept of approximants with correct degree satisfying "best possible" order conditions is discussed in Corollary 4.3.

COROLLARY 4.3. *Let each $\mathbf{P} \in \mathbb{K}^m[z]$ have finite order and let $\delta + \min\,\{n_1, \ldots, n_m\} \geq 0$. Consider the problem of finding "optimal" PHPAs $\mathbf{P}_1, \ldots, \mathbf{P}_\lambda, \lambda \leq m - s$ with*

(i) $\mathbf{P}_1(0), \ldots, \mathbf{P}_\lambda(0)$ *are linearly independent,*
(ii) $\operatorname{dct} \mathbf{P}_1 > -\delta, \ldots, \operatorname{dct} \mathbf{P}_\lambda > -\delta$,
(iii) *the number* ($\operatorname{ord} \mathbf{P}_1 + \cdots + \operatorname{ord} \mathbf{P}_\lambda$) *is maximal,*
(iv) $\operatorname{ord} \mathbf{P}_1 =: \sigma(1) > \operatorname{ord} \mathbf{P}_2 =: \sigma(2) > \cdots > \operatorname{ord} \mathbf{P}_\lambda =: \sigma(\lambda)$

(*it is easy to see that condition* (iv) *only implies a particular ordering for the* PHPAs *determined by* (i), (ii), (iii)). *A solution for this problem is given by*

$$\sigma(j) := \max\{\sigma : \operatorname{card}\ \{l \in L_\sigma\ :\ d_{l,\sigma} \geq -\delta\} \geq j\}$$

and $\mathbf{P}_j = \mathbf{P}_{\pi_{\sigma(j)},\sigma(j)}, j = 1, \ldots, \lambda$.

Corollary 4.3 is a canonical generalization of the optimal Hermite Padé form of type $\mathbf{n}(\delta)$ of [22] ($s = \lambda = 1$). Paszkowski [22] speaks of nonexistent optimal Hermite Padé forms if $\mathbf{P}_1$ is not unique, i.e., if there is a further (necessarily reducible) PHPA $\mathbf{P}_0$ with $\operatorname{dct} \mathbf{P}_0 > -\delta$ and $\operatorname{ord} \mathbf{P}_0 > \operatorname{ord} \mathbf{P}_1$. For $\lambda = m - s = 1$, for example, scalar simultaneous (partial) Padé approximation, our approach is closely connected to a concept proposed by de Bruin [8] for nonnormal solution tables. Note that although in view of Corollary 4.2.(b), the numbers $\sigma(1), \ldots, \sigma(\lambda)$ are unique, we might get several tuples of optimal PHPAs being essentially different. The significance of the integer $\sigma(\lambda)$ for matrix Padé approximation is discussed at the end of §5.

Following Corollary 4.3, we always find irreducible approximants with correct degree, but the order condition might be weakened. In contrast, Van Barel and Bultheel [24], [26] look for irreducible approximants with correct order and a type of minimal degree. More precisely, instead of (ii)–(iv), the conditions

(v) $\operatorname{ord} \mathbf{P}_1 \geq \sigma, \ldots, \operatorname{ord} \mathbf{P}_\lambda \geq \sigma$,
(vi) the number ($\operatorname{dct} \mathbf{P}_1 + \cdots + \operatorname{dct} \mathbf{P}_\lambda$) is maximal

are imposed. As above, this problem will not generally have a unique solution. However, the method FPHPS also gives a solution for this problem: due to Corollary 4.2(b) we can take those λ approximants $\mathbf{P}_{l,\sigma}, l \in L_\sigma$ with maximal defect.

The problem of uniqueness for both concepts is illustrated in Example 4.4.

Example 4.4. Let

$$m = 4, \qquad s = 2, \qquad \mathbf{n} = (2,2,2,2), \qquad \delta = 0,$$
$$\mathbf{F}(z) = \left(1, z, \frac{z}{1 - z^4} + z^{10}, \frac{z}{1 + z^4} + z^{12}\right)^T + \mathcal{O}\left(z^{16}\right).$$

An application of FPHPS gives the values $\pi_0, \pi_1, \ldots, \pi_{13} = 1,2,1,2,1,3,1,3,1,4,2,4,3,4$. In particular, we obtain a σ-basis for $\sigma = 10$ (output in matrix form with the rows $\mathbf{P}_{1,10}, \mathbf{P}_{2,10}, \mathbf{P}_{3,10}$, and $\mathbf{P}_{4,10}$ as the basis elements) as

$$\mathbf{P}_{10}(z) = \begin{bmatrix} z^5 & 0 & 0 & 0 \\ 0 & z^2 & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 - z^2 & -\frac{1}{2} + z^2 & -\frac{1}{2} \\ 0 & -2z & z & z \end{bmatrix},$$

with s-residuals

$$\mathbf{P}_{10}(z^2) \cdot \mathbf{F}(z) = \begin{bmatrix} z^{10} + \mathcal{O}\left(z^{26}\right) \\ -\frac{z^{10}}{2} + \frac{z^{12}}{2} - z^{13} + \mathcal{O}\left(z^{16}\right) \\ -\frac{z^{10}}{2} - \frac{z^{12}}{2} + z^{13} + z^{14} + \mathcal{O}\left(z^{16}\right) \\ 2z^{11} + z^{12} + z^{14} + \mathcal{O}\left(z^{18}\right) \end{bmatrix}.$$

The defects for this basis are $-2, 1, 1$ and 2, respectively. Hence, $\mathcal{L}_0^{10}$ does not have dimension 2 (as expected from comparing the number of equations and unknowns) but 4. For $\lambda = 1$, a particular solution with "minimal degree" (satisfying conditions (i), (v), and (vi) above) is given by $a \cdot \mathbf{P}_{2,10} + b \cdot \mathbf{P}_{3,10} + (cz + d) \cdot \mathbf{P}_{4,10}$ with arbitrary constants $a, b, c, d, |a| + |b| \neq 0$. A particular solution with "maximal order" $\sigma(1) = 12$ (satisfying conditions (i), (ii), (iii), and (iv) above) is given by $a \cdot \mathbf{P}_{3,12} + b \cdot \mathbf{P}_{4,12} = a \cdot (\mathbf{P}_{3,10} - \mathbf{P}_{2,10}) + b \cdot z \cdot \mathbf{P}_{4,10}$ with arbitrary constants $a, b, a \neq 0$ (the solution proposed in Corollary 4.3 equals $\mathbf{P}_{3,12}$).

