On the robustness of exponential base terms and the Padé denominator in some least squares sense

Ferre Knaepkens[1*] and Annie Cuyt[1]

[1]Department of Computer Science, Universiteit Antwerpen, Middelheimlaan 1, Antwerpen, Belgium

To the 80-th birthday of Claude Brezinski, our very respected colleague in Padé approximation research.

*Corresponding author(s). E-mail(s): ferre.knaepkens@uantwerpen.be; Contributing authors: annie.cuyt@uantwerpen.be;

Abstract

The exponential analysis of $2n$ uniformly collected samples from an n-term exponential sum, is equivalent to the reconstruction of a rational function of degree $n - 1$ over n. The latter is by computing the Padé approximant of the z-transform of the sequence of samples.

In practice, the samples are often noisy and $2n$ is replaced by $N > 2\nu$ with $\nu > n$, leading to a least squares computation of the Padé approximant of degree $\nu - 1$ over ν.

We show that the latter is a perturbed version of the one of degree $n - 1$ over n and that the n exponential base terms can still be retrieved reliably. This has remained an open problem for many years, despite the fact that the least squares computation was used in most applications.

Keywords: Exponential analysis, Padé approximation, perturbation analysis, conditioning, robustness

1 Introduction

Padé approximation is a rational approximation method that, among other things, can reconstruct a rational function $F(z)$ of degree m in the numerator and degree n in the denominator from its first $m + n + 1$ Taylor series coefficients. In practical situations, where the Taylor series coefficients stem from measurements or earlier computations, one needs to analyze and understand the effect of such a perturbation on the true mathematical coefficients.

Let us introduce the notation $[\mu/\nu]_F$ for the Padé approximant of degree μ over ν to a given function $F(z)$. In case $m = n - 1$, it is proved in [1, 2] that a random perturbation of $\mathcal{O}(\epsilon)$ with $\epsilon \ll 1$ to the first $\mu + \nu + 1$ Taylor series coefficients with $\mu - m = \nu - n = k > 0$, disturbs the Padé approximants of a rational function $F(z)$ of degree m over n in the following way. The procedure adds k zeroes and poles to the numerator and denominator polynomials of $F(z)$, with the additional factors in the numerator of $[\mu/\nu]_F$ almost cancelling the additional factors in its denominator. This near-cancelling instead of the exact cancelling of the additional factors is caused by the presence of some $\mathcal{O}(\epsilon)$ quantities in numerator and denominator of $[\mu/\nu]_F$. So the process behaves quite as expected when approximating such a rational function $F(z)$ by a Padé approximant $[\mu/\nu]_F$ of too high degree. Another robustness result is formulated in [3] for a variant of Padé approximation, computed from the same first $\mu + \nu + 1$ Taylor coefficients, but with the aim to filter out these disturbances.

A different approach when dealing with perturbations is to switch to the computation of the Padé approximant in some least squares sense. In [4] a total-least-squares style approach to the problem statement is probably presented for the first time. In [5] the Padé denominator is computed by solving the linear system of equations defining its coefficients in the least squares sense, while the numerator coefficients are still computed from the denominator coefficients in the traditional way. In exponential analysis though, it is standard to compute both numerator and denominator in a least squares sense, as is done here. In Section 2 we summarize what the reader needs to know about Padé approximants and their connection to the exponential analysis problem statement. In Section 3 we further analyze in detail, for the first time, the effect of an $\mathcal{O}(\epsilon)$ perturbation on the full least squares problem statement. In Section 4 we examine how sensitive the rooting of the denominator polynomial is in the considered least squares setting and which parameters influence this sensitivity.

2 Exponential analysis and Padé approximation

2.1 Exponential analysis

Let the signal $f(x)$ be given by

$$f(x) = \sum_{j=1}^{n} \alpha_j \exp(\phi_j x),$$

where the coefficients $\alpha_j \in \mathbb{C}$ and exponents $\phi_j \in \mathbb{C}$ are unknown parameters. Note that in general the number of exponential terms n, called the sparsity, is also unknown. The objective is to obtain the values of all these unknown parameters from a limited number of equidistant samples

$$f_k = f(k\Delta) = \sum_{j=1}^{n} \alpha_j \exp(\phi_j k\Delta) = \sum_{j=1}^{n} \alpha_j \Phi_j^k, \qquad k = 0, 1, \ldots \qquad (1)$$

with sampling step $\Delta \neq 0$. Here we introduce the notation $\Phi_j = \exp(\phi_j \Delta)$, $j = 1, \ldots, n$, which are called the base terms. Hence, the exponential analysis problem consists in finding the coefficients α_j and base terms Φ_j such that the non-linear interpolation conditions (1) are satisfied. Under the assumption that $|\Im(\phi_j)| < \pi/|\Delta|$, one then easily retrieves the exponent values ϕ_j from the base terms Φ_j.

Already in 1795, it was shown by de Prony that only $2n$ equidistant samples are required to solve the exponential analysis problem if the sparsity n is known. In a more modern version the interpolation problem is expressed as a generalized eigenvalue problem. Let the samples f_k be the elements of the square Hankel matrices

$$H_n^{(s)} = \begin{pmatrix} f_s & f_{s+1} & \cdots & f_{s+n-1} \\ f_{s+1} & f_{s+2} & \cdots & f_{s+n} \\ \vdots & \vdots & & \vdots \\ f_{s+n-1} & f_{s+n} & \cdots & f_{s+2n-2} \end{pmatrix}, \qquad s \geq 0.$$

These matrices factorise as [6]

$$H_n^{(s)} = V_n D_\alpha D_\Phi^s V_n^T,$$

where the matrices D_α and D_Φ are respectively the diagonal matrices $\text{diag}(\alpha_1, \ldots, \alpha_n)$ and $\text{diag}(\Phi_1, \ldots, \Phi_n)$, and where V_n is the Vandermonde

matrix

$$
\begin{pmatrix}
1 & \cdots & 1 \\
\Phi_1 & \cdots & \Phi_n \\
\vdots & & \vdots \\
\Phi_1^{n-1} & \cdots & \Phi_n^{n-1}
\end{pmatrix},
$$

with the base terms Φ_j as the generators. From this factorisation we obtain

$$
H_n^{(1)} - \lambda H_n^{(0)} = V_n D_\alpha \left(D_\Phi - \lambda I_n \right) V_n^T,
$$

which means that the base terms Φ_j are obtained as the generalized eigenvalues λ_j of the generalized eigenvalue problem

$$
H_n^{(1)} v_j = \lambda_j H_n^{(0)} v_j. \tag{2}
$$

After the base terms Φ_j are recovered, the coefficients α_j are computed from the linear interpolation conditions (1), which translate to the Vandermonde structured linear system

$$
\begin{pmatrix}
1 & \cdots & 1 \\
\Phi_1 & \cdots & \Phi_n \\
\vdots & & \vdots \\
\Phi_1^{2n-1} & \cdots & \Phi_n^{2n-1}
\end{pmatrix}
\begin{pmatrix}
\alpha_1 \\
\alpha_2 \\
\vdots \\
\alpha_n
\end{pmatrix}
=
\begin{pmatrix}
f_0 \\
f_1 \\
\vdots \\
f_{2n-1}
\end{pmatrix}.
$$

In an exact noisefree context the first n are linearly independent, thus the remaining n equations are not required. In a noisy setting, it is advised to solve the linear system in a least-squares sense for all available interpolation conditions.

