JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

ELSEVIER Journal of Computational and Applied Mathematics 121 (2000) 197–219

www.elsevier.nl/locate/cam

Multivariate Padé approximation

Philippe Guillaume*, Alain Huard

UMR MIP 5640, Département de Mathématiques, INSA, Complexe Scientifique de Rangueil, 31077 Toulouse Cedex, France

Received 11 November 1999; received in revised form 20 February 2000

Abstract

This paper is a survey on the multivariate Padé approximation. Two types of approximants are considered: those which can approximate general meromorphic functions $f = h/g$ where both h and g are holomorphic, and those which are specialized to the approximation of functions of the same form where g is a polynomial. Algorithms are described, together with the different techniques used for proving convergence. © 2000 Elsevier Science B.V. All rights reserved.

MSC: 41A21; 41A63; 65D15

Keywords: Multivariate Padé approximants; Convergence of multivariate Padé approximants; Homogeneous approximants; Nonhomogeneous approximants; Padé $\circ$ Padé approximants; Nested Padé approximants; Least-squares Padé approximants

1. Introduction: from univariate to multivariate Padé approximation

Let $f(z)$ be a function defined on a subset of the complex plane. In many applications, the function is known through the first terms of its power series expansion. For example, in Electromagnetics or in Mechanics, the electric field (respectively the displacement) is the solution to a square linear system which depends on a parameter z (e.g. the frequency):

$$(A + Bz + Cz^2)\, f(z) = b(z).$$

If the matrix $A \in \mathscr{M}_n(\mathbb{C})$ is invertible and if $b(z)$ is holomorphic, then the solution $f(z)$ is holomorphic around the origin, and has a power series expansion

$$f(z) = \sum_{k \geqslant 0} c_k z^k, \quad c_k \in \mathbb{C}^n. \tag{1}$$

* Corresponding author.

E-mail addresses: guillaum@gmm.insa-tlse.fr (P. Guillaume), huard@gmm.insa-tlse.fr (A. Huard).

0377-0427/00/$ - see front matter © 2000 Elsevier Science B.V. All rights reserved.
PII: S0377-0427(00)00337-X

The vectors c_k in this series can easily be computed by solving successively the systems

$$Ac_0 = b_0,$$
$$Ac_1 = -Bc_0 + b_1,$$
$$Ac_k = -Bc_{k-1} - Cc_{k-2} + b_k, \quad k \geqslant 2,$$

where $b(z) = \sum_{k\geqslant 0} b_k z^k$ [13]. These systems are obtained by the identification of the coefficients of z^k in $(A + Bz + Cz^2)\sum_{k\geqslant 0} c_k z^k = \sum_{k\geqslant 0} b_k z^k$. Of course, only a finite number of coefficients c_k, $0 \leqslant k \leqslant N$, are computed and a good approximation of $f(z)$ may be obtained by a Taylor polynomial

$$f(z) \simeq \sum_{k=0}^{N} c_k z^k.$$

However, such an approximation will be accurate if series (1) itself converges, that is, if $|z| < \rho$, where ρ is the convergence radius of the series. Unfortunately, ρ is often finite because there are complex numbers z_i, $i = 1, 2, \ldots$, such that $\det(A + Bz_i + Cz_i^2) = 0$, for which $f(z_i)$ is not defined. Hence, the function $f(z)$ is usually a meromorphic function with poles z_i and a convergence radius $\rho = \min_i |z_i|$.

In such a case, it is well known that a Padé approximation can be far more accurate than a Taylor approximation. Essentially, it is a consequence of the famous Montessus de Ballore theorem, who established in 1902 the uniform convergence of Padé approximants on compact subsets excluding the poles. Particularly, a good approximation of $f(z)$ can be obtained outside the disk of convergence of series (1), where the Taylor expansion fails to converge.

The above-mentioned example depended on a single variable z. Often, there are, in fact, some other parameters like shape variables, material properties (Hookes law, electromagnetic properties), boundary conditions, etc. For such cases, it would be desirable to construct a multivariate Padé approximation. However, the problem in several variables is much more difficult than in one variable, and many research has been done in the last 30 years in order to find a generalization with good convergence properties.

In order to understand these difficulties, and the different solutions which have been proposed to overcome them, first let us consider the univariate Padé approximant of a scalar and meromorphic function f defined on the complex plane,

$$f(z) = \frac{u(z)}{v(z)}, \tag{2}$$

where u and v are both holomorphic functions on $\mathbb{C}$ and $v(0) \neq 0$. For given integers m and n, let $p(z)/q(z)$, $\deg p \leqslant m$, $\deg q \leqslant n$, be a nontrivial solution to the homogeneous and linear system

$$q(z)f(z) - p(z) = \mathrm{O}(z^{m+n+1}). \tag{3}$$

The proof of the Montessus de Ballore theorem is essentially based on the fact that the zeros of the function v form an at most countable set $Z = \{z_1, z_2, z_3, \ldots\}$ of isolated points. This property, which is particular to the univariate case, allows to rewrite the function f on a given disk $D(0, r)$ as a fraction with a polynomial denominator:

$$f(z) = \frac{h(z)}{g(z)}, \qquad g(z) = \prod_{|z_i|<r} (z - z_i)^{\alpha_i}, \quad \forall z \in D(0, r) \backslash Z, \tag{4}$$

where α_i is the order of the pole z_i, and $h(z) = u(z)g(z)/v(z)$ is holomorphic on $D(0,r)$. Let n be the degree of the polynomial g. Using this form of the function f, the Montessus de Ballore theorem states that the Padé approximants $p(z)/q(z)$ solution to (3) converge uniformly to f on all compact subset of $D(0,r)\backslash Z$ when $m \to \infty$. A natural generalization of this theorem would ask for multivariate rational approximants to converge to f uniformly on compact subsets excluding the zero set of the denominator.

There are mainly two difficulties in the generalization to several variables. The first one is that the substitution of a polynomial for v in Eq. (2) is no longer possible for a multivariate meromorphic function. The reason is that usually the zero set of an holomorphic function in $\mathbb{C}^d$, $d > 1$, does not coincide, even locally, with the zero set of a polynomial. This implies that for a general meromorphic function, one cannot hope from a sequence p/q of fractions where the degree of q remains bounded, to converge to f. The first way for overcoming this difficulty was proposed in 1988 by Chaffy who introduced the Padé $\circ$ Padé approximants [3]. In the case of two complex variables x, y, they are obtained in two steps. The first one consists in the computation of the Padé approximant of the function $f_y : x \mapsto f(x,y)$ with respect to the variable x. The second step consists in the computation of the Padé approximant of the resulting function with respect to the variable y. Using a similar approach, one of the authors introduced the nested Padé approximants [11]. The difference lies in the second step, where the Padé approximants of the coefficients of the first step result are computed. These two approximants are rational approximants of f in the field $\mathbb{C}(x,y)$ of fractions in x and y with coefficients in $\mathbb{C}$, but they are computed in the field $\mathbb{C}(x)(y)$ (or $\mathbb{C}(y)(x)$ for the second one) of fractions in y with coefficients in $\mathbb{C}(x)$. Of course, these fields are the same, but the representation is changed, and particularly the number of coefficients is not the same for a given degree. It may seem that such a representation is not as elegant as the usual one. For example, the symmetry in x, y is lost, though this fact may be used when the two variables are of different nature. However, what is gained through this representation is that convergence can be obtained for a large class of meromorphic functions, even with non polynomial denominators. The proofs of convergence are based on the fact that under some suitable assumptions the zeros of a function $v(x,y)$ can locally be identified with the zeros of a function $g(x,y)$ which is a polynomial with respect to one variable, that is of the form $g(x,y) = \prod_{i=1}^{n}(x - x_i(y))$. Then univariate techniques can be applied to this form.

In order to get around the previous difficulty, many authors have concentrated their attention on the functions f which can be written in the form

$$f(z) = \frac{h(z)}{g(z)}, \quad z = (z_1, z_2, \ldots, z_d) \in \mathbb{C}^d, \tag{5}$$

where h is holomorphic and g is a polynomial of degree n.

The second difficulty, which appears when the field $\mathbb{C}(z_1, z_2, \ldots, z_d)$ is used for the approximation, is that "*no formal equation analogous to* (the univariate case) *gives the correct number of linear equations to determine the coefficients*" [4]. Thus, several choices have been made in order to define multivariate Padé approximants. One of the first definitions was proposed by Chisholm in 1973 [4]. A few years later, a most general definition was given by Levin [16], from which a particular case was studied by Cuyt [5], the so-called homogeneous Padé approximants (see also [15] for a definition based on orthogonal polynomials). They are closely related to the univariate approximation, and they allowed Cuyt to obtain in 1985 the first uniform convergence result for the multivariate case [6], recently improved in [7]. For numerical applications, these approximants suffer from a lack of convergence on the complex lines where $t \mapsto g(tz)$ has less than n roots. Particularly, they have a singularity at the origin, which has been carefully studied by Werner [19]. Up to now, no convergence has been obtained for the general definition proposed by Levin. The standard proofs of consistency and convergence break down for the same reason that there are not enough equations to uniquely determine a Padé approximant (see, for example, (11)). In the Padé approximation theory, consistency usually means that if $f = h/g$ is a rational fraction, then its Padé approximant P/Q should be equal to f if the degrees of P and Q are correctly chosen. The least-squares Padé approximants introduced by the authors in [12] have allowed to obtain consistency and uniform convergence on compact subsets excluding the zero set of g. In the univariate case, the latter formulation provides an alternative to the classical Padé approximation, and coincides with it for a particular choice of the interpolation set.

