Fraction-Free Computation of Simultaneous Padé Approximants

Bernhard Beckermann,
Laboratoire Painlevé UMR 8524 (ANO-EDP),
Université des Sciences et Technologies de Lille,
F-59655 Villeneuve d'Ascq CEDEX, France,
bbecker@math.univ-lille1.fr

George Labahn
Cheriton School of Computer Science
University of Waterloo,
Waterloo, Ontario, Canada
glabahn@uwaterloo.ca

ABSTRACT

In this paper we give a new, fast algorithm for solving the simultaneous Padé approximation problem. The algorithm is fraction-free and is suitable for computation in domains where growth of coefficients in intermediate computations are a central concern. The algorithm gives significant improvement on previous fraction-free methods, in particular when solved via the use of vector Hermite-Padé approximation using the FFFG order basis algorithm previously done by the authors. The improvements are both in terms of bit complexity and in reduced size of the intermediate quantities.

Categories and Subject Descriptors: G.4 [Mathematics of Computing]: Mathematical Software

General Terms: Algorithms.

Keywords: Rational approximation, Fraction-free arithmetic, Order bases

1. INTRODUCTION

Let $\mathbf{F} = (f^{(1)}, \ldots, f^{(m)})$ (with $m \geq 2$) be an m-tuple of formal power series and $\vec{n} = (\vec{n}^{(1)}, \ldots, \vec{n}^{(m)})$ an m-tuple of nonnegative integers. We will always suppose that at least one of the power series $f^{(\ell)}$ starts with a non-zero term and, after reordering, assume that $f^{(1)}(0) \neq 0$.

A *Hermite-Padé approximant* for $\mathbf{F}$ of index $\vec{n}$ (also known as a type 1 or Latin approximant) is a vector of polynomials $\vec{p} = (p^{(1)}, ..., p^{(m)})$ satisfying

$$deg\, p^{(\ell)}(z) \leq \vec{n}^{(\ell)} - 1, \quad \ell = 1, 2, ..., m, \tag{1.1}$$

$$f^{(1)}(z)p^{(1)}(z) + \cdots + f^{(m)}(z)p^{(m)}(z) = z^{|\vec{n}|-1} r(z) \tag{1.2}$$

while a *simultaneous Padé approximant* for $\mathbf{F}$ of index $\vec{n}$ (also known as a type 2 or German approximant) is a vector of polynomials $\mathfrak{p} = (\mathfrak{p}^{(1)}, ..., \mathfrak{p}^{(m)})$ satisfying

$$deg\, \mathfrak{p}^{(\ell)} \leq |\vec{n}| - \vec{n}^{(\ell)}, \quad \ell = 1, 2, ..., m, \tag{1.3}$$

$$f^{(1)}(z)\mathfrak{p}^{(\ell)}(z) - f^{(\ell)}(z)\mathfrak{p}^{(1)}(z) = z^{|\vec{n}|+1} \mathfrak{r}^{(\ell)}(z) \tag{1.4}$$

for each ℓ. Here r and $\mathfrak{r}^{(\ell)}$ are all formal power series and $|\vec{n}| = n^{(1)} + ... + n^{(m)}$. The case $m = 2$ covers the classic case of Padé approximation. Since we have assumed that $f^{(1)}(0) \neq 0$, it follows from (1.4) that

$$f^{(j)}(z)\mathfrak{p}^{(\ell)}(z) - f^{(\ell)}(z)\mathfrak{p}^{(j)}(z) = \mathcal{O}(z^{|\vec{n}|+1})$$

for all $j, \ell = 1, ..., m$. When $f^{(1)}(z) = 1$ then simultaneous Padé approximants roughly model the property

$$f^{(\ell)}(z) \approx \frac{\mathfrak{p}^{(\ell)}(z)}{\mathfrak{p}^{(1)}(z)}$$

as $z \to 0$. That is, near 0 we have rational approximations to the power series $f^{(\ell)}$ all having a common denominator.

Historically these concepts date back to 1873 when Hermite used simultaneous rational approximations to power series for $e^x, e^{2x}, \ldots, e^{mx}$ to prove the transcendence of e [11]. Twenty years later, Padé [16], a student of Hermite, expanded on Hermite's earlier work and formalized the notion of best possible formal rational approximations to analytic functions. In a seminal paper of 1932 (but only published in 1968) K. Mahler [14] formalized the concepts to arbitrary analytic functions and interpolation points and studied relationships and recursions between the matrix of neighbours of both types of approximants.

Both types of approximants occur frequently in computer algebra. Hermite-Padé approximants are used in the van Hoeij factorization algorithm for differential operators [17] while simultaneous Padé approximants are used in solving linear systems of equations [15]. Both approximants appear in inversion formulae for striped and block Hankel matrices [13] which in turn are used in fast methods for exact solving of sparse linear systems [8].

There have been two main types of approaches for computing these approximants. The first approach looks at the associated linear system for the coefficients of the polynomials defined by the order conditions (1.2) or (1.4) – $|\vec{n}|$ unknowns for Hermite-Padé and $(m-1)|\vec{n}| + m$ unknowns for simultaneous Padé. Methods for efficient computation take advantage of special reductions of the structured matrix for the linear system [6, 7]. A second approach takes a module theoretic point of view and efficiently finds all solutions for specific order problems [2]. In this case the simultaneous Padé problem is represented as a vector Hermite-Padé problem and solved using a fast sigma-basis algorithm

ISSAC'09, July 28–31, 2009, Seoul, Republic of Korea.

(cf [2] or [10]). For a vector of non-negative integers $\vec{n}$ the method of [6] usually (but not always) finds a simultaneous Padé approximant of type $\vec{n}$ with arithmetic operation cost $\mathcal{O}(m^3|\vec{n}|\log^2|\vec{n}|)$. Sigma-basis methods finds *all* simultaneous Padé approximants of index $\vec{n}$ with cost $\mathcal{O}(m^3|\vec{n}|\log^2|m\vec{n}|)$ arithmetic operations (using the method of [2]) or $\mathcal{O}(m^\omega|\vec{n}|\log^2|m\vec{n}|)$ arithmetic operations (using the method of [10]) with ω the exponent of fast matrix multiplication. Assuming strong conditions on the input, Olesh and Storjohann [15] gave an algorithm to compute a simultaneous Padé approximant of type $\vec{n}$ with complexity $\mathcal{O}(m|\vec{n}|\log^2|\vec{n}|)$.

In this paper we focus on efficient computation when the coefficients of the power series's come from an integral domain $\mathbb{D}$ (such as the integers or domains with parameters such as $\mathbb{Q}[a_1,\ldots,a_k]$) rather than a field. Our intend is to use fraction-free methods which avoid expensive gcd computation in the coefficient arithmetic domain and allows us to avoid going through the field $\mathbb{K}$ of fractions. It is important in this case to be able to divide by a priori known common factors, in order to prevent exponential growth of coefficients during our computations. A typical example is the Bareiss variant [1, 9] of Gaussian elimination, where instead of the solution x of the linear system $Ax = b$ one computes the Cramer solution $det(A)x$. In this case all quantities at the kth elimination step are divisible by the pivot of step $(k-1)$ without passing through fractions.

Using the associated linear system approach one can obtain a fraction-free algorithm by using fraction-free Gaussian elimination. If $f^{(\ell)}$ are of size $\mathcal{O}(\kappa)$ then this gives a complexity of $\mathcal{O}(m^5|\vec{n}|^5\kappa^2)$. This does not, however, make use of the special structure of the matrix of the linear system. Since one can represent simultaneous Padé approximants as vector Hermite-Padé approximants one can use the fraction-free order basis algorithm (FFFG) of [4] to compute both types of approximants. In this case the FFFG algorithm computes Hermite-Padé approximants in bit complexity $\mathcal{O}(m|\vec{n}|^4\kappa^2)$ (see [4, Theorem 6.3]) and simultaneous Padé approximants with $\mathcal{O}(m^5|\vec{n}|^4\kappa^2)$ bit operations. The aim of the present paper is to present a new algorithm having two improvements over the direct use of FFFG: first the term m^5 can be replaced by m^2, and secondly the intermediate quantities have a smaller size.