2.2 Padé approximation

A Padé approximant is a rational function of which the Taylor series expansion coincides with a given formal power series up to a specified order. More formally, given

$$
F(z) = \sum_{k=0}^{\infty} c_k z^k,
$$

with $c_0 \neq 0$, the Padé approximation problem consists in finding the polynomials

$$
p(z) = \sum_{k=0}^{m} a_j z^k, \qquad q(z) = \sum_{k=0}^{n} b_j z^k
$$

such that

$$F(z) - \frac{p(z)}{q(z)} = \mathcal{O}\left(z^{m+n+1}\right). \tag{3}$$

More commonly, the linearised version

$$F(z)q(z) - p(z) = \mathcal{O}\left(z^{m+n+1}\right)$$

is used instead. Since the approximating function $p(z)/q(z)$ is a rational function, there is a degree of freedom in the representation of $p(z)$ and $q(z)$. In general, the Padé approximant is defined as the irreducible rational function $p(z)/q(z)$ which satisfies (3) and for which $q(0) = 1$. This rational approximant is then denoted as $[m/n]_F$.

One can represent this Padé approximant using determinant formulas. Let $F_\ell(z)$ denote the partial sum

$$F_\ell(z) = \sum_{k=0}^{\ell} c_k z^k,$$

with the convention that $F_\ell(z) = 0$ if $\ell < 0$. Then the Padé approximant is the irreducible form of the rational function $p(z)/q(z)$ where

$$q(z) = \frac{1}{D}\begin{vmatrix} 1 & z & \dots & z^n \\ c_{m+1} & & & \\ \vdots & & D_{m,n} & \\ c_{m+n} & & & \end{vmatrix}, p(z) = \frac{1}{D}\begin{vmatrix} F_m(z) & zF_{m-1}(z) & \dots & z^n F_{m-n}(z) \\ c_{m+1} & & & \\ \vdots & & D_{m,n} & \\ c_{m+n} & & & \end{vmatrix}, \tag{4}$$

under the condition that

$$D = |D_{m,n}| = \begin{vmatrix} c_m & c_{m-1} & \dots & c_{m-n+1} \\ c_{m+1} & c_m & \dots & c_{m-n+1} \\ \vdots & \vdots & & \vdots \\ c_{m+n-1} & c_{m+n} & \dots & c_m \end{vmatrix} \neq 0.$$

2.3 Connection

At first sight it seems that exponential analysis and Padé approximation are completely different subjects. They are, however, most certainly connected [7–9]. Instead of arranging the samples f_k in the Hankel matrices $H_n^{(s)}$, one can use them as coefficients in the formal power series

$$F(z) = \sum_{k=0}^{\infty} f_k z^k.$$

When plugging in expression (1) for the samples f_k , we obtain

$$F(z) = \sum_{k=0}^{\infty} f_k z^k = \sum_{k=0}^{\infty} \left(\sum_{j=1}^{n} \alpha_j \Phi_j^k \right) z^k = \sum_{j=1}^{n} \alpha_j \sum_{k=0}^{\infty} (\Phi_j z)^k$$

$$= \sum_{j=1}^{n} \frac{\alpha_j}{1 - \Phi_j z} = \frac{\pi_{n-1}(z)}{\prod_{j=1}^{n} (1 - \Phi_j z)}. \tag{5}$$

Hence, the power series $F(z)$ stems from a rational function of degree $n-1$ in the numerator and degree n in the denominator. Due to the consistency property of Padé approximants, we find that the Padé approximant $[n-1/n]_F = F(z)$ because it reconstructs rational functions of degree $n-1$ over n. In addition, the base terms Φ_j are also obtained as the reciprocals of the poles of the Padé approximant $[n-1/n]_F$. Subsequently, the coefficients α_j are obtained from the partial fraction decomposition of the same Padé approximant.

  The generalized eigenvalue problem formulation (2) of exponential analysis and the determinant formulas (4) of the Padé approximant are also closely connected. Since the poles of the Padé approximant $[n-1/n]_F$ equal the reciprocals of the base terms Φ_j, we consider the reverse of the generalized eigenvalue problem (2),

$$H_n^{(0)} v_j = \frac{1}{\lambda_j} H_n^{(1)} v_j.$$

Solving this reverse eigenvalue problem is equivalent to finding the zeroes of the characteristic polynomial which is given by $|H_n^{(0)} - z H_n^{(1)}|$, which equals

$$\begin{vmatrix} f_0 - z f_1 & f_1 - z f_2 & \cdots & f_{n-1} - z f_n \\ f_1 - z f_2 & f_2 - z f_3 & \cdots & f_n - z f_{n+1} \\ \vdots & \vdots & & \vdots \\ f_{n-1} - z f_n & f_n - z f_{n+1} & \cdots & f_{2n-2} - z f_{2n-1} \end{vmatrix} \tag{6}$$

Note that the polynomial given in (6) is the reverse of the so-called Prony polynomial [10, p. 458].

  When splitting the columns, performing the proper column subtractions, taking out factors z and flipping all the columns from left to right, we obtain

$$\pm \begin{vmatrix} f_{n-1} & \cdots & f_1 & f_0 \\ f_n & \cdots & f_2 & f_1 \\ \vdots & & \vdots & \vdots \\ f_{2n-2} & \cdots & f_n & f_{n-1} \end{vmatrix} \pm (-1)z \begin{vmatrix} f_n & f_{n-2} & \cdots & f_0 \\ f_{n+1} & f_{n-1} & \cdots & f_1 \\ \vdots & & \vdots & \vdots \\ f_{2n-2} & f_{2n-4} & \cdots & f_{n-1} \end{vmatrix} \pm \ldots$$

$$\cdots \pm (-1)^n z^n \begin{vmatrix} f_n & \cdots & f_2 & f_1 \\ f_{n+1} & \cdots & f_3 & f_2 \\ \vdots & & \vdots & \vdots \\ f_{2n-1} & \cdots & f_{n+1} & f_n \end{vmatrix}.$$

This expression is also obtained when expanding, along the first row, the determinant

$$\pm \begin{vmatrix} 1 & z & \cdots & z^n \\ f_n & f_{n-1} & \cdots & f_0 \\ \vdots & \vdots & & \vdots \\ f_{2n-1} & f_{2n-2} & \cdots & f_{n-1} \end{vmatrix},$$

which equals $\pm D q(z)$ for $m = n - 1$ in (4). Hence, computing the zeroes of the Padé denominator and solving the generalized eigenvalue problem are mathematically equivalent.

3 Least squares computation

The above manipulation of the determinant expressions proves to be very useful in studying the effect of noise on the computation of the base terms Φ_j in the exponential analysis problem. We start from what is known about the effect of noise in the Padé approximation of rational functions such as (5). In [2, p. 292] it is shown that for the perturbed power series

$$\widetilde{F}(z) = \sum_{k=0}^{\infty} (f_k + \epsilon r_k) z^k,$$

where the r_k are random noise terms following some distribution law and $\epsilon \ll 1$, the Padé approximant $[\nu - 1/\nu]_{\widetilde{F}}$ for $\nu \geq n$, is expressed as

$$[\nu - 1/\nu]_{\widetilde{F}} = \frac{p(z)}{q(z)} = \frac{\pi_{n-1}(z)K_{\nu-n}(z) + \mathcal{O}(\epsilon P_{\nu-1}(z))}{\prod_{j=1}^{n}(1 - \Phi_j z)K_{\nu-n}(z) + \mathcal{O}(\epsilon P_\nu(z))}, \qquad q(0) = 1, \quad (7)$$

where $K_{\nu-n}(z)$ is a polynomial of degree $\nu - n$ of which the coefficients depend on the random noise terms r_k and where the notation $\mathcal{O}(\epsilon P_m(z))$ represents a finite sum of the form

$$\epsilon p_{m,1}(z) + \epsilon^2 p_{m,2}(z) + \epsilon^3 p_{m,3}(z) + \ldots + \epsilon^s p_{m,s}(z), \qquad \deg(p_{m,k}(z)) \leq m.$$

Because of the connection between Padé approximation and exponential analysis, the denominator of (7) allows to study the computation of the base terms Φ_j from noisy data. Formula (7) shows that the coefficients of the denominator

polynomial lie at a distance of order ϵ from those of a polynomial of which n zeroes are the reciprocals of the base terms Φ_j and the remaining $\nu - n$ zeroes are dictated by the noise terms.

