Adv Comput Math (2018) 44:987–1002
https://doi.org/10.1007/s10444-017-9570-8

Multivariate exponential analysis from the minimal number of samples

Annie Cuyt[1] · **Wen-shin Lee**[1]

Received: 16 January 2017 / Accepted: 2 November 2017 /
Published online: 16 November 2017

Abstract The problem of multivariate exponential analysis or sparse interpolation has received a lot of attention, especially with respect to the number of samples required to solve it unambiguously. In this paper we show how to bring the number of samples down to the absolute minimum of $(d+1)n$ where d is the dimension of the problem and n is the number of exponential terms. To this end we present a fundamentally different approach for the multivariate problem statement. We combine a one-dimensional exponential analysis method such as ESPRIT, MUSIC, the matrix pencil or any Prony-like method, with some linear systems of equations because the multivariate exponents are inner products and thus linear expressions in the parameters.

Keywords Exponential sum · Multivariate · Prony's method

Mathematics Subject Classification (2010) 42B99 · 42A15

1 Introduction

Multivariate exponential analysis is a classical problem at the basis of many application domains (such as, for instance, [13, 14, 25, 27]) that recently has gained a lot

Communicated by: Yuesheng Xu

✉ Wen-shin Lee
wen-shin.lee@uantwerpen.be

Annie Cuyt
annie.cuyt@uantwerpen.be

[1] Department of Mathematics and Computer Science, University of Antwerp (CMI), Middelheimlaan 1, 2020 Antwerpen, Belgium

of attention. The problem statement is that of recovering the vectors $\phi_j \in \mathbb{C}^d$, $j = 1, \ldots, n$ and the coefficients $\alpha_j \in \mathbb{C}$, $j = 1, \ldots, n$ in the d-variate n-sparse sum

$$f(x) := f(x_1, \ldots, x_d) = \sum_{j=1}^{n} \alpha_j \exp\left(\langle \phi_j, x \rangle\right),$$

$$x = (x_1, \ldots, x_d), \qquad \phi_j = (\phi_{j1}, \ldots, \phi_{jd}), \qquad \langle \phi_j, x \rangle = \sum_{i=1}^{d} \phi_{ji} x_i,$$

from $(d+1)n$ samples of $f(x_1, \ldots, x_d)$, which is the minimal number of samples because it equals the number of parameters in the problem statement.

When $d = 1$ then the problem can be solved using a variety of Prony-based algorithms [2, 10, 20, 23], in which the identification of the ϕ_j and α_j is separated and taken care of in two stages. The frequencies ϕ_j, $j = 1, \ldots, n$ are obtained from a generalized eigenvalue or polynomial rooting problem, while the linear coefficients α_j, $j = 1, \ldots, n$ are computed from a Vandermonde system of linear equations [8, pp. 378–382]. Input to these algorithms are $2n$ samples of $f(x)$ at some equidistant points $f(s\Delta)$, $s = 0, \ldots, 2n-1$. This number of samples is minimal if n is known. Otherwise at least one more sample is required to identify the sparsity n. For more details on the latter we refer to [5, 11].

Several computational methods were developed to solve the problem also when $d > 1$, from straightforward generalizations to more sophisticated approaches, all of them using more than a minimum of $(d+1)n$ samples though. It should be obvious to the reader that the challenge is not to recover inner products $\langle \phi_j, x \rangle$ and the associated coefficients α_j for $j = 1, \ldots, n$, from $2n$ equidistant samples in higher-dimensional space. Under modest conditions this can be achieved using the univariate techniques mentioned above. Instead, the challenge is to recover the individual ϕ_{ji}, $j = 1, \ldots, n$, $i = 1, \ldots, d$ and the coefficients α_j. We describe the state of the art in multivariate exponential analysis and explain how our approach differs from it.

The one-dimensional matrix pencil method was generalized to the 2-dimensional matrix enhancement and matrix pencil method (MEMP) [9] and can be extended to higher dimensions in a straightforward manner. It uses a Hankel-block-Hankel matrix to decompose the 2-dimensional problem into two one-dimensional problems reflecting each dimension. This decomposition introduces an additional challenge though, namely that of matching or pairing the information computed from the one-dimensional problems [19]. Moreover, when constructing a uniform d-dimensional grid of sample points, the amount of information is $O(n^d)$.

Solving the problem along some one-dimensional subspace, in other words computing some projection such as in [17, 18] requires only $O(n)$ samples. Using an adaptive sampling scheme and under some mild condition on the coefficients, this remains valid in the 2-dimensional case [26]. However, in [6] is shown that there is no finite set of (independently of f) predefined lines for which the bivariate reconstruction problem has a unique solution. A lower bound for the number of samples in the reconstruction when $d = 2$ is $O(n^2)$. In order to solve the pairing problem,

[6] reformulates the problem as a non-convex optimization problem, which is not computationally feasible for practical purposes.

Rather than projecting on one-dimensional subspaces, a symbolic approach based on [22] is developed in [21] making use of constructive ideal theory and multivariate polynomial interpolation. The largest number of required samples in this setting is estimated to be $O((d+1)n^2 \log^{2d-2} n)$. In the same corner one finds [12] and [16] which obtain the multivariate exponents as common roots of a finite system of d-variate polynomials. Still making use of $O(n^d)$ samples however, algebraic geometry theory now guarantees the correct pairing and recovery.

The method we propose differs significantly from all of the above, not only in its informational usage which can be as low as $(d+1)n$, but also in its approach which only makes use of a 1-dimensional Prony technique combined with some linear systems of equations because the individual ϕ_{ji} appear linearly in the $\langle \phi_j, x \rangle$. The presented multivariate exponential analysis technique results from ideas that were initially formulated in [3, 4]: a so-called identification shift in the sampling strategy allows to overcome any ambiguity in the exponential analysis.

After this state of the art of the literature, Sections 2 and 3 deal with the ideal case where some mild assumptions are verified and only $(d+1)n$ evaluations are necessary, thus generalizing Prony's result where $2n$ samples solve a univariate exponential analysis problem. In Section 4 the most general case is detailed, requiring slightly more samples because the assumptions do not hold. An analysis of the worst case scenario and an algorithm for the detection of n is presented in Section 5. Finally, the new algorithm is illustrated with an example in Section 6.

2 Multivariate exponential analysis

As surveyed in the introduction, up to now computational methods require more samples than the minimal number, for one or other reason. We now explain how the problem statement can also be solved in the multivariate setting using the minimal number $(d+1)n$ of samples. The trick to achieve this is to split the set of samples in two subsets, namely $2n$ equidistant samples and another $(d-1)n$ samples that may but need not be equidistant in the higher-dimensional space (they cannot be entirely unstructured though). We discuss the use of the $2n$ equidistant samples in this section and that of the additional $(d-1)n$ samples in Section 3. For now we assume in the multivariate setting that the value of n is known. How to detect n is further discussed in Section 5.