The remainder of this paper is organized as follows. The next section includes preliminary material involving the associated linear system of a simultaneous Padé problem, its Cramer solutions, the FFFG algorithm and the new determinantal representation for simultaneous Padé approximants. Section 3 introduces the fundamental building blocks for our constructions, called *type 2* Mahler systems. In the next section we introduce the notion of normal indices and show that type 2 Mahler systems at normal indices generate *all* solutions to a simultaneous Padé problem of a given type for any order (that is these are *order bases*). Section 5 gives a fraction-free recursion with the complexity given in the following section. The paper ends with an example and then a conclusion along with topics for future research.

2. PRELIMINARIES : LINEAR SYSTEMS

Type 1 and 2 approximation problems both have associated linear systems in terms of the unknown coefficients of the polynomials. Cramer solutions to these systems then give polynomials free of fractions. In this case the Cramer solutions are also represented as determinantal polynomials as in [4] (and as found in the subresultant theory of scalar polynomials). While it is possible to look at the linear system (and corresponding Cramer's solution and determinantal representation) associated to type 2 approximants via its representation as a vector Hermite Padé approximant, in this section we give a new determinantal representation which (in the case $f^{(1)}(0) \neq \pm 1$) has considerably smaller bit size.

For a power series

$$g(z) = g_0 + g_1 z + ...,$$

with $g_j = 0$ for $j < 0$, define the rectangular $u \times v$ Toeplitz matrices

$$T^{(j)}_{u,v}(g) = \begin{bmatrix} g_j & g_{j-1} & \cdots & g_{j-v+1} \\ g_{j+1} & g_j & \cdots & g_{j-v+2} \\ \vdots & \vdots & & \vdots \\ g_{j+u-1} & g_{j+u-2} & \cdots & g_{j+u-v+1} \end{bmatrix}.$$

Setting $T_{u,v}(g) = T^{(0)}_{u,v}(g)$, let

$$K(\vec{n},\sigma) = [\ T_{\sigma,\vec{n}^{(1)}}(f^{(1)}) \ \big|\ \ldots\ \big|\ T_{\sigma,\vec{n}^{(m)}}(f^{(m)})\ ]$$

be the striped Sylvester matrices (having m column blocks) of size $\sigma \times |\vec{n}|$ and multigradients

$$d(\vec{n}) = det\, K(\vec{n},|\vec{n}|).$$

In addition, following [4], define $E^{(\ell)}(\vec{n},z)$ to be

$$E^{(\ell)}(\vec{n},z) = [0,\ldots,0,1,\ldots,z^{\vec{n}^{(\ell)}-1},0,\ldots,0]$$

a row vector of length $|\vec{n}|$ with blocks of size $\vec{n}^{(j)}$, $j = 1,...,m$, with the ℓth block containing $1, z, z^2, ..., z^{\vec{n}^{(\ell)}-1}$, and the others 0.

From [4] we know that the Cramer solutions of the homogeneous systems of equations behind the approximants of type 1 are given by the column vector

$$p(\vec{n},z) = \left[p^{(\ell)}(\vec{n},z)\right]_{\ell=1,\ldots,m},$$

with

$$p^{(\ell)}(\vec{n},z) = det\begin{bmatrix} K(\vec{n},|\vec{n}|-1) \\ E^{(\ell)}(\vec{n},z) \end{bmatrix}.$$

It is not difficult to check (see [4, Lemma 5.1]) that

$$p^{(\ell)}(\vec{n},z) = \pm d(\vec{n}-\vec{e}_\ell)z^{\vec{n}^{(\ell)}-1} + \text{ lower terms}$$

and

$$\sum_\ell f^{(\ell)}(z)p^{(\ell)}(\vec{n},z) = \pm d(\vec{n})z^{|\vec{n}|-1} + \text{ higher terms} \quad (2.1)$$

so that (1.1) and (1.2) hold. Here and in what follows $\vec{e}_\ell$ denotes the ℓth canonical vector, and $\vec{e} = (1,1,...,1)$.

It is possible to represent simultaneous Padé approximants as a vector Hermite-Padé problem

$$\begin{bmatrix} f^{(2)}(z) \\ \vdots \\ f^{(m)}(z) \end{bmatrix} \mathfrak{p}^{(1)}(z) + \begin{bmatrix} -f^{(1)}(z) \\ \vdots \\ 0 \end{bmatrix} \mathfrak{p}^{(2)}(z) + \cdots + \begin{bmatrix} 0 \\ \vdots \\ -f^{(1)}(z) \end{bmatrix} \mathfrak{p}^{(m)}(z) = O(z^{|\vec{n}|+1})$$

and obtain both a linear system and determinantal representations for the type 2 problem similar to those given above.

However, in this case we prefer to give a new, alternate determinantal representation of type 2 approximants.

LEMMA 2.1. *Consider the column vector*

$$\mathfrak{p}(\vec{n},z)=\left[\mathfrak{p}^{(\ell)}(\vec{n},z)\right]_{\ell=1,\dots,m},$$

defined by

$$\mathfrak{p}^{(\ell)}(\vec{n},z)=(-1)^{\ell}\det\begin{bmatrix}K(\vec{n}+\vec{e},|\vec{n}|+1)\\ E^{(1)}(\vec{n}+\vec{e},z)\\ \cdots\\ E^{(\ell-1)}(\vec{n}+\vec{e},z)\\ E^{(\ell+1)}(\vec{n}+\vec{e},z)\\ \cdots\\ E^{(m)}(\vec{n}+\vec{e},z)\end{bmatrix}.$$

Then $\mathfrak{p}(\vec{n},z)$ is a type 2 approximant satisfying (1.3) and (1.4). More precisely

$$\mathfrak{p}^{(\ell)}(\vec{n},z)=\pm d(\vec{n}+\vec{e}_\ell)z^{|\vec{n}|-\vec{n}^{(\ell)}}+\ \textit{lower terms},\qquad(2.2)$$

and for $\ell\geq 2$

$$\begin{aligned}&f^{(1)}(z)\mathfrak{p}^{(\ell)}(\vec{n},z)-f^{(\ell)}(z)\mathfrak{p}^{(1)}(\vec{n},z)\qquad(2.3)\\ &=\pm\det\begin{bmatrix}K(\vec{n}+\vec{e},|\vec{n}|+2)\\ E^{(2)}(\vec{n}+\vec{e},0)\\ \cdots\\ E^{(\ell-1)}(\vec{n}+\vec{e},0)\\ E^{(\ell+1)}(\vec{n}+\vec{e},0)\\ \cdots\\ E^{(m)}(\vec{n}+\vec{e},0)\end{bmatrix}z^{|\vec{n}|+1}+\ \textit{higher terms}\,.\end{aligned}$$

Note that $E^{(j)}(\vec{n}+\vec{e},0)$ has a 1 in column $\vec{n}^{(1)}+\cdots+\vec{n}^{(j-1)}+1$ and is 0 elsewhere.

PROOF. The degree bound $deg\,\mathfrak{p}^{(\ell)}(\vec{n},z)\leq|\vec{n}|-\vec{n}^{(\ell)}$ immediately follows from expanding the determinant in the definition of $\mathfrak{p}^{(\ell)}(\vec{n},z)$ with respect to the last $m-1$ rows. Formula (2.2) for the coefficient in front of $z^{|\vec{n}|-\vec{n}^{(\ell)}}$ is obtained by expanding with respect to the last $m-1$ rows and the columns containing the highest powers of z, that is, the last columns of blocks $\lambda=1,2,...,m$, $\lambda\neq\ell$. Thus the matrix $K(\vec{n}+\vec{e},|\vec{n}|+1)$ becomes $K(\vec{n}+\vec{e}_\ell,|\vec{n}|+1)$, i.e., we obtain the expression stated in (2.2).

