Structured low rank decomposition of multivariate Hankel matrices

Jouhayna Harmouch, Houssam Khalil, Bernard Mourrain

Structured low rank decomposition of multivariate Hankel matrices

J. Harmouch[a,b], H. Khalil[b], B. Mourrain[a]

[a] UCA, Inria, Aromath, Sophia Antipolis, France
[b] Laboratory of mathematics and its applications LaMa-Lebanon, Lebanese University

Abstract

We study the decomposition of a multivariate Hankel matrix H_σ as a sum of Hankel matrices of small rank in correlation with the decomposition of its symbol σ as a sum of polynomial-exponential series. We present a new algorithm to compute the low rank decomposition of the Hankel operator and the decomposition of its symbol exploiting the properties of the associated Artinian Gorenstein quotient algebra $\mathcal{A}_\sigma$. A basis of $\mathcal{A}_\sigma$ is computed from the Singular Value Decomposition of a sub-matrix of the Hankel matrix H_σ. The frequencies and the weights are deduced from the generalized eigenvectors of pencils of shifted sub-matrices of H_σ. Explicit formula for the weights in terms of the eigenvectors avoid us to solve a Vandermonde system. This new method is a multivariate generalization of the so-called Pencil method for solving Prony-type decomposition problems. We analyse its numerical behaviour in the presence of noisy input moments, and describe a rescaling technique which improves the numerical quality of the reconstruction for frequencies of high amplitudes. We also present a new Newton iteration, which converges locally to the closest multivariate Hankel matrix of low rank and show its impact for correcting errors on input moments.

1. Introduction

Structured matrices such as Toeplitz or Hankel matrices appear in many problems. They are naturally associated to operations on polynomials or series [Fuh12]. The correlation with polynomial algebra can be exploited to accelerate matrix computations [BP94]. The associated algebraic model provides useful information on the problem to be solved or the phenomena to be analysed. Understanding its structure often yields a better insight on the problem and its solution. In many cases, an efficient way to analyze the structure of the underlying models is to decompose the structured matrix into a sum of low rank matrices of the same structure. This low rank decomposition has applications in many domains [Mar12] and appears under different formulations [Lan11, BCMT10, BBCM13].

In this paper, we study specifically the class of Hankel matrices. We investigate the problem of decomposing a Hankel matrix as a sum of indecomposable Hankel matrices of low rank. Natural questions arise. What are the indecomposable Hankel matrices? Are they necessarily of rank 1 ? How to compute a decomposition of a Hankel matrix as a sum of indecomposable Hankel matrices? Is the structured low rank decomposition of a Hankel matrix unique ?

These questions have simple answers for non-structured or dense matrices: The indecomposable dense matrices are the matrices of rank one, which are the tensor product of two vectors. The Singular Value Decomposition of a dense matrix yields a decomposition as a minimal sum of rank one matrices, but this decomposition is not unique.

It turns out that for the Hankel structure, the answers to these questions are not so direct and involve the analysis of the so-called symbol associated to the Hankel matrix. The symbol is a formal power series defined from the coefficients of the Hankel matrix. As we will see, the structured decomposition of an Hankel matrix is closely related to the decomposition of the symbol as a sum of polynomial-exponential series.

The decomposition of the symbol of a Hankel matrix is a problem, which has a long history. The first work on this problem is probably due to Gaspard-Clair-François-Marie Riche de Prony [Bar95]. He proposed a method to reconstruct a sum of exponentials from the values at equally spaced data points, by computing a polynomial in the kernel of a Hankel matrix, and deducing the decomposition from the roots of this polynomial. Since then, many works have been developed to address the decomposition problem in the univariate case, using linear algebra tools on Hankel matrices such as Pencil method [HS90], ESPRIT method [RK89] or MUSIC method [SK92]. Other methods such as [GP03], approximate Prony Method [BM05], [PT11] use minimization techniques, to recover the frequencies or the weights in the sum of exponential functions. See [PS12][chap. 1] for a survey on some of these approaches.

The numerical behavior of these methods has also been studied. The condition number of univariate Hankel matrices, which decomposition involves real points has been investigated in [Tyr94], [Bec97]. It is shown that it grows exponentially with the dimension of the matrix. The condition number of Vandermonde matrices of general complex points has been studied recently in [Pan16]. In [BGL07], the numerical sensitivity of the generalized eigenvalues of pencils of Hankel matrices appearing in Prony's method has been analysed.

The development of multivariate decomposition methods is more recent. Extension of the univariate approaches have been considered e.g. in [ACd10], [PT13], [PPS15]. These methods project the problem in one dimension and solve several univariate decomposition problems to recover the multivariate decomposition by least square minimization from a grid of frequencies. In [PPS15], [KPRv16], the decomposition problem is transformed into the solution of an overdetermined polynomial system associated to the kernel of these Hankel matrices, and the frequencies of the exponential terms are found by general polynomial solver. These methods involves Hankel matrices of size exponential in the number of variables of the problem or moments of order at least twice the number of terms in the decomposition. In [Sau16], an H-basis of the ideal defining the frequencies is computed from Hankel matrices built from moments of big enough order. Tables of multiplications are deduced from the H-basis and their eigenvalues yield the frequencies of the exponential terms. The weights are computed as the solution of a Vandermonde linear system. Moments of order bigger than twice the degree of an H-basis are involved in this construction.

Contributions. We study the decomposition of a multivariate Hankel matrix as a sum of Hankel matrices of small rank in correlation with the decomposition of its symbol σ as a sum of polynomial-exponential series. We show how to recover efficiently this decomposition from the structure of the quotient algebra $\mathcal{A}_\sigma$ of polynomials modulo the kernel of the corresponding Hankel operator H_σ. In particular, a basis of $\mathcal{A}_\sigma$ can be extracted from any maximal non-zero minor of the matrix of H_σ. We also show how to compute the matrix of multiplication by a variable in the basis of $\mathcal{A}_\sigma$ from sub-matrices of the matrix of H_σ. We describe how the frequencies of the polynomial-exponential decomposition of the symbol can be deduced from the eigenvectors of these matrices. Exploiting properties of these multiplication operators, we show that the weights of the decomposition can be recovered directly from these eigenvectors, avoiding the solution of a Vandermonde system. We present a new algorithm to compute the low rank decomposition of H_σ and the decomposition of its symbol as a sum of polynomial-exponential series from sub-matrices of the matrix of H_σ. A basis of $\mathcal{A}_\sigma$ is computed from the Singular Value Decomposition of a sub-matrix. The frequencies and the weights are deduced from the generalized eigenvectors of pencils of sub-matrices of H_σ. This new method is a multivariate generalization of the so-called Pencil method for solving Prony-type decomposition problems. It can be used to decompose series as sums of polynomial-exponential functions from moments and provides structured low rank decomposition of multivariate Hankel matrices. We analyse its numerical behaviour in the presence of noisy input moments, for different numbers of variables, of exponential terms of the symbol and different amplitudes of the frequencies. We present a rescaling technique, which improves the numerical quality of the reconstruction for frequencies of high amplitudes. We also present a new Newton iteration, which converges locally to the multivariate Hankel matrix of a given rank the closest to a given input Hankel matrix. Numerical experimentations show that the Newton iteration combined the decomposition method allows to compute accurately and efficiently the polynomial-exponential decomposition of the symbol, even for noisy input moments.

Structure of the paper. The next section describes multivariate Hankel operators, their symbol and the generalization of Kronecker theorem, which establishes a correspondence between Hankel operators of finite rank and polynomial-exponential series. In Section 3, we recall techniques exploiting the properties of multiplication operators for solving polynomial systems and show how they can be used for the Artinian Gorenstein algebra associated to Hankel operators of finite rank. In Section 4, we describe in details the decomposition algorithm. Finally in section 5, we present numerical experimentations, showing the numerical behaviour of the decomposition method for noisy input moments and the improvements obtained by rescaling and by an iterative projection method.

2. Hankel matrices and operators

Hankel matrices are structured matrices of the form

$$H = [\sigma_{i+j}]_{0 \le i \le l, 0 \le j \le m}$$

where the entry σ_{i+j} of the ith row and the jth columns depends only on the sum $i + j$. By reversing the order of the columns or the rows, we obtain *Toeplitz matrices*, which entries depend on the difference of the row and column indices. Exploiting their structure leads to superfast methods for many linear algebra operations such as matrix-vector product, solution of linear systems, ... (see e.g. [BP94]).

A Hankel matrix is a sub-matrix of the matrix of the *Hankel operator* associated to a sequence $\sigma = (\sigma_k) \in \mathbb{C}^{\mathbb{N}}$:

$$
\begin{aligned}
H_\sigma : L_0(\mathbb{C}^{\mathbb{N}}) &\rightarrow \mathbb{C}^{\mathbb{N}} \\
(p_k)_k &\mapsto \left(\sum_k p_k \sigma_{k+l}\right)_{l \in \mathbb{N}}
\end{aligned}
$$

where $L_0(\mathbb{C}^{\mathbb{N}})$ is the set of sequences of $\mathbb{C}^{\mathbb{N}}$ with a finite support.

Multivariate Hankel matrices have a similar structure of the form

$$H = [\sigma_{\alpha+\beta}]_{\alpha \in A, \beta \in B}$$

where $A, B \subset \mathbb{N}^n$ are subsets of multi-indices $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ $\beta = (\beta_1, \ldots, \beta_n) \in \mathbb{N}^n$ indexing respectively the rows and columns. *Multivariate Hankel operators* are associated to multi-index sequences $\sigma = (\sigma_\alpha)_{\alpha \in \mathbb{N}^n} \in \mathbb{C}^{\mathbb{N}^n}$:

$$
\begin{aligned}
H_\sigma : L_0(\mathbb{C}^{\mathbb{N}^n}) &\rightarrow \mathbb{C}^{\mathbb{N}^n} \\
(p_\alpha)_\alpha &\mapsto \left(\sum_\alpha p_\alpha \sigma_{\alpha+\beta}\right)_{\beta \in \mathbb{N}^n}
\end{aligned}
\tag{1}
$$

where $L_0(\mathbb{C}^{\mathbb{N}^n})$ is the set of sequences of $\mathbb{C}^{\mathbb{N}^n}$ with a finite support. In order to describe the algebraic properties of Hankel operators, we will identify hereafter the space $L_0(\mathbb{C}^{\mathbb{N}^n})$ with the ring $\mathbb{C}[\mathbf{x}] = \mathbb{C}[x_1, \ldots, x_n]$ of polynomials in the variables $\mathbf{x} = (x_1, \ldots, x_n)$ with coefficients in $\mathbb{C}$.