Let $\Delta = (\Delta_1, \ldots, \Delta_d) \neq (0, \ldots, 0)$ and $|\Im \phi_{ji}| < \pi/|\Delta_i|$, $j = 1, \ldots, n, i = 1, \ldots, d$ [15, 24], where the function `i` returns the imaginary part of a complex number. Let us sample $f(x_1, \ldots, x_d)$ at the points $s\Delta$, $s = 0, \ldots, 2n-1$:

$$F_s := f(s\Delta_1, \ldots, s\Delta_d), \qquad s = 0, \ldots, 2n-1. \tag{1}$$

For the time being, we also assume that the sampling direction Δ is such that the values $\exp(\langle \phi_j, \Delta \rangle)$, $j = 1, \ldots, n$ are mutually distinct. How to deal with collisions in these values is described in Section 4.

Following the univariate scheme [8, pp. 378–382] the coefficients $\beta_i, i = 0, \ldots, n-1$ of the polynomial

$$B(z) = \prod_{j=1}^{n} \left(z - \exp\left(\langle \phi_j, \Delta \rangle\right)\right) = z^n + \beta_{n-1} z^{n-1} + \ldots + \beta_0 \tag{2}$$

can be obtained from the $n \times n$ Hankel system of linear equations

$$\begin{pmatrix} F_0 & F_1 & \cdots & F_{n-1} \\ F_1 & \cdots & & F_n \\ \vdots & & & \vdots \\ F_{n-1} & F_n & \cdots & F_{2n-2} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_{n-1} \end{pmatrix} = - \begin{pmatrix} F_n \\ \vdots \\ F_{2n-1} \end{pmatrix}, \tag{3}$$

or the roots $\exp(\langle \phi_j, \Delta \rangle)$, $j = 1, \ldots, n$ of $B(z)$ can be found as the generalized eigenvalues λ of the problem

$$\begin{pmatrix} F_1 & F_2 & \cdots & F_n \\ F_2 & \cdots & & F_{n+1} \\ \vdots & & & \vdots \\ F_n & F_{n+1} & \cdots & F_{2n-1} \end{pmatrix} v = \lambda \begin{pmatrix} F_0 & F_1 & \cdots & F_{n-1} \\ F_1 & \cdots & & F_n \\ \vdots & & & \vdots \\ F_{n-1} & F_n & \cdots & F_{2n-2} \end{pmatrix} v, \qquad v \in \mathbb{C}^n. \tag{4}$$

So we can recover the expressions $\exp(\Phi_j)$ where

$$\Phi_j = \langle \phi_j, \Delta \rangle, \qquad j = 1, \ldots, n. \tag{5}$$

Although we have not yet identified the individual $\phi_{ji}, j = 1, \ldots, n, i = 1, \ldots, d$, nothing prevents us from already computing the linear coefficients α_j from one of the $n \times n$ Vandermonde systems

$$\begin{pmatrix} \exp(k\Phi_1) & \exp(k\Phi_2) & \cdots & \exp(k\Phi_n) \\ \vdots & & & \vdots \\ \exp((k+n-1)\Phi_1) & \exp((k+n-1)\Phi_2) & \cdots & \exp((k+n-1)\Phi_n) \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$
$$= \begin{pmatrix} F_k \\ \vdots \\ F_{k+n-1} \end{pmatrix}, \qquad 0 \le k \le n. \tag{6}$$

The latter can also be replaced by the $2n \times n$ Vandermonde system involving all samples, which is then solved in the least squares sense, as recommended in the case of real-life and hence noisy data.

3 Identification shifts

In order to extract the $\phi_{ji}, j = 1, \ldots, n, i = 1, \ldots, d$ from the $\Phi_j, j = 1, \ldots, n$, still under the assumption that the values $\exp(\Phi_j), j = 1, \ldots, n$ are mutually distinct, some additional samples are required. For this purpose we choose a set

$\{\Delta, \delta_1, \ldots, \delta_{d-1}\}$ of d linearly independent vectors in $\mathbb{C}^d$. The additional samples are taken along a linear combination of Δ and some δ_i, $i = 1, \ldots, d-1$:

$$\begin{aligned} F_{\ell i} &:= f(\kappa_{\ell i}\Delta + \delta_i) = f(\kappa_{\ell i}\Delta_1 + \delta_{i1}, \ldots, \kappa_{\ell i}\Delta_d + \delta_{id}), \\ \ell &= 1, \ldots, n, \quad i = 1, \ldots, d-1 \end{aligned} \tag{7}$$

where the $\kappa_{\ell i}, \ell = 1, \ldots, n$ for fixed i are taken to be mutually distinct. A simple choice for $\kappa_{\ell i}$ for all i is $\kappa_{\ell i} = \ell - 1$. Then the additional samples are taken equidistantly along independent shifts δ_i with respect to the original vector Δ, in other words $F_{\ell i} = f((\ell-1)\Delta + \delta_i)$. At the same time we assume that

$$|\Im\langle \phi_j, \delta_i/||\delta_i||\rangle| < \pi/||\delta_i||, \qquad j = 1, \ldots, n, \quad i = 1, \ldots, d$$

in order to comply with the Shannon-Nyquist conditions formulated in [15, 24]. We call these vectors $\delta_i, i = 1, \ldots, d-1$ identification shifts for reasons that will become apparent: they allow to identify the individual ϕ_{ji} from the computed Φ_j. For this identification we exploit the fact that the ϕ_{ji} appear linearly in the Φ_j and hence we turn our attention to systems of linear equations rather than to multivariate polynomial root solving or structured generalized eigenvalue problems.