Consider now the problem of determining the complexity of the FPHPS algorithm. For simplicity, we still impose the conditions before Theorem 4.1 (otherwise, the complexity will be still smaller). As seen in §2, in most applications one must determine σ-bases of PHPAs for $\sigma \approx \|\mathbf{n}\|$. To determine the number of arithmetic operations (AO) required for the computation of a $\|\mathbf{n}\|$-basis, we essentially only have to take into account the computation of $c_{1,\sigma}, \ldots, c_{m,\sigma}$ and of $\mathbf{P}_{1,\sigma+1}, \ldots, \mathbf{P}_{m,\sigma+1}, 0 \leq \sigma < \|\mathbf{n}\|$. Here the complexity strongly depends on the parameters $\mathbf{F}$ and s.

THEOREM 4.5 (Complexity). *The* FPHPS *algorithm for computing* PHPAs *of order $\sigma = 0, 1, \ldots, \|\mathbf{n}\|$ has a complexity of at most*

$$4(m - s) \cdot \|\mathbf{n}\|^2 + \mathcal{O}\left(m^2 \cdot \|\mathbf{n}\|\right) \text{AO}, \tag{13}$$

roughly half additions and half multiplications plus $\mathcal{O}(m \cdot \|\mathbf{n}\|)$ divisions. At least for the case $\mathbf{n} = (n, \ldots, n)$, we obtain the sharper bound

$$\left(1 - \frac{s}{m}\right) \cdot (2m - \text{card } L) \cdot \|\mathbf{n}\|^2 + \mathcal{O}\left(m^2 \cdot \|\mathbf{n}\|\right) \text{AO}, \tag{14}$$

where $L = \{l : f_l(z) = z^j$ with a $j \in \mathbb{N}_0\}$.

Proof. Since $c_{\pi_{\sigma-s},\sigma} = c_{\pi_{\sigma-s},\sigma-s}$ and $P_{\pi_\sigma,\sigma+1}$ can be easily determined by shifting some coefficients, for the complexity it remains to consider the computation of at most $c_{l,\sigma}$ and $\mathbf{P}_{l,\sigma+1}$ for $l \in L_{\sigma+1}$. In addition, we are not interested in PHPAs with dct $\mathbf{P}_{l,\sigma} \leq 0$, since they do not occur in the solution sets $\mathcal{L}_\delta^\sigma, \delta \leq 0$ (cf. (6)). Therefore, the degree of the λth component of $\mathbf{P}_{l,\sigma}$ is bounded by $n_\lambda - d_{l,\sigma} \leq n_\lambda$ and we require for loop number σ the number of at most $2 \cdot \sum_{l \in L_{\sigma+1}} \sum_{\lambda=1}^{m} (n_\lambda + 1) + \mathcal{O}(m^2) = 2(m-s) \cdot \|\mathbf{n}\| + \mathcal{O}(m^2)$ additions/subtractions and the same number of multiplications that totally gives a complexity as stated in (13). For the case $\mathbf{n} = (n, \ldots, n)$, we can apply the relation

$$2 \cdot \sum_{\sigma=0}^{\|\mathbf{n}\|-1} \sum_{l \in L_{\sigma+1}} (d_{l,\sigma} + 1) \geq \cdots \geq \|\mathbf{n}\|^2 - 2 \cdot s \cdot \sum_{l=1}^{m} \sum_{j=0}^{n+1} j,$$

which by using similar arguments leads to (14). □

TABLE 2
Complexity for solving matrix-type Padé approximation problems.

Example	Complexity via (13)	Via (14), special case
Classical Hermite Padé	$4(m-1)\cdot\|\mathbf{n}\|^2+\mathcal{O}(m^2\cdot\|\mathbf{n}\|)$	for $n_1=\cdots=n_m$: $2(m-1)\cdot\|\mathbf{n}\|^2+\mathcal{O}(m^2\cdot\|\mathbf{n}\|)$
2.1	$4q[p(M+1)+q(N+1)]^2 +\mathcal{O}((p+q)^2\cdot(M+N))$	for $M=N$: $q(2q+p)(p+q)(M+1)^2 +\mathcal{O}((p+q)^2\cdot M)$
2.2	$4p[q(M+1)+p(N+1)]^2 +\mathcal{O}((p+q)^2\cdot(M+N))$	for $M=N$: $p(2p+q)(p+q)(M+1)^2 +\mathcal{O}((p+q)^2\cdot M)$
2.3	$4\mu p^3\rho^2+\mathcal{O}(\mu^2p^4\rho)$	for $\rho_0=\cdots=\rho_\mu$: $2\mu p^3\rho^2+\mathcal{O}(\mu^2p^4\rho)$
2.4 with $p=1, A_0=1$	$4\mu^2\rho^2+\mathcal{O}(\mu^4\rho)$	for $\rho_0=\cdots=\rho_\mu$: $\frac{\mu+2}{\mu+1}\mu^2\rho^2+\mathcal{O}(\mu^4\rho)$

It should be mentioned that our algorithm can be implemented very efficiently on a vector or on a parallel processor (with, e.g., m or $\|\mathbf{n}\|$ processors). The complexity of our algorithm for the examples of §2 is given in Table 2, whereas in Table 3 some solved subproblems and their corresponding PHPA solution space are listed.