The set of multi-index sequences $\mathbb{C}^{\mathbb{N}^n}$ can be identified with the ring of formal power series $\mathbb{C}[[y_1, \ldots, y_n]] = \mathbb{C}[[\mathbf{y}]]$. A sequence $\sigma = (\sigma_\alpha)_\alpha$ is identified with the series

$$\sigma(\mathbf{y}) = \sum_{\alpha \in \mathbb{N}^n} \sigma_\alpha \frac{\mathbf{y}^\alpha}{\alpha!} \in \mathbb{C}[[\mathbf{y}]]$$

where $\mathbf{y}^\alpha = y_1^{\alpha_1} \cdots y_n^{\alpha_n}$, $\alpha! = \prod_{i=1}^n \alpha_i!$ for $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$. It can also be interpreted as a linear functional on polynomials as follows:

$$
\begin{aligned}
\sigma : \mathbb{C}[\mathbf{x}] &\rightarrow \mathbb{C} \\
p = \sum_{\alpha \in A \subset \mathbb{N}^n} p_\alpha \mathbf{x}^\alpha &\mapsto \langle \sigma \mid p \rangle = \sum_{\alpha \in A \subset \mathbb{N}^n} p_\alpha \sigma_\alpha.
\end{aligned}
$$

The identification of σ with an element of $\mathbb{C}[\mathbf{x}]^* = \operatorname{Hom}_{\mathbb{C}}(\mathbb{C}[\mathbf{x}], \mathbb{C})$ is uniquely defined by its coefficients $\sigma_\alpha = \langle \sigma \mid \mathbf{x}^\alpha \rangle$ for $\alpha \in \mathbb{N}^n$, which are called the *moments* of σ. This allows us to identify the dual $\mathbb{C}[\mathbf{x}]^*$ with $\mathbb{C}[[\mathbf{y}]]$ or with the set of multi-index sequences $\mathbb{C}^{\mathbb{N}^n}$.

The dual space $\mathbb{C}[\mathbf{x}]^* \equiv \mathbb{C}[[\mathbf{y}]]$ has a natural structure of $\mathbb{C}[\mathbf{x}]$-module, defined as follows:

$$\forall \sigma \in \mathbb{C}[[\mathbf{y}]], \forall p, q \in \mathbb{C}[\mathbf{x}], \ \langle p \star \sigma \mid q \rangle = \langle \sigma \mid pq \rangle.$$

For a polynomial $p = \sum_{\alpha \in \mathbb{N}^n} p_\alpha \mathbf{x}^\alpha$ with $p_\alpha = 0$ for almost all $\alpha \in \mathbb{N}^n$, we have

$$p \star \sigma = \sum_{\beta \in \mathbb{N}^n} \Big(\sum_{\alpha \in \mathbb{N}^n} p_\alpha \sigma_{\alpha+\beta} \Big) \frac{\mathbf{y}^\alpha}{\alpha!}.$$

We check that $p \star \sigma = p(\partial_1, \ldots, \partial_n)(\sigma) = p(\boldsymbol{\partial})(\sigma)$ where $\boldsymbol{\partial} = (\partial_1, \ldots, \partial_n)$ and ∂_i is the derivation with respect to the variable y_i.

Identifying $L_0(\mathbb{C}^{\mathbb{N}^n})$ with $\mathbb{C}[\mathbf{x}]$, the Hankel operator (1) is nothing else than the operator of multiplication by σ:

$$H_\sigma : \mathbb{C}[\mathbf{x}] \ \rightarrow \ \mathbb{C}[[\mathbf{y}]]$$
$$p \ \mapsto \ p \star \sigma$$

Truncated Hankel operators are obtained by restriction of Hankel operators. For $A, B \subset \mathbb{N}^n$, let $\langle \mathbf{x}^B \rangle \subset \mathbb{C}[\mathbf{x}]$, $\langle \mathbf{y}^A \rangle \subset \mathbb{C}[[\mathbf{y}]]$ be the vector spaces spanned respectively by the monomials $\mathbf{x}^\beta$ for $\beta \in B$ and $\mathbf{y}^\alpha$ for $\alpha \in A$. The truncated Hankel operator of σ on A, B is

$$H_\sigma^{A,B} : \langle \mathbf{x}^B \rangle \ \rightarrow \ \langle \mathbf{y}^A \rangle$$
$$p = \textstyle\sum_{\beta \in B} p_\beta \mathbf{x}^\beta \ \mapsto \ \textstyle\sum_{\alpha \in A} \big(\sum_{\beta \in B} p_\alpha \sigma_{\alpha+\beta} \big) \frac{\mathbf{y}^\alpha}{\alpha!} = p \star \sigma_{|\langle \mathbf{x}^A \rangle}$$

The matrix of $H_\sigma^{A,B}$ in the bases $(\mathbf{x}^\beta)_{\beta \in B}$ and $(\frac{\mathbf{y}^\alpha}{\alpha!})_{\alpha \in A}$ is of the form:

$$H_\sigma^{A,B} = [\sigma_{\alpha+\beta}]_{\alpha \in A, \beta \in B}.$$

It is also called the *moment matrix* of σ. Multivariate Hankel matrices have a structure, which can be exploited to accelerate linear algebra operations (see e.g. [MP00] for more details).

Example. Consider the series $\sigma = 1 + 2y_1 + 3y_2 + 4\frac{y_1^2}{2} + 5y_1 y_2 + 6\frac{y_2^2}{2} + 7\frac{y_1^3}{6} + 8\frac{y_1^2 y_2}{2} + \cdots \in \mathbb{C}[[y_1, y_2]]$. Its truncated Hankel matrix on $A = [(0,0), (1,0), (0,1)]$ (corresponding to the monomials $1, x_1, x_2$), $B = [(0,0), (1,0), (0,1), (2,0)]$ (corresponding to the monomials $1, x_1, x_2, x_1^2$) is

$$H_\sigma^{A,B} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 5 & 7 \\ 3 & 5 & 6 & 8 \end{bmatrix}$$

For $d \in \mathbb{N}$, we denote by $\mathbb{C}[\mathbf{x}]_d$ the vector space of polynomials of degree $\leq d$. Its dimension is $s_d = \binom{n+d}{n}$. For $d, d' \in \mathbb{N}$, we denote by $H_\sigma^{d,d'}$ the Hankel matrix of σ on the subset of monomials in $\mathbf{x}$ respectively of degree $\leq d$ and $\leq d'$. We also denote by $H_\sigma^{d,d'}$ the corresponding truncated Hankel operator of H_σ from $\mathbb{C}[\mathbf{x}]_{d'}$ to $(\mathbb{C}[\mathbf{x}]_d)^*$. More generally, for $U = \{u_1, \ldots, u_l\} \subset \mathbb{C}[\mathbf{x}]$, $V = \{v_1, \ldots, v_m\} \subset \mathbb{C}[\mathbf{x}]$, the Hankel matrix of σ on U, V is $H_\sigma^{U,V} = (\langle \sigma \mid u_i v_j \rangle)_{1 \leq i \leq l, 1 \leq j \leq m}$. We use the same notation $H_\sigma^{U,V}$ for the truncated Hankel operator from $\langle V \rangle$ to $\langle U \rangle^*$.

2.1. Hankel operator of finite rank

We are interested in structured decompositions of Hankel matrices (resp. operators) as sums of Hankel matrices (resp. operators) of low rank. This raises the question of describing the Hankel operators of finite rank and leads to the problem of decomposing them into indecomposable Hankel operators of low rank.

A first answer is given by the celebrated theorem of Kronecker [Kro81].

4

Theorem 2.1 (Kronecker Theorem). *The Hankel operator*

$$H_\sigma : (p_k) \in L_0(\mathbb{C}^{\mathbb{N}}) \mapsto \left(\sum_k p_k \sigma_{k+l}\right)_{l \in \mathbb{N}} \in \mathbb{C}^{\mathbb{N}}$$

is of finite rank r, if and only if, there exist polynomials $\omega_1, \ldots, \omega_{r'} \in \mathbb{C}[y]$ and $\xi_1, \ldots, \xi_{r'} \in \mathbb{C}$ distinct s.t.

$$\sigma_n = \sum_{i=1}^{r'} \omega_i(n) \xi_i^n$$

with $\sum_{i=1}^{r'} (\deg(\omega_i) + 1) = r$.

This results says that the Hankel operator H_σ is of finite rank, if and only if, its symbol σ is of the form

$$\sigma(y) = \sum_{n \in \mathbb{N}} \sigma_n \frac{y^n}{n!} = \sum_{i=1}^{r'} \tilde{\omega}_i(y) e^{\xi_i y}$$

for some univariate polynomials $\tilde{\omega}_i(y) \in \mathbb{C}[y]$ and distinct complex numbers ξ_i $i = 1, \ldots, r'$. Moreover, the rank of H_σ is $r = \sum_{i=1}^{r'} (\deg(\tilde{\omega}_i) + 1)$.

The previous result admits a direct generalization to multivariate Hankel operators, using polynomial-exponential series.

Definition 2.2. *For $\xi = (\xi_1, \ldots, \xi_n) \in \mathbb{C}^n$, we denote $\mathbf{e}_\xi(\mathbf{y}) = e^{y_1 \xi_1 + \cdots + y_n \xi_n} = \sum_{\alpha \in \mathbb{N}^n} \xi^\alpha \frac{\mathbf{y}^\alpha}{\alpha!} \in \mathbb{C}[[\mathbf{y}]]$ where $\xi^\alpha = \xi_1^{\alpha_1} \cdots \xi_n^{\alpha_n}$ for $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$.*

Let $\mathcal{POLYEXP}(\mathbf{y}) = \{\sigma = \sum_{i=1}^r \omega_i(\mathbf{y}) \mathbf{e}_{\xi_i}(\mathbf{y}) \in \mathbb{C}[[\mathbf{y}]] \mid \xi_i \in \mathbb{C}^n, \omega_i(\mathbf{y}) \in \mathbb{C}[\mathbf{y}]\}$ be the set of polynomial-exponential series. The polynomials $\omega_i(\mathbf{y})$ are called the weights of σ and ξ_i the frequencies.

For $\omega(\mathbf{y}) \in \mathbb{C}[\mathbf{y}]$, we denote by $\mu(\omega)$ the dimension of the vector space spanned by ω and its derivatives $\partial^\alpha \omega(\mathbf{y}) = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n} \omega(\mathbf{y})$ of any order for $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$.

