Nonlinear approximation by sums of nonincreasing exponentials

Daniel Potts[*] and Manfred Tasche[†]

Dedicated to Professor P.L. Butzer on the occasion of his 80th birthday

Many applications in electrical engineering, signal processing, and mathematical physics lead to following approximation problem: Let h be a short linear combination of nonincreasing exponentials with complex exponents. Determine all exponents, all coefficients, and the number of summands from finitely many equispaced sampled data of h. This is a nonlinear inverse problem. This paper is an extension of [11], where only the case of sums of complex exponentials with real frequencies is considered. In the following, we present new results on an approximate Prony method (APM) which is based on [2]. In contrast to [2], we apply perturbation theory for a singular value decomposition of a rectangular Hankel matrix such that we can describe the properties and the numerical behavior of APM in detail. For this inverse problem, the number of sampled data acts as regularization parameter. The first part of APM recovers the exponents. The second part computes the coefficients by an overdetermined linear Vandermonde–type system. Numerical experiments show the performance of our method.

Key words and phrases: nonlinear approximation, exponential sum, approximate Prony method, singular value decomposition, matrix perturbation theory, perturbed rectangular Hankel matrix, Vandermonde–type matrix.

AMS *Subject Classifications*: 41A30, 15A18, 65F15, 65F20, 94A12.

[*]potts@mathematik.tu-chemnitz.de, Chemnitz University of Technology, Department of Mathematics, D–09107 Chemnitz, Germany

[†]manfred.tasche@uni-rostock.de, University of Rostock, Institute for Mathematics, D–18051 Rostock, Germany

1 Introduction

Let h be a finite sum of nonincreasing exponentials

$$h(x) = \sum_{j=1}^{M} c_j \, \mathrm{e}^{f_j x} \quad (x \in \mathbb{R}) \tag{1.1}$$

with complex coefficients $c_j \neq 0$ and complex exponents $f_j \in \mathbb{F}$, where the rectangle $\mathbb{F}$ is defined by

$$\mathbb{F} := \{f \in \mathbb{C};\, -\alpha \leq \operatorname{Re} f \leq 0,\, -\pi < \operatorname{Im} f \leq \pi\}$$

with $\alpha \geq 0$. Then all values $z_j := \mathrm{e}^{f_j}$ belong to the circular ring $\mathbb{D} := \{z \in \mathbb{C};\, \mathrm{e}^{-\alpha} \leq |z| \leq 1\}$. Assume that all values z_j $(j = 1, \ldots, M)$ are pairwise different. In the case $\operatorname{Re} f_j < 0$ $(j = 1, \ldots, M)$, the function h is a sum of decaying exponentials.
We consider the following *nonlinear approximation problem*: For given sampled data $h(k)$ $(k = 0, \ldots, 2N)$ with $N \geq M$, determine the positive integer M, the exponents $f_j \in \mathbb{F}$ and the coefficients $c_j \in \mathbb{C} \setminus \{0\}$ $(j = 1, \ldots, M)$ such that

$$h(k) = \sum_{j=1}^{M} c_j \, \mathrm{e}^{f_j k} \quad (k = 0, \ldots, 2N)\,. \tag{1.2}$$

By our assumption $N \geq M$, the overmodeling factor $(2N + 1)/M$ is larger than 2. Here $2N+1$ is the number of sampled data and M is the number of exponential terms. Further let L be an a priori known upper bound of M. The above approximation problem is a nonlinear inverse problem which can be simplified by original ideas of G.R. de Prony (see Section 2). But the classical Prony method is notorious for its sensitivity to noise such that numerous modifications were attempted to improve its numerical behavior. The main drawback of the Prony method is the property of the exact rectangular Hankel matrix $\mathbf{H} = \big(h(k + l)\big)_{k,l=0}^{2N-L,L}$ that 0 is a singular value (see Lemma 2.1 or step 1 of Algorithm 2.3). But in practice, only noisy data are given such that this property is not fulfilled. This paper is an extension of recent results [11] on parameter estimation of exponential sums. Our results are based on the paper [2] of G. Beylkin and L. Monzón. The nonlinear approximation problem of finding the exponents and coefficients can be split into two problems. To obtain the exponents, we solve a singular value problem of the rectangular Hankel matrix $\mathbf{H}$ and find the exponents via roots of a convenient polynomial of degree L. To obtain the coefficients, we use the computed exponents and solve an overdetermined linear Vandermonde–type system. In contrast to [2], we present an *approximate Prony method* (APM) by means of matrix perturbation theory such that we can describe the properties and the numerical behavior of the APM in detail. The first part of APM recovers all exponents. The second part computes all coefficients by an overdetermined linear Vandermonde–type system in a stable way.
In applications, perturbed values $\tilde{h}_k \in \mathbb{R}$ of the exact sampled data $h(k)$ are only known with the property