5. An example of matrix Padé approximation. In this section, we give an example of a matrix Padé approximation problem computed using the FPHPS algorithm. Let

$$A(z)=\begin{bmatrix}1+z^2+2z^4-z^5+z^6+\mathcal{O}\left(z^8\right) & z^7+\mathcal{O}\left(z^8\right)\\ -z^5+\mathcal{O}\left(z^8\right) & 1+z^2+z^4+z^7+\mathcal{O}\left(z^8\right)\end{bmatrix}$$

and consider the problem of determining a (2, 3) right-hand matrix Padé form for $A(z)$. Thus, we are looking for 2×2 matrix polynomials P and Q of degree at most 2 and 3, respectively, such that

$$A(z)\cdot Q(z)-P(z)=z^6\cdot R(z)$$

for some matrix power series R. The suitable choice of the parameters is stated in Table 1, row 2. Note that, for any PHPA (P_1,P_2,P_3,P_4) of type $(M,M,N,N),2(M+N+1),2)$, the components P_1 and P_2 correspond to a column of an (M,N) right-hand matrix Padé numerator, whereas P_3,P_4 correspond to a column of the denominator (cf. Table 3, for left-hand matrix Padé approximation, P_1,P_2 and P_3,P_4 correspond to rows of numerator and denominator, respectively).

Setting $s=2, \mathbf{n}=(2,2,3,3)$ and

$$\begin{aligned}\mathbf{F}^T(z)&=[1,z]\cdot\left[\mathbf{I},\,-A\left(z^2\right)\right]\\ &=\left[1,z,-1-z^4-2z^8+z^{10}+z^{11}-z^{12}+\mathcal{O}\left(z^{16}\right),\right.\\ &\qquad\left.-z-z^5-z^9-z^{14}-z^{15}+\mathcal{O}\left(z^{16}\right)\right]\end{aligned}$$

and using the FPHPS algorithm gives a σ-basis for $\sigma=12$ (output in matrix form with the rows as the basis elements) as

$$\begin{bmatrix}-z-z^2+z^3 & 0 & -z-z^2+2z^3+z^4 & 0\\ 1+z-z^2 & z^3 & 1+z-2z^2-z^3 & z^3\\ -z^2 & 0 & -z^2+z^4 & 0\\ 0 & -1 & 0 & -1+z^2\end{bmatrix}.$$

TABLE 3

Some matrix-type Padé subproblems solved by FPHPS *and their corresponding* PHPA *solution spaces, parameters* $m, \mathbf{n}, s, \sigma, \mathbf{F}$ *as in Table* 1.

Example	Type of subproblem	PHPA solution space
Classical Hermite Padé	$(n_1 - j, \ldots, n_m - j), j \leq \min\{n_l + 1\}$	$\mathcal{L}_{-j}^{\sigma - j \cdot m}$
2.1	$(M - j, N - j), j \leq \min\{M + 1, N + 1\}$	$\mathcal{L}_{-j}^{\sigma - 2 \cdot j \cdot s}$ contains rows of (P^T, Q^T)
2.2	$(M - j, N - j), j \leq \min\{M + 1, N + 1\}$	$\mathcal{L}_{-j}^{\sigma - 2 \cdot j \cdot s}$ contains rows of (P, Q)
2.3	$(\rho_0 - j, \ldots, \rho_\mu - j), j \leq \min\{\rho_l\}$	$\mathcal{L}_{-j}^{\sigma - j \cdot m}$ contains rows of $(P_0, \ldots, P_\mu)$
2.4 with $p = 1, A_0 = 1$	$(\rho_0 - j, \ldots, \rho_\mu - j), j \leq \min\{\rho_l\}$	$\mathcal{L}_{-j \cdot s}^{\sigma - j \cdot s \cdot m}$

The defects for this basis are 0, 0, 0, and 2, respectively. Therefore, a basis for the solution space $\mathcal{L}_0^{12}$, as a finite-dimensional space over $\mathbb{K}$, is given by $(a + b \cdot z) \cdot \mathbf{P}_{4,12} = (a + b \cdot z) \cdot [0, -1, 0, -1 + z^2]$, with a and b being arbitrary constants. Translating the solution space basis into matrix form implies that the columns of P and Q are generated by

$$(a + bz) \cdot \begin{bmatrix} 0 \\ -1 \end{bmatrix} \quad \text{and} \quad (a + bz) \cdot \begin{bmatrix} 0 \\ -1 + z^2 \end{bmatrix},$$

respectively. This gives a right matrix Padé form of type (2, 3) for $A(z)$ as

$$P(z) = \begin{bmatrix} 0 & 0 \\ -1 & -z \end{bmatrix} \quad \text{and} \quad Q(z) = \begin{bmatrix} 0 & 0 \\ -1 + z^2 & -z + z^3 \end{bmatrix}.$$

In this case, such a matrix Padé form is unique up to multiplication on the right by a nonsingular 2×2 matrix. In particular, note that it is not possible to construct a right matrix Padé fraction of type (2, 3) in this instance.