The next theorem characterizes the multivariate Hankel operators of finite rank in terms of their symbol [Mou16]:

Theorem 2.3 (Generalized Kronecker Theorem). *Let $\sigma(\mathbf{y}) \in \mathbb{C}[[\mathbf{y}]]$. Then $\operatorname{rank} H_\sigma = r < \infty$, if and only if, $\sigma(\mathbf{y}) = \sum_{i=1}^{r'} \omega_i(\mathbf{y}) \mathbf{e}_{\xi_i}(\mathbf{y}) \in \mathcal{POLYEXP}(\mathbf{y})$ with $\omega_i(\mathbf{y}) \in \mathbb{C}[\mathbf{y}] \smallsetminus \{0\}$ and $\xi_i \in \mathbb{C}^n$ pairwise distinct, with $r = \sum_{i=1}^{r'} \mu(\omega_i)$ where $\mu(w_i)$ is the dimension of the inverse system spanned by $\omega_i(\mathbf{y})$ and all its derivatives $\partial^\alpha \omega_i(\mathbf{y})$ for $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$.*

Example 2.4. For $\xi \in \mathbb{C}^n$, the series $\mathbf{e}_\xi(\mathbf{y}) = \sum_{\alpha \in \mathbb{N}^n} \xi^\alpha \frac{\mathbf{y}^\alpha}{\alpha!} = e^{\mathbf{y} \cdot \xi}$ represents the linear functional corresponding to the evaluation at ξ:

$$\forall p \in R, \langle \mathbf{e}_\xi | p \rangle = \sum_{\alpha \in \mathbb{N}^n} p_\alpha \xi^\alpha = p(\xi).$$

The Hankel operator $H_{\mathbf{e}_\xi} : p \mapsto p \star \mathbf{e}_\xi = p(\xi) \mathbf{e}_\xi$ is of rank 1, since its image is spanned by $\mathbf{e}_\xi$. For $A, B \subset \mathbb{N}^n$, the Hankel matrix of $\mathbf{e}_\xi$ is $H_\xi^{A,B} = [\xi^{\beta + \alpha}]_{\beta \in B, \alpha \in A}$. If $H_\xi^{A,B} \neq 0$, it is a matrix of rank 1.

Hankel operators associated to evaluations $\mathbf{e}_\xi$ are of rank 1. As shown in the next example, a Hankel operator of rank > 1 is not necessarily the sum of such Hankel operators of rank 1.

Example 2.5. For $n = 1$ and $\sigma = y$, we check that H_y is of rank 2, but it cannot be decomposed as a sum of two rank-one Hankel operators. If $A = \{1, x, x^2\}$, we have

$$H_y^{A,A} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \neq \lambda_1 \begin{bmatrix} 1 & \xi_1 & \xi_1^2 \\ \xi_1 & \xi_1^2 & \xi_1^3 \\ \xi_1^2 & \xi_1^3 & \xi_1^4 \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 & \xi_2 & \xi_2^2 \\ \xi_2 & \xi_2^2 & \xi_2^3 \\ \xi_2^2 & \xi_2^3 & \xi_2^4 \end{bmatrix}$$

for $\lambda_1, \lambda_2, \xi_1, \xi_2 \in \mathbb{C}$. This shows that the symbol y is indecomposable as a sum of polynomial-exponential series, though it defines an Hankel operator of rank 2.

Definition 2.6. *For $\sigma \in \mathbb{C}[[\mathbf{y}]]$, we say that σ is indecomposable if σ cannot be written as a sum $\sigma = \sigma_1 + \sigma_2$ with $\operatorname{img} H_\sigma = \operatorname{img} H_{\sigma_1} \oplus \operatorname{img} H_{\sigma_2}$.*

Proposition 2.7. *The series $\omega(\mathbf{y})\, e_\xi(\mathbf{y})$ with $\omega(\mathbf{y}) \in \mathbb{C}[\mathbf{y}] \smallsetminus \{0\}$ and $\xi \in \mathbb{C}^n$ is indecomposable.*

Proof. Let $\sigma = \omega\, e_\xi$ and $r = \mu(\omega)$ the rank of H_σ. Suppose that $\sigma = \sigma_1 + \sigma_2$ with $\operatorname{img} H_\sigma = \operatorname{img} H_{\sigma_1} \oplus \operatorname{img} H_{\sigma_2}$. We assume that the rank of H_{σ_1} is minimal. By the Generalized Kronecker Theorem, $\sigma_1 = \sum_{i=1}^{r_1} \omega_{1,i}\, e_{\xi_{1,i}}$, $\sigma_2 = \sum_{i=1}^{r_2} \omega_{2,i}\, e_{\xi_{2,i}}$ with $\omega_{l,i} \in \mathbb{C}[\mathbf{x}]$, $\xi_{l,i} \in \mathbb{C}^n$ and

$$\omega\, e_\xi = \sum_{i=1}^{r_1} \omega_{1,i}\, e_{\xi_{1,i}} + \sum_{i=1}^{r_2} \omega_{2,i}\, e_{\xi_{2,i}}.$$

By the independence of the polynomial-exponential series [Mou16][Lem. 2.7], we can assume that $\xi_{1,1} = \xi_{2,1} = \xi$ and that $\omega = \omega_{1,1} + \omega_{2,1}$ (possibly with $\omega_{2,1} = 0$) and that $\omega_{1,i} = -\omega_{2,i}$ for $i = 2, \ldots, r_1 = r_2$. As $\operatorname{rank} H_{\sigma_1} = \sum_{i=1}^{r_1} \mu(\omega_{1,i})$ is minimal, we can assume moreover that $\omega_{1,i} = 0$ for $i = 2, \ldots, r_1$, that is, $r_1 = r_2 = 1$. Then, we have $\sigma = \omega\, e_\xi$, $\sigma_1 = \omega_1\, e_\xi$ $\sigma_2 = \omega_2\, e_\xi$ with $\omega = \omega_1 + \omega_2$. As $\operatorname{img} H_{\sigma_i} = \langle \partial^\alpha(\omega_i)\, e_\xi \rangle$, $i = 1, 2$, we have $\operatorname{img} H_{\sigma_1} \cap \operatorname{img} H_{\sigma_2} \ni e_\xi$ and $\operatorname{img} H_\sigma$ is not the direct sum of $\operatorname{img} H_{\sigma_1}$ and $\operatorname{img} H_{\sigma_2}$. This shows that σ is indecomposable. $\qquad\square$

The goal of this paper is to present a method to decompose the symbol of a Hankel operator as a sum of indecomposable polynomial-exponential series from truncated Hankel matrices.

3. Structured decomposition of Hankel matrices

In this section, we show how the decomposition of the symbol σ of a Hankel operator H_σ as a sum of polynomial-exponential series reduces to the solution of polynomial equations. This corresponds to the decomposition of H_σ as a sum of Hankel matrices of low rank. We first recall classical techniques for solving polynomial systems and show how these methods can be applied on the Hankel matrix H_σ, to compute the decomposition.

3.1. Solving polynomial equations by eigenvector computation

A quotient algebra $\mathcal{A} = \mathbb{C}[\mathbf{x}]/I$ is *Artinian* if it is of finite dimension over $\mathbb{C}$. In this case, the ideal I defines a finite number of roots $\mathcal{V}(I) = \{\xi_1, \ldots, \xi_{r'}\} = \{\xi \in \mathbb{C}^n \mid \forall p \in I, p(\xi) = 0\}$ and we have a decomposition of $\mathcal{A}$ as a sum of sub-algebras:

$$\mathcal{A} = \mathbb{C}[\mathbf{x}]/I = \mathcal{A}_1 \oplus \cdots \oplus \mathcal{A}_{r'}$$

where $\mathcal{A}_i = \mathbf{u}_{\xi_i} \mathcal{A} \sim \mathbb{C}[\mathbf{x}]/Q_i$ and Q_i is the primary component of I associated to the root $\xi_i \in \mathbb{C}^n$. The elements $\mathbf{u}_1, \ldots, \mathbf{u}_{r'}$ satisfy the relations

$$\mathbf{u}_{\xi_i}^2(\mathbf{x}) \equiv \mathbf{u}_{\xi_i}(\mathbf{x}), \ \sum_{i=1}^{r} \mathbf{u}_{\xi_i}(\mathbf{x}) \equiv 1.$$

The polynomials $\mathbf{u}_{\xi_1}, \ldots, \mathbf{u}_{\xi_{r'}}$ are called *idempotents* of $\mathcal{A}$. The dimension of $\mathcal{A}_i$ is the *multiplicity* of the point ξ_i. For more details, see [EM07][Chap. 4].

For $g \in \mathbb{C}[\mathbf{x}]$, the multiplication operator $\mathcal{M}_g$ is defined by

$$\begin{array}{rccc} \mathcal{M}_g: & \mathcal{A} & \to & \mathcal{A} \\ & h & \mapsto & \mathcal{M}_g(h) = g\,h. \end{array}$$

The transpose $\mathcal{M}_g^\mathsf{T}$ of the multiplication operator $\mathcal{M}_g$ is

$$\begin{array}{rccc} \mathcal{M}_g^\mathsf{T}: & \mathcal{A}^* & \to & \mathcal{A}^* \\ & \Lambda & \mapsto & \mathcal{M}_g^\mathsf{T}(\Lambda) = \Lambda \circ \mathcal{M}_g = g \star \Lambda. \end{array}$$

The main property that we will use to recover the roots is the following [EM07][Thm. 4.23]:

Proposition 3.1. *Let I be an ideal of $\mathbb{C}[\mathbf{x}]$ and suppose that $\mathcal{V}(I) = \{\xi_1, \xi_2, \ldots, \xi_{r'}\}$. Then*

- *for all $g \in \mathcal{A}$, the eigenvalues of $\mathcal{M}_g$ and $\mathcal{M}_g^\mathsf{T}$ are the values $g(\xi_1), \ldots, g(\xi_{r'})$ of the polynomial g at the roots with multiplicities $\mu_i = \dim \mathcal{A}_i$.*

- *The eigenvectors common to all $\mathcal{M}_g^\mathsf{T}$ with $g \in \mathcal{A}$ are - up to a scalar - the evaluations $\mathbf{e}_{\xi_1}, \ldots, \mathbf{e}_{\xi_{r'}}$.*

If $B = \{b_1, \ldots, b_r\}$ is a basis of $\mathcal{A}$, then the coefficient vector of the evaluation $\mathbf{e}_{\xi_i}$ in the dual basis of B is $[\langle \mathbf{e}_{\xi_i} | b_j \rangle]_{\beta \in B} = [b_j(\xi_i)]_{i=1\ldots r} = B(\xi_i)$. The previous proposition says that if M_g is the matrix of $\mathcal{M}_g$ in the basis B of $\mathcal{A}$, then

$$M_g^\mathsf{T} B(\xi_i) = g(\xi_i) B(\xi_i).$$

If moreover the basis B contains the monomials $1, x_1, x_2, \ldots, x_n$, then the common eigenvectors of M_g^T are of the form $\mathbf{v}_i = c[1, \xi_{i,1}, \ldots, \xi_{i,n}, \ldots]$ and the root ξ_i can be computed from the coefficients of $\mathbf{v}_i$ by taking the ratio of the coefficients of the monomials $x_1, \ldots, x_n$ by the coefficient of 1: $\xi_{i,k} = \frac{\mathbf{v}_{i,k+1}}{\mathbf{v}_{i,1}}$. Thus computing the common eigenvectors of all the matrices M_g^T for $g \in \mathcal{A}$ yield the roots ξ_i $(i = 1, \ldots, r)$.