This paper is organized as follows. The next section discusses the multivariate Padé approximation of functions $f = h/g$ where h is holomorphic and g is a polynomial, and Section 3 considers the more general case where g is holomorphic. A simple algorithm is given for computing each kind of multivariate Padé approximant. We also describe their convergence properties and show how the proofs of convergence are closely related to the univariate case.

2. Multivariate Padé approximants of f/g with g polynomial

Many definitions have been proposed for the multivariate approximation of a function $f(z) = h(z)/g(z)$, $z \in \mathbb{C}^d$, $d \geqslant 1$, where h is holomorphic and g is a polynomial of degree n. We will focus our attention to the two consistent approximations for which uniform convergence has been proved: the homogeneous (HPA) and the least-squares (LSPA) multivariate Padé approximants. It seems that although some other approximants may have some historical interest, their lack of convergence is a serious handicap for numerical applications.

In the HPA, the coefficients of the approximant P/Q are defined by a linear system which is over-determined for more than two variables. Due the particular choice of the degrees and the interpolation indices set, this system can be solved exactly. In the LSPA, the over-determined linear system defining the coefficients is solved in a weighted least-squares sense.

First some notation is introduced, and the consistency of a multivariate Padé approximation is discussed. Next, we describe the HPA and the LSPA. The proofs of convergence are very similar, and reported at the end of the section.

2.1. Notation

For a given finite subset $M \subset \mathbb{N}^d$, the set of polynomials $P \in \mathbb{C}[z]$ having the form $P(z) = \sum_{\alpha \in M} P_\alpha z^\alpha$ is denoted by $\mathbb{P}_M$. The standard notation $z^\alpha = \prod_{i=1}^d z_i^{\alpha_i}$ is used for $\alpha \in \mathbb{N}^d$ and $z \in \mathbb{C}^d$.

A polynomial $P \in \mathbb{P}_M$ is said *M-maximal* if for all polynomial $Q \in \mathbb{C}[z]$, the condition $PQ \in \mathbb{P}_M$ implies $Q \in \mathbb{C}$. The degree of a polynomial $P = \sum_\alpha P_\alpha z^\alpha$ is $\max\{|\alpha|, P_\alpha \neq 0\}$ where $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_d$, and the valuation of a series $S = \sum_\alpha S_\alpha z^\alpha$ is $\min\{|\alpha|,\ S_\alpha \neq 0\}$.

A subset $M \subset \mathbb{N}^d$ has the *rectangular inclusion property* if the conditions $\alpha \in \mathbb{N}^d$, $\beta \in M$ and $\alpha \leqslant \beta$ imply $\alpha \in M$. The standard partial order of $\mathbb{N}^d$ is used, that is, $\alpha \leqslant \beta$ means $\alpha_i \leqslant \beta_i$ for $1 \leqslant i \leqslant d$.

The number of elements of a finite subset M is denoted by $|M|$, and $M+N$ denotes the set $\{\alpha+\beta;\ \alpha\in M,\ \beta\in N\}$. If $P\in\mathbb{P}_M$ and $Q\in\mathbb{P}_N$, then $PQ\in\mathbb{P}_{M+N}$.

For a sequence $(M_m)_{m\geqslant 0}$, $M_m\subset\mathbb{N}^d$, we say that $\lim_{m\to\infty} M_m=\infty$ if for all bounded subset B of $\mathbb{N}^d$, there exists an integer k such that $B\subset M_m$ for all $m\geqslant k$. For the sake of simplicity, we will omit the subscript m and write $M\to\infty$.

For a function f which is holomorphic around the origin, the coefficient of z^α in its power series expansion is denoted by f_α, that is, $f(z)=\sum_{\alpha\geqslant 0} f_\alpha z^\alpha$, and for $E\subset\mathbb{N}^d$, $f_E(z)$ denotes the partial sum $f_E(z)=\sum_{\alpha\in E} f_\alpha z^\alpha$.

2.2. Consistency of a rational approximation

Here we suppose that f is a fraction,

$$f(z)=\frac{h(z)}{g(z)},\quad h\in\mathbb{P}_R,\ g\in\mathbb{P}_S,$$

where R and S are finite subsets of $\mathbb{N}^d$, and $g(0)\neq 0$. Consider three finite subsets $M, N, E\subset\mathbb{N}^d$, and a fraction P/Q, $P\in\mathbb{P}_M$, $Q\in\mathbb{P}_N$, such that

$$(f-P/Q)_E=0. \tag{6}$$

If the requirements $M\subset E$ and $|E|=|M|+|N|-1$ are added, then this equation corresponds to the general definition given by Levin [16].

Like in other domains of numerical analysis, consistency is almost necessary for convergence. The question is under which conditions on the sets M, N, E, does Eq. (6) define a consistent approximation, that is, Eq. (6) implies $P/Q=f$.

In the univariate case, one has $M=R=\{0,1,\ldots,m\}$, $N=S=\{0,1,\ldots,n\}$, and

$$E=M+N, \tag{7}$$

$$|E|=|M|+|N|-1. \tag{8}$$

An extra condition such as $Q(0)=1$ is usually added in order to avoid the zero solution. Then, Eq. (8) means that a square system is obtained for the free coefficients of P and Q. We will see that Eq. (7) is a consistency condition. When the sets M, N, E satisfy the rectangular inclusion property, the two conditions (7) and (8) are equivalent in the univariate case but incompatible in the multivariate case, because the identity $|M+N|=|M|+|N|-1$ holds only for $d=1$. Hence one of them must be abandoned.

Proposition 2.1. *Let $f=h/g$ be an irreducible fraction, where $h\in\mathbb{P}_R$, $g\in\mathbb{P}_S$, $g(0)\neq 0$. For a given $E\subset\mathbb{N}^d$ satisfying the inclusion property, let $(P,Q)\in\mathbb{P}_M\times\mathbb{P}_N$ be a nontrivial solution to the linear and homogeneous system*

$$(Qf-P)_E=0. \tag{9}$$

If

$$N+R\subset E\quad\text{and}\quad M+S\subset E, \tag{10}$$

then

$$\frac{P}{Q}=\frac{h}{g}.$$

Moreover, if $N=S$ *and if* g *is* N*-maximal, then there exists a constant* $c\in\mathbb{C}$ *such that*

$$P=ch,\quad Q=cg.$$

Proof. Due to $g(0)\neq 0$ and the rectangular inclusion property, system (9) is equivalent to $(Qh-Pg)_E=0$. Due to $Qh\in\mathbb{P}_{N+R}$, $Pg\in\mathbb{P}_{M+S}$, $N+R\subset E$ and $M+S\subset E$, we have $Qh=Pg$. It follows from the Gauss lemma that there exists a polynomial c such that $P=hc$, $Q=gc$, hence $P/Q=h/g$. If $Q\in\mathbb{P}_N$ and g is N-maximal, then $c\in\mathbb{C}$. □

This proposition shows that condition (7) is sufficient for consistency when $M=R$ and $N=S$. Without any special assumption on the function f, the latter condition is also necessary, as illustrated by the following example where trying to preserve (8) instead of (7) leads to divergence. Let

$$f(x,y)=\frac{1}{(1-x)(1-y)}. \tag{11}$$

The power series expansion of function f around the origin reads $f(x,y)=\sum_{ij}x^iy^j$. Let $M=R=\{(i,j)\in\mathbb{N}^2;\ i+j\leqslant 2\}$, $N=S=\{0,1\}^2$, $E=\{0,1,2\}^2$, $P(x,y)=1+x+y+x^2+y^2\in\mathbb{P}_M$ and $Q(x,y)=1-xy\in\mathbb{P}_N$. Then P and Q are solution to Eq. (9), condition (8) is fulfilled, but $P/Q\neq f$. Such an indetermination of the denominator coefficients appears also for higher degrees of the numerator, making it impossible to obtain uniform convergence of the Padé approximants to the function f on compact subsets of $\{(x,y)\in\mathbb{C}^2;\ (1-x)(1-y)\neq 0\}$ when M increases.

Observe that system (9) has the nontrivial solution (h,g), although it is usually strongly overdetermined. When h is not a polynomial, depending on the choice of the sets M, N, E, this system will not necessarily have some nontrivial solutions. In the homogeneous approach, system (9) can always be solved exactly by letting $|N|\to\infty$ when $M\to\infty$. In the least-squares approach, the set N is kept constant with $N=S$, and system (9) is solved in a least-squares sense.

2.3. The homogeneous multivariate Padé approximants

The HPA were introduced by Cuyt [5]. Let $f(z)=h(z)/g(z)$ be a meromorphic function where g is a polynomial of degree n with $g(0)\neq 0$.

2.3.1. Definition

The polynomials P and Q and the interpolation set E are chosen in the following way. For a given $m\geqslant 0$, consider the three sets

$$M=\{\alpha\in\mathbb{N}^d,\ mn\leqslant|\alpha|\leqslant mn+m\}, \tag{12}$$

$$N=\{\alpha\in\mathbb{N}^d,\ mn\leqslant|\alpha|\leqslant mn+n\}, \tag{13}$$

$$E=\{\alpha\in\mathbb{N}^d,\ 0\leqslant|\alpha|\leqslant mn+m+n\}. \tag{14}$$

We look for polynomials $P \in \mathbb{P}_M$ and $Q \in \mathbb{P}_N$ such that

$$(Qf - P)_E = 0. \tag{15}$$

For $R = \{\alpha \in \mathbb{N}^d,\ 0 \leqslant |\alpha| \leqslant m\}$ and $S = \{\alpha \in \mathbb{N}^d,\ 0 \leqslant |\alpha| \leqslant n\}$, one has $M + S = N + R = E$, and it follows from Proposition 2.1 that this approximation is consistent.