So the characteristic polynomial of the generalized eigenvalue problem (2) for $\nu \geq n$ terms lies at a distance of order ϵ from the polynomial with the reciprocals of the base terms Φ_j as n of its zeroes and the remaining $\nu - n$ roots dictated by the noise. Now, the question remains if the same holds true when an overdetermined exponential analysis problem is considered. By overdetermined we mean an exponential analysis problem of sparsity n for which more than the minimally required $2n$ samples are collected and used for the computation of the unknown parameters.

Let $f_0, f_1, \ldots, f_{N-1}$, for $N \geq 2n$ be the number of collected samples for the exponential analysis problem. Let us arrange these samples in the rectangular Hankel matrices

$$H_{m,n}^{(s)} = \begin{pmatrix} f_s & f_{s+1} & \cdots & f_{s+n-1} \\ f_{s+1} & f_{s+2} & \cdots & f_{s+n} \\ \vdots & \vdots & & \vdots \\ f_{s+m-1} & f_{s+} & \cdots & f_{s+m+n-2} \end{pmatrix},$$

$$m \geq n, \quad 0 \leq s \leq N - m - n + 1.$$

Similar to their square counterparts, these Hankel matrices factorise as

$$H_{m,n}^{(s)} = V_m D_\alpha D_\Phi^s V_n^T,$$

where the subscript of the Vandermonde matrices denotes the number of rows and where the number of columns equals n, i.e. $V_\ell = \left(\Phi_j^{i-1} \right)_{i=1, j=1}^{\ell, n}$. Consequently, the eigenvalues of the overdetermined generalized eigenvalue problem

$$H_{m,n}^{(1)} v_j = \lambda_j H_{m,n}^{(0)} v_j \tag{8}$$

are once again the base terms Φ_j. The most natural way to interpret such an overdetermined eigenvalue problem is based on the pseudo-inverse, which is used to transform the rectangular problem (8) to a square one. Assuming that $H_{m,n}^{(0)}$ has rank n, we let

$$\left(H_{m,n}^{(0)} \right)^\dagger = \left(\left(H_{m,n}^{(0)} \right)^* H_{m,n}^{(0)} \right)^{-1} \left(H_{m,n}^{(0)} \right)^*$$

denotes the pseudo-inverse of the matrix $H_{m,n}^{(0)}$ and the superscript $*$ the conjugate transpose of the same matrix. Then the rectangular generalized eigenvalue

problem can be reformulated as the square generalized eigenvalue problem

$$\left(H_{m,n}^{(0)}\right)^* H_{m,n}^{(1)} v_j = \lambda_j \left(H_{m,n}^{(0)}\right)^* H_{m,n}^{(0)} v_j \tag{9}$$

or its reverse

$$\left(H_{m,n}^{(0)}\right)^* H_{m,n}^{(0)} v_j = \frac{1}{\lambda_j} \left(H_{m,n}^{(0)}\right)^* H_{m,n}^{(1)} v_j. \tag{10}$$

When the number of terms is overestimated by $\nu \geq n$, we assume that the pseudo-inverse of the matrix $H_{m,\nu}^{(0)}$ still exists due to the randomness of the noise. The objective is now to show that the characteristic polynomial corresponding to the generalized eigenvalue problem (10) with n overestimated by ν, so the problem

$$\left(H_{m,\nu}^{(0)}\right)^* H_{m,\nu}^{(0)} v_j = \frac{1}{\lambda_j} \left(H_{m,\nu}^{(0)}\right)^* H_{m,\nu}^{(1)} v_j. \tag{11}$$

is also of the form

$$\prod_{j=1}^{n}(1 - \Phi_j z)K_{\nu-n}(z) + \mathcal{O}(\epsilon P_\nu(z)),$$

where again $K_{\nu-n}(z)$ depends on the noise terms added to the samples f_k. We first take a detailed look at the case of only one base term, i.e. $n = 1$ and afterwards discuss how to tackle the general case.

3.1 Case of one base term

When $n = 1$, the samples equal

$$f_k = \alpha \Phi^k + \epsilon r_k, \qquad k = 0, \ldots, N - 1$$

and the $m \times \nu$ Hankel matrices contain the elements

$$\left(H_{m,\nu}^{(0)}\right)_{i,j} = \alpha \Phi^{i+j-2} + \epsilon r_{i+j-2}, \qquad \left(H_{m,\nu}^{(1)}\right)_{i,j} = \alpha \Phi^{i+j-1} + \epsilon r_{i+j-1},$$

where $\nu \geq n$ and $m = N - \nu \geq \nu$. Then the elements in the matrix on the left-hand-side of the generalized eigenvalue problem (11) are given by

$$\left(\left(H_{m,\nu}^{(0)}\right)^* H_{m,\nu}^{(0)}\right)_{i,j} = \sum_{k=1}^{m} \left(\overline{H}_{m,\nu}^{(0)}\right)_{k,i} \left(H_{m,\nu}^{(0)}\right)_{k,j}$$

$$= \sum_{k=1}^{m} \left(\overline{\alpha}\overline{\Phi}^{i+k-2} + \epsilon \overline{r}_{i+k-2}\right) \left(\alpha \Phi^{j+k-2} + \epsilon r_{j+k-2}\right)$$

$$= A\overline{\Phi}^{i-1}\Phi^{j-1} + \epsilon\left(\overline{B}_j\overline{\Phi}^{i-1} + B_i\Phi^{j-1}\right) + \epsilon^2 C_{i,j},$$

where $\overline{x}$ denotes the complex conjugate of x and

$$A = |\alpha|^2 \sum_{k=1}^m |\Phi|^{2k-2}, \quad B_j = \alpha \sum_{k=1}^m \Phi^{k-1}\overline{r}_{j+k-2}, \quad C_{i,j} = \sum_{k=1}^m \overline{r}_{i+k-2}r_{j+k-2}.$$

Analogously, we find

$$\left(\left(H_{m,\nu}^{(0)}\right)^* H_{m,\nu}^{(1)}\right)_{i,j} = A\overline{\Phi}^{i-1}\Phi^j + \epsilon\left(\overline{B}_{j+1}\overline{\Phi}^{i-1} + B_i\Phi^j\right) + \epsilon^2 C_{i,j+1}.$$

By defining

$$b_j(z) = \overline{B}_j - z\overline{B}_{j+1}, \qquad c_{i,j}(z) = C_{i,j} - zC_{i,j+1},$$

we find that

$$\left(\left(H_{m,\nu}^{(0)}\right)^* H_{m,\nu}^{(0)} - z\left(H_{m,\nu}^{(0)}\right)^* H_{m,\nu}^{(1)}\right)_{i,j}$$
$$= A\overline{\Phi}^{i-1}\Phi^{j-1}(1 - z\Phi) + \epsilon b_j(z)\overline{\Phi}^{i-1} + \epsilon B_i\Phi^{j-1}(1 - z\Phi) + \epsilon^2 c_{i,j}(z).$$

Furthermore, let

$$g_j(z) = b_{j+1}(z) - \Phi b_j(z), \qquad h_{i,j}(z) = c_{i,j+1}(z) - \Phi c_{i,j}(z).$$