In practice, it is enough to compute the common eigenvectors of $M_{x_1}^\mathsf{T}, \ldots, M_{x_n}^\mathsf{T}$, since $\forall g \in \mathbb{C}[\mathbf{x}], M_g^\mathsf{T} = g(M_{x_1}^\mathsf{T}, \ldots, M_{x_n}^\mathsf{T})$. Therefore, the common eigenvectors $M_{x_1}^\mathsf{T}, \ldots, M_{x_n}^\mathsf{T}$ are also eigenvectors of any M_g^T.

The multiplicity structure, that is the dual $Q_i^\perp$ of each primary component Q_i of I, also called the *inverse system* of the point ξ_i can be deduced by linear algebra tools (see e.g. [Mou96]).

In the case of simple roots, we have the following property [EM07][Chap. 4]:

Proposition 3.2. *If the roots $\{\xi_1, \xi_2, \ldots, \xi_r\}$ of I are simple (i.e. $\mu_i = \dim \mathcal{A}_i = 1$) then we have the following:*

- $\mathbb{u} = \{\mathbf{u}_{\xi_1}, \ldots, \mathbf{u}_{\xi_r}\}$ *is a basis of $\mathcal{A}$.*

- *The polynomials $\mathbf{u}_{\xi_1}, \ldots, \mathbf{u}_{\xi_r}$ are interpolation polynomials at the roots ξ_i: $\mathbf{u}_{\xi_i}(\xi_j) = 1$ if $i = j$ and 0 otherwise.*

- *The matrix of $\mathcal{M}_g$ in the basis $\mathbb{u}$ is the diagonal matrix $\mathrm{diag}(g(\xi_1), \ldots, g(\xi_r))$.*

This proposition tells us that if g is separating the roots, i.e. $g(\xi_i) \neq g(\xi_j)$ for $i \neq j$, then the eigenvectors of $\mathcal{M}_g$ are, up to a scalar, interpolation polynomials at the roots.

3.2. Artinian Gorenstein algebra of a multivariate Hankel operator

We associate to a Hankel operator H_σ, the quotient $\mathcal{A}_\sigma = \mathbb{C}[\mathbf{x}]/I_\sigma$ of the polynomial ring $\mathbb{C}[\mathbf{x}]$ modulo the kernel $I_\sigma = \{p \in \mathbb{C}[\mathbf{x}] \mid \forall q \in R, \langle \sigma \mid pq \rangle = 0\}$ of H_σ. We check that I_σ is an ideal of $\mathbb{C}[\mathbf{x}]$, so that $\mathcal{A}_\sigma$ is an algebra.

As $\mathcal{A}_\sigma = \mathbb{C}[\mathbf{x}]/I_\sigma \sim \mathrm{img}\, H_\sigma$, the operator H_σ is of finite rank r, if and only if, $\mathcal{A}_\sigma$ is Artinian of dimension $\dim_\mathbb{C} \mathcal{A}_\sigma = r$.

A quotient algebra $\mathcal{A}$ is called *Gorenstein* if its dual $\mathcal{A}^* = \mathrm{Hom}_\mathbb{C}(\mathcal{A}, \mathbb{C})$ is a free $\mathcal{A}$-module generated by one element.

In our context, we have the following equivalent properties [Mou16]:

- $\sigma = \sum_{i=1}^{r'} \omega_i(\mathbf{y}) \mathbf{e}_{\xi_i}(\mathbf{y})$ with $\omega_i \in \mathbb{C}[\mathbf{y}]$, $\xi_i \in \mathbb{C}^n$ and $\sum_{i=1}^{r'} \mu(\omega_i) = r$,

- H_σ is of rank r,

- $\mathcal{A}_\sigma$ is an Artinian Gorenstein algebra of dimension r.

The following proposition shows that the frequencies ξ_i and the weights ω_i can be recovered from the ideal I_σ (see [Mou16] for more details):

Proposition 3.3. *If $\sigma(\mathbf{y}) = \sum_{i=1}^{r'} \omega_i(\mathbf{y}) \mathbf{e}_{\xi_i}(\mathbf{y})$ with $\omega_i(\mathbf{y}) \in \mathbb{C}[\mathbf{y}] \smallsetminus \{0\}$ and $\xi_i \in \mathbb{C}^n$ pairwise distinct, then we have the following properties:*

- *The points $\xi_1, \xi_2, \ldots, \xi_{r'} \in \mathbb{C}^n$ are the common roots of the polynomials in $I_\sigma = \ker H_\sigma = \{p \in \mathbb{C}[\mathbf{x}] \mid \forall q \in \mathbb{C}[\mathbf{x}], \langle \sigma | pq \rangle = 0\}$.*

- *The series $\omega_i(\mathbf{y})\mathbf{e}_{\xi_i}$ is a generator of the inverse system of $Q_i^\perp$, where Q_i is the primary component of I_σ associated to ξ_i such that $\dim \mathbb{C}[\mathbf{x}]/Q_i = \mu(\omega_i)$.*

This result tells us that the problem of decomposing σ as a sum of polynomial-exponential series reduces to the solution of the polynomial equations $p = 0$ for p in the kernel I_σ of H_σ.

Another property that will be helpful to determine a basis of $\mathcal{A}_\sigma$ is the following:

Lemma 3.4. *Let $B = \{b_1, \ldots, b_r\}$, $B' = \{b'_1, \ldots, b'_r\} \subset \mathbb{C}[\mathbf{x}]$. If the matrix $H_\sigma^{B,B'} = (\langle \sigma | b_i b'_j \rangle)_{1 \le i,j \le r}$ is invertible, then B and B' are linearly independent in $\mathcal{A}_\sigma$.*

Proof. Suppose that $H_\sigma^{B,B'}$ is invertible. If there exists $p = \sum_i \lambda_i b_i$ ($\lambda_i \in \mathbb{C}$) such that $p \equiv 0$ in $\mathcal{A}_\sigma$. Then $p \star \sigma = 0$ and $\forall q \in R, \langle \sigma | pq \rangle = 0$. In particular, for $j = 1, \ldots, r$ we have

$$\sum_{i=1}^r \langle \sigma | b_i b'_j \rangle \lambda_i = 0.$$

As $H_\sigma^{B,B'}$ is invertible, $\lambda_i = 0$ for $i = 1, \ldots, r$ and B is a family of linearly independent elements in $\mathcal{A}_\sigma$. Since we have $(H_\sigma^{B,B'})^\mathsf{T} = H_\sigma^{B',B}$, we prove by a similar argument that $H_\sigma^{B,B'}$ invertible also implies that B' is linearly independent in $\mathcal{A}_\sigma$. $\qquad\square$

By this Lemma, bases of $\mathcal{A}_\sigma$ can be computed by identifying non-zero minors of maximal size of the matrix of H_σ.

Proposition 3.5. *Let B, B' be basis of $\mathcal{A}_\sigma$ and $g \in \mathbb{C}[\mathbf{x}]$. We have*

$$H_{g \star \sigma}^{B,B'} = (M_g^B)^\mathsf{T} H_\sigma^{B,B'} = H_\sigma^{B,B'} M_g^{B'}. \tag{2}$$

where M_g^B (resp. $M_g^{B'}$) is the matrix of the multiplication by g in the basis B (resp. B') of $\mathcal{A}_\sigma$.

Proof. Let $B = \{b_1, \ldots, b_r\}, B' = \{b'_1, \ldots, b'_r\}$ be two bases of $\mathcal{A}_\sigma$. We have $g b_j = \sum_{i=1}^r m_{i,j} b'_i + \kappa$ where $m_{i,j}$ is the (i,j) entry of the matrix M_g^B of multiplication by g in the basis B and $\kappa \in I_\sigma$. Then,

$$(H_{g \star \sigma}^{B,B'})_{[i,j]} = \langle \sigma \mid g\, b_i\, b'_j \rangle = \langle \sigma \mid \sum_{l=1}^r m_{l,i} b_l b'_j \rangle + \langle \sigma \mid \kappa b_j \rangle = \sum_{l=1}^r m'_{l,i} \langle \sigma \mid b_l b'_j \rangle = ((M_g^B)^t H_\sigma^{B,B'})_{[i,j]}.$$

Similarly, we have $g b'_j = \sum_{i=1}^r m'_{i,j} b'_i + \kappa'$ where $m'_{i,j}$ is the (i,j) entry of the matrix $M_g^{B'}$ of multiplication by g in the basis B' and $\kappa' \in I_\sigma$. For $1 \le i,j \le r$, the entry (i,j) of $H_{g \star \sigma}^{B,B'}$ is

$$(H_{g \star \sigma}^{B,B'})_{[i,j]} = \langle \sigma \mid b_i\, g\, b'_j \rangle = \langle \sigma \mid \sum_{l=1}^r m_{l,j} b_i b'_l \rangle + \langle \sigma \mid b_i \kappa' \rangle = \sum_{l=1}^r \langle \sigma \mid b_i b'_l \rangle m_{l,j} = (H_\sigma^{B,B'} M_g^B)_{[i,j]}.$$

This concludes the proof of the relations (2). $\qquad\square$

We deduce the following property:

Proposition 3.6. *Let $\sigma(\mathbf{y}) = \sum_{i=1}^r \omega_i(\mathbf{y})\mathbf{e}_{\xi_i}(\mathbf{y})$ with $\omega_i \in \mathbb{C}[\mathbf{y}] \smallsetminus \{0\}$ and $\xi_i \in \mathbb{C}^n$ distinct and let B, B' be bases of $\mathcal{A}_\sigma$. We have the following properties:*

- *For $g \in \mathbb{C}[\mathbf{x}]$, $M_g^{B'} = (H_\sigma^{B,B'})^{-1} H_{g \star \sigma}^{B,B'}$, $(M_g^B)^\mathsf{T} = H_{g \star \sigma}^{B,B'} (H_\sigma^{B,B'})^{-1}$.*

- *For $g \in \mathbb{C}[\mathbf{x}]$, the generalized eigenvalues of $(H_{g \star \sigma}^{B,B'}, H_\sigma^{B,B'})$ are $g(\xi_i)$ with multiplicity $\mu_i = \mu(\omega_i)$, $i = 1, \ldots, r$.*

- *The generalized eigenvectors common to all $(H_{g\star\sigma}^{B,B'}, H_\sigma^{B,B'})$ for $g \in \mathbb{C}[\mathbf{x}]$ are - up to a scalar - $(H_\sigma^{B,B'})^{-1} B(\xi_i)$, $i = 1, \ldots, r$.*

Proof. The two first points are direct consequences of Propositions 3.5 and 3.1. The third point is also a consequence of Proposition 3.1, since the coordinate vector of the evaluation $\mathbf{e}_{\xi_i}$ in the dual basis of B is $B(\xi_i)$ for $i = 1, \ldots, r$. $\qquad\square$

This proposition shows that the matrices of multiplication by an element g in $\mathcal{A}$, and thus the roots $\{\xi_1, \ldots, \xi_r\}$ and their multiplicity structure, can be computed from truncated Hankel matrices, provided we can determine bases B, B' of $\mathcal{A}_\sigma$. In practice, it is enough to compute the generalized eigenvectors common to $(H_{x_i\star\sigma}^{B,B'}, H_\sigma^{B,B'})$ for $i = 1, \ldots, n$ to recover the roots. As $H_{x_i\star\sigma}^{B,B'} = H_\sigma^{x_i B, B'} = H_\sigma^{B, x_i B'}$, the decomposition can be computed from sub-matrices of H_σ^{B,B'^+} or $H_\sigma^{B^+,B'}$ where $B^+ = B \cup x_1 B \cup \cdots \cup x_n B$, $B'^+ = B' \cup x_1 B' \cup \cdots \cup x_n B'$.