The idea is to introduce the univariate polynomials p and q defined, for fixed $z \in \mathbb{C}^d$, by $p(t) = t^{-mn}P(tz)$ and $q(t) = t^{-mn}Q(tz)$, $t \in \mathbb{C}$, of respective degrees m and n (at most). The coefficient of t^k in $P(tz)$ is an homogeneous polynomial in z of degree k, which explains the name given to these approximants. One can consider p and q as elements of the ring $\mathbb{C}[z][t]$, and compute a univariate Padé approximant solution to the linear system

$$q(t)f(tz) - p(t) = \mathrm{O}(t^{m+n+1}). \tag{16}$$

This system with $m+n+2$ unknown coefficients and $m+n+1$ equations has always a nontrivial solution in $(\mathbb{C}[z])^{m+n+2}$. Moreover, if the denominator below is not zero, a solution is given explicitly by Jacobi's determinant representation

$$\frac{p(t)}{q(t)} = \frac{\begin{vmatrix} t^n F_{m-n}(t) & t^{n-1}F_{m-n+1}(t) & \cdots & F_m(t) \\ f_{m-n+1} & f_{m-n+2} & \cdots & f_{m+1} \\ \vdots & \vdots & & \vdots \\ f_m & f_{m+1} & \cdots & f_{m+n} \end{vmatrix}}{\begin{vmatrix} t^n & t^{n-1} & \cdots & 1 \\ f_{m-n+1} & f_{m-n+2} & \cdots & f_{m+1} \\ \vdots & \vdots & & \vdots \\ f_m & f_{m+1} & \cdots & f_{m+n} \end{vmatrix}} \tag{17}$$

with

$$f_k = \sum_{|\alpha|=k} f_\alpha z^\alpha \quad \text{if } k \geqslant 0, \qquad f_k = 0 \quad \text{if } k < 0, \qquad F_k(t) = \sum_{i=0}^{k} f_i t^i.$$

It follows from (16) that $t^{mn}q(t)f(tz) - t^{mn}p(t) = \mathrm{O}(t^{mn+m+n+1})$ in $\mathbb{C}[z][[t]]$, the ring of power series in t with coefficients in $\mathbb{C}[z]$. By construction, the coefficient of t^k in $t^{mn}q(t)f(tz) - t^{mn}p(t)$ is homogeneous in z of degree k. This implies that $(q(1)f - p(1))_E = 0$. Moreover, $P(z) = p(1)$ and $Q(z) = q(1)$ are polynomials in z with valuations at least mn and respective degrees at most $mn+m$ and $mn+n$, thus P and Q are a solution to Eq. (15).

Observe that as $Qf - P$ has valuation at least mn, Eq. (15) can also be written

$$(Qf - P)_{E'} = 0, \quad E' = \{\alpha \in \mathbb{N}^d,\ mn \leqslant |\alpha| \leqslant mn + m + n\}. \tag{18}$$

For the particular dimension $d = 2$, $(Qf - P)_{E'} = 0$ is an homogeneous and linear system with

$$s = \binom{mn+m+n+2}{2} - \binom{mn+1}{2}$$

equations and $s+1$ unknowns. Hence it is not surprising to find a nontrivial solution. However, for $d > 2$, the latter system becomes over-determined, but still has some nontrivial solutions.

The solution P/Q found above is not necessarily an irreducible fraction. If P_1/Q_1 is another solution to Eq. (15), then $PQ_1 - P_1Q = 0$. Hence the irreducible form $P_{(m,n)}/Q_{(m,n)}$ of P/Q is unique.

Definition 2.1. The $[m,n]$ homogeneous multivariate approximant of the function f is the irreducible form $P_{(m,n)}/Q_{(m,n)}$ of P/Q where P and Q satisfy (15).

When P/Q is not irreducible, it may happen that the valuation of the polynomial $Q/Q_{(m,n)}$ has a positive valuation s, and there is a "backward shift" s on the valuations of P and Q. In that case, the polynomials $P_{(m,n)}$ and $Q_{(m,n)}$ do not necessarily satisfy Eq. (15). For example, if the backward shift is mn, then one can only guaranty $(Q_{(m,n)}f - P_{(m,n)})_F = 0$ where $F = \{\alpha \in \mathbb{N}^d,\ 0 \leqslant |\alpha| \leqslant m+n\}$.

In a more algebraic presentation, and following Brezinski's univariate theory, Kida defines in [16] the same multivariate approximant as a particular case of a Padé-type approximant for which the generating polynomial is precisely $q(t)$ in (17). However, one can observe that the substitution of tz for z, which allows to use the univariate construction, is made possible because of the particular choice of the sets E, M, N (12)–(14).

2.3.2. Computation

The $[m,n]$ homogeneous multivariate approximant can be computed in several ways. One possibility consists in solving directly (18). Here the unknowns are complex numbers. Although overdetermined if $d > 2$, this system has always some nontrivial solutions. The algorithm is the following.

Algorithm 1: HPA computation

1. Choose three enumerations $(\alpha_i)_{1\leqslant i\leqslant |M|}$, $(\beta_i)_{1\leqslant i\leqslant |N|}$, $(\gamma_i)_{1\leqslant i\leqslant |E'\backslash M|}$, of the respective subsets $M = \{\alpha \in \mathbb{N}^d,\ mn \leqslant |\alpha| \leqslant mn + m\}$, $N = \{\beta \in \mathbb{N}^d,\ mn \leqslant |\beta| \leqslant mn + n\}$ and $E'\backslash M = \{\gamma \in \mathbb{N}^d,\ mn + m < |\gamma| \leqslant mn + m + n\}$.
2. Compute the coefficients of the $|E'\backslash M| \times |N|$ complex matrix A:
$$A_{ij} = f_{\gamma_i - \beta_j}, \quad 1 \leqslant i \leqslant |E'\backslash M|,\ 1 \leqslant j \leqslant |N|,$$
where $f_\alpha := 0$ if $0 \nleqslant \alpha$.
3. Compute a nontrivial solution $v \in \mathbb{C}^{|N|}$ to the system $Av = 0$, and define the polynomial Q by
$$Q(z) = \sum_{i=1}^{|N|} v_i z^{\beta_i}.$$
4. Compute
$$w_i = \sum_{j=1}^{|N|} v_j f_{\alpha_i - \beta_j}, \quad 1 \leqslant i \leqslant |M|$$
and define the polynomial P by
$$P(z) = \sum_{i=1}^{|M|} w_i z^{\alpha_i}.$$
5. The fraction P/Q is the $[m,n]$ HPA of the function f.

A second possibility is to use symbolic computation for solving (16) in the ring $\mathbb{C}[z]$. Here the unknowns are complex polynomials, and this system is always under-determined (size $(m+n+1)\times(m+n+2)$). Such an approach is also used for computing the Padé ∘ Padé approximants (cf. Section 3.3).

Finally, one can also take advantage of the construction of the homogeneous approximant. For $z\in\mathbb{C}^d$ such that $t\mapsto g(tz)$, $t\in\mathbb{C}$, has exactly n roots, the rational fraction $t\mapsto P_{(m,n)}(tz)/Q_{(m,n)}(tz)$ is the $[m,n]$ univariate Padé approximant of the function $f_z(t):=f(tz)$. Hence all the algorithms developed in the univariate case can be applied to compute $P_{(m,n)}(tz)/Q_{(m,n)}(tz)=[m,n]_{f_z}(t)$. Particularly, computation of staircase sequences like in the ε-algorithm or the qd-algorithm can be used. For a description of these two algorithms, we refer the reader to [8].

2.3.3. Convergence

Here we consider sequences of $[m,n]$ HPA where n is fixed and $m\to\infty$. The degrees and the valuations of the numerator and the denominator in the nonreduced form P/Q increase when $m\to\infty$, and the convergence is obtained on compact subsets excluding the zero set of g if there exists a subsequence of approximants $P_{(m(k),n)}/Q_{(m(k),n)}$ such that $Q_{(m(k),n)}\neq 0$. This implies that the backward shift in the denominator Q of $[m,n]$ must be at least mn. For a given function f, the existence of such a subsequence remains to our knowledge an open question. When it exists, it can be interpreted as a particular case of the LSPA for which the underlying system is solved exactly. In the general case, the HPA converge on a smaller subset obtained by removing also the complex cone formed by the vectors z for which $t\mapsto g(tz)$, $t\in\mathbb{C}$, has less than n roots (Theorem 2.3).

Suppose that $f=h/g$ is meromorphic on a neighborhood of a polydisc $\bar{D}(0,\rho)=\{z\in\mathbb{C}^d;\,|z_i|\leqslant\rho_i,\ i=1,\dots,d\}$, and h is holomorphic on a neighborhood of $\bar{D}(0,\rho)$. The polynomial g is normalized in such a way that $\sum_{\alpha\in N}|g_\alpha|^2=1$. Its decomposition into irreducible factors reads

$$g=\prod_{i=1}^{l} g_i^{\tau_i}$$

and the associated algebraic set G with its decomposition into irreducible components G_i are:

$$G=\{z\in\mathbb{C}^d;\ g(z)=0\},$$

$$G_i=\{z\in\mathbb{C}^d;\ g_i(z)=0\}.$$

Suppose also that $G_i\cap D(0,\rho)\neq\emptyset$ for $1\leqslant i\leqslant l$, and that $h(z)\neq 0$ on a dense subset of $G\cap D(0,\rho)$. The following theorem was proved in [6].