By subtracting the j-th column multiplied by Φ from the $(j+1)$-th column, for $j = 1, \ldots, \nu - 1$, and subsequently splitting the first column, we obtain for the characteristic polynomial of (11) the expression

$$\epsilon^{\nu-1}(1 - z\Phi)\begin{vmatrix} A+\epsilon B_1 & g_1(z)+\epsilon h_{1,1}(z) & \cdots & g_{\nu-1}(z)+\epsilon h_{1,\nu-1}(z) \\ \vdots & \vdots & & \vdots \\ A\overline{\Phi}^{\nu-1}+\epsilon B_1\overline{\Phi}^{\nu-1} & g_1(z)+\epsilon h_{\nu,1}(z) & \cdots & \overline{\Phi}^{\nu-1}g_{\nu-1}(z)+\epsilon h_{\nu,\nu-1}(z) \end{vmatrix} \qquad (12)$$

$$+ \epsilon^\nu\begin{vmatrix} b_1(z)+\epsilon c_{1,1}(z) & g_1(z)+\epsilon h_{1,1}(z) & \cdots & g_{\nu-1}(z)+\epsilon h_{1,\nu-1}(z) \\ \vdots & \vdots & & \vdots \\ \overline{\Phi}^{\nu-1}b_1(z)+\epsilon c_{\nu,1}(z) & \overline{\Phi}^{\nu-1}g_1(z)+\epsilon h_{\nu,1}(z) & \cdots & \overline{\Phi}^{\nu-1}g_{\nu-1}(z)+\epsilon h_{\nu,\nu-1}(z) \end{vmatrix}. \qquad (13)$$

In the following, the determinant expressions in (12) and (13) are handled separately. For (12) we subtract the i-th row multiplied by $\overline{\Phi}$ from the $(i+1)$-th row and then split the first row to rewrite (12) as

$$\epsilon^{\nu-1} \underbrace{\begin{vmatrix} A & g_1(z) & \dots & g_{\nu-1}(z) \\ B_1(1-\overline{\Phi}) & s_{1,1}(z) & \dots & s_{1,\nu-1}(z) \\ \vdots & \vdots & & \vdots \\ B_1(1-\overline{\Phi}) & s_{\nu-1,1}(z) & \dots & s_{\nu-1,\nu-1}(z) \end{vmatrix}}_{(I)}$$

$$+ \epsilon^{\nu} \underbrace{\begin{vmatrix} B_1 & h_{1,1}(z) & \dots & h_{1,\nu-1}(z) \\ B_1(1-\overline{\Phi}) & s_{1,1}(z) & \dots & s_{1,\nu-1}(z) \\ \vdots & \vdots & & \vdots \\ B_1(1-\overline{\Phi}) & s_{\nu-1,1}(z) & \dots & s_{\nu-1,\nu-1}(z) \end{vmatrix}}_{(II)},$$

where we denote

$$s_{i,j}(z) = h_{i+1,j}(z) - \overline{\Phi}h_{i,j}(z).$$

The second part (II) of this expression is simplified further by adding row 1 multiplied by $\overline{\Phi}$ to the second row, then adding row 2 multiplied by $\overline{\Phi}$ to the third row and so on. This process returns

$$(II) = \begin{vmatrix} B_1 & h_{1,1}(z) & \dots & h_{1,\nu-1}(z) \\ B_1 & h_{2,1}(z) & \dots & h_{2,\nu-1}(z) \\ \vdots & \vdots & & \vdots \\ B_1 & h_{\nu,1}(z) & \dots & h_{\nu,\nu-1}(z) \end{vmatrix}.$$

Since $g_j(z)$, $h_{i,j}(z)$ and $s_{i,j}(z)$ are all polynomials of degree 1, both determinants (I) and (II) are polynomials of degree $\nu-1$. We denote these polynomials respectively by $K_{\nu-1}(z)$ and $p_{\nu-1}^{(1)}(z)$.

For expression (13) the same steps are repeated to obtain

$$\epsilon^{\nu-1} \underbrace{\begin{vmatrix} b_1(z) & g_1(z) & \dots & g_{\nu-1}(z) \\ t_1(z) & s_{1,1}(z) & \dots & s_{1,\nu-1}(z) \\ \vdots & \vdots & & \vdots \\ t_{\nu-1}(z) & s_{\nu-1,1}(z) & \dots & s_{\nu-1,\nu-1}(z) \end{vmatrix}}_{(III)} + \epsilon^{\nu} \underbrace{\begin{vmatrix} c_{1,1}(z) & h_{1,1}(z) & \dots & h_{1,\nu-1}(z) \\ c_{2,1}(z) & h_{2,1}(z) & \dots & h_{2,\nu-1}(z) \\ \vdots & \vdots & & \vdots \\ c_{\nu,1}(z) & h_{\nu,1}(z) & \dots & h_{\nu,\nu-1}(z) \end{vmatrix}}_{(IV)},$$

where

$$t_i(z) = c_{i+1,1}(z) - \overline{\Phi}c_{i,j}(z).$$

Both determinants (III) and (IV) are polynomials of degree ν since all involved functions are polynomials of degree 1. These polynomials are respectively denoted by $p_{\nu}^{(2)}(z)$ and $p_{\nu}^{(3)}(z)$.

By combining the results for (12) and (13), we find for the characteristic polynomial

$$\epsilon^{2\nu-2}(1-z\Phi)K_{\nu-1}(z) + \epsilon^{2\nu-1}\left((1-z\Phi)p^{(1)}_{\nu-1}(z) + p^{(2)}_{\nu}(z)\right)$$
$$+ \epsilon^{2\nu}p^{(3)}_{\nu}(z). \quad (14)$$

We now normalize (14) so that it equals 1 at 0: this means dividing (14) by $\left|\left(H^{(0)}_{m,\nu}\right)^* H^{(0)}_{m,\nu}\right|$. We take the liberty to re-use the notation $K_{\nu-1}(z)$ after normalization (although its coefficients have been touched) in order to not overload the already lavish notation. When denoting $p^{(4)}_{\nu}(z) = (1 - z\Phi)p^{(1)}_{\nu-1}(z) + p^{(2)}_{\nu}(z)$, we finally obtain the result

$$q(z) = (1-z\Phi)K_{\nu-1}(z) + \epsilon p^{(4)}_{\nu}(z) + \epsilon^2 p^{(3)}_{\nu}(z).$$

Hence, the normalized characteristic polynomial $q(z)$ lies at a distance of order ϵ from the polynomial with the reciprocal of the base term Φ_j as its root, while the remaining zeroes, of the polynomial $K_{\nu-1}(z)$, depend on the random noise terms r_k.

3.2 General case with more base terms

In order to avoid the introduction of unnecessarily complicated notation, the general case is discussed in less detail. Consider the noisy variant of the multi-exponential samples in (1):

$$f_k = \sum_{j=1}^{n} \alpha_j \Phi_j^k + \epsilon r_k, \qquad k = 0, \ldots, N-1.$$

Then for $\nu \geq n$ and $m = N - \nu \geq \nu$, the elements of the Hankel matrices are given by

$$\left(H^{(0)}_{m,\nu}\right)_{i,j} = \sum_{\ell=1}^{n} \alpha_\ell \Phi_\ell^{i+j-2} + \epsilon r_{i+j-2}, \quad \left(H^{(1)}_{m,\nu}\right)_{i,j} = \sum_{\ell=1}^{n} \alpha_\ell \Phi_\ell^{i+j-1} + \epsilon r_{i+j-1}.$$