For a proof of (2.3), we first shift the rows of the matrix underlying $\mathfrak{p}^{(1)}(\vec{n},z)$ such that $E^{(\ell)}(\vec{n}+\vec{e},z)$ becomes the new $(m-1)$st last row. Then

$$\begin{aligned}&f^{(1)}(z)\mathfrak{p}^{(\ell)}(\vec{n},z)-f^{(\ell)}(z)\mathfrak{p}^{(1)}(\vec{n},z)\\ &=\pm det\begin{bmatrix}K(\vec{n}+\vec{e},|\vec{n}|+1)\\ f^{(1)}(z)E^{(1)}(\vec{n}+\vec{e},z)+f^{(\ell)}(z)E^{(\ell)}(\vec{n}+\vec{e},z)\\ E^{(2)}(\vec{n}+\vec{e},z)\\ \cdots\\ E^{(\ell-1)}(\vec{n}+\vec{e},z)\\ E^{(\ell+1)}(\vec{n}+\vec{e},z)\\ \cdots\\ E^{(m)}(\vec{n}+\vec{e},z)\end{bmatrix}.\end{aligned}$$

Adding multiples of the last $(m-2)$ rows to the $(m-1)$st last row, this row becomes

$$\begin{aligned}&\sum_{\lambda=1}^{m}f^{(\lambda)}(z)E^{(\lambda)}(\vec{n}+\vec{e},z)\\ &=\sum_{j=0}^{\infty}z^j\left[\ T^{(j)}_{1,\vec{n}^{(1)}+1}(f^{(1)})\ \middle|\ \cdots\ \middle|\ T^{(j)}_{1,\vec{n}^{(m)}+1}(f^{(m)})\ \right].\end{aligned}$$

For $j=0,...,|\vec{n}|$, the row vector in the sum on the right occurs already in the matrix $K(\vec{n}+\vec{e},|\vec{n}|+1)$. Hence, by linearity of the determinant, $f^{(1)}(z)\mathfrak{p}^{(\ell)}(\vec{n},z)-f^{(\ell)}(z)\mathfrak{p}^{(1)}(\vec{n},z)$ is equal (up to a sign) to the expression

$$det\begin{bmatrix}K(\vec{n}+\vec{e},|\vec{n}|+2)\\ E^{(2)}(\vec{n}+\vec{e},0)\\ \cdots\\ E^{(\ell-1)}(\vec{n}+\vec{e},0)\\ E^{(\ell+1)}(\vec{n}+\vec{e},0)\\ \cdots\\ E^{(m)}(\vec{n}+\vec{e},0)\end{bmatrix}z^{|\vec{n}|+1}$$

plus higher order terms, as claimed in (2.3). □

EXAMPLE 2.2. *Let $\vec{n}=(1,2,1,2)$. Then $\mathfrak{p}^{(3)}(\vec{n},z)$ is given by minus the determinant of the matrix*

$$\left[\begin{array}{cc|ccc|ccc|ccc}
f_0^{(1)}&0&f_0^{(2)}&0&0&f_0^{(3)}&0&f_0^{(4)}&0&0\\
f_1^{(1)}&f_0^{(1)}&f_1^{(2)}&f_0^{(2)}&0&f_1^{(3)}&f_0^{(3)}&f_1^{(4)}&f_0^{(4)}&0\\
f_2^{(1)}&f_1^{(1)}&f_2^{(2)}&f_1^{(2)}&f_0^{(2)}&f_2^{(3)}&f_1^{(3)}&f_2^{(4)}&f_1^{(4)}&f_0^{(4)}\\
f_3^{(1)}&f_2^{(1)}&f_3^{(2)}&f_2^{(2)}&f_1^{(2)}&f_3^{(3)}&f_2^{(3)}&f_3^{(4)}&f_2^{(4)}&f_1^{(4)}\\
f_4^{(1)}&f_3^{(1)}&f_4^{(2)}&f_3^{(2)}&f_2^{(2)}&f_4^{(3)}&f_3^{(3)}&f_4^{(4)}&f_3^{(4)}&f_2^{(4)}\\
f_5^{(1)}&f_4^{(1)}&f_5^{(2)}&f_4^{(2)}&f_3^{(2)}&f_5^{(3)}&f_4^{(3)}&f_5^{(4)}&f_4^{(4)}&f_3^{(4)}\\
f_6^{(1)}&f_5^{(1)}&f_6^{(2)}&f_5^{(2)}&f_4^{(2)}&f_6^{(3)}&f_5^{(3)}&f_6^{(4)}&f_5^{(4)}&f_4^{(4)}\\
1&z&0&0&0&0&0&0&0&0\\
0&0&1&z&z^2&0&0&0&0&0\\
0&0&0&0&0&0&0&1&z&z^2
\end{array}\right].$$

Expanding determinants by the least row and the determinants of the submatrices along their last rows and so on gives a polynomial of degree at most 5. The leading coefficient in this case then comes from the determinant of the matrix obtained by removing the last three rows and columns 2, 5 and 10. This is the same as the determinant of $K((1,2,2,2),7)$.

In our case, we will only consider indices where the simultaneous Padé approximants are (up to normalization) uniquely given by the objects defined in Lemma 2.1.

LEMMA 2.3. *Suppose that $d(\vec{n})\neq 0$. Then, for $\ell=1,...,m$, type 2 approximants of index $\vec{n}-\vec{e}_\ell$ are scalar multiples of $\mathfrak{p}(\vec{n}-\vec{e}_\ell,z)\neq 0$.*

PROOF. Set $\vec{\nu}=\vec{n}-\vec{e}_\ell$. Since $f^{(1)}(0)\neq 0$, we may divide (1.4) by $f^{(1)}$, and thus suppose without loss of generality that $f^{(1)}=1$. In this case, $\mathfrak{p}$ is a type 2 approximant of index $\vec{\nu}$ if and only if its first component $\mathfrak{p}^{(1)}(z)=a_0+a_1z+...+a_Nz^N$, $N=|\vec{\nu}|-\vec{\nu}^{(1)}$, is such that the expansion of $f^{(j)}\mathfrak{p}^{(1)}$ does not contain the powers z^k for $k=|\vec{\nu}|-\vec{\nu}^{(j)}+1,|\vec{\nu}|-\vec{\nu}^{(j)}+2,...,|\vec{\nu}|$ and $j=2,3,...,m$, and the other components of $\mathfrak{p}$ are obtained by taking the lower powers of $f^{(j)}\mathfrak{p}^{(1)}$. This leads to the homogeneous system with N equations and $N+1$ unknowns $[a_N,a_{N-1},...,a_0]K=0$, with

$$K=\left[\ T^{(\vec{\nu}^{(1)})}_{N+1,\vec{\nu}^{(2)}}(f^{(2)})\ \middle|\ \cdots\ \middle|\ T^{(\vec{\nu}^{(1)})}_{N+1,\vec{\nu}^{(m)}}(f^{(m)})\ \right].$$

Since we may partition into a block triangular matrix

$$K(\vec{\nu},|\vec{\nu}|+1)=\begin{bmatrix}I&*\\0&K\end{bmatrix},$$

and since $K(\vec{\nu},|\vec{\nu}|+1)$ is obtained by dropping one column from the invertible matrix $K(\vec{n},|\vec{n}|)$, we find that K is of full column rank, and therefore has a left nullspace of dimension 1. Hence type 2 approximants of index $\vec{\nu}=\vec{n}-\vec{e}_\ell$ are unique up to multiplication with a scalar, and can be generated from $\mathfrak{p}(\vec{n}-\vec{e}_\ell,z)$, the latter being non-trivial according to (2.2). □

Assume that our arithmetic follows the bit complexity model:

$$\begin{array}{lcl} \text{size}(a+b) & = & \mathcal{O}(max\{\text{size}(a),\text{size}(b)\}),\\ \text{size}(a\cdot b) & = & \mathcal{O}(\text{size}(a)+\text{size}(b)),\\ \text{cost}(a+b) & = & \mathcal{O}(1),\\ \text{cost}(a\cdot b) & = & \mathcal{O}(\text{size}(a)\cdot\text{size}(b)).\end{array}$$

LEMMA 2.4. *Suppose that the coefficients of the input power series $f^{(\ell)}$ are of size $\mathcal{O}(\kappa)$. Then using the complexity model above we have*

$$size(\mathfrak{p}^{(\ell)}(\vec{n},z)) = \mathcal{O}(|\vec{n}|\kappa). \tag{2.4}$$

PROOF. The result follows by expanding the determinants in Lemma 2.1 and by using the Hadamard inequality. □

The representation of type 2 approximants in [4] gives an additional factor m for this size. We claim without proof that the type 2 approximants of [4] reduce to our type 2 approximants after a division by $f^{(1)}(0)^{(m-2)(|\vec{n}|+1)}$. In particular, they are the same (up to a sign) in the classic case of Padé approximants (when $m=2$).