Theorem 2.2 (Cuyt [6]). *Let $(P_{(m(k),n)}/Q_{(m(k),n)})_{k\geqslant 0}$ be a subsequence of homogeneous multivariate approximants such that $Q_{(m(k),n)}(0)\neq 0$ for all $k\geqslant 0$. Then*

$$\lim_{k\to\infty} P_{(m(k),n)}/Q_{(m(k),n)}(z)=f(z)$$

uniformly on all compact subsets of $\{z\in D(0,\rho);\ g(z)\neq 0\}$. Moreover, the subsequence $(Q_{(m(k),n)})_{k\geqslant 0}$ converges to $g(z)$ uniformly on all compact subsets of $D(0,\rho)$.

The following result has been obtained in [7] where more general sets than the polydisc $D(0,\rho)$ are considered. Let Λ be the set of vectors $z\in\partial D(0,\rho)$ for which the polynomial $t\mapsto g(tz)$, $t\in\mathbb{C}$, has less than n roots counted with multiplicity in $\bar{D}(0,1)$, and denote by E_Λ the cone $\{tz;\,t\in\mathbb{C},\,z\in\Lambda\}$.

Theorem 2.3 (Cuyt and Lubinsky [7]). *If $h(z)\neq 0$ for $z\in G$, then*

$$\lim_{m\to\infty} P_{(m,n)}/Q_{(m,n)}(z)=f(z)$$

uniformly on all compact subsets of $\{z\in D(0,\rho);\ z\notin E_A,\ g(z)\neq 0\}$.

2.4. The least-squares multivariate Padé approximants

Studying least-squares orthogonal polynomials, Brezinski proposed recently a least-squares formulation for univariate Padé approximants [2]. This idea has been generalized to the multivariate case in [12]. The formulation [2] did not involve any particular weights in the least-squares approximation, whereas some weights were introduced in [12], which have an important role as it can be seen from the proof of Theorem 2.4.

2.4.1. Definition

The requirements on the function f are the same than in Section 2.3.3. The norm on $\mathbb{P}_N$ is defined by $||Q||=(\sum_{\alpha\in N}|Q_\alpha|^2)^{1/2}$. For $P\in\mathbb{P}_M$, $Q\in\mathbb{P}_N$, and a finite set $E\subset\mathbb{N}^d$, consider the function

$$j(P,Q)=\left(\sum_{\alpha\in E}\rho^{2\alpha}|(Qf-P)_\alpha|^2\right)^{1/2}. \tag{19}$$

Definition 2.2. Let M, N, $E\subset\mathbb{N}^d$ be three finite subsets such that $E\supset M+N$ and E satisfies the inclusion property. A least-squares multivariate Padé approximant of the function f is a fraction P/Q with $(P,Q)\in\mathbb{P}_M\times\mathbb{P}_N$, $||Q||=1$, and

$$j(P,Q)\leqslant j(R,S),\quad \forall(R,S)\in\mathbb{P}_M\times\mathbb{P}_N,\quad ||S||=1. \tag{20}$$

A solution to this problem is denoted by $[M,N]_f$.

For $d=1$ and $E=M+N$, this definition coincides with the standard definition of the univariate Padé approximation. For $d=1$ and $E\supsetneq M+N$, the least-squares formulation provides an alternative to the exact Padé interpolation.

Observe that $j(P,Q)=0$ if $(g,h)\in\mathbb{P}_M\times\mathbb{P}_N$, and it follows from Proposition 2.1 that this approximation is consistent. Although there may exist several LSPA for given M, N and E (even if one considers the irreducible form), the next theorem shows that it has no incidence on the convergence.

2.4.2. Computation

In order to solve (20), first the coefficients of Q are computed, then the coefficients of P are recovered by expanding and truncating the product Qf, that is, $P=(Qf)_M$. The coefficients of Q are solution to

$$\min_{||Q||=1}\sum_{\alpha\in E\setminus M}\rho^{2\alpha}|(Qf)_\alpha|^2,$$

which can be written in the form

$$\min_{||s||^2=1} s^*A^*As, \tag{21}$$

where the vector $s\in\mathbb{C}^{|N|}$ contains the coefficients of Q, A is an $|E\backslash M|\times|N|$ complex matrix and A^* is the conjugate transpose of A. The optimality condition reads

$$A^*As=\lambda s,\quad \lambda\in\mathbb{R}.$$

Hence $s^*A^*As=\lambda\geqslant 0$, and a solution is given by any normalized eigenvector associated to the smallest eigenvalue value of A^*A. The algorithm is the following.

Algorithm 2: LSPA computation

1. Choose three enumerations $(\alpha_i)_{1\leqslant i\leqslant|M|}$, $(\beta_i)_{1\leqslant i\leqslant|N|}, (\gamma_i)_{1\leqslant i\leqslant|E\backslash M|}$, of the respective subsets M, N and $E\backslash M$.
2. Choose $\rho\in\mathbb{R}_+^d$ and compute the coefficients of the matrix A:
$$A_{ij}=\rho^{\gamma_i}f_{\gamma_i-\beta_j},\quad 1\leqslant i\leqslant|E\backslash M|,\ 1\leqslant j\leqslant|N|,$$
where $f_\alpha:=0$ if $0\nleqslant\alpha$.
3. Compute an eigenvector $v\in\mathbb{C}^{|N|}$ associated to the smallest eigenvalue of A^*A, and define the polynomial Q by
$$Q(z)=\sum_{i=1}^{|N|}v_iz^{\beta_i}.$$
4. Compute
$$w_i=\sum_{j=1}^{|N|}v_jf_{\alpha_i-\beta_j},\quad 1\leqslant i\leqslant|M|$$
and define the polynomial P by
$$P(z)=\sum_{i=1}^{|M|}w_iz^{\alpha_i}.$$
5. The fraction P/Q is an $[M,N]$ LSPA of the function f.

2.4.3. Convergence

In the following theorem [12], the set N is fixed, and can be any finite subset of $\mathbb{N}^d$ such that $g\in\mathbb{P}_N$ and g is N-maximal (cf. Section 2.1).

Theorem 2.4 (Guillaume et al. [12]). *Let $[M,N]_f=P^M/Q^M$ be a sequence of least-squares multivariate Padé approximants with $M\to\infty$. Then*

$$\lim_{M\to\infty}[M,N]_f(z)=f(z)$$

uniformly on all compact subsets of $\{z\in D(0,\rho);\ g(z)\neq 0\}$. Moreover, the sequence $Q^M(z)$ converges to $g(z)$ uniformly on all compact subsets of $D(0,\rho)$.

2.5. Proof of Theorems 2.2, 2.3 and 2.4

The proofs are an extension of the beautiful technique introduced by Karlsson and Wallin in the univariate case [14], which is based on the uniform convergence to zero of the function H^M in (22).

2.5.1. Proof of Theorems 2.2 and 2.4

Theorem 2.2 can be seen as a particular case of Theorem 2.4 by using the set $E' = \{\alpha \in \mathbb{N}^d, 0 \leqslant |\alpha| \leqslant m + n + mn - s\}$ instead of E (where s, $mn \leqslant s \leqslant mn + n$, is the backward shift on the valuation of Q, coming from the assumption $Q_{(m,n)}(0) \neq 0$), the notation $P^M = P_{(m,n)}$, $Q^M = Q_{(m,n)}$, and the fact that $(Q^M f - P^M)_{E'} = 0$, that is, the least-squares approximation is exact for the HPA. We outline the proof of Theorem 2.4.

Let $(P^M, Q^M) \in \mathbb{P}_M \times \mathbb{P}_N$ be a solution to problem (20) and consider the function

$$H^M = g(Q^M f - P^M), \tag{22}$$

which is holomorphic on a neighborhood of $\bar{D}(0,\rho)$. The keystone of the proof is the following lemma, whose proof is given at the end of the section.

Lemma 2.5. *One has*

$$\lim_{M\to\infty} H^M(z) = 0 \tag{23}$$

uniformly on all compact subsets of $D(0,\rho)$.

According to Definition 2.2, the sequence $(Q^M)_M$ is bounded in $\mathbb{P}_N$. Consider an arbitrary subsequence, still denoted by $(Q^M)_M$ for simplicity, which converges to a polynomial $Q \in \mathbb{P}_N$ with $||Q|| = 1$. The subsequence $(Q^M)_M$ converges also to Q, uniformly on all compact subsets of $\mathbb{C}^d$ when $M \to \infty$.