The left matrix Padé forms of type (2, 3) for $A(z)$ can also be computed by the FPHPS procedure. Setting $s = 2, \mathbf{n} = (2, 2, 3, 3)$ and

$$\mathbf{F}(z) = \begin{bmatrix} \mathbf{I} \\ -A\left(z^2\right) \end{bmatrix} \cdot \begin{bmatrix} 1 \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ z \\ -1 - z^4 - 2z^8 + z^{10} + \mathcal{O}\left(z^{12}\right) \\ -z - z^5 - z^9 + z^{10} + \mathcal{O}\left(z^{12}\right) \end{bmatrix},$$

and computing the σ-basis for $\sigma = 12$ gives

$$\begin{bmatrix} -1 + z^2 & 1 & -1 + 2z^2 & 1 - z^2 \\ 0 & z^4 & 0 & z^4 \\ -z & -1 & -z + z^3 & -1 + z^2 \\ 0 & -z & 0 & -z + z^3 \end{bmatrix}.$$

In this case the defects are $1, -1, 1$, and 1, respectively, so the solution space $\mathcal{L}_0^{12}$ is of the form $a \cdot \mathbf{P}_{1,12} + b \cdot \mathbf{P}_{3,12} + c \cdot \mathbf{P}_{4,12}$ with a, b, and c arbitrary constants. Again translating the basis information to matrix form implies that the rows of P and Q are generated by

$$a \cdot \left[-1 + z^2, 1\right] + b \cdot [-z, -1] + c \cdot [0, -z]$$

and

$$a \cdot \left[-1+2z^2, 1-z^2\right] + b \cdot \left[-z+z^3, -1+z^2\right] + c \cdot \left[0, -z+z^3\right],$$

respectively. Unlike the previous example, there is not one Padé form that is unique up to left multiplication by a nonsingular matrix of scalars. One possibility for a left matrix Padé form in this case is

$$P(z) = \begin{bmatrix} -z & -1 \\ -1+z^2 & 1 \end{bmatrix} \quad \text{and} \quad Q(z) = \begin{bmatrix} -z+z^3 & -1+z^2 \\ -1+2z^2 & 1-z^2 \end{bmatrix}.$$

Note that the denominator has a nonzero determinant, indeed that $Q(0)$ is nonsingular. Therefore, unlike the case for approximants on the right, one can always form the rational expression $Q(z)^{-1} \cdot P(z)$.

Using the FPHPS algorithm in the above example also determines, at no added cost, the σ-basis for $\mathcal{L}^4_{-2}$ and $\mathcal{L}^8_{-1}$. Hence, the right matrix Padé forms of type (0, 1)

$$P(z) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad Q(z) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and (1, 2)

$$P(z) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad Q(z) = \begin{bmatrix} -1+z^2 & 0 \\ 0 & -1+z^2 \end{bmatrix}$$

(determined uniquely up to matrix multiplication on the right in both cases) are byproducts of the previous computation. In addition, one can continue the computation to determine the matrix Padé form of type (3, 4) since the σ-basis for $\mathcal{L}^{12}_0$ can be used to determine the σ-basis for $\mathcal{L}^{16}_1$. In the case of the right matrix Padé form of type (3, 4) this gives (again unique up to matrix multiplication on the right)

$$P(z) = \begin{bmatrix} 0 & 1+z/5+11/5z^2+4/5z^3 \\ -z^2 & 1/5-2/5z+z^3 \end{bmatrix},$$

$$Q(z) = \begin{bmatrix} 0 & 1+z/5+6/5z^2+3/5z^3-16/5z^4 \\ -z^2+z^4 & -1/5-2/5z+1/5z^2+7/5z^3 \end{bmatrix},$$

an example where the denominator matrix polynomial Q is nonsingular but has a singular leading term $Q(0)$.

Our example shows that, in general, the matrix Padé approximation problem does not have a unique rational solution as in the scalar case. Moreover, there are three distinct and possible forms of a denominator matrix polynomial Q. First, the case occurs when $Q(z)$ is singular for all z and hence no matrix rational form exists; this type of degeneracy is not found in the scalar case. Second, it is possible that $Q(0)$ is nonsingular (cf. Corollary 4.2(a) and (b) and the following remarks). Here we can form $P(z) \cdot Q(z)^{-1}$ and its matrix power series agrees with $A(z)$ to the full order condition. Finally, if $Q(z)$ is nonsingular for some z, but $Q(0)$ is singular, we can cancel P and Q by a common matrix polynomial factor on the right. Here, similarly to the degenerate case found in scalar Padé approximation, the resulting matrix rational form $P(z) \cdot Q(z)^{-1}$ does not agree anymore with $A(z)$ to the full order condition.

Note that the concept as proposed in Corollary 4.3 ($2\lambda = 2s = m$) always leads to a matrix Padé-like form with correct degree and maximal order $[\sigma(s)/s]$ (perhaps less than $(M+N+1)$ as required for matrix Padé approximants) where by forming the rational function $P(z) \cdot Q(z)^{-1}$ we do not obtain an additional order deflation. In fact one can show that there is no other rational function of the form $P(z) \cdot Q(z)^{-1}$ satisfying the degree constraints and having an order greater than $[\sigma(s)/s]$.

6. A superfast PHPA solver. In §4 we have shown that the FPHPS algorithm computes a σ-basis with quadratic complexity. This is better than using methods such as Gaussian elimination and is optimal in special cases for arbitrary fields $\mathbb{K}$. However, when the field $\mathbb{K}$ allows for fast polynomial multiplication via the use of the FFT (cf. [15]), then there are faster methods in special cases. For example, when $s = 1$ and $m = 2$ (i.e., the case of Padé approximation) the algorithms of Brent, Gustavson, and Yun [9] and Cabay and Choi [10] compute these approximants with the superfast complexity $\mathcal{O}(\sigma \log^2 \sigma)$. Similarly, a recent algorithm of Cabay and Labahn [12] also solves the Hermite Padé and simultaneous Padé problems with superfast complexity. In this section we describe a second algorithm that takes advantage of fast polynomial multiplication when solving the PHPA problem. The new algorithm has the advantage of always being superfast—the algorithm of [12] sometimes slows down to quadratic or even cubic complexity (if most of the subproblems of type $\mathbf{n}(\delta), \delta < 0$ do not have a unique solution), although in practical problems this is rare.