4. Decomposition algorithm

We are given the first moments $\sigma_\alpha, |\alpha| \le d$ of the series $\sigma(\mathbf{y}) = \sum_{i=1}^r \omega_i \mathbf{e}_{\xi_i}(\mathbf{y})$ with $\omega_i \in \mathbb{C}\backslash(0)$ and $\xi_i \in \mathbb{C}^n$. The goal is to recover the number of terms r, the constant weights ω_i and the frequencies ξ_i of the series $\sigma(\mathbf{y})$.

4.1. Computation of a basis

The first problem is to find automatically bases B_1 and B_2 of the quotient algebra $\mathcal{A}_\sigma$, of maximal sizes such that $H_\sigma^{B_1,B_2}$ is invertible. Using Proposition 3.6, we will compute the multiplication matrices $M_g^{B_2}$ for $g = x_i, i = 1, \ldots, n$. The frequencies ξ_j and the weights ω_j, $j = 1, \ldots, r$ will be deduced from their eigenvectors, as described in section 4.2.

Given the set of moments $(\sigma_\alpha)_{|\alpha|\le d}$, we create two sets $A_1 = (\mathbf{x}^\alpha)_{|\alpha|\le d_1}$ and $A_2 = (\mathbf{x}^\beta)_{|\beta|\le d_2}$ of monomials such that α and β are multi-indices in $\mathbb{N}^n$ with $|\alpha| \le d_1$ and $|\beta| \le d_2$. The degrees d_1 and d_2 are chosen such that $d_1 + d_2 < d$. Let $N_1 = |A_1|$ and $N_2 = |A_2|$. The truncated Hankel operator associated to σ is:

$$
\begin{array}{cccc}
H_\sigma^{d_1,d_2} & : & \mathbb{C}[\mathbf{x}]_{d_2} & \to & (\mathbb{C}[\mathbf{x}]_{d_1})^* \\
& & p & \mapsto & p \star \sigma
\end{array}
$$

The Hankel matrix in these two monomial sets A_1 and A_2 is defined by $H_\sigma^{d_1,d_2} = [\sigma_{(\alpha+\beta)}]_{\substack{|\alpha|\le d_1 \\ |\beta|\le d_2}}$.

Computing the singular value decomposition of $H_\sigma^{d_1,d_2}$, we obtain

$$H_\sigma^{d_1,d_2} = U S V^\top$$

where S is the diagonal matrix of all singular values of $H_\sigma^{d_1,d_2}$ arranged in a decreasing order, U is an unitary matrix whose columns are the left singular vectors of $H_\sigma^{d_1,d_2}$, V is an unitary matrix whose columns are the right singular vectors of $H_\sigma^{d_1,d_2}$. We denote by U^H the hermitian transpose of U and $\overline{V}$ the conjugate of V.

Let $u_i = [u_{\alpha,i}]_{\alpha\in A_1}$ and $v_j = [v_{\beta,j}]_{\beta\in A_2}$ be respectively the ith and j^{th} columns of U^H and $\overline{V}$. They are vectors respectively in $\mathbb{C}^{N_1}$ and $\mathbb{C}^{N_2}$. We denote by $u_i(\mathbf{x}) = u_i^\top A_1 = \sum_{|\alpha|\le d_1} u_{\alpha,i}\mathbf{x}^\alpha$ and $v_j(\mathbf{x}) = v_j^\top A_2 = \sum_{|\beta|\le d_2} v_{\beta,j}\mathbf{x}^\beta$ the corresponding polynomials. The bases formed by these first r polynomials are denoted $U_r^\mathsf{H} := (u_i(\mathbf{x}))_{i=1,\ldots,r}$ and $\overline{V}_r := (v_j(\mathbf{x}))_{j=1,\ldots,r}$. We will also denote by U_r^H (resp. $\overline{V}_r$) the corresponding coefficient matrix, formed by the first rows (resp. columns) of U^H (resp. $\overline{V}$). We denote by S_r the diagonal matrix of the first r rows and columns of S, formed by the first r singular values.

Proposition 4.1. *Let $\sigma = \sum_{i=1}^{r'} \omega_i(\mathbf{y})\mathbf{e}_{\xi_i}$ with $\omega_i \in \mathbb{C}[\mathbf{y}]$, $\xi_i \in \mathbb{C}^n$ and $\sum_{i=1}^{r'} \mu(\omega_i) = r$. If $\operatorname{rank} H_\sigma^{d_1,d_2} = r$, then the sets of polynomials U_r^H and $\overline{V}_r$ are bases of $\mathcal{A}_\sigma$. The matrix $M_{x_i}^{\overline{V}_r}$ associated to the multiplication operator by x_i in the basis $\overline{V}_r$ of $\mathcal{A}_\sigma$ is $M_{x_i}^{\overline{V}_r} = S_r^{-1} U_r^\mathsf{H} H_{x_i\star\sigma}^{d_1,d_2} \overline{V}_r$ $i = 1, \ldots, n$.*

Proof. The (i,j) entry of the matrix $H_\sigma^{U_r^{\mathsf{H}},\overline{V}_r}$ of the truncated Hankel operator of σ with respect to U_r^{H} and $\overline{V}_r$ is equal to:

$$
\begin{aligned}
(H_\sigma^{U_r^{\mathsf{H}},\overline{V}_r})_{[i,j]} &= \langle \sigma | u_i(\mathbf{x}) v_j(\mathbf{x}) \rangle \\
&= \langle \sigma | \big(\sum_{|\alpha| \le d_1} u_{\alpha,i} \mathbf{x}^\alpha \big) \big(\sum_{|\beta| \le d_2} v_{\beta,j} \mathbf{x}^\beta \big) \rangle = \sum_{|\alpha| \le d_1} u_{\alpha,i} \sum_{|\beta| \le d_2} \langle \sigma | \mathbf{x}^\alpha \mathbf{x}^\beta \rangle v_{\beta,j} \\
&= [U_r^{\mathsf{H}} H_\sigma^{d_1,d_2} \overline{V}_r]_{[i,j]}.
\end{aligned}
\tag{3}
$$

Using the SVD decomposition of $H_\sigma^{d_1,d_2}$, we have

$$
H_\sigma^{U_r^{\mathsf{H}},\overline{V}_r} = U_r^{\mathsf{H}} H_\sigma^{d_1,d_2} \overline{V}_r = U_r^{\mathsf{H}} U S V^{\mathsf{T}} \overline{V}_r = S_r,
$$

since $U^H U = \mathrm{Id}_{N_1}$, $V^T \overline{V} = \mathrm{Id}_{N_2}$. As $r = \mathrm{rank}\, H_\sigma^{d_1,d_2}$, S_r is invertible and by Lemma 3.4, U_r^{H} and $\overline{V}_r$ are linearly independent in $\mathcal{A}_\sigma$, which is a vector space of dimension r. Thus they are bases of $\mathcal{A}_\sigma$.

Let $H_{x_i \star \sigma}^{U_r^{\mathsf{H}},\overline{V}_r}$ be the matrix of the truncated Hankel operator of $x_i \star \sigma$ on the two bases U_r^{H} and $\overline{V}_r$. A similar computation yields $H_{x_i \star \sigma}^{U_r^{\mathsf{H}},\overline{V}_r} = U_r^{\mathsf{H}} H_{x_i \star \sigma}^{d_1,d_2} \overline{V}_r$, where $H_{x_i \star \sigma}^{d_1,d_2}$ is the matrix of the truncated Hankel operator of $x_i \star \sigma$ in the bases A_1 and A_2 for all $i = 1, \ldots, n$. Since S_r is an invertible matrix, by Proposition 3.6 we obtain $M_{\mathbf{x}_i}^{\overline{V}_r} = (H_\sigma^{U_r^{\mathsf{H}},\overline{V}_r})^{-1} H_{x_i \star \sigma}^{U_r^{\mathsf{H}},\overline{V}_r} = S_r^{-1} U_r^{\mathsf{H}} H_{x_i \star \sigma}^{d_1,d_2} \overline{V}_r$. $\qquad \square$

By this proposition U_r^{H} and $\overline{V}_r$) are bases of $\mathcal{A}_\sigma$. By Proposition 3.6, the eigenvalues of $M_{x_i}^{\overline{V}_r}$ are the ith coordinates $x_i(\xi_j) = \xi_{j,i}$ of the roots ξ_j for $i = 1, \ldots, n, j = 1, \ldots, r$.

4.2. Computation of the weights

The weight $\omega_i, i = 1, \ldots, r$ of the decomposition of σ can be easily computed using the eigenvectors of all $M_{x_j}^{\overline{V}_r}, j = 1, \ldots, n$ as follows.

Proposition 4.2. *Let $\sigma = \sum_{i=1}^r \omega_i \mathbf{e}_{\xi_i}$ with $\omega_i \in \mathbb{C} \setminus \{0\}$, $\xi_i = (\xi_{i,1}, \ldots, \xi_{i,n}) \in \mathbb{C}^n$ distinct and let $M_{x_j}^{\overline{V}_r}$ be the matrix of multiplication by x_j in the basis $\overline{V}_r$. Let $\mathbf{v}_i$ be a common eigenvector of $M_{x_j}^{\overline{V}_r}$, $j = 1, .., n$ for the eigenvalues $\xi_{i,j}$. Then the weight of $\mathbf{e}_{\xi_i}$ in the decomposition of σ is*

$$
\omega_i = \frac{[1]^{\mathsf{T}} H_\sigma^{d_1,d_2} \overline{V}_r \mathbf{v}_i}{[\xi_i^\alpha]_{\alpha \in A_2}^{\mathsf{T}} \overline{V}_r \mathbf{v}_i}.
\tag{4}
$$