The set $G_i \cap D(0,\rho)$ was supposed nonempty. For $z \in G_i \cap D(0,\rho)$, one has $H^M(z) = h(z)Q^M(z)$ and $(H^M)_M$ converges to 0 on $D(0,\rho)$, thus $h(z)Q(z) = 0$, and $Q(z) = 0$ because $h(z) \neq 0$ on a dense subset of $G \cap D(0,\rho)$. The set of regular points of G_i is open, connected and dense in G_i, thus $Q = 0$ on G_i and g_i divides Q [1,18]. Similarly $g_i^{\tau_i}$ divides Q (consider the partial derivatives of H^M), which implies that g divides Q. One has $Q \in \mathbb{P}_N$ and g is N-maximal with $||g|| = 1$, thus

$$Q = cg, \quad |c| = 1.$$

Hence $\lim_{M\to\infty} Q^M = cg$ uniformly on all compact subsets of $\mathbb{C}^d$, and after division of (22) by gQ^M, one obtains with (23)

$$\lim_{M\to\infty} \left(f - \frac{P^M}{Q^M} \right)(z) = 0$$

uniformly on all compact subsets of $\{z \in D(0,\rho);\ g(z) \neq 0\}$. As this holds for all convergent subsequences of the bounded sequence $(Q^M)_M$, the whole sequence $f - P^M/Q^M$ converges to zero in the same way. □

Proof of the Lemma. The main line is the following. One has $gP^M \in \mathbb{P}_{M+N}$ and $E \supset M + N$, thus

$$H_\alpha^M = \begin{cases} (hQ^M)_\alpha & \text{if } \alpha \notin E, \\ (g(Q^M f - P^M)_E)_\alpha & \text{if } \alpha \in E \end{cases}$$

(and $H_\alpha^M = 0$ for the HPA if $\alpha \in E'$). The Cauchy integral yields

$$H_\alpha^M = \frac{1}{(2i\pi)^d} \int_{\Gamma_+} \frac{hQ^M}{z^{\alpha+\mathbf{1}}}\, \mathrm{d}z \quad \text{if } \alpha \notin E, \tag{24}$$

$$H_\alpha^M = \frac{1}{(2i\pi)^d}\int_{\Gamma_+}\frac{g(Q^M f - P^M)_E}{z^{\alpha+\mathbf{1}}}\,dz \quad \text{if } \alpha\in E, \tag{25}$$

where $\alpha+\mathbf{1}=(\alpha_1+1,\dots,\alpha_d+1)$. The sequence $(Q^M)_M$ is bounded in $\mathbb{P}_N$ and h is continuous on Γ_+, thus

$$|H_\alpha^M| \leqslant \frac{c}{\rho^\alpha} \quad \text{if } \alpha\notin E. \tag{26}$$

The change of variable $z=(\rho_1\exp(2i\pi\theta_1),\dots,\rho_d\exp(2i\pi\theta_d))$ in Eq. (25) yields

$$H_\alpha^M = \int_{[0,1]^d}\frac{g(Q^M f - P^M)_E}{z^\alpha}\,d\theta.$$

The Cauchy–Schwarz inequality and Parseval's formula give

$$|H_\alpha^M| \leqslant \left(\int_{[0,1]^d}\frac{|g|^2}{\rho^{2\alpha}}\,d\theta\right)^{1/2}\left(\sum_{\alpha\in E}\rho^{2\alpha}|(Q^M f - P^M)_\alpha|^2\right)^{1/2}.$$

Thus, using definition (19) of the function j, one has (possibly with a different c)

$$|H_\alpha^M| \leqslant \frac{c}{\rho^\alpha}\, j(P^M,Q^M) \quad \text{if } \alpha\in E. \tag{27}$$

Due to the definition of P_M, Q_M and to $(h_M,g)\in\mathbb{P}_M\times\mathbb{P}_N$, one has $j(P^M,Q^M)\leqslant j(h_M,g)$, and gathering Eqs. (26) and (27), one obtains

$$|H^M(z)| \leqslant c\left(j(h_M,g)\sum_{\alpha\in E}\left|\frac{z}{\rho}\right|^\alpha + \sum_{\alpha\notin E}\left|\frac{z}{\rho}\right|^\alpha\right), \tag{28}$$

where $|z/\rho|^\alpha = |z_1/\rho_1|^{\alpha_1}\cdots|z_d/\rho_d|^{\alpha_d}$. It follows from

$$j(h_M,g) = \left(\frac{1}{(2i\pi)^d}\int_{\Gamma_+}\frac{|h_{E\setminus M}|^2}{z}\,dz\right)^{1/2},$$

that $\lim_{M\to\infty} j(h_M,g)=0$, hence $\lim_{M\to\infty}|H^M(z)|=0$, uniformly on all compact subsets of $D(0,\rho)$. □

2.5.2. *Proof of Theorem 2.3*

The proof is based on a univariate projection, which allows to shift the degrees of $P_{(m,n)}$ and $Q_{(m,n)}$, that is, for a given $z\in\partial D(0,\rho)$, the univariate Padé approximant of $t\mapsto f(tz)$ reads

$$P_{m,z}(t)/Q_{m,z}(t) = P_{(m,n)}(tz)/Q_{(m,n)}(tz),$$

where $\deg P_{m,z}\leqslant m$, $\deg Q_{m,z}\leqslant n$. Here again, the key idea is to show that the function

$$\begin{aligned} H_{m,z}(t) &= g(tz)(Q_{m,z}(t)f(tz) - P_{m,z}(t)) \\ &= Q_{m,z}(t)h(tz) - P_{m,z}(t)g(tz), \end{aligned}$$

converges uniformly to zero. A local extension of an estimation similar to (26) is obtained, which leads to the local convergence (in z) of $Q_{m,z}$ to a polynomial Q_z, $\deg Q_z\leqslant n$. Particularly, if $g(tz)=0$, then $H_{m,z}(tz)=Q_{m,z}(t)h(tz)$. Taking the limit, it follows from $h(tz)\neq 0$ that $Q_z(t)=0$. Hence Q_z has exactly the same n roots than the polynomial $t\mapsto g(tz)$, which allows to complete the proof after division by gQ_z. □

3. Multivariate Padé approximants of f/g with g holomorphic

In this section are presented the Padé $\circ$ Padé approximants (PRPA) and the nested Padé approximants (NPA) for a meromorphic function $f(z)=h(z)/g(z)$. Both of them have a natural recursive structure and can be defined for $z\in\mathbb{C}^d$, $d>1$. For the sake of simplicity they are here presented in the case of two complex variables x and y. For more variables $z_1,z_2,\dots,z_d$, one substitutes z_1 for x and $(z_2,\dots,z_d)$ for y. Both PRPA and NPA are consistent and convergent, and start for fixed y with the univariate Padé approximant $[m,n]_{f_y}(x)$ of the function $f_y: x\mapsto f(x,y)$. The fraction $[m,n]_{f_y}(x)$ is an element of $\mathbb{C}[[y]](x)$, and is nothing else than a parameterized univariate Padé approximant.

In the PRPA, one computes in $\mathbb{C}(x)[[y]]$ the power series expansion of $[m,n]_{f_y}(x)$, and then, for fixed x, one computes the univariate Padé approximant of the function $x\mapsto[m,n]_{f_y}(x)$. The calculations are done in the field $\mathbb{C}(x)$, hence a good way of doing them is to use symbolic computation. In the NPA, one computes directly the univariate Padé approximants of the coefficients of $[m,n]_{f_y}(x)$, which belong to $\mathbb{C}[[y]]$. The computation does not need symbolic computation.

We point out the fact that in both cases only univariate Padé approximants are computed, for which much knowledge has been accumulated. Also noteworthy is that this kind of approximation can be applied to a larger class of functions than the approximations described in Section 2 because g needs not to be a polynomial.

Due to their construction, the convergence of PRPA or NPA cannot be obtained on all compact subset excluding the singularity of f. The complex lines (x,y) such that $[m,n]_{f_y}$ is not defined must also be removed from the convergence set. It is a sort of intermediate situation between the HPA and the LSPA, where the extra singularities of the HPA have been shifted away: instead of complex lines passing through the origin, these lines are here parallel to the x-axis.

First some notation is introduced. Then we describe the first step and give an intermediate convergence result which will be used for the convergence analysis of both PRPA and NPA.

3.1. Notation

Recall that in the univariate case, if the following linear system

$$q(x)u(x)-p(x)=\mathrm{O}(x^{m+n+1}),\quad q(0)=1 \tag{29}$$

has a unique solution, then the fraction p/q is irreducible and is called the $[m,n]$ Padé approximant of the function $u=\sum_{k\geqslant 0}u_kx^k$. This fraction is denoted by $[m,n]_u$. The Hankel matrix corresponding to this system is denoted by $H(u,m,n)$, and the right member by $C(u,m,n)$:

$$H(u,m,n)=\begin{pmatrix} u_{m-n+1} & \cdots & u_m \\ \vdots & & \vdots \\ u_m & \cdots & u_{m+n-1}\end{pmatrix},\quad C(u,m,n)=-\begin{pmatrix} u_{m+1} \\ \vdots \\ u_{m+n}\end{pmatrix},$$

where $u_i:=0$ if $i<0$. The coefficients $S=(q_n,\dots,q_1)^{\mathrm{T}}$ are solution to the system

$$H(u,m,n)S=C(u,m,n) \tag{30}$$

and the other coefficients p_i, $0\leqslant i\leqslant m$, are recovered by expanding the product $u(x)q(x)$.

3.2. First step (a parameterized Padé approximant)

Let f be a meromorphic function on a neighborhood of a polydisc $\bar{D}(0,\rho_1,\rho_2)$,

$$f(x,y)=\frac{u(x,y)}{v(x,y)},$$

where the functions u and v are holomorphic on a neighborhood of $\bar{D}(0,\rho_1,\rho_2)$. For the sake of simplicity, we make the following assumption: $v(x,y)=\sum_{i=0}^{n} v_i(y)x^i$ is a polynomial in x such that $x\mapsto v(x,0)$ has n simple roots with $v(0,0)\neq 0$. A particular case is when v is a polynomial in the two variables x and y. In the general case, the set where $v(x,y)$ vanishes is not necessarily algebraic.