Consider for fixed $i = 1, \ldots, d-1$, meaning for a chosen linearly independent shift vector δ_i, the following Vandermonde-like system of linear equations:

$$\begin{pmatrix} \exp(\kappa_{1i}\Phi_1) & \exp(\kappa_{1i}\Phi_2) & \cdots & \exp(\kappa_{1i}\Phi_n) \\ \exp(\kappa_{2i}\Phi_1) & & \cdots & \exp(\kappa_{2i}\Phi_n) \\ \vdots & & & \vdots \\ \exp(\kappa_{ni}\Phi_1) & & \cdots & \exp(\kappa_{ni}\Phi_n)) \end{pmatrix} \begin{pmatrix} A_{1i} \\ \vdots \\ A_{ni} \end{pmatrix} = \begin{pmatrix} F_{1i} \\ \vdots \\ F_{ni} \end{pmatrix}. \tag{8}$$

Since we know $\exp(\Phi_j)$, $j = 1, \ldots, n$ and have chosen $\kappa_{\ell i}, \ell = 1, \ldots, n$, with i fixed, the Vandermonde-like coefficient matrix can easily be composed. Note that for the choice $\kappa_{\ell i} = \ell - 1$ the Vandermonde-like coefficient matrix coincides with that of (6) where $k = 0$. The unknowns A_{ji}, $j = 1, \ldots, n$ come from a reinterpretation of the samples $F_{\ell i}$ as

$$F_{\ell i} = f(\kappa_{\ell i}\Delta + \delta_i) = \sum_{j=1}^{n} \alpha_j \exp\left(\langle \phi_j, \delta_i\rangle\right) \exp\left(\langle \phi_j, \kappa_{\ell i}\Delta\rangle\right), \qquad \ell = 1, \ldots, n$$

with

$$A_{ji} = \alpha_j \exp\left(\langle \phi_j, \delta_i\rangle\right), \qquad j = 1, \ldots, n$$

and

$$\exp(\kappa_{\ell i}\Phi_j) = \exp\left(\langle \phi_j, \kappa_{\ell i}\Delta\rangle\right), \qquad \ell, j = 1, \ldots, n.$$

The values A_{ji}/α_j equal

$$\frac{A_{ji}}{\alpha_j} = \exp\left(\langle \phi_j, \delta_i\rangle\right), \qquad j = 1, \ldots, n,$$

which we denote by

$$\exp(\Phi_{ji}) := \exp\left(\langle \phi_j, \delta_i\rangle\right), \qquad j = 1, \ldots, n.$$

Here the index i is still fixed. Note that we have no problem to pair the Φ_{ji} to the Φ_j, $j = 1, \ldots, n$ since for each i the A_{ji} are paired to the α_j, $j = 1, \ldots, n$ through the Vandermonde-like systems (6) and (8).

By setting up (8) for each $i = 1, \ldots, d-1$ and pairing its solution with Φ_j in (5), we obtain for fixed $j = 1, \ldots, n$ the linear system of equations

$$\begin{pmatrix} \Delta_1 & \cdots & \Delta_d \\ \delta_{11} & \cdots & \delta_{1d} \\ \vdots & & \vdots \\ \delta_{d-1,1} & \cdots & \delta_{d-1,d} \end{pmatrix} \begin{pmatrix} \phi_{j1} \\ \vdots \\ \phi_{jd} \end{pmatrix} = \begin{pmatrix} \Phi_j \\ \Phi_{j1} \\ \vdots \\ \Phi_{j,d-1} \end{pmatrix}. \tag{9}$$

Since the vectors Δ and δ_i, $i = 1, \ldots, d-1$ are linearly independent, the coefficient matrix of (9) is regular and so the individual ϕ_{ji}, $j = 1, \ldots, n, i = 1, \ldots, d$ can be computed, at the expense of $2n$ evaluations F_s in (1) and $(d-1)n$ evaluations $F_{\ell i}$ in (7).

Before we continue we point out that (as is clear from the semantics of the formulas) we can also denote Δ as δ_0, F_s as F_{s0} and Φ_j as Φ_{j0}.

4 Disentangling collisions

We now turn our attention to the situation in which the first batch of samples F_s at multiples of the vector Δ does not reveal all individual terms because some values $\exp(\Phi_j)$, $j = 1 \ldots, n$ collide and the exponential sum shrinks to $\nu < n$ terms. For ease of notation, but without loss of generality, we take the colliding terms to be successive, for instance: $\exp(\Phi_1) = \ldots = \exp(\Phi_{h_1}), \exp(\Phi_{h_1+1}) = \ldots = \exp(\Phi_{h_2}), \ldots, \exp(\Phi_{h_{\nu-1}+1}) = \ldots = \exp(\Phi_n)$. Assume that with $0 \le s \le 2\nu - 1$, $\nu \le n$ the exponential samples break down into

$$F_s = \sum_{j=1}^{\nu} \left(\alpha_{h_{j-1}+1} + \ldots + \alpha_{h_j}\right) \exp\left(\langle \phi_{h_j}, s\Delta \rangle\right), \quad h_0 = 0, \quad h_j < h_{j+1}, \quad h_\nu = n \tag{10}$$

because

$$\exp(\Phi_{h_{j-1}+1}) = \ldots = \exp(\Phi_{h_j}), \qquad j = 1, \ldots, \nu.$$

Since $|\Im \phi_{ji}| < \pi/|\Delta_i|$, $j = 1, \ldots, n, i = 1, \ldots, d$, we actually have

$$\Phi_{h_{j-1}+1} = \cdots = \Phi_{h_j}, \qquad j = 1, \ldots, \nu.$$

The Vandermonde system (6) now becomes

$$\begin{pmatrix} \exp(k\Phi_{h_1}) & \cdots & \exp(k\Phi_{h_\nu}) \\ \vdots & & \vdots \\ \exp((k+\nu-1)\Phi_{h_1}) & \cdots & \exp((k+\nu-1)\Phi_{h_\nu}) \end{pmatrix} \begin{pmatrix} \alpha_1 + \ldots + \alpha_{h_1} \\ \vdots \\ \alpha_{h_{\nu-1}+1} + \ldots + \alpha_{h_\nu} \end{pmatrix}$$
$$= \begin{pmatrix} F_k \\ \vdots \\ F_{k+\nu-1} \end{pmatrix}, \qquad 0 \le k \le \nu. \tag{11}$$

Note that at the same time, the degree of the polynomial $B(z)$ in (2) is only ν. How this is detected and how the true n is revealed is discussed in the next section. To proceed we denote

$$A_j := \alpha_{h_{j-1}+1} + \ldots + \alpha_{h_j}, \qquad j = 1, \ldots, \nu. \tag{12}$$

To disentangle the collisions in the exponential sum, we need additional evaluations besides the minimal number $(d+1)n$. At the end of Section 5 we also explain how these additional evaluations allow to deal with the situation where some $A_j = 0$.