Proof. According to Proposition 3.2, the eigenvectors of the multiplication operator $\mathcal{M}_{x_i}$ are, up to scalar, the interpolation polynomials $\mathbf{u}_i(\mathbf{x})$ at the roots. Let $\mathbf{u}_{\xi_i}$ be the coefficient vector associated to $\mathbf{u}_{\xi_i}(\mathbf{x})$ in the basis $\overline{V}_r$ of $\mathcal{A}_\sigma$. Let $\mathbf{v}_i = \lambda \mathbf{u}_{\xi_i}$ be the eigenvector of $M_{x_i}^{\overline{V}_r}$ associated to the eigenvalue $\xi_{j,i}$ for $j = 1, \ldots, r, i = 1, \ldots, n$ such that $\mathbf{v}_i(\mathbf{x}) = A_2^{\mathsf{T}} \tilde{\mathbf{v}}_i = \sum_{|\beta| \le d_2} \tilde{\mathbf{v}}_{i\beta} \mathbf{x}^\beta$ where $\tilde{\mathbf{v}}_i = \overline{V}_r \mathbf{v}_i$. Applying the series on all the idempotents, we obtain

$$
\langle \sigma \mid \mathbf{u}_{\xi_i}(\mathbf{x}) \rangle = \langle \sum_{j=1}^r \omega_j \mathbf{e}_{\xi_j}(\mathbf{y}) \mid \mathbf{u}_{\xi_i}(\mathbf{x}) \rangle = \omega_i \mathbf{u}_{\xi_i}(\xi_i) = \omega_i.
$$

Therefore, we have $\omega_i = \frac{\langle \sigma | \lambda \mathbf{u}_{\xi_i}(\mathbf{x}) \rangle}{\lambda} = \frac{\langle \sigma | \mathbf{v}_i(\mathbf{x}) \rangle}{\lambda} = \frac{\langle \sigma | \mathbf{v}_i(\mathbf{x}) \rangle}{\mathbf{v}_i(\xi_i)}$ because of $\mathbf{v}_i(\xi_i) = (\lambda \mathbf{u}_{\xi_i})(\xi_i) = \lambda$. Then

$$
< \sigma \mid \mathbf{v}_i(\mathbf{x}) > = [1]^{\mathsf{T}} H_\sigma^{d_1,d_2} \tilde{\mathbf{v}}_i = [1]^{\mathsf{T}} H_\sigma^{d_1,d_2} \overline{V}_r \mathbf{v}_i,
$$

where $[1]$ is the vector of coefficients of the polynomial 1 in the monomial basis $A_1 = (\mathbf{x}^\alpha)_{|\alpha| \le d_1}$ and

$$
\mathbf{v}_i(\xi_i) = [\xi_i^\alpha]_{\alpha \in A_2}^{\mathsf{T}} \tilde{\mathbf{v}}_i = [\xi_i^\alpha]_{\alpha \in A_2}^{\mathsf{T}} \overline{V}_r \mathbf{v}_i.
$$

We deduce that $\omega_i = \frac{[1]^{\mathsf{T}} H_\sigma^{d_1,d_2} \overline{V}_r \mathbf{v}_i}{[\xi_i^\alpha]_{\alpha \in A_2}^{\mathsf{T}} \overline{V}_r \mathbf{v}_i}$. $\qquad \square$

We describe now the algorithm to recover the sum $\sigma(\mathbf{y}) = \sum_{j=1}^{r} \omega_j \, \mathbf{e}_{\xi_j}(\mathbf{y})$, $\omega_j \in \mathbb{C} \smallsetminus \{0\}, \xi_j \in \mathbb{C}^n$, from the first coefficients $(\sigma_\alpha)_{|\alpha| \leq d}$ of the formal power series $\sigma = \sum_\alpha \sigma_\alpha \dfrac{\mathbf{y}^\alpha}{\alpha!}$.

Algorithm 4.1: Decomposition of polynomial-exponential series with constant weights

Input: the moments σ_α of σ for $|\alpha| \leq d$.

Let d_1 and d_2 be positive integers such that $d_1 + d_2 + 1 = d$, for example $d_1 := \lceil \frac{d-1}{2} \rceil$ and $d_2 := \lfloor \frac{d-1}{2} \rfloor$.

1. Compute the Hankel matrix $H_\sigma^{d_1,d_2} = [\sigma_{(\alpha+\beta)}]_{\substack{|\alpha| \leq d_1 \\ |\beta| \leq d_2}}$ of σ in for the monomial sets $A_1 = (\mathbf{x}^\alpha)_{|\alpha| \leq d_1}$ and $A_2 = (\mathbf{x}^\beta)_{|\beta| \leq d_2}$.

2. Compute the singular value decomposition of $H_\sigma^{d_1,d_2} = USV^\mathsf{T}$ with singular values $s_1 \geq s_2 \geq \cdots \geq s_m \geq 0$.

3. Determine its numerical rank, that is, the largest integer r such that $\frac{s_r}{s_1} \geq \epsilon$.

4. Form the matrices $M_{x_i}^{\overline{V}_r} = S_r^{-1} U_r^{\mathsf{H}} H_{x_i \star \sigma}^{d_1,d_2} \overline{V}_r, i = 1, \ldots, n$, where $H_{x_i \star \sigma}^{d_1,d_2}$ is the Hankel matrix associated to $x_i \star \sigma$.

5. Compute the eigenvectors $\mathbf{v}_j$ of $\sum_i^n l_i M_{x_i}$ for a random choice of l_i in $[-1,1]$, $i = 1, \ldots, n$ and for each $j = 1, \ldots, r$ do the following:

   - Compute $\xi_{j,i}$ such that $M_i \mathbf{v}_j = \xi_{j,i} \mathbf{v}_j$ for $i = 1, \ldots, n$ and deduce the point $\xi_j := (\xi_{j,1}, \ldots, \xi_{j,n})$.

   - Compute $\omega_j = \dfrac{\langle \sigma | \mathbf{v}_j(\mathbf{x}) \rangle}{\mathbf{v}_j(\xi_j)} = \dfrac{[1]^\mathsf{T} H_\sigma^{d_1,d_2} \overline{V}_r \mathbf{v}_j}{[\xi_i^\alpha]_{\alpha \in A_2}^\mathsf{T} \overline{V}_r \mathbf{v}_j}$ where $[1]$ is the coefficient vector of 1 in the basis A_1.

Output: $r \in \mathbb{N}$, $\omega_j \in \mathbb{C}\backslash(0)$, $\xi_j \in \mathbb{C}^n$, j=1,$\ldots$,r such that $\sigma(\mathbf{y}) = \sum_{j=1}^{r} \omega_j \, \mathbf{e}_{\xi_j}(\mathbf{y})$ up to degree d.

---

5. Experimentation

In this section, we present numerical experimentations for the decomposition of $\sigma = \sum_{\alpha \in \mathbb{N}} \sigma_\alpha \frac{\mathbf{y}^\alpha}{\alpha!}$ from its moments σ_α. For a given number of variables n and a fixed degree d, we compute the coefficients $\sigma_\alpha = \sigma(\mathbf{x}^\alpha) = \sum_{i=1}^{r} \omega_i \xi_i^\alpha$ such that $|\sigma_\alpha| \leq d$ where $\omega_j \in \mathbb{C}\backslash(0)$ and $\xi_i = (\xi_{i,1}, \ldots, \xi_{i,n}), i = 1, \ldots, r$ have random uniform distributions such that $0.5 \, M \leq |\xi_{i,j}| \leq 1.5 \, M$, $-\pi \leq arg(\xi_{i,j}) \leq \pi$, $0.5 \leq |\omega_i| \leq 1$ and $-\pi \leq arg(\omega_i) \leq \pi$, for $M \geq 1$. To analyse the numerical behaviour of our method, we compare the results with the known frequencies and weights used to compute the moments of σ.

We use Maple 16 to implement the algorithms. The arithmetic operations are performed on complex numbers, with a numerical precision fixed to $Digits = 15$.

5.1. Numerical behavior against perturbation

We apply random perturbations on the moments of the form $\sigma_\alpha + \epsilon(p_\alpha + \boldsymbol{i}\, q_\alpha)$ where p_α and q_α are two random numbers in $[-1,1]$ with a uniform distribution, and $\varepsilon = 10^{-e}$ where e is a fixed positive integer.

To measure the consistency of our algorithm, we compute the maximum error between the input frequencies ξ_i and the output frequencies ξ_i', and between the input weights ω_i and the output weights ω_i':

$$\text{err} = \max(err(\xi_i, \xi_i'), \text{err}(\omega_i, \omega_i')) \text{ where } \text{err}(\omega_i, \omega_i') = \max_{1 \leq i \leq r} |\omega_i - \omega_i'| \text{ and } \text{err}(\xi_i, \xi_i') = \max_{1 \leq i \leq r} \|\xi_i - \xi_i'\|_2. \quad (5)$$

In each computation, we compute the average of the maximum errors resulting from 10 tests.

In Figures 1a and 1b, we study the evolution of the error in terms of the perturbation $\varepsilon = 10^{(-e)}$, for a fixed degree $d = 10$, a number of variables $n = 3$, different ranks $r = 5, 10, 20, 30$ and for two different amplitudes of the frequencies $M = 1$ and $M = 100$.

In Figure 1a for $M = 1$, the lower error is for the lower rank $r = 5$. Between $\varepsilon \approx 10^{-12}$ and $\varepsilon = 1$, the error err increases in terms of the perturbation as err= $exp(t\,e)$ for some slope $t \approx 1$. The slope t remains approximately constant but the error increases slightly with the rank r. Before $\varepsilon = 10^{-13}$, it is approximately
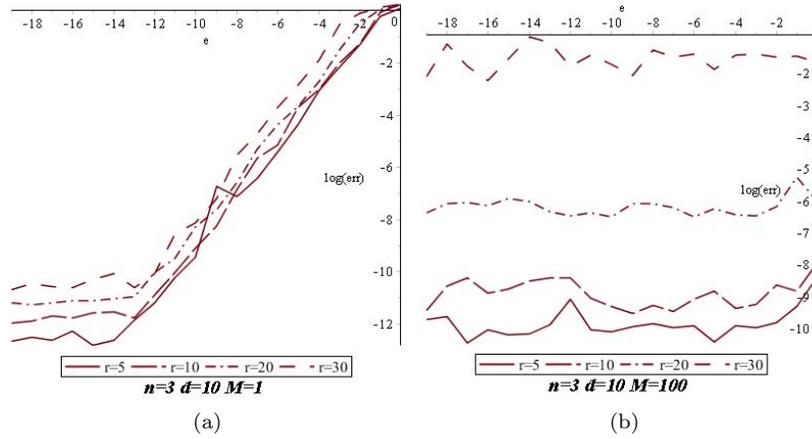
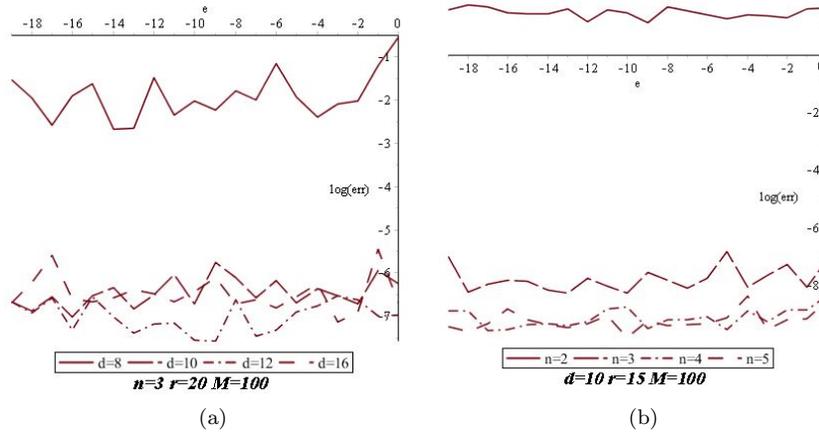Figure 1: The influence of the amplitude of the frequencies on the maximum error.