Let $\mathscr{Y}_0\subset D(0,\rho_2)$ be an open subset where the function $y\mapsto f(0,y)$ is holomorphic and the determinant of $H(f_y,m,n)$ is nonzero, and suppose that $0\in\mathscr{Y}_0$. For a fixed $y\in\mathscr{Y}_0$, we can consider the $[m,n]$ Padé approximant of the function $f_y: x\mapsto f(x,y)$,

$$[m,n]_{f_y}(x)=\frac{U^m(x,y)}{V^m(x,y)}=\frac{\sum_{i=0}^{m} s_i^m(y)x^i}{1+\sum_{i=1}^{n} s_{m+i}^m(y)x^i}. \tag{31}$$

The subscript m indicates the dependence on m, whereas n is fixed once for all.

3.2.1. Computation

The coefficients of this parameterized Padé approximant can be computed in $\mathbb{C}[[y]]$ in the following way. For $y\in\mathscr{Y}_0$, the vector $S(y)=(s_{m+n}^m(y),\ldots,s_{m+1}^m(y))^{\mathrm{T}}$ is the unique solution to the linear system

$$H(f_y,m,n)S(y)=C(f_y,m,n). \tag{32}$$

Due to the assumption $\det H(f_0,m,n)\neq 0$, the vector-valued function $S(y)$ is holomorphic around zero and has a power series expansion

$$S(y)=\sum_{j\geqslant 0} S_j y^j,\quad S_j\in\mathbb{C}^n.$$

The power series expansion of H and C read

$$H(f_y,m,n)=\sum_{j\geqslant 0} H_j y^j,\quad H_j\in\mathbb{C}^{n\times n},$$

$$C(f_y,m,n)=\sum_{j\geqslant 0} C_j y^j,\quad C_j\in\mathbb{C}^n.$$

It follows from (32) that the vectors S_j are solution to the systems

$$H_0S_0=C_0, \tag{33}$$

$$H_0S_j=-\sum_{k=1}^{j} H_kS_{j-k}+C_j,\quad j\geqslant 1, \tag{34}$$

which all have the same matrix. Like in the univariate case, the other coefficients $s_i^m(y)$, $0\leqslant i\leqslant m$ are obtained by expanding in x the product $f_y(x)V^m(x,y)$. The pseudo-algorithm is the following (series are here considered, which will be later truncated in Algorithms 4 and 5).

Algorithm 3: Intermediate Padé approximant computation

1. Compute

$$H_j = \begin{pmatrix} f_{m-n+1,j} & \dots & f_{m,j} \\ \vdots & & \vdots \\ f_{m,j} & \dots & f_{m+n-1,j} \end{pmatrix}, \qquad C_j = -\begin{pmatrix} f_{m+1,j} \\ \vdots \\ f_{m+n,j} \end{pmatrix}, \quad j \geqslant 0,$$

where $f(x,y) = \sum_{i,j\geqslant 0} f_{i,j} x^i y^j$ and $f_{i,j} := 0$ if $i < 0$.

2. Solve (33) and (34) for $j \geqslant 1$. Using the numbering $S_j = (s^m_{m+n,j}, \dots, s^m_{m+1,j})^{\mathrm{T}}$, define

$$s^m_{m+i}(y) = \sum_{j\geqslant 0} s^m_{m+i,j} y^j, \quad 1 \leqslant i \leqslant n,$$

$$V(x,y) = 1 + \sum_{i=1}^{n} s^m_{m+i}(y) x^i.$$

3. Compute

$$s^m_{i,j} = \sum_{k=0}^{n} s^m_{m+k,j} f_{i-k,j}, \quad 0 \leqslant i \leqslant m, \ j \geqslant 0,$$

where $s^m_{m,0} := 1$ and $s^m_{m,j} := 0$ for $j > 0$, and define

$$s^m_i(y) = \sum_{j\geqslant 0} s^m_{i,j} y^j, \quad 0 \leqslant i \leqslant m,$$

$$U(x,y) = \sum_{i=0}^{m} s^m_i(y) x^i.$$

4. The function $U(x,y)/V(x,y)$ is the $[m,n]$ intermediate Padé approximant of f.

3.2.2. Intermediate convergence

The convergence is a direct consequence of the theory developed in the univariate case. Let $\mathscr{Y} \subset D(0,\rho_2)$ be an open subset with $0 \in \mathscr{Y}$ such that for all $y \in \mathscr{Y}$:

- $v_0(y) \neq 0, \ v_n(y) \neq 0$,
- the polynomial $x \mapsto v(x,y)$ has n simple roots $\alpha_i(y), \ 1 \leqslant i \leqslant n, \ |\alpha_i(y)| < \rho_1$, the functions α_i being holomorphic on $\mathscr{Y}$ (simple roots can be replaced by roots of constant multiplicity),
- $u(x,y) \neq 0$ if $v(x,y) = 0$.

The following lemma was proved in [3].

Lemma 3.1 (Chaffy-Camus [3]). *For all compact subsets $\mathscr{K}_y \subset \mathscr{Y}$, there is an integer m_0 such that for all $m \geqslant m_0$ and all $y \in \mathscr{K}_y$, there is a unique intermediate Padé approximant $[m,n]_{f_y} = U^m(x,y)/V^m(x,y)$. Let $\mathscr{O}$ be the open subset*

$$\mathscr{O} = \{(x,y) \in D(0,\rho_1,\rho_2), \ y \in \mathscr{Y}, \ v(x,y) \neq 0\}.$$

The sequence $([m,n]_{f_y})_{m\geqslant m_0}$ converges uniformly to f on all compact subsets of $(D(0,\rho_1) \times \mathscr{Y}) \cap \mathscr{O}$. Moreover, the sequence $(V^m(x,y))_{m\geqslant 0}$ converges to $v(x,y)$ uniformly on all compact subsets of $\mathbb{C} \times \mathscr{Y}$.

Proof. We give the main line of the proof, which is adapted from the very elegant technique used by Saff in the univariate case [17]. Define

$$\tilde{U}^m(x,y) = \frac{U^m(x,y)}{s^m_{m+n}(y)}, \quad \tilde{V}^m(x,y) = \frac{V^m(x,y)}{s^m_{m+n}(y)}.$$

After dividing in f the numerator and the denominator by the function v_n (which does not vanish on $\mathscr{Y}$), the function f can be put in the following form which fits the form $\tilde{U}^m/\tilde{V}^m$ of $[m,n]_{f_y}$:

$$f(x,y) = \frac{h(x,y)}{g(x,y)}, \quad g(x,y) = \sum_{i=0}^{n-1} g_i(y)x^i + x^n,$$

where the functions h and g are meromorphic on $D(0,\rho_1,\rho_2)$ and holomorphic on $D(0,\rho_1)\times\mathscr{Y}$. The idea is to search $\tilde{V}^m(x,y)$ under the form

$$\tilde{V}^m(x,y) = g(x,y) + \sum_{k=0}^{n-1} w_k^m(y)W_k(x,y),$$

where $W_0 \equiv 1$, $W_k(x,y) = (x-\alpha_1(y))\cdots(x-\alpha_k(y))$ is a polynomial in x of degree k, holomorphic on $\mathbb{C}\times\mathscr{Y}$, and to reformulate the problem as follows.

For fixed $y\in\mathscr{Y}$, let $\pi_m(x,y)$ be the Taylor expansion of degree $m+n$ at $x=0$ of the function $x \mapsto \tilde{V}^m(x,y)h(x,y)$. The coefficients $w_k^m(y)$ are chosen in such a way that the polynomial in x, $\pi_m(x,y)$ vanishes at the n roots $\alpha_k(y)$ of $g(.,y)$. Hence, there exists a polynomial in x, $\tilde{U}^m(x,y)$ such that $\pi_m(x,y)=\tilde{U}^m(x,y)g(x,y)$, and it follows that $(h\tilde{V}^m - g\tilde{U}^m)(x,y)=\mathrm{O}(x^{m+n+1})$. If $\tilde{V}^m(0,y)\neq 0$, these conditions coincide with the ones defining U^m and V^m.

Owing to Hermite's formula

$$\pi_m(x,y) = \frac{1}{2\mathrm{i}\pi}\int_{|z|=\rho_1}\left(1-\left(\frac{x}{z}\right)^{m+n+1}\right)\frac{\tilde{V}^m(z,y)h(z,y)}{z-x}\,\mathrm{d}z, \tag{35}$$

the coefficients $w_0^m(y),\dots,w_{n-1}^m(y)$ are solution to the system

$$\sum_{k=0}^{n-1} A_{jk}^m(y)w_k^m(y) = B_j^m(y), \quad j=1,2,\dots,n,$$

$$A_{jk}^m(y) = \frac{1}{2\mathrm{i}\pi}\int_{|z|=\rho_1}\left(1-\left(\frac{\alpha_j(y)}{z}\right)^{m+n+1}\right)\frac{W_k(z,y)h(z,y)}{z-\alpha_j(y)}\,\mathrm{d}z,$$

$$B_j^m(y) = \frac{1}{2\mathrm{i}\pi}\int_{|z|=\rho_1}\left(\frac{\alpha_j(y)}{z}\right)^{m+n+1}\frac{g(z,y)h(z,y)}{z-\alpha_j(y)}\,\mathrm{d}z,$$

which converges uniformly on $\mathscr{K}_y$ to a triangular, homogeneous and invertible system. Thus, for $m\geqslant m_0$ sufficiently large and $y\in\mathscr{K}_y$, the coefficients $w_0^m(y),\dots,w_{n-1}^m(y)$ are uniquely determined, holomorphic in y, they converge uniformly to zero, and $\tilde{V}^m$ converges uniformly to g on all compact subsets of $\mathbb{C}\times\mathscr{K}_y$. The proof is completed as in Section 2.5. □

3.3. The Padé∘Padé approximants

The Padé ∘ Padé approximants were introduced by Chaffy [3]. In Section 3.2 were defined the intermediate Padé approximant $[m,n]_{f_y}(x) = U^m(x,y)/V^m(x,y)$ of the function $f = u/v$, where U^m and V^m belong to $\mathbb{C}[[y]][x]$, and have respective degrees m and n in x. The basic idea is now to compute an $[m',n']$ Padé approximant of $[m,n]_{f_y}(x)$ with respect to the variable y.