Hence the elements of the matrix product in the left-hand-side of the generalized eigenvalue problem (11) are

$$\left(\left(H_{m,\nu}^{(0)}\right)^* H_{m,\nu}^{(0)}\right)_{i,j}$$

$$= \sum_{k=1}^{m} \left(\overline{H}_{m,\nu}^{(0)}\right)_{k,i} \left(H_{m,\nu}^{(0)}\right)_{k,j}$$

$$= \sum_{k=1}^{m} \left(\sum_{\ell=1}^{n} \overline{\alpha}_\ell \overline{\Phi}_\ell^{i+k-2} + \epsilon \overline{r}_{i+k-2}\right) \left(\sum_{\ell=1}^{n} \alpha_\ell \Phi_\ell^{j+k-2} + \epsilon r_{j+k-2}\right)$$

$$= \sum_{k=1}^{m} \sum_{\ell_1=1}^{n} \sum_{\ell_2=1}^{n} \overline{\alpha}_{\ell_1} \alpha_{\ell_2} \overline{\Phi}_{\ell_1}^{i+k-2} \Phi_{\ell_2}^{j+k-2} + \mathcal{O}(\epsilon)$$

$$= \sum_{\ell_1=1}^{n} \sum_{\ell_2=1}^{n} \overline{\alpha}_{\ell_1} \alpha_{\ell_2} \overline{\Phi}_{\ell_1}^{i-1} \Phi_{\ell_2}^{j-1} \sum_{k=1}^{m} \overline{\Phi}_{\ell_1}^{k-1} \Phi_{\ell_2}^{k-1} + \mathcal{O}(\epsilon)$$

$$= \sum_{\ell_1=1}^{n} \sum_{\ell_2=1}^{n} A_{\ell_1,\ell_2} \overline{\Phi}_{\ell_1}^{i-1} \Phi_{\ell_2}^{j-1} + \mathcal{O}(\epsilon),$$

with

$$A_{\ell_1,\ell_2} = \overline{\alpha}_{\ell_1} \alpha_{\ell_2} \sum_{k=1}^{m} \overline{\Phi}_{\ell_1}^{k-1} \Phi_{\ell_2}^{k-1}.$$

Analogously, we find

$$\left(\left(H_{m,\nu}^{(0)}\right)^* H_{m,\nu}^{(1)}\right)_{i,j} = \sum_{\ell_1=1}^{n} \sum_{\ell_2=1}^{n} A_{\ell_1,\ell_2} \overline{\Phi}_{\ell_1}^{i-1} \Phi_{\ell_2}^{j} + \mathcal{O}(\epsilon).$$

For the elements of the matrix of which the determinant defines the characteristic polynomial linked to (11) we have

$$\left(\left(H_{m,\nu}^{(0)}\right)^* H_{m,\nu}^{(0)} - z \left(H_{m,\nu}^{(0)}\right)^* H_{m,\nu}^{(1)}\right)_{i,j}$$

$$= \sum_{\ell_1=1}^{n} \sum_{\ell_2=1}^{n} A_{\ell_1,\ell_2} \overline{\Phi}_{\ell_1}^{i-1} \Phi_{\ell_2}^{j-1} (1 - z\Phi_{\ell_2}) + \mathcal{O}(\epsilon P_1(z))$$

$$= \sum_{\ell_2=1}^{n} (1 - z\Phi_{\ell_2}) \Phi_{\ell_2}^{j-1} \sum_{\ell_1=1}^{n} A_{\ell_1,\ell_2} \overline{\Phi}_{\ell_1}^{i-1} + \mathcal{O}(\epsilon P_1(z))$$

$$= \sum_{\ell=1}^{n} B_{\ell,i} D_{\ell,j}^{(0)} (1 - z\Phi_\ell) + \mathcal{O}(\epsilon P_1(z)),$$

where

$$B_{\ell,i} = \sum_{\ell_1=1}^{n} A_{\ell_1,\ell_2} \overline{\Phi}_{\ell_1}^{i-1}, \qquad D_{\ell,j}^{(0)} = \Phi_{\ell}^{j-1}.$$

Note that $B_{\ell,i}$ and $D_{\ell,j}^{(0)}$ are defined separately and not combined into only one coefficient since they respectively only depend on the row or column of the element. This distinction is required for the following operations on the determinant.

The characteristic polynomial is given by

$$\begin{vmatrix} \sum\limits_{\ell=1}^{n} B_{\ell,1} D_{\ell,1}^{(0)}(1-z\Phi_\ell)+\mathcal{O}(\epsilon P_1(z)) & \sum\limits_{\ell=1}^{n} B_{\ell,1} D_{\ell,2}^{(0)}(1-z\Phi_\ell)+\mathcal{O}(\epsilon P_1(z)) \\ \vdots & \vdots \\ \sum\limits_{\ell=1}^{n} B_{\ell,\nu} D_{\ell,1}^{(0)}(1-z\Phi_\ell)+\mathcal{O}(\epsilon P_1(z)) & \sum\limits_{\ell=1}^{n} B_{\ell,\nu} D_{\ell,2}^{(0)}(1-z\Phi_\ell)+\mathcal{O}(\epsilon P_1(z)) \\ & \cdots & \sum\limits_{\ell=1}^{n} B_{\ell,1} D_{\ell,\nu}^{(0)}(1-z\Phi_\ell)+\mathcal{O}(\epsilon P_1(z)) \\ & & \vdots \\ & \cdots & \sum\limits_{\ell=1}^{n} B_{\ell,\nu} D_{\ell,\nu}^{(0)}(1-z\Phi_\ell)+\mathcal{O}(\epsilon P_1(z)) \end{vmatrix}.$$

From columns $j = 2,\ldots,\nu$ we subtract column 1 multiplied by $D_{1,j}^{(0)}/D_{1,1}^{(0)}$ to eliminate the coefficient of the factor $(1-z\Phi_1)$ in all columns except the first one. This yields

$$\begin{vmatrix} \sum\limits_{\ell=1}^{n} B_{\ell,1} D_{\ell,1}^{(1)}(1-z\Phi_\ell)+\mathcal{O}(\epsilon P_1(z)) & \sum\limits_{\ell=2}^{n} B_{\ell,1} D_{\ell,2}^{(1)}(1-z\Phi_\ell)+\mathcal{O}(\epsilon P_1(z)) \\ \vdots & \vdots \\ \sum\limits_{\ell=1}^{n} B_{\ell,\nu} D_{\ell,1}^{(1)}(1-z\Phi_\ell)+\mathcal{O}(\epsilon P_1(z)) & \sum\limits_{\ell=2}^{n} B_{\ell,\nu} D_{\ell,2}^{(1)}(1-z\Phi_\ell)+\mathcal{O}(\epsilon P_1(z)) \\ & \cdots & \sum\limits_{\ell=2}^{n} B_{\ell,1} D_{\ell,\nu}^{(1)}(1-z\Phi_\ell)+\mathcal{O}(\epsilon P_1(z)) \\ & & \vdots \\ & \cdots & \sum\limits_{\ell=2}^{n} B_{\ell,\nu} D_{\ell,\nu}^{(1)}(1-z\Phi_\ell)+\mathcal{O}(\epsilon P_1(z)) \end{vmatrix},$$

where

$$D_{\ell,j}^{(1)} = \begin{cases} D_{\ell,1}^{(0)}, & j = 1, \\[2ex] D_{\ell,j}^{(0)} - \dfrac{D_{1,j}^{(0)}}{D_{1,1}^{(0)}} D_{\ell,1}^{(0)}, & j \neq 1. \end{cases}$$