3. TYPE 2 MAHLER SYSTEMS

In this section we describe the building blocks for our new algorithm, which we call *type 2 Mahler systems*. As was the case with the (type 1) Mahler systems from [4, Definition 5.2] these appear naturally from the associated linear system of the Padé problem. For a given vector $\vec{n}$ of non-negative integers, Mahler systems are $m\times m$ matrix polynomials whose columns are approximants of order $|\vec{n}|$. Type 2 Mahler systems are implicitly obtained by forming cofactors of the (type 1) Mahler systems. This duality, first observed in a special case and with an alternate normalization by Mahler [14] and then later in the general case by the authors [3], will be heavily exploited.

Type 1 Mahler systems are given by

$$\begin{array}{lcl}\mathbf{M}(\vec{n},z) & = & \left[\pm p(\vec{n}+\vec{e}_j,z)\right]_{j=1,\ldots,m}\\ & = & \left[\mathbf{M}^{(\lambda,j)}(\vec{n},z)\right]_{\lambda,j=1,\ldots,m}\end{array}$$

where the signs are chosen such that all $\mathbf{M}^{(\ell,\ell)}(\vec{n},z)$ have the same leading coefficient $d(\vec{n})$ (or $-d(\vec{n})$) which is assumed to be different from 0. These Mahler systems differ from those considered by Mahler in [14] only by normalization: the multiplication with $d(\vec{n})$ insures that no fractions occur.

In the same way, we define for indices $\vec{n}$ with $d(\vec{n})\neq 0$ the type 2 Mahler systems

$$\begin{array}{lcl}\mathfrak{M}(\vec{n},z) & = & \left[\pm\mathfrak{p}(\vec{n}-\vec{e}_j,z)\right]_{j=1,\ldots,m}\\ & = & \left[\mathfrak{M}^{(\lambda,j)}(\vec{n},z)\right]_{\lambda,j=1,\ldots,m}\end{array}$$

where the signs are chosen such that all $\mathfrak{M}^{(\ell,\ell)}(\vec{n},z)$ have the same leading coefficient $d(\vec{n})\neq 0$ (or $-d(\vec{n})$), compare with (2.2). It follows from [14, §25] or directly from the Sylvester determinantal identity that

$$\text{cof }\mathbf{M}(\vec{n},z) = \pm d(\vec{n})^{m-2}\,\mathfrak{M}(\vec{n},z), \tag{3.1}$$

where as in [3, Section 6] the cofactor of a square matrix A is defined by $\text{cof }A = (\text{adj }A)^T = (A^{-1})^T \det A$. This identity is the motivation for our introduction of type 2 Mahler systems. Indeed, from [14, 4] we know that such duality relations are quite helpful in the recursive computation of Hermite-Padé or matrix Padé approximations.

4. NON-PERFECT SYSTEMS AND CLOSEST NORMAL INDICES

When the multigradient $d(\vec{n})$ is non-zero, the proof of Lemma 2.3 implies that the matrix of coefficents for the linear system for simultaneous Padé approximants of index $\vec{n}$ has a unique solution up to normalization. The situation is considerably more involved if any of the multigradients vanish. In [4] for the case of Hermite-Padé (i.e. type 1) approximants, we have introduced so-called *closest normal indices* and show how to describe the module of all vector polynomials of a given order. In this section we show the same can be said using the type 2 Mahler systems introduced in the previous section.

THEOREM 4.1. *Given a multi-index $\vec{n}$, define the closest normal indices $\vec{\nu}_\sigma$ (with $|\vec{\nu}_\sigma|=\sigma$ and $d(\vec{\nu}_\sigma)\neq 0$) by $\vec{\nu}_0=\vec{0}$, and for $\sigma\geq 0$*

$$\vec{\nu}_{\sigma+1} = \vec{\nu}_\sigma + \vec{e}_{\pi(\sigma)},$$

where $\pi=\pi(\sigma)\in\{1,\ldots,m\}$ is defined by

$$\pi = \arg\max_\ell\{\vec{n}^{(\ell)}-\vec{\nu}_\sigma^{(\ell)} : \ell=1,\ldots,m,\ d(\vec{\nu}_\sigma+\vec{e}_\ell)\neq 0\}.$$

Then for all integers $d,\sigma\geq 0$

$$\begin{array}{l}\Big\{p\in\mathbb{K}[z]^m : \sum_{\ell=1}^m f^{(\ell)}(z)p^{(\ell)}(z) = \mathcal{O}(z^\sigma), \forall\ell :\\ \qquad\qquad \deg p^{(\ell)}\leq \vec{n}^{(\ell)}+d\Big\}\\ = \Big\{\mathbf{M}(\vec{\nu}_\sigma,z)q : q\in\mathbb{K}[z]^m,\\ \qquad\qquad \forall\lambda : \deg q^{(\lambda)}\leq \vec{n}^{(\lambda)}-\vec{\nu}_\sigma^{(\lambda)}+d\Big\}.\end{array}$$

PROOF. We show by recurrence on σ that $\pi=\pi(\sigma)$ is well-defined, that is, there is at least one index ℓ with $d(\vec{\nu}_\sigma+\vec{e}_\ell)\neq 0$, compare with [3, Lemma 2.7]. In this case, the claimed parametrization of the set of solutions with specified order and degree constraints have been shown in the algorithm FFFG given in [4, Table 2] together with [4, Theorem 7.2 and Theorem 7.3] (we notice that the sequence of multi–indices $(\vec{n}_k)_k$ defined in [4, Table 2] satisfies $\vec{n}_{|\vec{n}|+md}=\vec{n}+d\vec{e}$).

For $\sigma=0$, the assertion follows from observing that

$$d(\vec{\nu}_\sigma+\vec{e}_1) = d(\vec{e}_1) = f^{(1)}(0)\neq 0.$$

Suppose now that the multi-index $\vec{\nu}_\sigma$ with $|\vec{\nu}_\sigma|=\sigma$ is well-defined with $d(\vec{\nu}_\sigma)\neq 0$, but $d(\vec{\nu}_\sigma+\vec{e}_\ell)=0$ for $\ell=1,2,\ldots,m$. Then, by definition, the type 1 Mahler system $\mathbf{M}(\vec{\nu}_\sigma,z)$ has columns of order at least σ, and, in view of (2.1), these columns have order at least $\sigma+1$. Taking into account the above representation of polynomial column vectors of order σ, we may conclude that there is no such vector of precise order σ. This is however in contradiction with the observation that z^σ times the first canonical vector has precise order σ. □

Being able to describe the set of type 1 approximants allows us to apply a duality argument in order to deduce a parametrization of the set of type 2 aproximants.