Some restrictions are made which are more technical than really necessary. For instance, a disk $D(0,\rho_3)$ is substituted for $\mathscr{Y}$ because, in order to apply once more Lemma 3.1, one needs to be sure that the function $y \mapsto v(x_0, y)$ has exactly n' simple zeros in $D(0,\rho_3)$. The Padé ∘ Padé approximant is defined locally in x in the following way.

Let $\rho_3 > 0$ be such that $D(0,\rho_3) \subset \mathscr{Y}$ and $v(0,y) \neq 0$ for all $y \in D(0,\rho_3)$. Let $x_0 \in D(0,\rho_1)$ be fixed such that $v(x_0,0) \neq 0$, and suppose that the function $y \mapsto v(x_0,y)$ has exactly n' simple zeros in $D(0,\rho_3)$. The case of zeros of constant multiplicity could also be considered.

Recall that $V^m(x,y)$ converges to $v(x,y)$ uniformly on all compact subsets of $\mathbb{C} \times \mathscr{Y}$. It follows from the implicit functions theorem, Lemma 3.1 and Rouché's theorem, that there exists an integer m_0 and a neighborhood $\mathscr{V}(x_0)$ of x_0 such that for $m \geqslant m_0$ and $(x,y) \in \mathscr{V}(x_0) \times D(0,\rho_3)$, $v(x,y)$ and $V^m(x,y)$ can be written under the form

$$v(x,y) = c\prod_{i=1}^{n'}(y - \beta_i(x))w(x,y),$$

$$V^m(x,y) = c_m\prod_{i=1}^{n'}(y - \beta_i^m(x))W^m(x,y),$$

where the functions β_i and β_i^m are holomorphic and do not vanish on $\mathscr{V}(x_0)$, the functions w and W^m are holomorphic and do not vanish on $\mathscr{V}(x_0) \times \mathscr{Y}$. Moreover $\lim_{m\to\infty} \beta_i^m(x) = \beta_i(x)$ uniformly on all compact subsets of $\mathscr{V}(x_0)$.

Definition 3.1. For fixed $m \geqslant 0$, let $s(x,y) = [m,n]_{f_y}(x)$. The Padé ∘ Padé approximant $[m',n']_y \circ [m,n]_x(f)$, if it exists, is defined on $\mathscr{V}(x_0)$ by

$$[m',n']_y \circ [m,n]_x(f)(x,y) = [m',n']_{s_x}(y).$$

Remark 3.1. The rational fraction $r(x,y) = [m',n']_y \circ [m,n]_x(f)$ has the following interpolation property:

$$\partial_x^k r(0,0) = \partial_x^k f(0,0), \quad 0 \leqslant k \leqslant m+n,$$

$$\partial_y^k r(x,0) = \partial_y^k s(x,0), \quad 0 \leqslant k \leqslant m'+n',\ \forall x \in \mathscr{V}(x_0).$$

If $0 \in \mathscr{V}(x_0)$ (for example if $x_0 = 0$), then the Padé ∘ Padé satisfies at the origin the usual interpolation property. This follows from $s(0,y) = [m,n]_{f_y}(0) = f(0,y)$ for small $|y|$.

3.3.1. Convergence

We are now in the same position than the beginning of the first step if we exchange the variables x and y, and substitute $[m,n]_{f_y}(x)$ for the function f. The next theorem [3] follows from Lemma 3.1 applied first to the function $[m,n]_{f_y}(x)$, next to the function f_y.

Theorem 3.2 (Chaffy-Camus [3]). *There exists an integer m'_0 and a neighborhood $\mathscr{V}(x_0)$ of x_0 such that for all $m' \geqslant m'_0$ and all $x \in \mathscr{V}(x_0)$, there is a unique Padé approximant $[m', n']_y \circ [m, n]_x(f)$. The Padé $\circ$ Padé approximants converge in the following sense*:

$$\lim_{m\to\infty}\left(\lim_{m'\to\infty}[m', n']_y \circ [m, n]_x(f)(x, y)\right) = f(x, y)$$

uniformly on all compact subsets of $\{(x, y) \in \mathscr{V}(x_0) \times D(0, \rho_3),\ v(x, y) \neq 0\}$.

3.3.2. Computation

In order to compute $[m', n']_y \circ [m, n]_x(f)$, one first need to compute the double power series expansion of $[m, n]_{f_y}(x)$. Using symbolic computation, one expands fraction (31), $[m, n]_{f_y}(x) = (\sum_{i=0}^m s_i^m(y)x^i)/(1 + \sum_{i=1}^n s_{m+i}^m(y)x^i)$, in the form

$$[m, n]_{f_y}(x) = \sum_{j \geqslant 0} b_j(x)y^j, \tag{36}$$

where the b_j's are rational fractions, solution in $\mathbb{C}(x)$ to (37)–(38). Observe that $b_0(x)$ is a fraction with degrees at most $[m, n]$, and more generally $b_j(x)$ is a fraction with degrees at most $[m + jn, (j + 1)n]$. Next the Padé $\circ$ Padé approximant of f is obtained by computing the univariate Padé approximant of (36) with respect to y. Here again symbolic computation is used to solve the associated linear system in the field $\mathbb{C}(x)$, and the degrees in x will increase once more. The algorithm is the following.

Algorithm 4: PRPA computation

1. Use Algorithm 3 to compute the coefficients $s_{i,j}^m$, $0 \leqslant i \leqslant m + n$, $0 \leqslant j \leqslant m' + n'$ (in Algorithm 3, substitute $0 \leqslant j \leqslant m' + n'$ for $j \geqslant 0$).
2. Solve in $\mathbb{C}(x)$ the following triangular system (using symbolic computation):

$$\left(1 + \sum_{i=1}^{n} s_{m+i,0}^m x^i\right) b_0(x) = \sum_{i=0}^{m} s_{i,0}^m x^i, \tag{37}$$

$$\sum_{k=0}^{j}\left(\sum_{i=1}^{n} s_{m+i,j-k}^m x^i\right) b_k(x) = \sum_{i=0}^{m} s_{i,j}^m x^i, \quad j = 1, 2, \ldots, n' + m'. \tag{38}$$

3. Solve the system (using symbolic computation):

$$\begin{pmatrix} b_{m'-n'+1}(x) & \ldots & b_{m'}(x) \\ \vdots & & \vdots \\ b_{m'}(x) & \ldots & b_{m'+n'-1}(x) \end{pmatrix} \begin{pmatrix} q_{n'}(x) \\ \vdots \\ q_1(x) \end{pmatrix} = -\begin{pmatrix} b_{m'+1}(x) \\ \vdots \\ b_{m'+n'}(x) \end{pmatrix},$$

where $b_i(x) := 0$ if $i < 0$, and define

$$Q(x, y) = 1 + \sum_{i=0}^{n'} q_i(x)y^i.$$

4. Compute

$$p_i(x) = \sum_{j=0}^{n'} q_j(x) b_{i-j}(x), \quad 0 \leqslant i \leqslant m',$$

where $q_0(x) := 1$, and define

$$P(x, y) = \sum_{i=0}^{m'} p_i(x) y^i.$$

5. The fraction $P(x, y)/Q(x, y)$ is the Padé ∘ Padé of the function f.

3.4. The nested multivariate approximants

The nested multivariate approximants were introduced in [10]. Consider a fraction $R \in \mathbb{C}(y)(x)$ of the form

$$R(x, y) = \frac{P(x, y)}{Q(x, y)} = \frac{\sum_{i=0}^{m} r_i(y) x^i}{1 + \sum_{i=1}^{n} r_{m+i}(y) x^i}, \tag{39}$$

where the $r_i(y)$ are also fractions:

$$r_i(y) = \frac{p_i(y)}{q_i(y)} = \frac{\sum_{j=0}^{m_i} p_{ij} y^j}{1 + \sum_{j=1}^{n_i} q_{ij} y^j}, \quad 0 \leqslant i \leqslant n + m \tag{40}$$

with

$$m + n = d_1, \qquad m_i + n_i = d_2, \quad 0 \leqslant i \leqslant m + n, \tag{41}$$

$$\deg Q(x, 0) = n, \qquad \deg q_i = n_i, \quad 0 \leqslant i \leqslant n + m. \tag{42}$$

Let $E(d_1, d_2) = \{0, 1, \ldots, d_1\} \times \{0, 1, \ldots, d_2\}$. For $\alpha = (\alpha_1, \alpha_2) \in \mathbb{N}^2$, we denote by ∂^α the usual differential operator $\partial^\alpha = \partial^{|\alpha|} / \partial_x^{\alpha_1} \partial_y^{\alpha_2}$.