We start with $i = 1$ and the identification shift vector δ_1. First we point out how the Vandermonde-like system (8) of Section 3 looks like in case of such collisions: in the coefficient matrix the value n is replaced by ν and in Φ_j the index j is replaced by h_j. With the collisions in (10), the unknowns A_{j1}, $j = 1, \ldots, \nu$ take the form

$$A_{j1} = \alpha_{h_{j-1}+1} \exp\left(\langle \phi_{h_{j-1}+1}, \delta_1 \rangle\right) + \ldots + \alpha_{h_j} \exp\left(\langle \phi_{h_j}, \delta_1 \rangle\right), \qquad j = 1, \ldots, \nu.$$

In the sequel we denote from here on the additional evaluations $F_{\ell 1}$ mentioned in Section 3 by

$$F_{1\ell 1} := F_{\ell 1} = f(\kappa_{\ell 1}\Delta + \delta_1), \qquad \ell = 1, \ldots, \nu$$

and we add, still with $i = 1$, the samples

$$F_{s\ell 1} := f(\kappa_{\ell 1}\Delta + s\delta_1), \quad s = 2, 3, \ldots, 2 \max_{1 \le j \le \nu} (h_j - h_{j-1}), \quad \ell = 1, \ldots, \nu, \quad i = 1.$$

The triple index expresses the shift vector multiple in the index s, the collision into ν piles of the Φ_j in the index ℓ, and the identification level in i (which is $i = 1$ here).

Since the values of h_j are actually unknown, the addition of samples is done further and interlaced with singularity checks of some Hankel matrices, as we explain now. The checks are performed for each collision or pile h_j and later repeated for each i. Collisions in the space spanned by Δ may not be fully disentangled in the space spanned by Δ and δ_1, but they are gradually being disentangled as we add independent vectors δ_i until we span the whole space. At the last stage, when dealing with the full basis $\Delta, \delta_1, \ldots, \delta_{d-1}$, the true n is revealed because in the end all collisions are taken apart, given enough additional samples. For the moment we continue with $i = 1$.

For each s separately, we set up in analogy with (8), the Vandermonde-like system

$$\begin{pmatrix} \exp(\kappa_{h_1 1}\Phi_{h_1}) & \exp(\kappa_{h_1 1}\Phi_{h_2}) & \cdots & \exp(\kappa_{h_1 1}\Phi_{h_\nu}) \\ \exp(\kappa_{h_2 1}\Phi_{h_1}) & & \cdots & \exp(\kappa_{h_2 1}\Phi_{h_\nu}) \\ \vdots & & & \vdots \\ \exp(\kappa_{h_\nu 1}\Phi_{h_1}) & & \cdots & \exp(\kappa_{h_\nu 1}\Phi_{h_\nu}) \end{pmatrix} \begin{pmatrix} A_{s11} \\ \vdots \\ A_{s\nu 1} \end{pmatrix} = \begin{pmatrix} F_{s11} \\ \vdots \\ F_{s\nu 1} \end{pmatrix} \tag{13}$$

where

$$A_{sj1} = \alpha_{h_{j-1}+1} \exp\left(\langle \phi_{h_{j-1}+1}, s\delta_1 \rangle\right) + \ldots + \alpha_{h_j} \exp\left(\langle \phi_{h_j}, s\delta_1 \rangle\right), \qquad j = 1, \ldots, \nu. \tag{14}$$

Note that the coefficient matrix is independent of s. Also, the former unknowns A_j and A_{j1} can as well be indexed as A_{0j1} and A_{1j1} respectively, and so (13) and

(14) remain valid for $s = 0, 1$, which is important for the sequel. The values A_j from (12) and $A_{sj1}, s \geq 1$ from (14) are actually equidistant samples of the function

$$A_{j1}(x) = A_{j1}(x_1, \ldots, x_d) = \alpha_{h_{j-1}+1} \exp\left(\langle \phi_{h_{j-1}+1}, x\rangle\right) + \ldots + \alpha_{h_j} \exp\left(\langle \phi_{h_j}, x\rangle\right), \quad j = 1, \ldots, \nu, \tag{15}$$

taken at $x = s\delta_1, s \geq 0$. For each fixed $j = 1, \ldots, \nu$ we now put together the Hankel matrix

$$\begin{pmatrix} A_{0j1} & A_{1j1} & A_{2j1} & A_{3j1} & \ldots \\ A_{1j1} & A_{2j1} & A_{3j1} & \ldots & \\ A_{2j1} & A_{3j1} & A_{4j1} & \ldots & \\ A_{3j1} & \vdots & \vdots & & \\ \vdots & & & & \end{pmatrix}. \tag{16}$$

Note that in order to enlarge (16) with one row and column for a particular j, one needs to solve (13) for two additional values of s, thereby obtaining the additional A_{sj1} for all $1 \leq j \leq \nu$.

It is known that the rank of any $(h_j - h_{j-1} + t) \times (h_j - h_{j-1} + t)$ submatrix for finite $t \geq 0$ is bounded by $h_j - h_{j-1}$ [5, 11] since $h_j - h_{j-1}$ equals the number of terms in each of the evaluations $A_j, A_{sj1}, s \geq 1$. The actual rank r_j of the $(h_j - h_{j-1}) \times (h_j - h_{j-1})$ submatrix with A_j in the top left corner tells us (with high probability [11]) how many of the $h_j - h_{j-1}$ terms in $A_j(x)$ can indeed be separated at the current level ($i = 1$) where identification shift δ_1 is brought into the picture. The value of r_j is discovered as one adds samples $F_{s\ell 1}$, solves (13) and enlarges (16) step by step. This explains why we need to add samples $F_{s\ell 1}$ until s reaches $2 \max_j (h_j - h_{j-1})$ or until for all j the rank r_j is known. How do we proceed to extract the coefficients and exponential parameters from (15) and disentangle the collisions?

For j fixed, r_j of the individual terms

$$\alpha_{h_{j-1}+k} \exp(\langle \phi_{h_{j-1}+k}, \delta_1\rangle), \qquad k = 1, \ldots, h_j - h_{j-1}, 1 \leq r_j \leq h_j - h_{j-1}$$

of $A_{j1}(x)$ can be deduced from the samples $A_j, A_{sj1}, s \geq 1$ of $A_{j1}(x)$ using one of the Prony-like methods [2, 10, 20, 23] which were already mentioned to solve for (5) from (3) or (4) and compute the coefficients from (6). We remark that when $r_j < h_j - h_{j-1}$ then some collisions in $A_{j1}(x)$ still remain indistinguishable in the space spanned by Δ and δ_1.