$$\tilde{h}_k = h(k) + e_k\,, \quad |e_k| \leq \varepsilon_1 \quad (k = 0, \ldots, 2N),$$

where the error terms e_k are bounded by certain accuracy $\varepsilon_1 > 0$. Also if the sampled values $h(k)$ are accurately determined, then we still have a small roundoff error due to the use of floating point arithmetic. Furthermore we assume that $|c_j| \gg \varepsilon_1$ $(j = 1, \ldots, M)$. The special case $|z_j| = 1$ $(j = 1, \ldots, M)$ was recently analyzed by the authors [11].

This paper is organized as follows. In Section 2, we sketch the classical Prony method. Then in Section 3, we present four algorithms of APM. Using matrix perturbation theory, we discuss the properties of small singular values and related right/left singular vectors of a rectangular Hankel matrix formed by given noisy data. We emphasize that the Algorithm 2.3 is only of theoretical interest. Using Lemma 2.1 and matrix perturbation theory, we can show the existence of small singular values of the perturbed rectangular Hankel matrix $\tilde{\mathbf{H}} = \big(\tilde{h}_{k+l}\big)_{k,l=0}^{2N-L,L}$ in Section 3. Finally, some numerical examples are presented in Section 4.

2 Classical Prony method

The classical Prony method works with exact sampled data. Following an idea of G. R. de Prony from 1795 (see e.g. [8, pp. 303–310]), we regard the sampled data $h(k)$ $(k = 0, \ldots, 2N)$ as solution of a *homogeneous linear difference equation with constant coefficients*. If

$$h(k) = \sum_{j=1}^{M} c_j \left(\mathrm{e}^{f_j}\right)^k = \sum_{j=1}^{M} c_j\, z_j^k$$

is a solution of certain homogeneous linear difference equation with constant coefficients, then z_j $(j = 1, \ldots, M)$ must be zeros of the corresponding characteristic polynomial. Thus

$$P_0(z) = \prod_{j=1}^{M} (z - z_j) = \sum_{l=0}^{M} p_l\, z^l \quad (z \in \mathbb{C}) \tag{2.1}$$

with $p_M = 1$ is the monic characteristic polynomial of minimal degree. With these complex coefficients p_k $(k = 0, \ldots, M)$ and complex unknowns x_l , we compose the homogeneous linear difference equation

$$\sum_{l=0}^{M} x_{l+m}\, p_l = 0 \quad (m = 0,\, 1,\, \ldots), \tag{2.2}$$

which obviously has P_0 as characteristic polynomial. Consequently, (2.2) has the complex general solution

$$x_k = \sum_{j=1}^{M} c_j\, z_j^k \quad (k = 0,\, 1,\, \ldots)$$

with arbitrary coefficients $c_j \in \mathbb{C}$ $(j = 1, \ldots, M)$. Then we determine c_j $(j = 1, \ldots, M)$ in such a way that $x_k \approx h(k)$ $(k = 0, \ldots, 2N)$. To this end, we compute the least squares solution of the *overdetermined linear Vandermonde–type system*

$$\sum_{j=1}^{M} c_j\, z_j^k = h(k) \quad (k = 0, \ldots, 2N).$$

Let $L \in \mathbb{N}$ be a convenient upper bound of M, i.e., $M \le L \le N$. In applications, such an upper bound L is mostly known a priori. If this is not the case, then one can choose $L = N$.
The idea of G.R. de Prony is based on the separation of the unknown exponents f_j from the unknown coefficients c_j by means of a homogeneous linear difference equation (2.2). With the $2N+1$ sampled data $h(k) \in \mathbb{C}$ we form the *rectangular Hankel matrix*

$$\mathbf{H} := \big(h(l+m)\big)_{l,m=0}^{2N-L,L} \in \mathbb{C}^{(2N-L+1)\times(L+1)}\,. \tag{2.3}$$