The FPHPS algorithm of §3 provides a σ-basis $\mathbf{P}_1, \ldots, \mathbf{P}_m$ with respect to given $\mathbf{F}$, $\mathbf{n}$, and σ (and a fixed parameter s). For convenience, we arrange the $\mathbf{P}_l = (P_{l,1}, \ldots, P_{l,m})$ in a matrix

$$\mathbf{P} = (P_{l,\lambda})_{l=1,\ldots,m}^{\lambda=1,\ldots,m} .$$

Then with $\mathbf{d} := (d_1, \ldots, d_m), d_l :=$ dct $\mathbf{P}_l - 1$, we can symbolize the procedure as follows

$$(\mathbf{P}, \mathbf{d}) \leftarrow \text{FPHPS } (\mathbf{F}, \sigma, \mathbf{n}).$$

Note that, in general, the choice of π_σ and therefore the output of FPHPS is not unique, but uniqueness could be easily obtained, for instance, by the additional restriction that π_σ must be as small as possible.

The basic step of a divide-and-conquer version is described in Theorem 6.1.

THEOREM 6.1. *Let ρ, σ be integers with $0 \leq \rho \leq \sigma$. Suppose that we have iterated $\rho \leq \sigma$ times the recursive step of* FPHPS

$$(\mathbf{P}^{(1)}, \mathbf{d}^{(1)}) \leftarrow \text{FPHPS } (\mathbf{F}, \rho, \mathbf{n}),$$

and then continue iterating

$$(\mathbf{P}^{(3)}, \mathbf{d}^{(3)}) \leftarrow \text{FPHPS } (\mathbf{F}, \sigma, \mathbf{n}).$$

Suppose further that we restart the procedure with new initializations

$$\left(\mathbf{P}^{(2)}, \mathbf{d}^{(2)}\right) \leftarrow \text{FPHPS } \left(\mathbf{F}^{(1)}, \sigma - \rho, \mathbf{d}^{(1)}\right), \quad \textit{where } \mathbf{F}^{(1)}(z) := z^{-\rho} \cdot \mathbf{P}^{(1)}(z^s) \cdot \mathbf{F}(z),$$

where we always use the above uniqueness condition for the values π_σ. Then

$$\mathbf{P}^{(3)} = \mathbf{P}^{(2)} \cdot \mathbf{P}^{(1)} \quad \textit{and} \quad \mathbf{d}^{(3)} = \mathbf{d}^{(2)}.$$

Proof. We show Theorem 6.1 by induction on $(\sigma - \rho)$. Extending our notation slightly, set

$$\left(\mathbf{P}_{\rho}^{(1)}, \mathbf{d}_{\rho}^{(1)}\right) = \left(\mathbf{P}^{(1)}, \mathbf{d}^{(1)}\right), \qquad \left(\mathbf{P}_{\sigma}^{(3)}, \mathbf{d}_{\sigma}^{(3)}\right) = \left(\mathbf{P}^{(3)}, \mathbf{d}^{(3)}\right),$$
$$\left(\mathbf{P}_{\sigma-\rho}^{(2)}, \mathbf{d}_{\sigma-\rho}^{(2)}\right) = \left(\mathbf{P}^{(2)}, \mathbf{d}^{(2)}\right).$$

Note that Theorem 6.1 is trivially true for $\sigma - \rho = 0$. Assume now that the result is true for $\sigma - \rho \geq 0$. Then

$$\mathbf{P}_{\sigma}^{(3)} = \mathbf{P}_{\sigma-\rho}^{(2)} \cdot \mathbf{P}_{\rho}^{(1)} \quad \text{and} \quad \mathbf{d}_{\sigma}^{(3)} = \mathbf{d}_{\sigma-\rho}^{(2)}. \tag{15}$$

Consequently, the corresponding s-residuals

$$\mathbf{R}_{\sigma}^{(3)}(z) = z^{-\sigma} \cdot \mathbf{P}_{\sigma}^{(3)}(z^s) \cdot \mathbf{F}(z) \quad \text{and} \quad \mathbf{R}_{\sigma-\rho}^{(2)}(z) = z^{-\sigma+\rho} \cdot \mathbf{P}_{\sigma-\rho}^{(2)}(z^s) \cdot \mathbf{F}^{(1)}(z)$$

are equal. Hence, in both cases we must take the same value π and the assertion (15) with σ replaced by $(\sigma + 1)$ follows. □

The basic step of a divide-and-conquer version (15) yields the *superfast power Hermite Padé solver* (SPHPS), a reliable algorithm for computing a σ-basis of PHPAs with complexity $\mathcal{O}(\sigma \cdot \log^2 \sigma)$. The reason for the improvement in complexity results from the use of fast Fourier transform (FFT) techniques for fast polynomial multiplication. Such techniques consist of converting to a new coordinate representation via polynomial evaluation at roots of unity, computing the arithmetic operations in these new coordinates and transfering the results back to the original computation domain via polynomial interpolation. For purposes of efficiency we describe our superfast algorithm in both coordinate representations. Hence, we require some FFT details needed for our implementation. Additional details of the FFT procedure can be found in many texts (cf. [15]).