Now we subtract column 2 multiplied by $D_{2,j}^{(0)}/D_{2,2}^{(0)}$ from the columns $j = 1, 3, \ldots, \nu$ to eliminate the coefficients of the factor $(1 - z\Phi_2)$. The result is

$$
\begin{vmatrix}
\sum\limits_{\substack{\ell=1 \\ \ell\neq 2}}^{n} B_{\ell,1} D_{\ell,1}^{(2)}(1-z\Phi_\ell)+\mathcal{O}(\epsilon P_1(z)) & \sum\limits_{\ell=2}^{n} B_{\ell,1} D_{\ell,2}^{(2)}(1-z\Phi_\ell)+\mathcal{O}(\epsilon P_1(z)) \\
\vdots & \vdots \\
\sum\limits_{\substack{\ell=1 \\ \ell\neq 2}}^{n} B_{\ell,\nu} D_{\ell,1}^{(2)}(1-z\Phi_\ell)+\mathcal{O}(\epsilon P_1(z)) & \sum\limits_{\ell=2}^{n} B_{\ell,\nu} D_{\ell,2}^{(2)}(1-z\Phi_\ell)+\mathcal{O}(\epsilon P_1(z))
\end{vmatrix}
$$

$$
\begin{matrix}
\cdots & \sum\limits_{\ell=3}^{n} B_{\ell,1} D_{\ell,\nu}^{(2)}(1-z\Phi_\ell)+\mathcal{O}(\epsilon P_1(z)) \\
& \vdots \\
\cdots & \sum\limits_{\ell=3}^{n} B_{\ell,\nu} D_{\ell,\nu}^{(2)}(1-z\Phi_\ell)+\mathcal{O}(\epsilon P_1(z))
\end{matrix} \Bigg| ,
$$

where now

$$
D_{\ell,j}^{(2)} = \begin{cases} D_{\ell,1}^{(1)}, & j = 2, \\[2em] D_{\ell,j}^{(1)} - \dfrac{D_{1,j}^{(1)}}{D_{1,1}^{(1)}} D_{\ell,1}^{(1)}, & j \neq 2. \end{cases}
$$

This process is repeated to eliminate the coefficients of the factors $(1 - z\Phi_j)$ for $j = 3, \ldots, n$ which is possible since $\nu \geq n$. We obtain

$$
\begin{vmatrix}
B_{\ell,1} D_{\ell,1}^{(n)}(1-z\Phi_1)+\mathcal{O}(\epsilon P_1(z)) & \cdots & B_{\ell,1} D_{\ell,n}^{(n)}(1-z\Phi_n)+\mathcal{O}(\epsilon P_1(z)) \\
\vdots & & \vdots \\
B_{\ell,\nu} D_{\ell,1}^{(n)}(1-z\Phi_1)+\mathcal{O}(\epsilon P_1(z)) & \cdots & B_{\ell,\nu} D_{\ell,n}^{(n)}(1-z\Phi_n)+\mathcal{O}(\epsilon P_1(z))
\end{vmatrix}
$$

$$
\begin{matrix}
\mathcal{O}(\epsilon P_1(z)) & \cdots & \mathcal{O}(\epsilon P_1(z)) \\
\vdots & & \vdots \\
\mathcal{O}(\epsilon P_1(z)) & \cdots & \mathcal{O}(\epsilon P_1(z))
\end{matrix} \Bigg| .
$$

Splitting the first column and subsequently taking out the constant ϵ everywhere in the last $\nu - n$ columns, results in

$$
\epsilon^{\nu-n}(1 - z\Phi_1) \begin{vmatrix}
B_{\ell,1} D_{\ell,1}^{(n)} & B_{\ell,1} D_{\ell,2}^{(n)}(1-z\Phi_2)+\mathcal{O}(\epsilon P_1(z)) & \cdots \\
\vdots & \vdots & \\
B_{\ell,\nu} D_{\ell,1}^{(n)} & B_{\ell,\nu} D_{\ell,2}^{(n)}(1-z\Phi_2)+\mathcal{O}(\epsilon P_1(z)) & \cdots
\end{vmatrix}
$$

$$
\begin{matrix}
B_{\ell,1} D_{\ell,n}^{(n)}(1-z\Phi_n)+\mathcal{O}(\epsilon P_1(z)) & \mathcal{O}(P_1(z)) & \cdots & \mathcal{O}(P_1(z)) \\
\vdots & \vdots & & \vdots \\
B_{\ell,\nu} D_{\ell,n}^{(n)}(1-z\Phi_n)+\mathcal{O}(\epsilon P_1(z)) & \mathcal{O}(P_1(z)) & \cdots & \mathcal{O}(P_1(z))
\end{matrix} \Bigg| .
$$

$$+ \epsilon^{\nu-n+1} \begin{vmatrix} \mathcal{O}(P_1(z)) & B_{\ell,1}D_{\ell,2}^{(n)}(1-z\Phi_2)+\mathcal{O}(\epsilon P_1(z)) & ... \\ \vdots & \vdots & \\ \mathcal{O}(P_1(z)) & B_{\ell,\nu}D_{\ell,2}^{(n)}(1-z\Phi_2)+\mathcal{O}(\epsilon P_1(z)) & ... \end{vmatrix}$$

$$\begin{vmatrix} B_{\ell,1}D_{\ell,n}^{(n)}(1-z\Phi_n)+\mathcal{O}(\epsilon P_1(z)) & \mathcal{O}(P_1(z)) & ... & \mathcal{O}(P_1(z)) \\ \vdots & \vdots & & \vdots \\ B_{\ell,\nu}D_{\ell,n}^{(n)}(1-z\Phi_n)+\mathcal{O}(\epsilon P_1(z)) & \mathcal{O}(P_1(z)) & ... & \mathcal{O}(P_1(z)) \end{vmatrix}.$$

Since the second term has a higher power of ϵ we only continue with the first term. As above, the subsequent columns are split, which finally gives

$$\epsilon^{\nu-n}\prod_{j=1}^{n}(1-z\Phi_j) \begin{vmatrix} B_{\ell,1}D_{\ell,1}^{(n)} & B_{\ell,1}D_{\ell,2}^{(n)} & ... & B_{\ell,1}D_{\ell,n}^{(n)} & \mathcal{O}(P_1(z)) & ... & \mathcal{O}(P_1(z)) \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ B_{\ell,\nu}D_{\ell,1}^{(n)} & B_{\ell,\nu}D_{\ell,2}^{(n)} & ... & B_{\ell,\nu}D_{\ell,n}^{(n)} & \mathcal{O}(P_1(z)) & ... & \mathcal{O}(P_1(z)) \end{vmatrix}$$
$$+ \mathcal{O}\left(\epsilon^{\nu-n+1}P_\nu(z)\right).$$

Hence, there exists a polynomial $K_{\nu-n}(z)$ of degree $\nu - n$ such that the normalized characteristic polynomial takes the form

$$q(z) = \prod_{j=1}^{n}(1-z\Phi_j)K_{\nu-n}(z) + \mathcal{O}(\epsilon P_\nu(z)).$$

This means that also in the general case, the coefficients of the characteristic polynomial lie at a distance of order ϵ from the polynomial with the reciprocals of the base terms as n of its zeroes and the remaining $\nu - n$ roots influenced by the random noise terms r_k.

4 Sensitivity analysis and experiments

In the previous sections we showed that the Padé denominator of degree $\nu > n$, when computed in a least squares sense from noisy data as in Section 3, is very similar to the Padé denominator of the classical approximant $[\nu - 1/\nu]_F$ computed from the first 2ν perturbed Taylor coefficients of $F(z)$ as in [2]. In order to know that, besides the denominator polynomial, also the n base terms are returned reliably, we look at the sensitivity of the rooting of the least squares Padé denominator of degree $\nu > n$. This aspect was not discussed in [1, 2], but we will compute that classical case as well and compare to it.

It is known that polynomial rooting can be a very sensitive problem statement, even irrespective of the algorithm used to compute the roots. Although we are dealing with ν roots here, only the conditioning of the n authentic base terms interests us. We disregard the conditioning of the $\nu - n$ spurious ones.