Using the coefficients p_k $(k = 0, \ldots, M)$ of (2.1), we construct the vector $\mathbf{p} := (p_k)_{k=0}^{L}$, where $p_{M+1} = \ldots = p_L := 0$. By $\mathbf{S} := \big(\delta_{k-l-1}\big)_{k,l=0}^{L}$ we denote the forward shift matrix, where δ_k is the Kronecker symbol.

Lemma 2.1 *Let L, M, $N \in \mathbb{N}$ with $M \le L \le N$ be given. Furthermore let $h(k) \in \mathbb{C}$ be given by* (1.2) *with $c_j \in \mathbb{C} \setminus \{0\}$ and pairwise distinct $z_j \in \mathbb{D}$ $(j = 1, \ldots, M)$.*
Then the rectangular Hankel matrix (2.3) *has the singular value* 0*, where*

$$\ker \mathbf{H} = \operatorname{span}\{\mathbf{p}, \mathbf{S}\mathbf{p}, \ldots, \mathbf{S}^{L-M}\mathbf{p}\}$$

and

$$\dim(\ker \mathbf{H}) = L - M + 1\,.$$

Proof. 1. From

$$\sum_{l=0}^{M} h(l+m)\, p_l = 0 \quad (m = 0, \ldots, 2N-M)$$

it follows that

$$\mathbf{H}\big(\mathbf{S}^j\mathbf{p}\big) = \mathbf{o} \quad (j = 0, \ldots, L-M)\,,$$

where $\mathbf{o}$ denotes the zero vector. By $p_M = 1$ we see immediately that the vectors $\mathbf{S}^j\mathbf{p}$ $(j = 0, \ldots, L-M)$ are linearly independent and located in the kernel $\ker \mathbf{H}$.
2. Now we prove that $\ker \mathbf{H}$ is contained in the span of the vectors $\mathbf{S}^j\mathbf{p}$ $(j = 0, \ldots, L-M)$. Let $\mathbf{u} = (u_l)_{l=0}^{L} \in \mathbb{C}^{L+1}$ be an arbitrary right singular vector of $\mathbf{H}$ related to the singular value 0 and let P be the polynomial

$$P(z) := \sum_{l=0}^{L} u_l\, z^l \quad (z \in \mathbb{C})\,.$$

Using (1.2), we receive

$$\begin{aligned} 0 &= \sum_{l=0}^{L} h(l+m)\, u_l = \sum_{l=0}^{L} u_l \Big(\sum_{j=1}^{M} c_j\, z_j^{l+m}\Big) \\ &= \sum_{j=1}^{M} c_j\, z_j^m\, P(z_j) \quad (m = 0, \ldots, 2N-L) \end{aligned}$$

and hence

$$\big(z_j^m\big)_{m=0,j=1}^{2N-M,M} \big(c_j\, P(z_j)\big)_{j=1}^{M} = \mathbf{o}\,.$$

Since $z_j \in \mathbb{D}$ are pairwise different by assumption, $\big(z_j^m\big)_{m=0,j=1}^{M-1,M}$ is a regular Vandermonde matrix such that by $c_j \neq 0$ we obtain $P(z_j) = 0$ $(j = 1, \ldots, M)$. Thus it follows that $P(z) = P_0(z)\, P_1(z)$ with a certain polynomial

$$P_1(z) = \sum_{k=0}^{L-M} \beta_k\, z^k \quad (\beta_k \in \mathbb{C})\,.$$

But this means for the coefficients of P, P_0, and P_1 that

$$\mathbf{u} = \beta_0\, \mathbf{p} + \beta_1\, \mathbf{Sp} + \ldots + \beta_{L-M}\, \mathbf{S}^{L-M}\mathbf{p}\,.$$