Figure 2: The influence of the degree and dimension on the maximum error.

constant (approximately 10^{-12} for $r = 5$) This is due to the fact, in this range, the perturbation is lower than the numerical precision.

In Figure 1b for $M = 100$, the lower error is also for the lower rank. The error has almost a constant value when e varies. It is bigger than for $M = 1$ for small perturbations. For $r = 5, 10$ the error slightly increases between $e = -2$ and $e = 0$, with a similar slope. This figure clearly shows that the error degrades significantly from $M = 1$ to $M = 100$ and that the degradation increases rapidly with the rank r.

In Figure 2a, we fix the number of variables $n = 3$, the rank $r = 20$ and we change the degree d which induces a change in the dimensions of the Hankel matrices. For $e \in [-19, 0]$, the error decreases when we increase the degree from $d = 8$ to $d = 10$. It is slightly lower when $d = 12$ than when $d = 10$, and error is similar for $d = 10$ and $d = 16$. This increase of the precision with the degree can be related to ratio of number of moments by the number of values to recover in the decomposition.

In Figure 2b, we fix the degree $d = 10$ and the rank $r = 15$ and we change the number of variables $n = 2, 3, 4, 5$. The dimension of the matrices increases polynomially with n. We observe that the error decreases quickly with n. It shows that the precision improves significantly with the dimension.

5.2. Numerical Rank

To compute the number r of terms in the decomposition of σ, we arrange the diagonal entries in the decreasing order $s_1 \geq s_2 \geq \cdots \geq s_r > s_{r+1} \geq \cdots \geq s_{N_2}$ and we determine the numerical rank of $H_\sigma^{d_1,d_2}$ by fixing the largest integer r such that $s_r/s_1 \geq \epsilon$.

It is known that the ill-conditioning of the Hankel matrix associated to Prony's method is in the origin of a numerical instability with respect to perturbed measurements

$$\tilde{\sigma}_{\alpha+\beta} = \sigma_{\alpha+\beta} + \varepsilon_{\alpha+\beta} \qquad |\alpha + \beta| \leq d.$$

In our algorithm the computation of the numerical rank can be affected by this instability. We can explain this instability, using a reasoning close to [Sau16], as follows.

We denote by s_j (resp. $\tilde{s}_j$) the jth largest singular value of $H := H_\sigma^{d_1,d_2}$ (resp. $\tilde{H} := H_{\tilde{\sigma}}^{d_1,d_2}$). The perturbation result for singular values satisfies the estimate (see [GL96])

$$|s_j - \tilde{s}_j| \leq s_1(\varepsilon) = \|\varepsilon\|_2.$$

Then, as long as the perturbation is small relative to the conditioning of the problem, that is

$$\|\epsilon\|_2 \leq \tfrac{1}{2} s_r \quad \text{provided that } r = \text{rank}(H),$$

then $\|s_j - \tilde{s}_j\| \leq \tfrac{1}{2} s_r \ \forall j$ and therefore $\tilde{s}_r \geq \tfrac{1}{2} s_r$ and $\tilde{s}_{r+1} \leq \tfrac{1}{2} s_r$. Hence by taking $\varepsilon \leq \tfrac{1}{2} s_r$ as a threshold level we will be sure that the rank is calculated correctly.

But the problem may be badly ill-conditioned and then such a level will not be reasonable. In fact

$$H = (\sigma_{\alpha+\beta})_{\substack{|\alpha| \leq d_1 \\ |\beta| \leq d_2}} = \left(\sum_{i=1}^{r} \omega_i \xi_i^{\alpha+\beta} \right)_{\substack{|\alpha| \leq d_1 \\ |\beta| \leq d_2}} = \sum_{i=1}^{r} \omega_i v_{i,d_1} v_{i,d_2}^T,$$

where $v_{i,d_1} = (\xi_i^{|\alpha|})_{|\alpha| \leq d_1}$ (resp. $v_{i,d_2} = (\xi_i^{|\beta|})_{|\beta| \leq d_2}$) is the ith column of the Vandermonde matrix $V_{d_1} = (\xi_i^\alpha)_{\substack{1 \leq i \leq r \\ |\alpha| \leq d_1}}$ (resp. $V_{d_2} = (\xi_i^\beta)_{\substack{1 \leq i \leq r \\ |\beta| \leq d_2}}$). Then

$$H = \sum_{i=1}^{r} \omega_i V_{d_1} e_i e_i^T V_{d_2}^T = V_{d_1} \left(\sum_{i=1}^{r} \omega_i e_i e_i^T \right) V_{d_2}^T = V_{d_1} C V_{d_2}^T$$

where $C = \text{diag}\left((\omega_i)_{1 \leq i \leq r} \right)$ is the diagonal matrix with ω_i on the diagonal.

Now, using the fact that

$$s_r(H) = \min_{\substack{\|x\|=1 \\ Hx \neq 0}} \|Hx\|_2 = \min_{\substack{\|x\|=1 \\ Hx \neq 0}} \|V_{d_1} C V_{d_2}^T x\|_2,$$

we remark that if V_{d_2} (resp. V_{d_1}) is ill-conditioned then $\|V_{d_2} x\|_2$ (resp. $\|V_{d_1} C V_{d_2}^T x\|$) may be very small and $s_r(H)$ is small as well. This situation can also be produced if $\max_{1 \leq i \leq r} \omega_i$ is very small. In our numerical experiments, the ω_i are chosen randomly in $[0.5, 1]$ and then they don't cause any numerical instability.

On the other hand, the ξ_i vary in such a way that their amplitude can be large, which can generate very ill-conditioned Vandermonde matrices. In fact, it is known (see [Pan16]), that for a nonsingular univariate Vandermonde matrix $V = (a_i^j)_{0 \leq i,j \leq n-1}$, where $(a_i)_{0 \leq i \leq n-1}$ denotes a vector of n distinct knots, the condition number of V is exponential in n if $\max_{0 \leq i \leq n-1} |a_i| > 1$ or in k if $|a_i| < 1$ for at least k knots a_i. Therefore an $n \times n$ Vandermonde matrix is badly ill-conditioned unless all knots lie in or near the disc $D(0,1) = \{x : |x| \leq 1\}$ and unless they lie mostly on or near its boundary $C(0,1)$.

In the multivariate case, it appears that the condition number of multivariate Vandermonde matrices has the same behavior as in univariate case. That is, it is exponential in the highest degree of the entries.

According to the foregoing, when the amplitude M of the frequencies increases (even for moderate values of M) the numerical rank calculated by truncating the singular values of H will be different from the exact rank of H. An idea to remedy this problem is to rescale the frequencies ξ_i in order to obtain points with coordinates close to the unitary circle $C(0,1)$.

5.3. Rescaling

As we have seen in Figures 1a and 1b, the error increases significantly with the amplitude M. To remedy this issue, we present a rescaling technique and its numerical impact. It's done like this:

- For a chosen non-zero constant λ, we transform the input moments of the series as follows:

$$\sigma(\mathbf{y}) \coloneqq \sum_{\alpha \in N^n} \sigma_\alpha \frac{\mathbf{y}^\alpha}{\alpha!} \longrightarrow \tilde{\sigma}(\mathbf{y}) \coloneqq \sigma(\lambda \mathbf{y}) = \sum_{\alpha \in N^n} \lambda^{|\alpha|} \sigma_\alpha \frac{\mathbf{y}^\alpha}{\alpha!},$$

  which corresponds to the scaling on the frequencies $e_\xi(\lambda \mathbf{y}) = e_{\lambda \xi}(\mathbf{y})$.

- We compute decomposition of $\tilde{\sigma}(\mathbf{y}) = \sigma(\lambda \mathbf{y})$ from the moments $\tilde{\sigma}_\alpha = \lambda^{|\alpha|} \sigma_\alpha$.

- We apply the inverse scaling on the computed frequencies $\tilde{\xi}_i$ which gives $\xi_i = \frac{\tilde{\xi}_i}{\lambda} = \left(\frac{\tilde{\xi}_{i,1}}{\lambda}, \ldots, \frac{\tilde{\xi}_{i,n}}{\lambda} \right)$.

To determine the scaling factor λ, we use $\lambda \coloneqq \frac{1}{m}$ where $m = \frac{max_{|\alpha|=d}|\sigma_\alpha|}{max_{|\alpha|=d-1}|\sigma_\alpha|}$. This is justified as follows: If $|\omega_j| \le 1, j = 1, \ldots, r$, then $|\sigma_\alpha| = |\sum_{j=1}^r \omega_j {\xi_j}^\alpha| \simeq M^d$ for $|\alpha| = d$ big and for M is the highest modulus of frequencies. Similarly $|\sigma_{\alpha'}| \simeq M^{d-1}$ for $|\alpha'| = d - 1$. Then we have $m = \frac{max_{|\alpha|=d}|\sigma_\alpha|}{max_{|\alpha|=d-1}|\sigma_\alpha|} \approx M$.
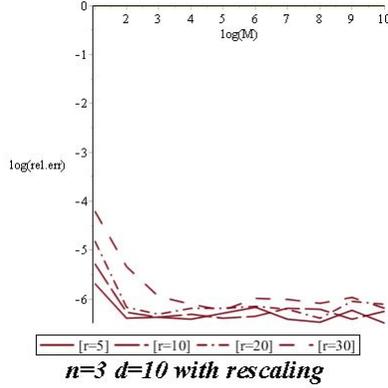

Figure 3: The rescaling influence

To study the numerical influence of the rescaling, we compute the maximum relative error between the input frequencies ξ_i and the output frequencies $\tilde{\xi}_i$, and the maximum error between the input weights ω_i and the output weights $\tilde{\omega}_i$, and we take their maximum:

$$\text{rel.err} = \max(\text{rel.err}(\xi_i, \tilde{\xi}_i), \text{err}(\omega_i, \tilde{\omega}_i)) \tag{6}$$

where $\text{err}(\omega_i, \tilde{\omega}_i) = \max_{1 \le i \le r} |\omega_i - \tilde{\omega}_i|$ and $\text{rel.err}(\xi_i, \tilde{\xi}_i) = \max_{1 \le i \le r} \frac{\|\xi_i - \tilde{\xi}_i\|_2}{\|\xi_i\|_2}$.

In Figure 3, we see the influence of the rescaling on the maximum relative error. The perturbation on the moments is of the order $\varepsilon = 10^{-6}$. Each curve for $r = 5, 10, 20, 30$, has almost a constant evolution with the increasing values of M between 10^2 and 10^{10}. The maximum relative error is lower when $M = 100$ than when $M = 1$ which is confirmed with the results shown in Figures 1a and 1b. When we increase r the maximum relative error decreases slightly.