In the sequel of this section we can therefore return to the case of $n = 1$ base term with $\nu - 1$ spurious roots. In the case of only one base term, the normalized characteristic polynomial which equals the denominator of the least

squares Padé approximant, takes the form

$$q(z) = (1 - z\Phi)K_{\nu-1}(z) + \mathcal{O}(\epsilon P_\nu(z)), \qquad q(0) = 1.$$

With exact data, so for $\epsilon = 0$, the classical Padé denominator $q(z)$ equals $1 - z\Phi$ because of the consistency of the Padé approximation process. With ϵ assumed small, we are interested in the sensitivity of the root $1/\Phi$ of $Q(z) := (1 - z\Phi)K_{\nu-1}(z)$ with respect to each of the coefficients of $Q(z)$. Let $q_j, j = 0, \ldots, \nu$ denote the coefficient of z^j in $Q(z)$. Then the root sensitivity with respect to q_j is given by

$$\left| \frac{\partial z}{\partial q_j} \big|_{z=1/\Phi} \right| = \left| \frac{-\partial Q/\partial q_j}{\partial Q/\partial z} \big|_{z=1/\Phi} \right| \tag{15}$$

$$= \left| \frac{z^j}{\Phi K_{\nu-1}(z) + (z\Phi - 1)K'_{\nu-1}(z)} \big|_{z=1/\Phi} \right| \tag{16}$$

$$= \left| \frac{1}{\Phi^{j+1} K_{\nu-1}(1/\Phi)} \right|. \tag{17}$$

So this sensitivity essentially depends on the location of the base term Φ, the number of rows m and the number of columns ν in the Hankel matrices, and the random numbers $r_j, j = 0, \ldots, N - 1$ appearing in $K_{\nu-1}(z)$. In our numerical experiments, the latter will follow a complex Gaussian distribution with mean 0 and variance 1. To illustrate the influence of the former, we perform 3 experiments illustrating the behaviour of (15), where:

- we vary $|\Phi|$ and $\arg(\Phi)$ while keeping m and ν fixed,
- we fix Φ and vary m and ν, exploring different ratios for m/ν,
- we study the distribution of the sensitivity for fixed Φ, m and ν.

In the Figures 1 and 2 illustrating the sensitivity experiments, the scale of the z-axis is logarithmic (Briggs). In the Figures 3, 4, 5 and 6 the scale of the y-axis is not. The value $K_{\nu-1}(1/\Phi)$ is obtained from the determinant expression associated with (14).

In our first experiment we choose $m = 10$ and $\nu = 5$ while varying Φ within the ring $0.75 < |\Phi| < 1.25$. This is an important region as many real-life applications deal with Φ on the unit circle in the complex plane. For each Φ we run 100 computations of the sensitivity, every time with different random values r_j and display the median of the maximal sensitivity over the various coefficients q_j in $Q(z)$. Observe from Figure 1 that primarily the modulus of the base term Φ influences the sensitivity, with larger moduli leading to a smaller sensitivity. That should not come as a total surprise because of the denominator Φ^{j+1} in $\partial z/\partial q_j$. The argument of Φ plays a lesser role, only causing some fluctuations due to the random nature of the r_j.

An especially interesting and frequently occurring case is when the base terms lie on the unit circle. For $|\Phi| = 1$ we read from Figure 1 that the median worst condition number with respect to the various coefficients of $Q(z)$

is approximately $10^{-0.514} \approx 0.306$ (with a small standard deviation ≈ 0.013). The mean value rounded to 3 significant digits is the same.
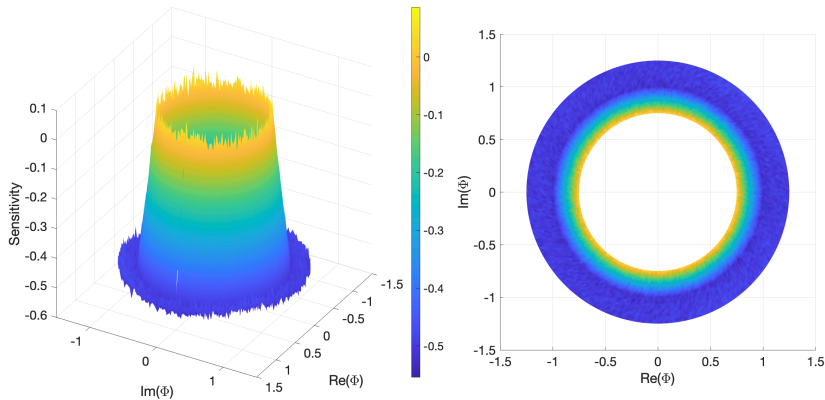

Fig. 1 Median of $\max_{j=0,\dots,\nu}$ sensitivity for different base terms.

In the second experiment we randomly fix the value of the base term $\Phi = \exp(1.8115838742\pi i)$ without any issue, as a consequence of the first experiment. We study the influence on the sensitivity of both the number of rows m and the number of columns ν of the Hankel matrices. Again 100 computations of the sensitivity are run, every time with different random values r_j and the median of the maximal sensitivity over the various coefficients is selected. The results are shown in Figure 2. In Figure 3 we zoom in on some specific ratios for m/ν, thereby comparing to the conditioning of the classical Padé denominator (where $m = \nu$), obtained in [1]. Both the square and the overdetermined problem statements are clearly similarly well-conditioned.

We conclude that the sensitivity primarily decreases with the number of columns, a recommendation that was already formulated in [8] and [11]. Rather than add a handful of terms to n to counter the noise, we recommend to choose ν a good deal larger than n.

In a final experiment, we again randomly fix the base term $\Phi = \exp(1.8115838742\pi i)$, and moreover fix the dimensions $\nu = 10$, $m = 20$. In this case 1 million different noise realizations are used. The left plot in Figure 4 depicts a histogram over all values of the sensitivity, while the right plot in Figure 4 is restricted to the interval $[0, 0.5]$. We observe that there are a few outliers present, but the majority of the values lie in the interval $[0, 0.5]$. Subsequently, we increase the value of m to 50, and repeat the experiment. The results are found in Figure 5. First of all, we notice that we do not find any outliers. Also, the distribution becomes narrower. However, the mean has slightly increased. This influence of the number of rows m is confirmed in Figure 6,
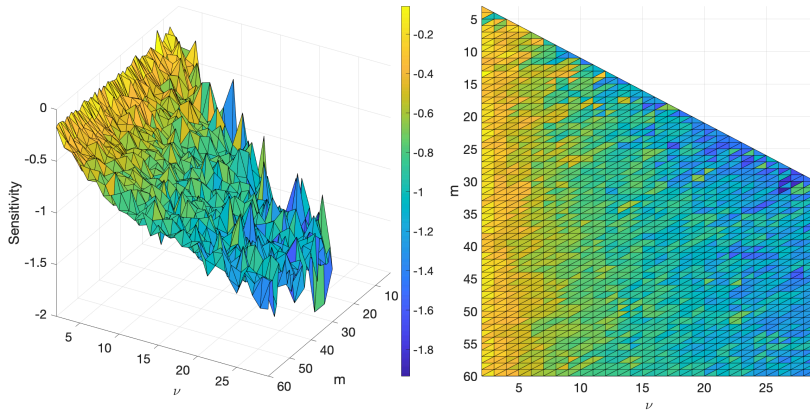
Fig. 2 Median of $\max_{j=0,\ldots,\nu}$ sensitivity in terms of m and ν.
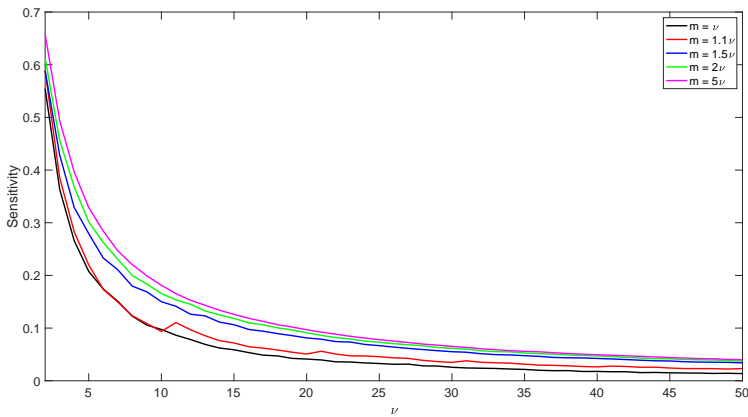

Fig. 3 Conditioning for various m/ν ratios, including $m = \nu$ from [1].

where we further increase m to 100. In this case, the distribution becomes even more narrow. Hence, we can conclude that increasing the number of rows does not decrease the sensitivity, but rather stabilizes this value.