Hence it follows that the vectors $\mathbf{S}^j\mathbf{p}$ $(j = 0, \ldots, L-M)$ compose a basis of $\ker \mathbf{H}$ and we obtain $\dim(\ker \mathbf{H}) = L - M + 1$. This completes the proof. ■

The Prony method is based on following

Lemma 2.2 *Let L, M, $N \in \mathbb{N}$ with $M \le L \le N$ be given. Let (1.2) be exact sampled data with $c_j \in \mathbb{C} \setminus \{0\}$ and pairwise distinct $z_j \in \mathbb{D}$.*
Then following assertions are equivalent:
(i) *The polynomial*

$$P(z) = \sum_{k=0}^{L} u_k\, z^k \quad (z \in \mathbb{C}) \tag{2.4}$$

with complex coefficients u_k $(k = 0, \ldots, L)$ has M different zeros z_j $(j = 1, \ldots, M)$.
(ii) 0 *is a singular value of the complex rectangular Hankel matrix* (2.3) *with a right singular vector $\mathbf{u} := (u_l)_{l=0}^{L} \in \mathbb{C}^{L+1}$.*

Proof. 1. Assume that $P(z_j) = 0$ $(j = 1, \ldots, M)$. We compute the sums

$$\sum_{l=0}^{L} h(l+m)\, u_l \quad (m = 0, \ldots, 2N-L)$$

by using (1.2) and obtain for $m = 0, \ldots, 2N-L$

$$\sum_{l=0}^{L} h(l+m)\, u_l = \sum_{l=0}^{L} u_l \Big(\sum_{j=1}^{M} c_j\, z_j^{l+m}\Big) = \sum_{j=1}^{M} c_j\, z_j^m\, P(z_j) = 0\,.$$

Therefore we get $\mathbf{Hu} = \mathbf{o}$.
2. Assume that $\mathbf{u}$ is a right singular vector of $\mathbf{H}$ related to the singular value 0. Following the same lines as in step 2 of the proof of Lemma 2.1, we can see that $P(z_j) = 0$ $(j = 1, \ldots, M)$. ■

Now we formulate Lemma 2.2 as

Algorithm 2.3 (*Classical Prony Method*)

Input: $L, N \in \mathbb{N}$ ($N \gg 1$, $3 \le L \le N$, L *is upper bound of the number of exponentials*), $h(k) \in \mathbb{C}$ $(k = 0, \ldots, 2N)$, $0 < \varepsilon \ll 1$.

1. *Compute a right singular vector* $\mathbf{u} = (u_l)_{l=0}^{L}$ *corresponding to the singular value* 0 *of the exact rectangular Hankel matrix* (2.3).
2. *Form the corresponding polynomial* (2.4) *and evaluate all zeros* $z_j \in \mathbb{D}$ $(j = 1, \ldots, \tilde{M})$. *Note that* $L \ge \tilde{M}$.
3. *Compute* $c_j \in \mathbb{C}$ $(j = 1, \ldots, \tilde{M})$ *as least squares solution of the overdetermined linear Vandermonde–type system*

$$\sum_{j=1}^{\tilde{M}} c_j\, z_j^k = h(k) \quad (k = 0, \ldots, 2N). \tag{2.5}$$

4. *Cancel all that pairs* (z_l, c_l) $(l \in \{1, \ldots, \tilde{M}\})$ *with* $|c_l| \le \varepsilon$ *and denote the remaining set by* $\{(z_j, c_j) : j = 1, \ldots, M\}$ *with* $M \le \tilde{M}$. *Form* $f_j := \log z_j$ $(j = 1, \ldots, M)$, *where* log *is the principal value of the complex logarithm.*

Output: $M \in \mathbb{N}$, $c_j \in \mathbb{C}$, $f_j \in \mathbb{F}$ $(j = 1, \ldots, M)$.