Definition 3.2. Consider the following equation:

$$\partial^\alpha R(0, 0) = \partial^\alpha f(0, 0) \quad \forall \alpha \in E(d_1, d_2). \tag{43}$$

If the fraction R (39) is the unique solution to this equation, it is called the $[m, n, (m_i), (n_i)]$ nested Padé approximant of the function f and denoted by $[m, n, (m_i), (n_i), x, y]_f$.

Similarly to the univariate case, existence occurs as soon as the degrees of the numerators are sufficiently large (cf. Theorem 3.4). A sufficient condition for uniqueness, which implies consistency, is the following [11]. The fraction R (39) is said irreducible if the fractions $x \mapsto R(x, 0)$ and r_i, $0 \leqslant i \leqslant n + m$ are irreducible.

Proposition 3.3. *If the fraction R (39) is a solution to (43) and is irreducible, then it is the unique solution to (43).*

3.4.1. Computation

Eq. (43) is a nonlinear system of $(d_1+1)(d_2+1)$ equations, with the same number of unknowns. However the solution of this system can be obtained in two steps by solving small linear systems. The first step has been described in Section 3.2 and Algorithm 3, where the coefficients $s_i^m(y)$ of the intermediate Padé were obtained. The second step is the following.

For a given m and $0\leqslant i\leqslant d_1$, degrees m_i and n_i are chosen in such a way that $m_i+n_i=d_2$ (see, e.g., [9] for the choice of the degrees). We suppose here that the following Padé approximants:

$$r_i(y)=[m_i,n_i]_{s_i^m},\quad 0\leqslant i\leqslant d_1,$$

exist in the sense of definition (29), that their denominators are of degree n_i, and that $r_{m+n}(0)\neq 0$. Let

$$R(x,y)=\frac{\sum_{i=0}^{m} r_i(y)x^i}{1+\sum_{i=1}^{n} r_{m+i}(y)x^i}. \tag{44}$$

It can easily be proved that if this fraction R is irreducible, then it is the nested Padé approximant of order $[m,n,(m_i),(n_i)]$ of the function f [10]. The algorithm is the following, which needs no symbolic computation.

Algorithm 5: NPA computation

1. Use Algorithm 3 to compute the coefficients $s_{i,j}^m$, $0\leqslant i\leqslant d_1$, $0\leqslant j\leqslant d_2$ (in Algorithm 3, substitute $0\leqslant j\leqslant d_2$ for $j\geqslant 0$).
2. For $i=1$ to d_1:
   - solve
   $$\begin{pmatrix} s_{i,m_i-n_i+1}^m & \cdots & s_{i,m_i}^m \\ \vdots & & \vdots \\ s_{i,m_i}^m & \cdots & s_{i,m_i+n_i-1}^m \end{pmatrix}\begin{pmatrix} q_{n_i} \\ \vdots \\ q_1\end{pmatrix} = -\begin{pmatrix} s_{i,m_i+1}^m \\ \vdots \\ s_{i,m_i+n_i}^m\end{pmatrix},$$
   where $s_{i,j}:=0$ if $j<0$, and compute
   $$p_j=\sum_{k=0}^{n_i} q_k s_{i,j-k},\quad 0\leqslant j\leqslant m_i,$$
   - define
   $$p_i(y)=\sum_{j=0}^{m_i} p_j\, y^j,\qquad q_i(y)=1+\sum_{j=1}^{n_i} q_j\, y^j,\qquad r_i(y)=\frac{p_i(y)}{q_i(y)}.$$
3. Define
$$P(x,y)=\sum_{i=0}^{m} r_i(y)x^i,$$
$$Q(x,y)=1+\sum_{i=1}^{n} r_{m+i}(y)x^i.$$
4. The fraction $P(x,y)/Q(x,y)$ is the nested Padé approximant of the function f.

3.4.2. Convergence

The next theorem was proved in [11].

Theorem 3.4 (Guillaume [11]). *The sequence of nested Padé approximants converges uniformly to f on all compact subsets of $\mathcal{O} = \{(x,y) \in D(0,\rho_1,\rho_2), y \in \mathcal{Y}, v(x,y) \neq 0\}$ in the following sense: for all $\varepsilon > 0$ and all compact subsets $\mathcal{K} \subset \mathcal{O}$, there is an integer m_0 such that for all $m \geqslant m_0$, there exist integers d_0 and $n_i^m \leqslant d_0$, $0 \leqslant i \leqslant m+n$, such that for all $d_2 \geqslant d_0$, the nested Padé approximant $[m,n,(d_2 - n_i^m),(n_i^m),x,y]_f$ of the function f is well defined and*

$$\sup_{(x,y) \in \mathcal{K}} |f(x,y) - [m,n,(d_2 - n_i^m),(n_i^m),x,y]_f(x,y)| < \varepsilon.$$

Each n_i^m can be chosen equal to the number of poles (counted with multiplicity) within the disk $D(0,\rho_2)$ of the function s_i^m, $0 \leqslant i \leqslant m+n$.

Proof. The main line is the following. It follows from Lemma 3.1 that for $m \geqslant m_0$ sufficiently large, V^m is well defined and holomorphic around the origin. Hence the functions $s_i^m(y)$ are holomorphic around zero. Due to their construction (32), they are meromorphic on $D(0,\rho_2)$ and have a finite number n_i^m of poles (counted with multiplicity) within this disk. Owing to the Montessus de Ballore theorem, there is an integer d_0 such that the Padé approximants $[d_2 - n_i^m, n_i^m]_{s_i^m}$ are well defined for $d_2 \geqslant d_0$, and each sequence $([d_2 - n_i^m, n_i^m]_{s_i^m})_{d_2}$ converges to s_i^m uniformly on $\mathcal{K}$ when $d_2 \to \infty$. □

Remark 3.2. Although the number of poles n_i^m of the functions s_i^m are not known, the technique described in [9] for counting the number of poles of meromorphic functions within a disk can be used here. Besides, the existence of an upper bound of the numbers n_i^m remains an open question, although numerical tests indicate they are bounded.

4. Final comments

An open question is whether convergence of more or less diagonal sequences in the LSPA table can be obtained for the approximation of general meromorphic functions. If such a result was obtained, it could be an improvement over the PRPA or NPA, where, due do their recursive construction, artificial singularities are present.

References

[1] M. Bôcher, Introduction to Higher Algebra, The Macmillan Company, New York, 1952.
[2] C. Brezinski, Least-squares orthogonal polynomials, J. Comput. Appl. Math. 46 (1993) 229–239.
[3] C. Chaffy-Camus, Convergence uniforme d'une nouvelle classe d'approximants de Padé à plusieurs variables, C. R. Acad. Sci. Paris, Sér. I 306 (1988) 387–392.
[4] J.S.R. Chisholm, Rational approximants defined from double power series, Math. Comp. 27 (1973) 841–848.
[5] A. Cuyt, Multivariate Padé approximants, J. Math. Anal. Appl. 96 (1983) 283–293.
[6] A. Cuyt, A Montessus de Ballore Theorem for Multivariate Padé Approximants, J. Approx. Theory 43 (1985) 43–52.
[7] A. Cuyt, D.S. Lubinsky, A de Montessus theorem for multivariate homogeneous Padé approximants, Ann. Numer. Math. 4 (1997) 217–228.

[8] A. Cuyt, L. Wuytack, in: Nonlinear Methods in Numerical Analysis, North-Holland Mathematics Studies, Vol. 136, Studies in Computational Mathematics, Vol. 1, North-Holland Publishing Co, Amsterdam, 1987.
[9] B. Gleyse, V. Kaliaguine, On algebraic computation of number of poles of meromorphic functions in the unit disk, in: A. Cuyt (Ed.), Nonlinear Numerical Methods and Rational Approximation II, Kluwer Academic Publishers, Netherlands, 1994, pp. 241–246.
[10] Ph. Guillaume, Nested Multivariate Padé Approximants, J. Comput. Appl. Math. 82 (1997) 149–158.
[11] Ph. Guillaume, Convergence of the nested multivariate Padé approximants, J. Approx. Theory 94 (3) (1998) 455–466.
[12] Ph. Guillaume, A. Huard, V. Robin, Generalized Multivariate Padé Approximants, J. Approx. Theory 95 (2) (1998) 203–214.
[13] Ph. Guillaume, M. Masmoudi, Solution to the time-harmonic Maxwell's equations in a waveguide, use of higher order derivatives for solving the discrete problem, SIAM J. Numer. Anal. 34 (4) (1997) 1306–1330.
[14] J. Karlsson, H. Wallin, Rational approximation by an interpolation procedure in several variables, in: E. Saaf, R. Varga (Eds.), Padé and Rational Approximation: Theory and Applications, Academic Press, New York, 1977, pp. 83–100.
[15] S. Kida, Padé-type and Padé approximants in several variables, Appl. Numer. Math. 6 (1989/90) 371–391.
[16] D. Levin, General order Padé-type rational approximants defined from double power series, J. Inst. Math. Appl. 18 (1976) 395–407.
[17] E.B. Saff, An extension of Montessus de Ballore's theorem on the convergence of interpolating rational functions, J. Approx. Theory 6 (1972) 63–67.
[18] B.L. Van der Waerden, Einführung in die algebraische Geometrie, Springer, Berlin, 1973.
[19] H. Werner, Multivariate Padé Approximation, Numer. Math. 48 (1986) 429–440.