Let ω_κ be the principal κth root of unity (e.g., if $\mathbb{K}$ is the complex numbers, then $\omega_\kappa := \cos(2\pi/\kappa) + i \cdot \sin(2\pi/\kappa)$) and let

$$(\xi_j)_{j=0,\ldots,2\kappa-1} \leftarrow \mathrm{DFT}_{2\kappa}(p(z))$$

denote the evaluation of $\xi_j := p(\omega_{2\kappa}^j)$, $j = 0, \ldots, 2\kappa - 1$. Then for the classical discrete FFT algorithm, we split p into its even and odd part $p(z) = p_e(z^2) + z \cdot p_o(z^2)$ and use the fact that for $j = 0, \ldots, \kappa - 1$ we have

$$\xi_j = \xi_j^{(e)} + \omega_{2\kappa}^j \cdot \xi_j^{(o)} \quad \text{and} \quad \xi_{\kappa+j} = \xi_j^{(e)} - \omega_{2\kappa}^j \cdot \xi_j^{(o)},$$

where $(\xi_j^{(e)})_{j=0,\ldots,\kappa-1} \leftarrow \mathrm{DFT}_\kappa(p_e(z))$ and $(\xi_j^{(o)})_{j=0,\ldots,\kappa-1} \leftarrow \mathrm{DFT}_\kappa(p_o(z))$. The "inverse" polynomial interpolation computation of

$$p(z) \leftarrow \mathrm{IDFT}_\kappa\left((\xi_j)_{j=0,\ldots,\kappa-1}\right),$$

i.e. of the uniquely defined polynomial p of degree less than κ with $\xi_j := p(\omega_\kappa^j)$, $j = 0, \ldots, \kappa - 1$, is done by

$$\hat{p}(z) := \sum_{j=0}^{\kappa-1} \xi_j z^{n-j}, (\hat{\xi}_j)_{j=0,\ldots,\kappa-1} \leftarrow \mathrm{DFT}_\kappa(\hat{p}(z)), \qquad \text{then } p(z) = \frac{1}{\kappa} \cdot \sum_{j=0}^{\kappa-1} \hat{\xi}_j z^j.$$

Polynomials are multiplied componentwise in the new coordinates, that is, if

$$\left(\xi_j^{(1)}\right)_{j=0,\ldots,\kappa-1} \leftarrow \mathrm{DFT}_\kappa\left(p_1(z)\right),$$
$$\left(\xi_j^{(2)}\right)_{j=0,\ldots,\kappa-1} \leftarrow \mathrm{DFT}_\kappa\left(p_2(z)\right), \quad \text{and}$$
$$p^{(3)}(z) \leftarrow \mathrm{IDFT}_\kappa\left(\left(\xi_j^{(1)}\cdot\xi_j^{(2)}\right)_{j=0,\ldots,\kappa-1}\right),$$

and if c denotes the leading coefficient of $p_1(z)\cdot p_2(z)$, then

$$\begin{aligned} p^{(3)}(z) &= p^{(1)}(z)\cdot p^{(2)}(z) \bmod z^\kappa, \\ &= p^{(1)}(z)\cdot p^{(2)}(z) + \begin{cases} 0 & \text{if } \deg\left(p^{(1)}\cdot p^{(2)}\right) < \kappa, \\ -c\cdot z^\kappa + c & \text{if } \deg\left(p^{(1)}\cdot p^{(2)}\right) = \kappa. \end{cases} \end{aligned}$$

For κ a power of two, the complexity of converting to the new coordinate representation and back again (via either DFT_κ or IDFT_κ) is at most $\frac{1}{2}\cdot\kappa\cdot\log\kappa + \mathcal{O}(\kappa)$ multiplications and $\kappa\cdot\log\kappa + \mathcal{O}(\kappa)$ additions (the logarithm taken with respect to the basis 2). Therefore, the polynomial multiplication is of complexity $\mathcal{O}(\kappa\cdot\log\kappa)$.

In the SPHPS algorithm, we use the notations $\mathbf{e}_l$ is the lth unit vector, $\mathbf{I}$ the unit matrix of size $(m\times m)$, and $\{\sum_{j=0}^\infty c_j z^j\}_\kappa := \sum_{j=0}^{\kappa-1} c_j z^j$ denotes a truncated power series.

SPHPS ALGORITHM $(\mathbf{F},\sigma,\kappa,\mathbf{n})$
INPUT: $\sigma,\kappa\in\mathbb{N}_0$, with $\sigma\le\kappa=2^k$ for a $k\in\mathbb{N}_0$,
$\mathbf{n}=(n_1,\ldots,n_m)$, vector of integers,
$\mathbf{F}=(f_1,\ldots,f_m)^T$ vector of truncated power series,
i.e., of polynomials of degree less than κ,
Let $\mathbf{G}\in\mathbb{K}^{m\times s}[z]$ be defined by $\mathbf{F}(z)=\mathbf{G}(z^s)\cdot(1,z,\ldots,z^{s-1})^T$
OUTPUT: $\mathbf{P},\xi$, and $\mathbf{d}$ where:
$\mathbf{d}=(d_1,\ldots,d_m)$, vector of integers,
$\mathbf{P}=(P_{l,\lambda})_{l=1,\ldots,m}^{\lambda=1,\ldots,m}$, consisting of rows
$\mathbf{P}_l=(P_{l,1},\ldots,P_{l,m})$ with dct $\mathbf{P}_l=d_l+1$,
$\deg P_{l,l}\le\kappa$ and for $l\ne\lambda: \deg P_{l,\lambda}<\kappa$,
for all $\delta\in\mathbb{Z}: \mathcal{L}_\delta^\sigma=\{\alpha_1\mathbf{P}_1+\cdots+\alpha_m\mathbf{P}_m : \deg\alpha_l\le d_l+\delta\}$
$\xi=(\xi_j)_{j=0,\ldots,2\kappa-1}$, each ξ_j an m by m matrix,
$(\xi_j)_{j=0,\ldots,2\kappa-1}\longleftarrow\mathrm{DFT}_{2\kappa}(\mathbf{P}(z))$

THE RECURSION
CASE ($\sigma=0$ and $\kappa\ge1$) or ($\sigma=\kappa=1$ and $f_1(0)=\cdots=f_m(0)=0$):
RETURN $(\mathbf{P},\xi,\mathbf{d})=(\mathbf{I},\underbrace{(\mathbf{I},\ldots,\mathbf{I})}_{2\kappa},\mathbf{n})$