THEOREM 4.2. *With the notations of Theorem 4.1,*

$$\Big\{\mathfrak{p}\in\mathbb{K}[z]^m : \forall \ell, j : f^{(\ell)}(z)\mathfrak{p}^{(j)}(z) - f^{(j)}(z)\mathfrak{p}^{(\ell)}(z) = \mathcal{O}(z^\sigma),$$
$$\deg \mathfrak{p}^{(\ell)} \le |\vec{n}| - \vec{n}^{(\ell)} + d\Big\}$$
$$= \Big\{\mathfrak{M}(\vec{\nu}_\sigma, z)\mathfrak{q} : \mathfrak{q}\in\mathbb{K}[z]^m,$$
$$\forall \lambda : \deg \mathfrak{q}^{(\lambda)} \le (|\vec{n}| - \vec{n}^{(\lambda)}) - (|\vec{\nu}_\sigma| - \vec{\nu}_\sigma^{(\lambda)}) + d\Big\}.$$

PROOF. Using the notation of [3, Section 2.2], we conclude from Theorem 4.1 that $\mathbf{M}(\vec{\nu}_\sigma, z)$ is an $(\mathbf{G}, (z^\sigma, 1, ..., 1))$ order basis, with the square matrix $\mathbf{G}$ obtained as in [3, Example 2.3] by bordering the row $\mathbf{F} = [f^{(1)}, ..., f^{(m)}]$ by the block $[0, I_{m-1}]$. It follows from [3, Corollary 6.3(a) and Example 6.4] that cof $\mathbf{M}(\vec{\nu}_\sigma, z)$ is a (cof $\mathbf{G}, (1, z^\sigma, ..., z^\sigma)$) order basis, or, in other words,

$$\Big\{\mathfrak{p}\in\mathbb{K}[z]^m : \forall \ell, j : f^{(\ell)}\mathfrak{p}^{(j)} - f^{(j)}\mathfrak{p}^{(\ell)} = \mathcal{O}(z^\sigma)\Big\}$$
$$= \Big\{\mathfrak{M}(\vec{\nu}_\sigma, z)\mathfrak{q} : \mathfrak{q}\in\mathbb{K}[z]^m\Big\},$$

where we have used (3.1) together with the fact that $d(\vec{\nu}_\sigma) \neq 0$.

In order to obtain the additional degree constrains stated in Theorem 4.2, we recall from [4, Theorem 7.2] that $\mathbf{M}(\vec{\nu}_\sigma, z)$ is in $\vec{n}$–Popov form. In particular (see [4, Definition 7.1]), with

$$z^{\vec{n}} = diag\,(z^{\vec{n}^{(1)}}, ..., z^{\vec{n}^{(m)}}),$$

the matrix Laurent polynomial $z^{-\vec{n}}\mathbf{M}(\vec{\nu}_\sigma, z)$ is column reduced, with column degree of the ℓth column given by $\vec{\nu}_\sigma^{(\ell)} - \vec{n}^{(\ell)}$. It follows from (3.1) (compare with [3, Lemma 6.2(c)]) that also $z^{-|\vec{n}|\vec{e}+\vec{n}}\mathfrak{M}(\vec{\nu}_\sigma, z)$ is column reduced, with column degree of the ℓth column given by $-\vec{\nu}_\sigma^{(\ell)} + \vec{n}^{(\ell)} + |\vec{\nu}_\sigma| - |\vec{n}|$. Hence, given $\mathfrak{p} = \mathfrak{M}(\vec{\nu}_\sigma, z)\mathfrak{q}$ with $deg\,\mathfrak{p}^{(\ell)} \le |\vec{n}| - \vec{n}^{(\ell)} + d$, we conclude that $z^{-|\vec{n}|\vec{e}+\vec{n}}\mathfrak{p}$ has degree $\le d$, and the predictable degree property of [12, Theorem 6.3-13, p. 387] tells us that $deg\,\mathfrak{q}^{(\lambda)} \le d - (-\vec{\nu}_\sigma^{(\lambda)} + \vec{n}^{(\lambda)} + |\vec{\nu}_\sigma| - |\vec{n}|)$, as required for one inclusion in the statement of Theorem 4.2. Conversely, with the above degree bounds for $\mathfrak{q}^{(\lambda)}$ (together with those for the entries of $\mathfrak{M}(\vec{\nu}_\sigma, z)$) one easily obtains the above degree bounds for the entries of $\mathfrak{p} = \mathfrak{M}(\vec{\nu}_\sigma, z)\mathfrak{q}$, showing Theorem 4.2. □

5. RECURSIVE COMPUTATION OF TYPE 2 MAHLER SYSTEMS

In his seminal work [14] Mahler showed how to recursively compute type 1 and type 2 Mahler systems under the assumption that all multigradients $d(\vec{n})$ are non-zero (or, in, other words, $\vec{n}$ is normal). Starting from $\mathbf{M}(\vec{0}, z) = I_m$ (or $\mathfrak{M}(\vec{0}, z) = I_m$), he explains in [14, §41 and §43] how to obtain $\mathbf{M}(\vec{n} + \vec{e}_j, z)$ by multiplication on the right of $\mathbf{M}(\vec{n}, z)$ by some "simple" degree 1 matrix polynomial, and how to obtain $\mathfrak{M}(\vec{n}+\vec{e}_j, z)$ by multiplication on the right of $\mathfrak{M}(\vec{n}, z)$ by some other "simple" degree 1 matrix polynomial.

In the non-normal case where some of the $d(\vec{n})$ are zero the methods of Mahler break down. Also in this case we want to describe all solutions for a given order (since solutions are no longer unique up to scalar constant). The previous section shows that Mahler systems of type 1 and type 2 at closest normal indices give module bases for order problems. In addition the FFFG algorithm of [4] gives a method for computing Mahler systems of type 1. In this section we show how to combine the above with the cofactor relation (3.1) and obtain an algorithm to compute type 2 Mahler systems.

Let us start by briefly recalling the idea of the FFFG algorithm of [4]: given the closest normal index $\vec{\nu}_k$ of order $k = |\vec{\nu}_k|$ and the corresponding $\mathbf{M}_k := \mathbf{M}(\vec{\nu}_k, z)$, one first computes the first term of the residuals $(r^{(1)}, ..., r^{(m)})$ defined by

$$(f^{(1)}, ..., f^{(m)})\mathbf{M}_k = (r^{(1)}, ..., r^{(m)})z^k + \text{ higher terms}$$

(where from (2.1) we know that $r^{(\ell)} = \pm d(\vec{\nu}_k + \vec{e}_\ell)$). Theorem 4.1 describes how to compute the pivot $\pi \in \{1, ..., m\}$ together with the normal multi-index $\vec{\nu}_{k+1} = \vec{\nu}_k + \vec{e}_\pi$ to give the next Mahler system. The new Mahler system is built as follows:

(i) *Increase order of all columns:* Apply $(m-1)$ elementary operations of subtracting $r^{(\ell)}$ times the pivot column from the pivot $r^{(\pi)}$ times the ℓth column of $\mathbf{M}_k$ for $\ell \neq \pi$. This increases the order of all nonpivot columns via cross multiplication. The pivot column is then multiplied by $zr^{(\pi)}$.

(ii) *Degree correction for correct normalization:* The overshoot in the degree of the ℓth component for $\ell = 1, ..., m$, $\ell \neq \pi$, is corrected by subtracting a suitable multiple of the ℓth new column from the pivot column. The resulting matrix polynomial contains the common factor $d(\vec{\nu}_k)$(the pivot in step $(k-1)$). After division we obtain our new Mahler system.