For the sake of completeness we explicitly give the generalized eigenvalue problems that lead to the identification of $1 \leq r_j \leq h_j - h_{j-1}$ terms in $A_j(x)$:

$$\begin{pmatrix} A_{1j1} & A_{2j1} & \cdots & A_{r_j,j1} \\ A_{2j1} & A_{3j1} & \cdots & A_{r_j+1,j1} \\ \vdots & & \vdots & \\ A_{r_j,j1} & A_{r_j+1,j1} & \cdots & A_{2r_j-1,j1} \end{pmatrix} v = \lambda \begin{pmatrix} A_{0j1} & A_{1j1} & \cdots & A_{r_j-1,j1} \\ A_{1j1} & A_{2j1} & \cdots & A_{r_j,j1} \\ \vdots & & \vdots & \\ A_{r_j-1,j1} & A_{r_j,j1} & \cdots & A_{2r_j-2,j1} \end{pmatrix} v,$$
$$v \in \mathbb{C}^{r_j}.$$

After disentangling at $i = 1$, at least partially, some of the collisions, we can update the number of terms in the exponential model from ν to $\mu \geq \nu$ and reduce the collisions to

$$F_s = \sum_{j=1}^{\mu} \left(\alpha_{g_{j-1}+1} + \ldots + \alpha_{g_j}\right) \exp\left(\langle \phi_{g_j}, s\Delta \rangle\right), \quad g_0 = 0, \quad g_j < g_{j+1}, \quad g_\mu = n.$$

It is clear that the previous indices h_j, $j = 1, \ldots, \nu$ are among the $g_k, k = 1, \ldots, \mu$ but remember that we don't know the values h_j or g_k explicitly. We only know that for some j a collision from index $h_{j-1} + 1$ to h_j, $1 \leq j \leq \nu$ may have split into separate piles indexed by some g_k and g_{k+1}, $1 \leq k \leq \mu$. At this moment in the procedure, namely at the completion of step $i = 1$, we have computed

$$\exp(\Phi_{g_k 1}), \qquad \Phi_{g_k 1} := \langle \phi_{g_k}, \delta_1 \rangle, \qquad k = 1, \ldots, \mu.$$

Because $|\Im\langle \phi_j, \delta_1/||\delta_1|| \rangle| < \pi/||\delta_1||$, $j = 1, \ldots, n$ we in fact obtained all the values

$$\Phi_{g_{k-1}+1,1} = \cdots = \Phi_{g_k,1}, \qquad k = 1, \ldots, \mu, \qquad g_0 = 0, \quad , g_\mu = n$$

which we need later on in combination with the

$$\Phi_{h_{j-1}+1} = \cdots = \Phi_{h_j}, \qquad j = 1, \ldots, \nu \leq \mu$$

to identify the individual ϕ_{ji} as in (9).

We now explain how to move from i to $i + 1$. The first thing is to find proper locations for the samples involving the next identification shift δ_2. Some care needs to be taken with respect to the regularity of the Vandermonde matrices involved. For $i = 2$ we collect

$$\begin{aligned} F_{s\ell 2} &:= f(\kappa_{\ell 2}(\Delta + \delta_1) + s\delta_2), \\ s &= 1, \ldots, 2 \max_{1 \leq j \leq \mu} (g_j - g_{j-1}), \quad \ell = 1, \ldots, \mu, \quad i = 2. \end{aligned} \tag{17}$$

Let us denote

$$\Omega_{g_j 1} := \Phi_{g_j} + \Phi_{g_j 1}, \qquad j = 1, \ldots, \mu.$$

Note that the sum is a direct consequence of the choice $\Delta + \delta_1$ in (17), which is briefly discussed below. Similarly to (13) we write down, for each s separately,

$$\begin{pmatrix} \exp(\kappa_{g_1 2}\Omega_{g_1 1}) & \exp(\kappa_{g_1 2}\Omega_{g_2 1}) & \cdots & \exp(\kappa_{g_1 2}\Omega_{g_\mu 1}) \\ \exp(\kappa_{g_2 2}\Omega_{g_1 1}) & \cdots & & \exp(\kappa_{g_2 2}\Omega_{g_\mu 1}) \\ \vdots & & & \vdots \\ \exp(\kappa_{g_\mu 2}\Omega_{g_1 1}) & \cdots & & \exp(\kappa_{g_\mu 2}\Omega_{g_\mu 1}) \end{pmatrix} \begin{pmatrix} A_{s12} \\ \vdots \\ A_{s\mu 2} \end{pmatrix} = \begin{pmatrix} F_{s12} \\ \vdots \\ F_{s\mu 2} \end{pmatrix} \tag{18}$$

where

$$A_{sj2} = \alpha_{g_{j-1}+1} \exp\left(\langle \phi_{g_{j-1}+1}, s\delta_2 \rangle\right) + \ldots + \alpha_{g_j} \exp\left(\langle \phi_{g_j}, s\delta_2 \rangle\right), \qquad j = 1, \ldots, \mu. \tag{19}$$

From here it is clear how to finalize the $i = 2$ phase and how to proceed to the next value of i. We point out that instead of the linear combination $\Delta + \delta_1$ in (17),

any linear combination $c\Delta + e\delta_1$ with $ce \neq 0$ that guarantees the regularity of the coefficient matrix in (18) can be used (then the definition of Ω_{g_j} also needs to be adapted). This option may be useful as it allows to control the location of the sample points for numeric purposes or so.

To round up this section, we summarize the algorithm that recovers the vectors ϕ_j and coefficients α_j for $j = 1, \ldots, n$ in case of possible collisions of inner products with the chosen directional vectors $\Delta, \delta_i, i = 1, \ldots, d-1$. Before we proceed, we further adapt our notation. Let

$$\begin{aligned} \delta_0 &:= \Delta, \\ \nu_{-1} &:= 0, \\ \nu_0 &:= \nu, \\ \nu_1 &:= \mu \end{aligned}$$

Our first aim is to identify all the inner products $\Phi_{ji} = \langle \phi_j, \delta_i \rangle$, $j = 1, \ldots, n, i = 0, \ldots, d-1$, including possible collisions. This is done by making use of successively collected samples, namely

$$\begin{aligned} F_{s\ell i} &= f\left(\kappa_{\ell i}(\delta_0 + \ldots + \delta_{i-1}) + s\delta_i\right), \\ s &= 0, 1, 2, \ldots \quad \ell = 1, 2, \ldots, \nu_{i-1}, \quad i = 0, \ldots, d-1, \end{aligned}$$

where we assume that empty sums equal zero and values in an empty range need not be specified. The samples are collected by fixing the indices from the right to the left: at identification level i, collision or pile ℓ is being sparsely interpolated using the samples collected at shift multiples s. Here ν_i indicates the number of non-coinciding inner products at identification level i. Remember that s is running up to twice the number of terms in expression $A_{\ell i}(x)$ at level i (for $i = 1$ this is given in (15) and it is straightforward to imagine how it looks like for general i). We remind the reader that only the evaluation at multiples of $\delta_i, i \geq 0$ needs to follow an equidistant scheme. The values $\kappa_{\ell i}$ need not be like that. We also mentioned earlier that the sum $\delta_0 + \ldots + \delta_{i-1}$ can be replaced by another linear combination. The only crucial element is that the $\delta_i, i \geq 0$ are linearly independent. The latter will precisely allow us to identify the vector components ϕ_{ji}, $j = 1, \ldots, n, i = 1, \ldots, d$ from the inner products $\langle \phi_j, \delta_i \rangle$, $j = 1, \ldots, n, i = 0, \ldots, d-1$ as in (9).