CASE $\sigma=\kappa=1, f_\pi(0)\ne0$ and for all l with $n_l>n_\pi:\ f_l(0)=0$:

$$P\leftarrow\begin{bmatrix} 1 & & & -f_1(0)/f_\pi(0) & & & \\ & \ddots & & \vdots & & & \\ & & 1 & -f_{\pi-1}(0)/f_\pi(0) & & & \\ & & & z & & & \\ & & & -f_{\pi+1}(0)/f_\pi(0) & 1 & & \\ & & & \vdots & & \ddots & \\ & & & -f_m(0)/f_\pi(0) & & & 1 \end{bmatrix}$$

RETURN $(\mathbf{P}, \xi, \mathbf{d}) = (\mathbf{P}, (\mathbf{P}(1), \mathbf{P}(-1)), \mathbf{n} - \mathbf{e}_\pi)$

CASE $\sigma \geq 1$ and $\kappa > 1$: (Divide-and-conquer step)

Compute basis to order $\kappa/2$:

$\overline{\kappa} \longleftarrow \kappa/2, \overline{\sigma} \longleftarrow \min\{\sigma, \overline{\kappa}\}; \mathbf{F}^{(1)}(z) \longleftarrow \{\mathbf{F}(z)\}_{\overline{\kappa}}$

$(\mathbf{P}^{(1)}, \xi^{(1)}, \mathbf{d}^{(1)}) \longleftarrow$ SPHPS $(\mathbf{F}^{(1)}, \overline{\sigma}, \overline{\kappa}, \mathbf{n})$

Compute basis to order κ:

$(\eta_j)_{j=0,\ldots,\kappa-1} \longleftarrow \mathrm{DFT}_\kappa(\mathbf{G}(z))$

$\mathbf{G}^{(2)}(z) \longleftarrow \mathrm{IDFT}_\kappa((\xi_j^{(1)} \cdot \eta_j)_{j=0,\ldots,\kappa-1})$

$\mathbf{F}^{(2)}(z) \longleftarrow \{z^{-\overline{\kappa}} \cdot \mathbf{G}^{(2)}(z^s) \cdot (1, z, \ldots, z^{s-1})^T\}_{\overline{\kappa}}$

$(\mathbf{P}^{(2)}, \xi^{(2)}, \mathbf{d}^{(2)}) \longleftarrow$ SPHPS $(\mathbf{F}^{(2)}, \sigma - \overline{\sigma}, \overline{\kappa}, \mathbf{d}^{(1)})$

Combine both parts:

$\xi_{2j}^{(3)} \longleftarrow \xi_j^{(2)} \cdot \xi_j^{(1)}$ for $j = 0, 1, \ldots, \kappa - 1$

$\mathbf{P}^{(3)}(z) \longleftarrow \mathrm{IDFT}_\kappa((\xi_{2j}^{(3)})_{j=0,\ldots,\kappa-1})$

If $\deg P_{l,l}^{(1)} = \deg P_{l,l}^{(2)} = \overline{\kappa}$, then $P_{l,l}^{(3)}(z) \longleftarrow P_{l,l}^{(3)}(z) - 1 + z^\kappa$

$(\xi_{2j+1}^{(3)})_{j=0,1,\ldots,\kappa-1} \longleftarrow \mathrm{DFT}_\kappa(\mathbf{P}^{(3)}(\omega_{2\kappa} \cdot z))$

RETURN $(\mathbf{P}, \xi, \mathbf{d}) = (\mathbf{P}^{(3)}, \xi^{(3)}, \mathbf{d}^{(2)})$

Consider now the problem of determining the complexity of the SPHPS algorithm. For simplicity, we still impose the conditions before Theorem 4.1 (otherwise, the complexity will be still smaller). As in §2, in most applications one must determine σ-bases of PHPAs for $\sigma \approx \|\mathbf{n}\|$.

THEOREM 6.2 (Complexity). *The* SPHPS *algorithm for computing* PHPAs *of order σ has a complexity of at most*

$$\frac{3}{2} \cdot (m+s) \cdot m \cdot \sigma \cdot \log^2 \sigma + \mathcal{O}(\sigma \cdot \log \sigma)\,\mathrm{AO}, \tag{16}$$

roughly half multiplications as additions.

Proof. Let $\Phi_A(\kappa)$ and $\Phi_M(\kappa)$ denote the number of additions/subtractions and multiplications/divisions required for the SPHPS algorithm with parameter κ, respectively. We easily obtain $\Phi_A(1) \leq 1$ and $\Phi_M(1) \leq m - 1$. Moreover, in the last case we call the subroutines DFT_κ or IDFT_κ at most $2(m+s)m$ times; hence,

$$\Phi_A(\kappa) \leq 2 \cdot \Phi_A(\kappa/2) + 2 \cdot (m+s) \cdot m \cdot \kappa \cdot \log \kappa + \mathcal{O}(\kappa)$$

and

$$\Phi_M(\kappa) \leq 2 \cdot \Phi_M(\kappa/2) + (m+s) \cdot m \cdot \kappa \cdot \log \kappa + \mathcal{O}(\kappa).$$

This gives the complexity result. □

We remark that, as was the case with method FPHPS, the complexity will be even less for some special cases. For example, for simultaneous Padé approximation, this number will be smaller if one carefully checks whether some entries of the matrix $\mathbf{G}$ always equal zero or 1.

7. Conclusions. In this paper we have studied the concept of a power Hermite Padé approximant. These approximants are shown to generalize a number of Padé approximation problems, including, for example, the classical Hermite Padé and simultaneous Padé approximation problems as well as matrix-type generalizations of common Padé approximation problems. A fast (and also a superfast), reliable algorithm to compute these approximants is given. In this way our work provides a uniform method of both describing and computing a wide variety of Padé and matrix-Padé approximation problems. As an immediate application, our work results in new and faster algorithms for a number of problems that rely on matrix-type Padé computation. For example, our algorithms, used in conjunction with the results of [17], gives faster algorithms for the inversion of striped or layered block Hankel (or Toeplitz) matrices. Similarly, the same algorithms combined with the results of [18] give similar improvements for the inversion of rectangular-block Hankel (or Toeplitz) matrices.