Let the matrix $\mathbf{A}$ represent the elementary column operations which increase the orders and the matrix $\mathbf{B}$ represent the elementary column operations which correct the degrees. Then in matrix form the above algorithm is given as

$$d(\vec{\nu}_k)\,\mathbf{M}_{k+1} = \mathbf{M}_k\mathbf{A}\mathbf{B}.$$

This form is particularly useful when one takes into consideration the duality of type 1 and type 2 approximants via the cofactor operation as given in (3.1). The result is the matrix equation

$$d(\vec{\nu}_k)\,\mathfrak{M}_{k+1} = \mathfrak{M}_k \text{cof } \mathbf{A} \text{cof } \mathbf{B}.$$

As an example, when the pivot $\pi = 1$ then the matrices $\mathbf{A}$ and $\mathbf{B}$ have the form

$$\mathbf{A} = \begin{bmatrix} 1 & * & \cdots & * \\ 0 & * & & \\ \vdots & & \ddots & \\ 0 & & & * \end{bmatrix} \text{ and } \mathbf{B} = \begin{bmatrix} z & 0 & \cdots & 0 \\ * & 1 & & \\ \vdots & & \ddots & \\ * & & & 1 \end{bmatrix}$$

which results in cofactors having the form

$$\text{cof } \mathbf{A} = \begin{bmatrix} * & & & \\ * & 1 & & \\ \vdots & & \ddots & \\ * & & & 1 \end{bmatrix}$$

and

$$\text{cof } \mathbf{B} = \begin{bmatrix} 1 & * & \cdots & * \\ & z & & \\ & & \ddots & \\ & & & z \end{bmatrix}$$

Translating cof $\mathbf{A}$ and cof $\mathbf{B}$ into elementary column operations implies that our new recurrence for type 2 Mahler systems is of the form:

(i) Compute the new πth pivot column by a suitable linear combination of all m columns, and then for all $\ell \neq \pi$ compute the new ℓth column by multiplying the old ℓth column by a scalar times z.

(ii) We now have an overshoot of degree for the πth component. This is corrected by subtracting a suitable multiple of the new pivot column.

In order to be more precise and to specify the scalars which appear in the recursion above, there are two obstacles which we overcome via the following two Theorems. First it is not obvious how to obtain the multigradients $d(\vec{\nu}_k + \vec{e}_\ell)$ from the type 2 Mahler systems $\mathfrak{M}_k = \mathfrak{M}(\vec{\nu}_k, z)$ (for the choice of the pivot index π). A second problem is that we require a formula for the above scalars which are not given in terms of the type 1 Mahler system $\mathbf{M}_k$, but rather given in terms of the type 2 Mahler system $\mathfrak{M}_k$.

THEOREM 5.1. *Let*

$$\mathfrak{F} = \begin{bmatrix} -f^{(2)} & f^{(1)} & 0 & \cdots & 0 \\ -f^{(3)} & 0 & f^{(1)} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ -f^{(m)} & 0 & \cdots & 0 & f^{(1)} \end{bmatrix}$$

and

$$C \;:=\; \left[C^{(j,\ell)}\right]_{j=2,\ldots,m}^{\ell=1,\ldots,m}$$

the $(m-1) \times m$ matrix satisfying

$$\mathfrak{F} \cdot \mathfrak{M}_k = C z^k + \text{ higher terms } .$$

Then the matrix C has full rank $m-1$. Furthermore its right nullspace is spanned by the vector $y = (y_\ell)_{\ell=1,2,\ldots,m}$, with $y_\ell = \pm d(\vec{\nu}_k + \vec{e}_\ell)$.

PROOF. We first have to fix the signs in the determinantal formulas for the entries of $\mathfrak{M}_k$, following the lines of the proof of Theorem 2.1. Denote by $E_0^{(\ell)}(\vec{n})$ the unit vector row of length $|\vec{n}|$ having in it's ℓth block of size $\vec{n}^{(\ell)}$ the entries $0, ..., 0, 1$ (and zeros otherwise). We then may border the matrix underlying $\mathfrak{M}_k^{(j,\ell)} = \pm\mathfrak{p}(\vec{\nu}_k - \vec{e}_\ell, z)$ by a new last column in the ℓth column block and the new row $E_0^{(\ell)}(\vec{\nu}_k + \vec{e})$, leading with some $\epsilon \in \{\pm 1\}$ to the formula

$$\epsilon\, \mathfrak{M}_k^{(j,\ell)} = det \begin{bmatrix} K(\vec{\nu}_k + \vec{e}, k) \\ E^{(1)}(\vec{\nu}_k + \vec{e}, z) \\ \cdots \\ E^{(j-1)}(\vec{\nu}_k + \vec{e}, z) \\ \\ E_0^{(\ell)}(\vec{\nu}_k + \vec{e}) \\ \\ E^{(j+1)}(\vec{\nu}_k + \vec{e}, z) \\ \cdots \\ E^{(m)}(\vec{\nu}_k + \vec{e}, z) \end{bmatrix}$$

In order to check that the sign is correct in the preceding formula, notice that in order to get the leading coefficient of $\mathfrak{M}_k^{(\ell,\ell)}$ (which should be independent of ℓ), one has to expand following the highest degree terms, that is, one expands with respect to the last m rows the determinant of the matrix obtained by replacing the rows $E^{(\lambda)}(\vec{\nu}_k + \vec{e}, z)$ by $E_0^{(\lambda)}(\vec{\nu}_k + \vec{e})$.

Hence

$$\begin{aligned} C^{(j,\ell)} &= \lim_{z\to 0} z^{-k} \left(-f^{(j)} \mathfrak{M}_k^{(1,\ell)} + f^{(1)} \mathfrak{M}_k^{(j,\ell)} \right) \\ &= \epsilon (-1)^{m-j} \lim_{z\to 0} z^{-k} \mathfrak{A}(z) \end{aligned}$$

where

$$\mathfrak{A}(z) \;=\; det \begin{bmatrix} K(\vec{\nu}_k + \vec{e}, k) \\ f^{(1)} E^{(1)}(\vec{\nu}_k + \vec{e}, z) + f^{(j)} E^{(j)}(\vec{\nu}_k + \vec{e}, z) \\ \left[E^{(\lambda)}(\vec{\nu}_k + \vec{e}, z)\right]_{\lambda=2,\ldots,m,\lambda\neq j} \\ E_0^{(\ell)}(\vec{\nu}_k + \vec{e}) \end{bmatrix}.$$

In order to evaluate the limit, we add $f^{(\lambda)} E^{(\lambda)}(\vec{\nu}_k + \vec{e}, z)$ to the second row block for $\lambda \neq 1, j$, and expand the resulting expression in powers of z. Permuting the first column of each block to the end, expanding with respect to the second last row block containing unit vectors, and taking into account the block Toeplitz structure leads for some $\widetilde{\epsilon} \in \{\pm 1\}$ to

$$\begin{aligned} C^{(j,\ell)} &= \epsilon\, (-1)^{m-j}\, det \begin{bmatrix} K(\vec{\nu}_k + \vec{e}, k+1) \\ E^{(2)}(\vec{\nu}_k + \vec{e}, 0) \\ \cdots \\ E^{(j-1)}(\vec{\nu}_k + \vec{e}, 0) \\ E^{(j+1)}(\vec{\nu}_k + \vec{e}, 0) \\ \cdots \\ E^{(m)}(\vec{\nu}_k + \vec{e}, 0) \\ E_0^{(\ell)}(\vec{\nu}_k + \vec{e}) \end{bmatrix} \\ &= \widetilde{\epsilon}\; det \begin{bmatrix} \begin{bmatrix} 0 \\ K(\vec{\nu}_k, k) \end{bmatrix} & K(\vec{e}_1 + \vec{e}_j, k+1) \\ E_0^{(\ell)}(\vec{\nu}_k) & 0 \end{bmatrix}, \end{aligned}$$

that is, a joint block of $(k+1)$ rows and columns with determinant being equal to $\pm f^{(1)} d(\vec{\nu}_k)$, bordered by a last row depending on $\ell \in \{1, 2, ..., m\}$, and a last column depending on $j \in \{2, 3, ..., m\}$.