5 Detecting the sparsity

The minimal number of $(d+1)n$ samples only delivers the parameters α_j, ϕ_{ji}, $j = 1, \ldots, n, i = 1, \ldots, d$ if the value of n is somehow known a priori and no collision of values $\exp(\Phi_j)$, $j = 1, \ldots, n$ occurs. In the previous section we described how to deal with eventual collisions. Here we detail how to detect the value of n should it not be given. In addition, we analyze how many samples are needed in the worst case when neither n is known nor the projections are collision free.

While collecting the samples $F_s = f(s\Delta)$ and building the Hankel matrices in (3) or (4), the rank of the Hankel matrix reveals (with high probability [11]) the number

ν of terms that do not collide when evaluating in the space spanned by the vector Δ. To this end we need at least $2\nu+1$ values so that we can compose the $(\nu+1)\times(\nu+1)$ Hankel matrix

$$\begin{pmatrix} F_0 & \dots & F_\nu \\ \vdots & & \vdots \\ F_\nu & \dots & F_{2\nu} \end{pmatrix}$$

and conclude that it is singular [1, 7, 11].

From ν and (3) or (4) we proceed to collect the samples $F_{1\ell 1}$ ($s = 1$) and $F_{2\ell 1}$ ($s = 2$), another 2ν in total ($\ell = 1, 2, \dots, \nu$). If all 2×2 Hankel matrices of the form (16) are singular, then every collision remains indistinguishable (unless the zero determinant was an unfortunate coincidence [11]) also in the space spanned by Δ and δ_1. However, if for some j the 2×2 matrix (16) is regular, then we have to proceed to the next values for s ($s = 3, 4$), collect another 2ν values in total, and find out how many terms actually can be revealed in the space spanned by Δ and δ_1. We proceed until we find no larger matrices of the form (16) that are regular. Only after working ourselves through all regular matrices of the form (16) with δ_1 ($i = 1$) we can update ν to $\mu \geq \nu$.

And then we bring the next identification shift vector δ_2 in the picture. We collect the samples $F_{s\ell 2}$ ($s = 1, 2$) as in (17) and compose matrices similar to (16) but now with the last index in the A_{sj1} replaced by $i = 2$ and with A_{sj2} defined as in (19). The inspection of the Hankel matrices containing the values computed for A_{sj2} is identical to the procedure described in the previous paragraph for $i = 1$. If required, as before, we add more samples for larger values of s.

Finally, by the time we reach $i = d - 1$ we can update the number of terms to the true value for n. Now how many samples has this cost us? When n is known a priori and we do not run into collisions or cancelations, which with high probability do not occur, the algorithm presented in Section 3 uses only

$$(d + 1)n$$

samples. Next, we look at the situation where collisions occur and Section 4 is put to work (how to deal with possible cancelations is dealt with at the end of this section). Also the sparsity n is not given. The A_j and Φ_{h_j} with $j = 1, \dots, \nu$ in (10) are retrieved from $O(\nu)$ samples where $\nu \leq n$. In A_j, $1 \leq j \leq \nu$ there are $h_j - h_{j-1}$ terms colliding, where each $h_j - h_{j-1} \leq n - \nu + 1$. To disentangle the terms in A_j we need $O(h_j - h_{j-1})$ samples and so we need at most $O(\nu(n - \nu + 1))$ samples to disentangle all A_j, $j = 1, \dots, \nu$. Note that we have overestimated each $h_j - h_{j-1}$ by $n - \nu + 1$, while if one $h_j - h_{j-1} = n - \nu + 1$, all others equal 1. The procedure is repeated when working with the identifications shifts $\delta_1, \dots, \delta_{d-1}$, leading us to a grand total of

$$O\left((d + 1) \max_{1\leq\nu\leq n} \nu(n - \nu + 1)\right). \tag{20}$$

Remains to discuss the issue of a vanishing A_{sji}. For simplicity, but without loss of generality, we discuss the situation where one of the coefficients A_j given by (12) vanishes, in other words $A_j = A_{j1}(0) = 0$ with $A_{j1}(x)$ given by (15). So besides encountering a collision, the result of the collision is now also zero.

If some $A_j = 0$ then the rank of the matrices in (4) is less than ν and will not reveal the correct value for ν. Of course, the accidental cancelation of a coefficient A_j happens only with very small probability. It suffices either to probe $f(x_1, \ldots, x_d)$ along another (random) choice for the vector Δ [11, 28], or if one absolutely wants to extract the information $\langle \phi_j, \Delta \rangle$ for the originally chosen Δ, to probe $f(x_1, \ldots, x_d)$ along one or more (random) parallel shifts of Δ, as in

$$F_s := f(s\Delta_1 + k\epsilon, \ldots, s\Delta_d + k\epsilon), \qquad s = 0, \ldots, n, \qquad k = 1, 2, \ldots \tag{21}$$

Such a shift affects the coefficient A_j in that it changes from $A_j(0)$ to

$$A_j(k\epsilon) = \alpha_{h_{j-1}+1} \exp\left(\langle \phi_{h_{j-1}+1}, k\epsilon \rangle\right) + \ldots + \alpha_{h_j} \exp\left(\langle \phi_{h_j}, k\epsilon \rangle\right).$$

The rank of the matrices in (4) when filled with the values in (21) either confirms the already computed rank ν or reveals a higher and more probably correct rank ν. The random probing or parallel translation can be added to every step $i = 0, \ldots, d-1$ in the procedure when selecting $\delta_0 = \Delta, \delta_1, \ldots, \delta_{d-1}$ without impacting our data usage analysis in (20).

All the above is now illustrated with an example in which we take the reader through the entire process, first collision-free, then including collision disentanglement.