There are a number of directions for new research in this area. Our algorithm follows an m-dimensional "diagonal" path. In special cases, however, fast, reliable algorithms are given (cf. [4]) that can succeed on arbitrary staircase paths in m-dimensional space. The methods of [4] could also be extended to compute the more general PHPAs on arbitrary staircase paths, leading to a method with smaller complexity (cf. [26]).

Our algorithm does not consider the problem of stability when the computations are to be done with floating point numbers. Recently, Cabay and Meleshko [13] presented a (weakly) stable algorithm for the case $s = 1$ and $m = 2$. We conjecture that such an algorithm is also possible for the PHPA problem with arbitrary s and m, though not necessarily using the same approach as used in this paper. Our algorithm assumes exact arithmetic and has been implemented in the Maple computer algebra system. However, it does not consider the problem of exponential growth of the coefficients resulting in our computations. It would be interesting to extend our algorithm to this case. This would be done by restricting $\mathbb{K}$ to be an integral domain rather than a field and perhaps using fraction-free methods similar to those used for solving polynomial greatest common divisor (gcd) problems (cf. [15]).

Finally, the concept of a PHPA is a scalar generalization of a Hermite Padé approximant used to solve matrix-like Padé approximation problems. For example, as shown in [5] this concept also allows for a description of the structures in a singular PHPA solution table by adapting the scalar techniques of [3]. For matrix-like rational interpolation problems (with arbitrary knots), a common framework is given by the vector M-Padé approximation as a canonical extension of Example 2.5 (see [26]). In contrast, we are interested in a scalar generalization of the M-Padé approximant that can be used for simple, fast, and efficient algorithms and that, following [3], [5], might also be helpful for obtaining results about the structure of the singular matrix rational interpolation table.

REFERENCES

[1] G. A. BAKER AND P. R. GRAVES-MORRIS, *Padé Approximants, Part* II, Addison-Wesley, Reading, MA, 1981.

[2] B. BECKERMANN, *Zur Interpolation mit polynomialen Linearkombinationen beliebiger Funktionen*, Ph.D. thesis, Dept. of Mathematics, University of Hannover, Germany, 1990.

[3] ———, *The structure of the singular solution table of the M-Padé approximation problem*, J. Comput. Appl. Math., 32 (1990), pp. 3–15.

[4] ———, *A reliable method for computing M-Padé approximants on arbitrary staircases*, J. Comput. Appl. Math., 40 (1992), pp. 19–42.

[5] B. BECKERMANN AND G. LABAHN, *A uniform approach for Hermite Padé and simultaneous Padé approximants and their matrix generalizations*, Numerical Algorithms, 3 (1992), pp. 45–54.

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

[7] M. G. DE BRUIN, *Some aspects of simultaneous rational approximation*, in Numerical Analysis and Mathematical Modelling, Banach center publications, Vol 24, PWN-Polish Scientific Publishers, Warsaw, 1990, pp. 51–84.

[8] M. G. DE BRUIN, *Simultaneous partial Padé approximants*, J. Comput. Appl. Math., 21 (1988), pp. 343–355.

[9] 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), pp. 259–295.

[10] S. CABAY AND D. K. CHOI, *Algebraic computations of scaled Padé fractions*, SIAM J. Comput., 15 (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., 39 (1992), pp. 295–313.

[12] S. CABAY AND G. LABAHN, *A superfast algorithm for multidimensional Padé systems*, Numerical Algorithms, 2 (1992), pp. 201–224.

[13] S. CABAY AND R. MELESHKO, *A weakly stable algorithm for the Padé approximants and the inversion of Hankel matrices*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 735–765.

[14] J. COATES, *On the algebraic approximation of functions*, Indag. Math., 28 (1966), pp. 421–461.

[15] K. O. GEDDES, S. R. CZAPOR, AND G. LABAHN, *Algorithms for Computer Algebra*, Kluwer, Boston, MA, 1992.

[16] H. JAGER, *A multidimensional generalization of the Padé table*, Indag. Math., 26 (1964), pp. 193–249.

[17] G. LABAHN, *Inversion components of block Hankel-like matrices*, Linear Algebra Appl., 177 (1992), pp. 7–48.

[18] ———, *Inversion Algorithms for Rectangular-block Hankel Matrices*, Research report CS-90-52, University of Waterloo, 1990.

[19] G. LABAHN AND S. CABAY, *Matrix Padé fractions and their computation*, SIAM J. Comput., 18 (1989), pp. 639–657.

[20] W. LÜBBE, *Über ein allgemeines Interpolationsproblem—Lineare Identitäten zwischen benachbarten Lösungssystemen*, Ph.D. thesis, University of Hannover, Germany 1983.

[21] K. MAHLER, *Perfect systems*, Compos. Math., 19 (1968), pp. 95–166.

[22] S. PASZKOWSKI, *Hermite Padé approximation: Basic notions and theorems*, J. Comput. Appl. Math., 32 (1990), pp. 229–236.

[23] R. E. SHAFER, *On quadratic approximation*, SIAM J. Numer. Anal., 11 (1974), pp. 447–460.

[24] M. VAN BAREL AND A. BULTHEEL, *A new approach to the rational interpolation problem*, J. Comput. Appl. Math., 32 (1990), pp. 281–289.

[25] ———, *The computation of nonperfect Padé–Hermite approximants*, Numerical Algorithms, 1 (1991), pp. 285–304.

[26] ———, *A general module theoretic framework for vector M-Padé and matrix rational interpolation*, Numerical Algorithms, 3 (1992), pp. 451–462.