Sylvester's determinantal identity shows that

$$\begin{aligned} &(-1)^{m-\ell}\, det \left[C^{(j,\lambda)}\right]_{j=2,\ldots,m}^{\lambda=1,\ldots,m,\lambda\neq\ell} \\ &= (-1)^{m-\ell}\; \widetilde{\epsilon}^{m-1} \left(f^{(1)} d(\vec{\nu}_k) \right)^{m-2} \\ &\quad det \begin{bmatrix} \begin{bmatrix} 0 \\ K(\vec{\nu}_k, k) \end{bmatrix} & K(\vec{e}, k+1) \\ \left[E_0^{(\lambda)}(\vec{\nu}_k)\right]_{\lambda=1,\ldots,m,\lambda\neq\ell} & 0 \end{bmatrix}, \end{aligned}$$

where the determinant on the right after expansion with respect to the last $(m-1)$ rows equals $\epsilon_\ell d(\vec{\nu}_k + \vec{e}_\ell) = \pm d(\vec{\nu}_k + \vec{e}_\ell)$. By Cramer's rule this quantity is the ℓth component of an element of the nullspace of C, which has a non-trivial $\pi(k)$th component by Theorem 4.1. Again from Cramer's rule we deduce that C has full row rank. Hence the vector y with components $y_\ell = (-1)^\ell \epsilon_\ell\, d(\vec{\nu}_k + \vec{e}_\ell)$ spans the nullspace of C, as claimed in Theorem 5.1. □

Nullspaces can be computed without fractions via fraction-free Gaussian elimination. However the determinantal representation of $C^{(j,\ell)}$ given in Theorem 5.1 implies that we can start with a divisor that is not 1.

COROLLARY 5.2. *Let y be the vector from Theorem 5.1. Then y can be obtained by fraction-free Gaussian elimination with column pivoting applied to the homogeneous system $Cy = 0$ and choosing a first pivot $C^{(1,0)} = \pm d(\vec{\nu}_k)\, f^{(1)}(0)$.*

With C rows numbered from 2 to m the algorithm is :

For $j = 1, 2, ..., m-1$
  Exchange columns j and $\lambda \geq j$ of $[\ C^T \mid y\ ]^T$
    in order to ensure $C^{(j+1,j)} \neq 0$
  For $\lambda = j+2, ..., m$
    For $\ell = j+1, ..., m$
      $C^{(\lambda,\ell)} \leftarrow \frac{C^{(\lambda,\ell)}C^{(j+1,j)} - C^{(j+1,\ell)}C^{(\lambda,j)}}{C^{(j,j-1)}}$
$y_m \leftarrow C^{(m,m-1)}$
For $\ell = m-1, m-2..., 1$
  $y_\ell \leftarrow -C^{(\ell+1,m)} y_m$
  For $\lambda = \ell+1, ..., m-1$: $y_\ell \leftarrow y_\ell - C^{(\ell+1,\lambda)} y_\lambda$
  $y_\ell \leftarrow y_\ell / C^{(\ell+1,\ell)}$

□

THEOREM 5.3. *Let $y = (y_1, \ldots, y_m)$ be as in Theorem 5.1, with $y_\pi \neq 0$. Let $a_k^{(\ell)} = coeff(\mathfrak{M}_k^{(\pi,\ell)}, z^{|\vec{\nu}_k| - \vec{\nu}_k^{(\pi)} - 1})$. Then for $j = 1, ..., m$ we have:*

$$\begin{aligned} d_k \mathfrak{M}_{k+1}^{(j,\pi)} &= \sum_{\ell=1}^{m} \mathfrak{M}_k^{(j,\ell)} y_\ell, && \text{if } \ell = \pi, \\ d_k \mathfrak{M}_{k+1}^{(j,\ell)} &= z\mathfrak{M}_k^{(j,\ell)} y_\pi - \mathfrak{M}_{k+1}^{(j,\pi)} a_k^{(\ell)}, && \text{if } \ell \neq \pi \\ d_{k+1} &= y_\pi. \end{aligned}$$

Here $d_0 = 1$ and the sign in $d_k = \pm d(\vec{\nu}_k)$ is chosen such that

$$\mathfrak{M}_k^{(\ell,\ell)} = d_k z^{|\vec{\nu}_k| - \vec{\nu}_k^{(\ell)}} + \text{ lower terms.}$$

PROOF. Consider the column vector polynomial

$$q = \sum_{\ell=1}^{m} \mathfrak{M}_k^{(\cdot,\ell)} y_\ell.$$

Since $\mathfrak{M}_k$ is a type 2 Mahler system for the index $\vec{\nu}_k$ with $|\vec{\nu}_k| = k$, the construction of y implies that q has order $k+1 = |\vec{\nu}_{k+1}|$. In addition

$$\deg\, q^{(j)} \begin{cases} = k - \vec{\nu}_k^{(\pi)} = |\vec{\nu}_{k+1}| - \vec{\nu}_{k+1}^{(\pi)} & \text{if } j = \pi \\ \leq k - \vec{\nu}_k^{(j)} = |\vec{\nu}_{k+1}| - 1 - \vec{\nu}_{k+1}^{(j)} & \text{otherwise,} \end{cases}$$

with the leading coefficient of the $\pi-$th row of q being equal to $d_k \cdot y_\pi = d_k \cdot d_{k+1}$. By the uniqueness of Lemma 2.3, this implies that $q = d_k \mathfrak{M}_{k+1}^{(\cdot,\pi)}$.

For the remaining columns note that if $\ell \neq \pi$ then

$$\deg\, z\mathfrak{M}_k^{(j,\ell)} \begin{cases} = k+1 - \vec{\nu}_k^{(\ell)} = |\vec{\nu}_{k+1}| - \vec{\nu}_{k+1}^{(\ell)} & \text{if } j = \ell \\ \leq k - \vec{\nu}_k^{(j)} = |\vec{\nu}_{k+1}| - 1 - \vec{\nu}_{k+1}^{(j)} & \text{if } j \neq \ell, \pi \\ \leq |\vec{\nu}_{k+1}| - \vec{\nu}_{k+1}^{(\pi)} & \text{if } j = \pi. \end{cases}$$

Thus all rows $j \neq \pi$ have correct degree bounds to be a type 2 approximant of degree bound $\vec{\nu}_{k+1} - \vec{e}_\ell$ except for row π where the degree is too large by 1. The degrees of this row are corrected by the second equation of the theorem since the coefficient of term $z^{|\vec{\nu}_{k+1}| - \vec{\nu}_{k+1}^{(\pi)}}$ then becomes

$$a_k \cdot y_\pi - d_{k+1} \cdot a_k = 0.$$

Finally, the leading coefficient of row ℓ in this case is given by $d_k y_\pi = d_k \cdot d_{k+1}$. Again by uniqueness the right hand side of the second equation above is the same as $d_k \mathfrak{M}_{k+1}^{(\cdot,\ell)}$ proving our result. □

In the normal case with a different, much simpler scaling, our formula for $\mathfrak{M}_{k+1}^{(j,\pi)}$ stated in Theorem 5.3 reduces to that found in Mahler [14, §24, third equation].

6. COMPLEXITY

In this section we show that our method provides a gain of m^3 over the fraction-free method found in [4].

THEOREM 6.1. *The complexity for computing a type 2 Mahler system $\mathfrak{M}_K$ for $K = |\vec{n}| \geq m$ and hence a basis for the set of simultaneous Padé approximants of index $\vec{n} - \vec{e}_\ell$ for input data of size $\mathcal{O}(\kappa)$ is $\mathcal{O}(m^2 K^4 \kappa^2)$.*

PROOF. Following the above approach, we have to compute recursively $\mathfrak{M}_k$ together with the first $K-k$ terms of the residuals $z^{-k}(f^{(j)}\mathfrak{M}_k^{(1,\ell)} - f^{(1)}\mathfrak{M}_k^{(j,\ell)})$ for $k = 1, 2, ..., K$. For the ℓth column of $\mathfrak{M}_k$ this gives

$$\sum_{\lambda} (k - \vec{\nu}_k^{(\lambda)}) = (m-1)k$$

scalar unknowns, plus an additional $(m-1)(K-k)$ scalar unknowns for the corresponding residuals, each of them of size $\mathcal{O}(k\kappa)$ according to (2.4). We can use the same updating formula for the residuals as for the Mahler systems, and hence the value at zero of these residuals gives us the corresponding matrix C of Theorem 5.1.