6 Numerical illustration

We take $d = 2$, write $u := x_1$, $v := x_2$, $x = (u, v)^t$ and consider the exponential sum

$$f(u, v) = \sum_{j=1}^{4} \alpha_j \exp(\langle \phi_j, x \rangle)$$

with

$$\begin{aligned}
\phi_1 &= (-0.5, 1 + \mathtt{i}2\pi \times 0.5), & \alpha_1 &= 1.7 \exp(\mathtt{i}2\pi/10),\\
\phi_2 &= (0.1 + \mathtt{i}2\pi \times 3.4, 1.5 + \mathtt{i}2\pi \times 5.2), & \alpha_2 &= 1.1 \exp(\mathtt{i}2\pi/20),\\
\phi_3 &= (0.1 + \mathtt{i}2\pi \times 3.4, -0.5 + \mathtt{i}2\pi \times 12.6), & \alpha_3 &= 0.9,\\
\phi_4 &= (-2.5 + \mathtt{i}2\pi \times 23.2, -10 + \mathtt{i}2\pi \times 82.3), & \alpha_4 &= 9.2 \exp(\mathtt{i}2\pi/2).
\end{aligned}$$

When outputting numerical results for this small scale example, we round all values to 4 significant digits (all relative errors are less than 5×10^{-4}). The numerical effect of the choice of the vectors Δ and δ_i throughout the process, and that of the underlying one-dimensional Prony-like method in use, is beyond the scope of this paper and will be the subject of further investigations.

First we show the simple case described in the Sections 2 and 3, where the number of terms $n = 4$ is known up front and no collisions of the inner products in the samples occur. Of course, the latter is hard to predict in practice.

We take $\Delta = (0.01, 0.01)$ and $\delta_1 = (-0.01, 0.01)$. Using 8 equidistant evaluations at $x = s\Delta, s = 0, \ldots, 7$, we obtain from (4) the values of $\exp(\Phi_j)$ and can deduce the Φ_j, $j = 1, \ldots, 4$ because $|\Im\phi_{ji}| < \pi/|\Delta_i|$:

$$\begin{aligned}
\Phi_1 &=< \phi_1, \Delta > \approx 0.005000 + 0.03142\mathtt{i},\\
\Phi_2 &=< \phi_2, \Delta > \approx 0.01600 + 0.5404\mathtt{i},\\
\Phi_3 &=< \phi_3, \Delta > \approx -0.004000 + 1.005\mathtt{i},\\
\Phi_4 &=< \phi_4, \Delta > \approx -0.1250 + 0.3456\mathtt{i}.
\end{aligned}$$

We obtain the coefficients α_j, $j = 1, \ldots, 4$ from (6):

$$\begin{aligned}
\alpha_1 &\approx 1.700 \exp(\mathtt{i}2\pi \times 0.1000),\\
\alpha_2 &\approx 1.100 \exp(\mathtt{i}2\pi \times 0.05000),\\
\alpha_3 &\approx 0.9000\\
\alpha_4 &\approx 9.200 \exp(\mathtt{i}2\pi \times 0.5000).
\end{aligned}$$

From 4 additional evaluations along the identification shift δ_1, we obtain the values of $\exp(\Phi_{11})$, $\exp(\Phi_{21})$, $\exp(\Phi_{31})$, $\exp(\Phi_{41})$ from (8). Their exponents are the projections of the vectors ϕ_j along δ_1:

$$\begin{aligned}
\Phi_{11} &=< \phi_1, \delta_1 > \approx 0.01500 + 0.03142\mathtt{i},\\
\Phi_{21} &=< \phi_2, \delta_1 > \approx 0.01400 + 0.1131\mathtt{i},\\
\Phi_{31} &=< \phi_3, \delta_1 > \approx -0.006000 + 0.5781\mathtt{i},\\
\Phi_{41} &=< \phi_4, \delta_1 > \approx -0.07500 + 3.713\mathtt{i}.
\end{aligned}$$

We finally obtain the values of $\phi_j = (\phi_{j1}, \phi_{j2})^t$ by solving for each $j = 1, \ldots, 4$

$$\begin{pmatrix} \Delta_1 & \Delta_2 \\ \delta_{11} & \delta_{12} \end{pmatrix} \begin{pmatrix} \phi_{j1} \\ \phi_{j2} \end{pmatrix} = \begin{pmatrix} \Phi_j \\ \Phi_{j1} \end{pmatrix}.$$

This leads to the following numerical approximations for the ϕ_j:

$$\begin{aligned}
\phi_1 &\approx (-0.5000, 1.000 + \mathtt{i}2\pi \times 0.5000),\\
\phi_2 &\approx (0.1000 + \mathtt{i}2\pi \times 3.400, 1.500 + \mathtt{i}2\pi \times 5.200),\\
\phi_3 &\approx (0.1000 + \mathtt{i}2\pi \times 3.400, -0.5000 + \mathtt{i}2\pi \times 12.60),\\
\phi_4 &\approx (-2.500 + \mathtt{i}2\pi \times 23.20, -10.00 + \mathtt{i}2\pi \times 82.30).
\end{aligned}$$

So far we have used 12 samples in total, which indeed equals $(d + 1)n$. Next we deal with the situation in which neither n is known, nor the assumption of the non-collision holds.

One additional evaluation in the first batch, at $x = 8\Delta$, would ideally (meaning that the numerical rank is easy to detect) and with high probability (meaning that we don't accidentally hit a root of the determinant) have revealed that $n = 4$, still under the assumption that no collisions occur at the inner products. But let us instead move to other directions Δ and δ_1 that get us in trouble because of colliding inner products.

Take $\Delta = (0.03, 0)$ and $\delta_1 = (0, 0.01)$. The projections of ϕ_2 and ϕ_3 along Δ clearly coincide. After 7 evaluations at $x = s\Delta, s = 0, \ldots, 6$ we found that $\nu_0 = 3$

and we obtain from (11) that (without actually knowing the values of the h_j which we list only to help the reader follow the example):

$$\begin{aligned}\Phi_{h_1} &=< \phi_1, \Delta >\approx -0.01500,\\ \Phi_{h_2} &=< \phi_3, \Delta >\approx 0.003000 + 0.6409\mathtt{i},\\ \Phi_{h_3} &=< \phi_4, \Delta >\approx -0.07500 + 4.373\mathtt{i}.\end{aligned}$$

We proceed without knowing n and without knowing whether and where some collisions have occurred. But we know, since $d = 2$, that after adding an independent shift vector δ_1, all terms will have revealed themselves.