Future work

In several exponential analysis algorithms the rooting of the characteristic polynomial $q(z)$ or its reverse $z^\nu q(1/z)$, which is called the Prony polynomial, is an essential step. Another approach to obtain the base terms $\Phi_j, j = 0, \ldots, n$ is through the solution of a generalized eigenvalue problem [12]. In a follow-up paper we plan to make a perturbation analysis of this alternative approach, when programmed in a least squares sense.
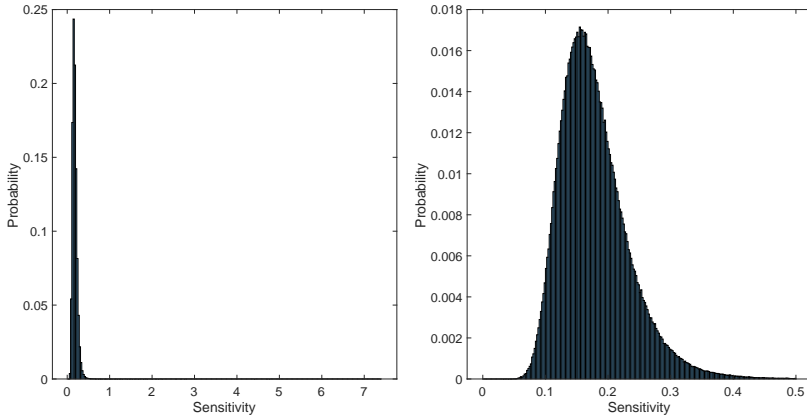
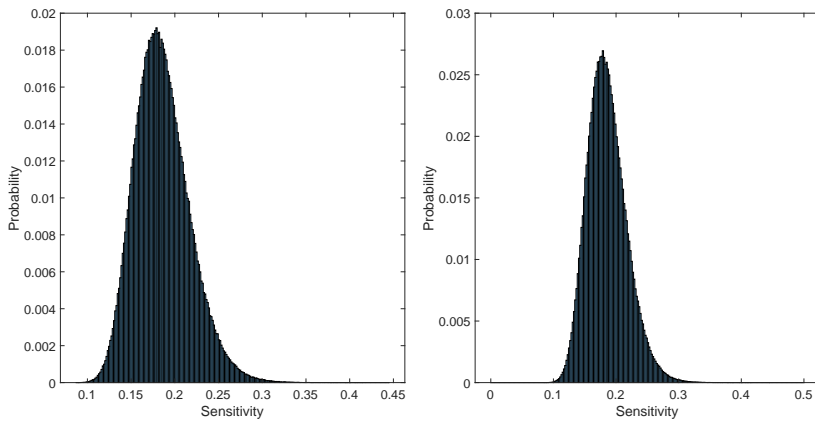**Fig. 4** Histogram of the sensitivity for $\nu = 10$, $m = 20$.

Fig. 5 Histogram of the sensitivity for $\nu = 10$, $m = 50$.

5 Acknowledgement

Declarations

The authors declare that they have no conflict of interest.
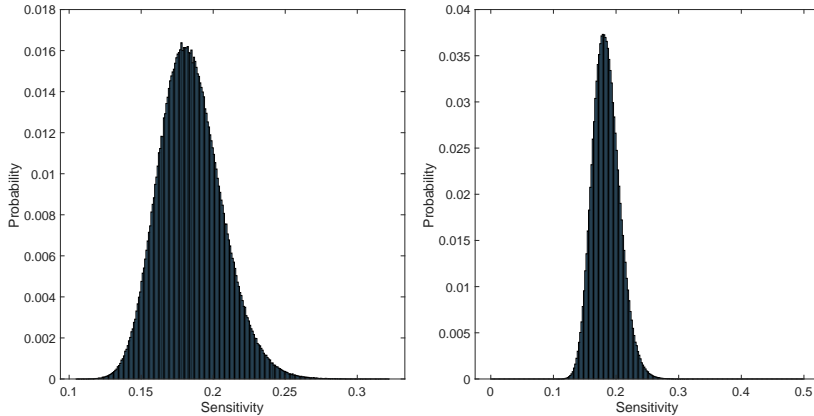The datasets generated during this study are available on the CEMath webpage at https://cemath.org/publications/papers/.

Fig. 6 Histogram of the sensitivity for $\nu = 10$, $m = 100$.

References

[1] Gilewicz, J., Pindor, M.: Padé approximants and noise: a case of geometric series. J. Comput. Appl. Math. **87**, 199–214 (1997). https://doi.org/10.1016/S0377-0427(97)00185-4

[2] Gilewicz, J., Pindor, M.: Padé approximants and noise: rational functions. J. Comput. Appl. Math. **105**, 285–297 (1999). https://doi.org/10.1016/S0377-0427(99)00041-2

[3] Gonnet, P., Güttel, S., Trefethen, L.N.: Robust Padé approximation via SVD. SIAM Rev. **55**, 101–117 (2013). https://doi.org/10.1137/110853236

[4] Householder, A.S.: On Prony's method of fitting exponential decay curves and multiple-hit survival curves. Report ORNL-455, Oak Ridge National Laboratory (1950)

[5] Guillaume, P., Huard, A.: Multivariate Padé approximation. J. Comput. Appl. Math. **121**, 197–219 (2000). https://doi.org/10.1016/S0377-0427(00)00337-X

[6] Ben-Or, M., Tiwari, P.: A deterministic algorithm for sparse multivariate polynomial interpolation. In: STOC '88: Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing, pp. 301–309. ACM, New York, NY, USA (1988). https://doi.org/10.1145/62212.62241

[7] Weiss, L., McDonough, R.N.: Prony's method, Z-transforms, and Padé approximation. SIAM Rev. **5**, 145–149 (1963)

[8] Cuyt, A., Lee, W.-s.: Sparse interpolation and rational approximation. In:

Hardin, D., Lubinsky, D., Simanek, B. (eds.) Contemporary Mathematics, vol. 661, pp. 229–242. American Mathematical Society, Providence, RI (2016). https://doi.org/10.1090/conm/661/13284

[9] Cuyt, A., Knaepkens, F., Lee, W.: From exponential analysis to Padé approximation and Tensor decomposition, in one and more dimensions. In: Gerdt, V.P., *et al.*(eds.) LNCS11077, Springer, Switzerland, pp. 116–130 (2018). https://doi.org/10.1007/978-3-319-99639-4. Proceedings CASC 2018, Lille (France)

[10] Hildebrand, F.B.: Introduction to Numerical Analysis, 2nd edn. Dover Publications, Inc., New York (1987)

[11] Cuyt, A., Tsai, M., Verhoye, M., Lee, W.-s.: Faint and clustered components in exponential analysis. Appl. Math. Comput. **327**, 93–103 (2018)

[12] Hua, Y., Sarkar, T.K.: Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. IEEE Trans. Acoust., Speech, Signal Process. **38**, 814–824 (1990). https://doi.org/10.1109/29.56027