We then first need to solve a $(m-1) \times m$ homogeneous system of linear equations by fraction-free Gaussian elimination but with an initial pivot. Thus at elimination step $j = 1, ..., m-1$ the entries are of size $\mathcal{O}(\kappa(k+j))$ and this leads to a complexity of

$$\mathcal{O}(\kappa^2 (k+m)^2 \sum_{j=1}^{m} (m-j)^2) = \mathcal{O}(\kappa^2 K^2 m^3).$$

Coefficients of elements not in the pivot column are updated using $\mathcal{O}(1)$ multiplications/additions of two elements of size $\mathcal{O}(\kappa k)$, whereas for coefficients of elements in the pivot column we update using $\mathcal{O}(m)$ multiplications/additions of two elements of size $\mathcal{O}(\kappa k)$. Thus the cost of updating all coefficients in step k is given by $\mathcal{O}(mKk^2\kappa^2)$ for the $O(m)$ non pivot columns, and $\mathcal{O}(m^2 K k^2 \kappa^2)$ for the pivot column. Summing $k = 1, ..., K$ gives the claimed complexity $\mathcal{O}((m^3K^3 + m^2K^4)\kappa^2)$. □

7. EXAMPLE

Consider the vector of power series whose first terms are given by

$$\begin{aligned} f^{(1)}(z) &= 3 + 3\,z + 6\,z^2 + 18\,z^3 + 72\,z^4 + 360\,z^5 + O(z^{11}) \\ f^{(2)}(z) &= 1 + 8\,z^3 + 64\,z^6 + 512\,z^9 + O(z^{11}) \\ f^{(3)}(z) &= 1 - z + z^2 - z^3 + z^4 - z^5 + O(z^{11}) \end{aligned}$$

with $\vec{n} = (3, 4, 3)$. In this case the algorithm determines the closest normal indices as by

$$[0,1,0], [1,1,0], [1,2,0], [1,2,1], [2,2,1],$$
$$[2,3,1], [2,3,2], [3,3,2], [3,4,2], [3,4,3]$$

After step 4 the closest normal index is $\vec{v}_4 = (1, 2, 1)$ with the type 2 Mahler system given by

$$\mathfrak{M}_4 = \begin{bmatrix} 48\,z^3 + 33\,z^2 + 30\,z + 3 & -9\,z^2 - 126\,z - 27 & -54\,z^2 - 36\,z - 18 \\ 9\,z + 1 & 48\,z^2 - 33\,z - 9 & -6\,z - 6 \\ -8\,z^2 + 8\,z + 1 & 72\,z^2 - 24\,z - 9 & 48\,z^3 - 6 \end{bmatrix}.$$

Each column has order 4 in this case and are simultaneus Padé approximants of index $(0, 2, 1)$, $(1, 1, 1)$ and $(1, 2, 0)$, respectively. To construct $\mathfrak{M}_5$ the C matrix for the next step is given by

$$\begin{bmatrix} -6 & 54 & -252 \\ 210 & -738 & -252 \end{bmatrix}.$$

Here the kernel is $y = (1386, 378, 48)^T$ and the pivot column is $\pi = 1$. After replacing column 1 of $\mathfrak{M}_4$ by

$$\mathfrak{M}_4^{(.,1)} y_1 + \mathfrak{M}_4^{(.,2)} y_2 + \mathfrak{M}_4^{(.,3)} y_3$$

we get a new column 1 which has order 5. Multiplying columns 2 and 3 by z then implies that all the columns now have order 5. In this case the degrees of the resulting matrix polynomial are

$$\begin{bmatrix} 3 & 3 & 3 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix}$$

and the leading coefficient of the diagonal terms is 48×1386. Eliminating the highest terms in row 1 of columns 2 and 3 using cross multiplication with the new first row of column 1 gives degrees of the form

$$\begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix}.$$

These are the correct degree bounds for the type 2 Mahler system for the closest normal point $\vec{v}_5 = (2, 2, 1)$. Dividing out by 48 then gives the leading coefficient of all the diagonal elements as 1386. The result is $\mathfrak{M}_5$.

The iterative algorithm is particularly simple to implement. An implementation in the Maple computer algebra system along with a set of examples is available at the website

`www/cs.uwaterloo.ca/~glabahn/pade-code`.

8. CONCLUSION

In this paper we have given a new algorithm for the computation of simultaneous Padé approximants. The algorithm is meant for exact domains where coefficient growth is an issue. The algorithm is fraction-free with elimination done in the domain of computation but without any need for gcd calculations. The complexity improvement is by a factor of m^2 over previous fraction-free algorithms. The quantities computed are also an order of magnitude smaller over those of previous methods.

There are a number of topics for future research in this area. We expect that our methods can be easily generalized to the case of power series defined by linear functionals having *special* elements as done in [4]. In particular this would give a fraction-free algorithm for power series determined by interpolation data. We are also interested in extending our work to more general matrix rational interpolation problems and more general paths of computation as done in [3]. We are interested in seeing how our method can be used for problems defined over noncommutative domains, in particular when the input $f^{(\ell)}$ are differential (or more generally Ore) operators. Finally, we will investigate how our approach can be used for efficient fraction-free computation of matrix normal forms as done in [5].

9. REFERENCES

[1] E. Bareiss. Sylvester's identity and multistep integer-preserving gaussian elimination. *Math. Comp.*, 22(103):565–578, 1968.

[2] B. Beckermann and G. Labahn. A uniform approach for fast computation of matrix-type Padé approximants. *SIAM J. Matrix Analysis and its Applications*, 15(3):804–823, July 1994.

[3] B. Beckermann and G. Labahn. Recursiveness in matrix rational interpolation problems. *Journal of Computational and Applied Math*, 77:5–34, 1997.

[4] B. Beckermann and G. Labahn. Fraction-free computation of matrix rational interpolants and matrix GCDs. *SIAM J. Matrix Analysis and its Applications*, 22(1):114–144, 2000.

[5] B. Beckermann, G. Labahn, and G. Villard. Shifted normal forms of polynomial matrices. In *Proceedings of ISSAC 1999*, pages 189–196, 1999.

[6] S. Cabay and G. Labahn. A superfast algorithm for multi-dimensional Padé systems. *Numerical Algorithms*, 2:201–224, 1992.

[7] S. Cabay, G. Labahn, and B. Beckermann. On the theory and computation of non-perfect Padé-Hermite approximants. *Journal of Computational and Applied Math*, 39:295–313, 1992.

[8] W. Eberly, M. Giesbrecht, P. Giorgi, A. Storjohann, and G. Villard. Solving sparse rational linear systems. In *Proceedings of ISSAC 2006*, pages 63–70, 2006.

[9] K. O. Geddes, S. R. Czapor, and G. Labahn. *Algorithms for Computer Algebra*. Kluwer Academic Publ., Boston, Massachusetts, USA, 1992.

[10] P. Giorgi, Claude-Pierre Jeannerod, and G. Villard. On the complexity of polynomial matrix computations. In *Proceedings of ISSAC 2003*, pages 135–142, 2003.

[11] C. Hermite. Sur la fonction exponentielle. *C.R. Acad. Sci.*, 77:18–24,74–79,226–233, 1873.

[12] T. Kailath. *Linear Systems*. Prentice-Hall, 1980.

[13] G. Labahn. Inversion components for block Hankel-like matrices. *Linear Algebra and its Applications*, 177:7–48, 1992.

[14] K. Mahler. Perfect systems. *Compos. Math*, 19:95–166, 1968.

[15] Z. Olesh and A. Storjohann. The vector rational function reconstruction problem. In *Proceedings of the Waterloo Workshop on Computer Algebra: devoted to the 60th birthday of Sergei Abramov (WWCA)*, pages 137–149. World Scientific, 2007.

[16] H. Padé. Sur la représentation approchée d'une fonction par des fractions rationnelles. *Ann. Sci. École Norm. Sup.*, 77:3–93, 1892.

[17] M. van Hoeij. Factorization of differential operators with rational function coefficients. *Journal of Symbolic Computation*, 24:537–561, 1997.