In conclusion, the rescaling has an important influence on the computation of the maximum relative error when the modulus M of points is quite big.

The scaling of moments by some computed factor λ also enhances the computation of the numerical rank r and leads to a better decomposition as we have seen in 5.2.

5.4. Newton iteration

Given a perturbation $\tilde{\sigma} = \sum_\alpha \tilde{\sigma}_\alpha \frac{\mathbf{y}^\alpha}{\alpha!}$ of a polynomial-exponential series $\sigma = \sum_{i=1}^r \omega_i \mathbf{e}_{\xi_i}(\mathbf{y})$, we want to remove the perturbation on $\tilde{\sigma}$ by computing the polynomial-exponential series of rank r, which is the closest to the perturbed series $\tilde{\sigma}$. Starting from an approximate decomposition, using the previous method on the perturbed data, we apply a Newton-type method to minimize the distance between the input series and a weighted sum of r exponential terms.

To evaluate the distance between the series, we use the first moments $\tilde{\sigma}_\alpha$ for $\alpha \in A$, where A is a finite subset of $\mathbb{N}^n$. For $\alpha \in A$, let $F_\alpha(\Xi) = \sum_{i=1} \omega_i \xi_i^\alpha - \tilde{\sigma}_\alpha$ be the error function for the moment $\tilde{\sigma}_\alpha$, where $\omega_i, \xi_{i,j}$ are variables. We denote by $\Xi = (\xi_{i,j})_{1 \le i \le r, 0 \le j \le n}$ this set of variables, with the convention that $\xi_{i,0} = \omega_i$ for $i = 1, \ldots, r$. Let $I = [1, r] \times [0, n] = \{(i,j) \mid 1 \le i \le r, 0 \le j \le n\}$ be the indices of the variables and $N = (n+1)r = |I|$. We denote by $F(\Xi) = (F_\alpha(\Xi))_{\alpha \in A}$ the vector of these error functions.

We want to minimize the distance

$$E(\Xi) = \frac{1}{2} \sum_{\alpha \in A} |F_\alpha(\Xi)|^2 = \frac{1}{2} \|F(\Xi)\|^2.$$

Let $M(\Xi_i) = [w_i \xi_i^\alpha]_{\alpha \in A}$. We denote by $V(\Xi) = (\partial_{(i,j)} M(\Xi_i))_{(i,j) \in I}$ the $|A| \times N$ Vandermonde-like matrix, which columns are the vectors $\partial_{(i,j)} M(\Xi_i)$. The gradient of $E(\Xi)$ is

$$\nabla E(\Xi) = (\langle \partial_{(i,j)} M(\Xi_i), F(\Xi) \rangle)_{(i,j) \in I} = V(\Xi)^T F(\Xi)$$

where $\partial_{(i,j)}$ is the derivation with respect to $\Xi_{i,j}$ for $(i,j) \in I$. We denote by $V(\Xi) = (\partial_{(i,j)} M(\Xi_i))_{(i,j) \in I}$ the $|A| \times N$ Vandermonde-like matrix, which columns are the vectors $\partial_{(i,j)} M(\Xi_i), (i,j) \in I$.

To find a local minimizer of $E(\Xi)$, we compute a solution of the system $\nabla E(\Xi) = 0$, by Newton method. The Jacobian of $\nabla E(\Xi)$ with respect to the variables Ξ is

$$J_\Xi(\nabla E) = \langle \partial_{(i,j)} M(\Xi_j), \partial_{(i',j')} M(\Xi_{j'}) \rangle + (\langle \partial_{(i,j)} \partial_{(i',j')} M(\Xi_i), F(\Xi) \rangle)_{(i,j) \in I, (i',j') \in I}$$

$$= V(\Xi)^\mathsf{T} V(\Xi) + (\langle \partial_{(i,j)} \partial_{(i',j')} M(\Xi_i), F(\Xi) \rangle)_{(i,j) \in I, (i',j') \in I}.$$

Notice that $\partial_{(i,j)} \partial_{(i',j')} M(\Xi_i) = 0$ if $i \ne i'$ so that the first matrix is a block diagonal matrix. Then, Newton iteration takes the form:

$$\Xi_{n+1} = \Xi_n - J_\Xi(\nabla E)^{-1} \nabla E(\Xi_n).$$

To study the numerical influence of Newton method, we compute the maximum absolute error between the input frequencies $\tilde{\xi}_i$ and the output frequencies $\tilde{\xi}_i$, and the maximum error between the input weights ω_i and the output weights $\tilde{\omega}_i$ as in (5).

Figures 4a and 4b show that Newton iterations improve the error. The error decreases by a factor of $\approx 10^2$ compared to the computation without Newton iterations. In Figure 4b for $M = 100$ the error is smaller than without Newton iterations by a similar order of magnitude (see in Figure 1b).

References

[ACd10] Fredrik Andersson, Marcus Carlsson, and Maarten V. de Hoop. Nonlinear approximation of functions in two dimensions by sums of exponential functions. *Applied and Computational Harmonic Analysis*, 29(2):156–181, September 2010.

[Bar95] Gaspard Riche Baron de Prony. Essai expérimental et analytique: Sur les lois de la dilatabilité de fluides élastique et sur celles de la force expansive de la vapeur de l'alcool, à différentes températures. *J. Ecole Polyt.*, 1:24–76, 1795.

[BBCM13] Alessandra Bernardi, Jérôme Brachat, Pierre Comon, and Bernard Mourrain. General tensor decomposition, moment matrices and applications. *Journal of Symbolic Computation*, 52:51–71, 2013.
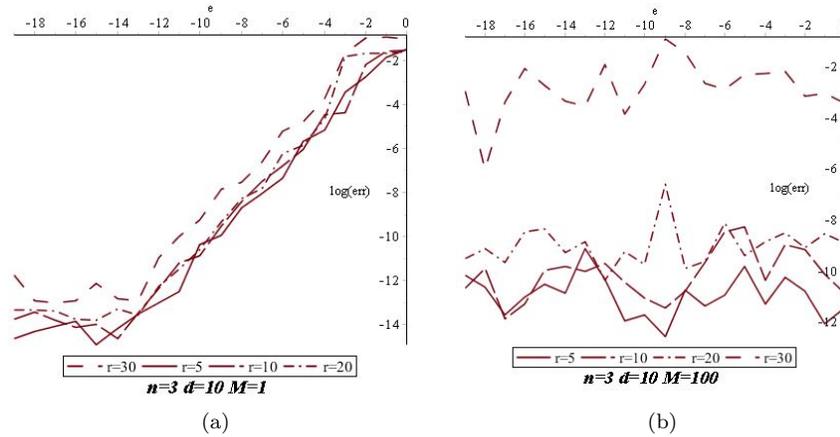
Figure 4: Newton influence with 5 iterations

[BCMT10] Jérôme Brachat, Pierre Comon, Bernard Mourrain, and Elias P. Tsigaridas. Symmetric tensor decomposition. *Linear Algebra and Applications*, 433(11-12):1851–1872, 2010.

[Bec97] Bernhard Beckermann. The Condition Number of real Vandermonde, Krylov and positive definite Hankel matrices. *Numerische Mathematik*, 85:553–577, 1997.

[BGL07] Bernhard Beckermann, Gene H. Golub, and George Labahn. On the numerical condition of a generalized Hankel eigenvalue problem. *Numerische Mathematik*, 106(1):41–68, 2007.

[BM05] Gregory Beylkin and Lucas Monzón. On approximation of functions by exponential sums. *Applied and Computational Harmonic Analysis*, 19(1):17–48, July 2005.

[BP94] Dario Bini and Victor Y. Pan. *Polynomial and Matrix Computations*. Birkhäuser Boston, Boston, MA, 1994.

[EM07] Mohamed Elkadi and Bernard Mourrain. *Introduction à la résolution des systèmes polynomiaux*, volume 59 of *Mathématiques et Applications*. Springer, 2007.

[Fuh12] Paul A. Fuhrmann. *A Polynomial Approach to Linear Algebra*. Universitext. Springer New York, New York, NY, 2012.

[GL96] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. JHU Press, 1996.

[GP03] Gene Golub and Victor Pereyra. Separable nonlinear least squares: The variable projection method and its applications. *Inverse Problems*, 19(2):R1–R26, 2003.

[HS90] Yingbo Hua and Tapan K. Sarkar. Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38(5):814–824, 1990.

[KPRv16] Stefan Kunis, Thomas Peter, Tim Römer, and Ulrich von der Ohe. A multivariate generalization of Prony's method. *Linear Algebra and its Applications*, 490:31–47, February 2016.

[Kro81] Leopold Kronecker. Zur Theorie der Elimination Einer Variabeln aus Zwei Algebraischen Gleichungen. pages 535–600., 1881.

[Lan11] Joseph M. Landsberg. *Tensors: Geometry and Applications*. Graduate studies in mathematics. American Mathematical Soc., 2011.

[Mar12] Ivan Markovsky. *Low Rank Approximation*. Communications and Control Engineering. Springer London, London, 2012.

[Mou96] Bernard Mourrain. Isolated points, duality and residues. *J. of Pure and Applied Algebra*, 117&118:469–493, 1996.

[Mou16] Bernard Mourrain. Polynomial-exponential decomposition from moments, 2016. hal-01367730, arXiv:1609.05720.

[MP00] Bernard Mourrain and Victor Y. Pan. Multivariate Polynomials, Duality, and Structured Matrices. *Journal of Complexity*, 16(1):110–180, March 2000.

[Pan16] Victor Y. Pan. How Bad Are Vandermonde Matrices? *SIAM Journal on Matrix Analysis and Applications*, 37(2):676–694, January 2016.

[PPS15] Thomas Peter, Gerlind Plonka, and Robert Schaback. Prony's Method for Multivariate Signals. *PAMM*, 15(1):665–666, 2015.

[PS12] V. Pereyra and G. Scherer, editors. *Exponential Data Fitting and Its Applications*. Bentham Science Publisher, 2012.

[PT11] Daniel Potts and Manfred Tasche. Nonlinear approximation by sums of nonincreasing exponentials. *Applicable Analysis*, 90(3-4):609–626, 2011.

[PT13] Daniel Potts and Manfred Tasche. Parameter estimation for multivariate exponential sums. *Electronic Transactions on Numerical Analysis*, 40:204–224, 2013.

[RK89] Richard Roy and Thomas Kailath. ESPRIT-estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(7):984–995, 1989.

[Sau16] Tomas Sauer. Prony's method in several variables. *Numerische Mathematik*, 2016. To appear.

[SK92] A. Lee Swindlehurst and Thomas Kailath. A performance analysis of subspace-based methods in the presence of model errors. I. The MUSIC algorithm. *IEEE Transactions on signal processing*, 40(7):1758–1774, 1992.

[Tyr94] Evgenij E. Tyrtyshnikov. How bad are Hankel matrices? *Numerische Mathematik*, 67(2):261–269, 1994.