So we add evaluations $F_{s\ell 1} = f(\kappa_{\ell 1}\Delta + s\delta_1)$ with $\ell = 1, 2, 3$ and $s = 1, 2, \ldots$ For simplicity we choose $\kappa_{\ell 1} = \ell - 1$. With $\ell = 1$ and $s = 1, 2$ we find that the matrix

$$\begin{pmatrix} A_1 & A_{111} \\ A_{111} & A_{211} \end{pmatrix},$$

where the A_{sji} are computed from (14), has rank 1 and so $h_1 = 1 = g_1$. With $\ell = 2$ and $s = 1, 2, 3, 4$ we find that the matrix

$$\begin{pmatrix} A_2 & A_{121} & A_{221} \\ A_{121} & A_{221} & A_{321} \\ A_{221} & A_{321} & A_{421} \end{pmatrix}$$

has rank 2. This indicates with high probability that there are 2 terms coinciding at Φ_{h_2} (hence $h_2 = 3$ and $g_2 = 2, g_3 = 3$). Remember that in order to obtain A_{s21}, $1 \le s \le 4$, we need to solve (13) which involves the samples F_{sj1}, $1 \le j \le 3$. Hence, continuing the sampling for $\ell = 2$ drags along $\ell = 1, 3$ at the same time. In other words, we are now spending 3×4 samples for $\ell = 1, 2, 3$ rather than only 4 samples for $\ell = 2$.

We now reveal $\langle \phi_2, \delta_1 \rangle$ and $\langle \phi_3, \delta_1 \rangle$ by solving the generalized eigenvalue problem

$$\begin{pmatrix} A_{121} & A_{221} \\ A_{221} & A_{321} \end{pmatrix} v = \lambda \begin{pmatrix} A_2 & A_{121} \\ A_{121} & A_{221} \end{pmatrix} v.$$

With $\ell = 3$ and $s = 1, 2$ we find the same conclusion as with $\ell = 1$, now for

$$\begin{pmatrix} A_3 & A_{131} \\ A_{131} & A_{231} \end{pmatrix},$$

and so $\nu_1 = 4$ with $h_3 = 4$, $g_4 = 4$.

At the expense of a total of $(2 \times 3 + 1) + 3 \times 4 = 19$ evaluations, we find that $n = 4$ and we can identify all ϕ_{ji} and α_j for $j = 1, \ldots, 4$ and $i = 1, 2$.

7 Conclusion

In 1795 the French scientist G. de Prony showed that a univariate linear combination of n exponential terms with unknown real but mutually distinct exponents could be fitted uniquely to $2n$ data samples. His result solves the $d = 1$ case of this paper.

The current paper is the first of its kind where this result is proven to hold for general $d > 1$: a multivariate linear combination of n exponential terms with unknown inner product exponents can, under mild conditions, be fitted using only $(d + 1)n$ data.

Acknowledgements This work was partially supported by a Research Grant of the FWO-Flanders (Flemish Science Foundation).

References

1. Baker, G., Graves-Morris, P.: Padé Approximants Part I: Basic Theory Encyclopedia of Mathematics and its Applications, vol. 13. Addison Wesley, London (1981)
2. 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 (1988)
3. Cuyt, A., Lee, W.-s.: Smart data sampling and data reconstruction. Patent US 61/611,899
4. Cuyt, A., Lee, W.-s.: Smart data sampling and data reconstruction. Patent PCT/EP2012/066204
5. Cuyt, A., Lee, W.-s.: Sparse interpolation and rational approximation. Contemp. Math. **661**, 229–242 (2016). American Mathematical Society
6. Diederichs, B., Iske, A.: Parameter estimation for bivariate exponential sums. In: IEEE International Conference Sampling Theory and Applications (SampTA2015), pp. 493–497 (2015)
7. Henrici, P.: Applied and Computational Complex Analysis I. Wiley, New York (1974)
8. Hildebrand, F.: Introduction to Numerical Analysis. McGraw-Hill, New York (1956)
9. Hua, Y.: Estimating two-dimensional frequencies by matrix enhancement and matrix pencil. IEEE Trans. Signal Process. **40**(9), 2267–2280 (1992)
10. 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**(5), 814–824 (1990)
11. Kaltofen, E., Lee, W.-s.: Early termination in sparse interpolation algorithms. J. Symb. Comput. **36**(3-4), 365–400 (2003)
12. Kunis, S., Peter, T., Römer, T., von der Ohe, U.: A multivariate generalization of Prony's method. Linear Algebra Appl. **490**, 31–47 (2016)
13. Li, Y., Razavilar, J., Liu, K.J.R.: A high-resolution technique for multidimensional NMR spectroscopy. IEEE Trans. Biomed. Eng. **45**(1), 78–86 (1998)
14. Mandelshtam, V.A.: The multidimensional filter diagonalization method: I. Theory and numerical implementation. J. Magn. Reson. **144**(2), 343–356 (2000)
15. Nyquist, H.: Certain topics in telegraph transmission theory. Trans. Am. Inst. Electr. Eng. **47**(2), 617–644 (1928)
16. Peter, T., Plonka, G., Schaback, R.: Reconstruction of multivariate signals via Prony's method. Proc. Appl. Math. Mech. to appear
17. Plonka, G., Wischerhoff, M.: How many Fourier samples are needed for real function reconstruction? J. Appl. Math. Comput. **42**(1-2), 117–137 (2013)
18. Potts, D., Tasche, M.: Parameter estimation for multivariate exponential sums. Electron. Trans. Numer. Anal. **40**, 204–224 (2013)
19. Rouquette, S., Najim, M.: Estimation of frequencies and damping factors by two-dimensional ESPRIT type methods. IEEE Trans. Signal Process. **49**(1), 237–245 (2001)
20. Roy, R., Kailath, T.: ESPRIT-Estimation of signal parameters via rotational invariance techniques. IEEE Trans. Acoust. Speech Signal Process. **37**(7), 984–995 (1989)
21. Sauer, T.: Prony's method in several variables: symbolic solutions by universal interpolation. J. Symb. Comput. **84**, 95–112 (2018)
22. Sauer, T.: Prony's method in several variables. Numer. Math. **136**(2), 411–438 (2017)
23. Schmidt, R.: Multiple emitter location and signal parameter estimation. IEEE Trans. Antennas Propag. **34**(3), 276–280 (1986)
24. Shannon, C.E.: Communication in the presence of noise. Proc. Inst. Radio Eng. **37**(1), 10–21 (1949)
25. Shukla, P., Dragotti, P.L.: Sampling schemes for multidimensional signals with finite rate of innovation. IEEE Trans. Signal Process. **55**(7), 3670–3686 (2007)

26. Wischerhoff, M., Plonka, G.: Reconstruction of polygonal shapes from sparse Fourier samples. J. Comput. Appl. Math. **297**, 117–131 (2016)
27. Yilmazer, N., Fernandez-Recio, R., Sarkar, T.K.: Matrix pencil method for simultaneously estimating azimuth and elevation angles of arrival along with the frequency of the incoming signals. Digital Signal Process. **16**(6), 796–816 (2006)
28. Zippel, R.: Probabilistic Algorithms for Sparse Polynomials. Ph.D. Thesis, Massachusetts Institute of Technology (1979)