We emphasize that the Algorithm 2.3 is only of theoretical interest. Now we can better explain the new approximate Prony method in Section 3.

Remark 2.4 *We can determine all roots of the polynomial* (2.4) *with* $u_L = 1$ *by computing the eigenvalues of the companion matrix*

$$\mathbf{U} := \begin{pmatrix} 0 & 0 & \ldots & 0 & -u_0 \\ 1 & 0 & \ldots & 0 & -u_1 \\ 0 & 1 & \ldots & 0 & -u_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -u_{L-1} \end{pmatrix} \in \mathbb{C}^{L \times L}.$$

This follows immediately from the fact $P(z) = \det(z\,\mathbf{I} - \mathbf{U})$.
Note that we consider a rectangular Hankel matrix (2.3) *with only* L *columns in order to determine the zeros of a polynomial* (2.4) *with relatively low degree* L *(see step 2 of Algorithm* 2.3*).* □

Remark 2.5 *Let* $N > M$. *If one knows* M *or a good upper bound of* M, *then one can use the following least squares Prony method (see e.g.* [4]*). Since the leading coefficient*

p_M of the characteristic polynomial P_0 is equal to 1, (2.2) *gives rise to the overdetermined linear system*

$$\sum_{l=0}^{M-1} h(l+m)\, p_l = -h(M+m)\, p_M = -h(M+m) \quad (m = 0, \ldots, 2N-M)\,, \tag{2.6}$$

which can be solved by a least squares method. See also the relation to the classic Yule–Walker system [3]. □

3 Approximate Prony method

In contrast to [2], we present a new approximate Prony method by means of perturbation theory for the singular value decomposition of a rectangular Hankel matrix. In praxis, only perturbed values $\tilde{h}_k := h(k) + e_k \in \mathbb{C}$ $(k = 0, \ldots, 2N)$ of the exact sampled data $h(k)$ of an exponential sum (1.1) are known. Here we assume that $|e_k| \le \varepsilon_1$ with certain accuracy $\varepsilon_1 > 0$ such that the error Hankel matrix

$$\mathbf{E} := \big(e_{k+l}\big)_{k,l=0}^{2N-L,L} \in \mathbb{C}^{(2N-L+1)\times(L+1)}$$

has a small spectral norm by

$$\|\mathbf{E}\|_2 \le \sqrt{\|\mathbf{E}\|_1\, \|\mathbf{E}\|_\infty} \le \sqrt{(L+1)\,(2N-L+1)}\,\varepsilon_1 \le (N+1)\,\varepsilon_1\,. \tag{3.1}$$

Then the *perturbed rectangular Hankel matrix* can be represented by

$$\tilde{\mathbf{H}} := \big(\tilde{h}_{k+l}\big)_{k,l=0}^{2N-L,L} = \mathbf{H} + \mathbf{E} \in \mathbb{C}^{(2N-L+1)\times(L+1)}\,. \tag{3.2}$$

By the singular value decomposition of the complex rectangular matrix $\tilde{\mathbf{H}}$ (see [6, pp. 414–415]), there exist two unitary matrices $\tilde{\mathbf{V}} \in \mathbb{C}^{(2N-L+1)\times(2N-L+1)}$, $\tilde{\mathbf{U}} \in \mathbb{C}^{(L+1)\times(L+1)}$ and a rectangular diagonal matrix $\tilde{\mathbf{D}} = \big(\tilde{\sigma}_k\, \delta_{j-k}\big)_{j,k=0}^{2N-L,L}$ with $\tilde{\sigma}_0 \ge \tilde{\sigma}_1 \ge \ldots \ge \tilde{\sigma}_L \ge 0$ such that

$$\tilde{\mathbf{H}} = \tilde{\mathbf{V}}\, \tilde{\mathbf{D}}\, \tilde{\mathbf{U}}^{\mathrm{H}}\,. \tag{3.3}$$

By (3.3), the orthonormal columns $\tilde{\mathbf{v}}_k \in \mathbb{C}^{2N-L+1}$ $(k = 0, \ldots, 2N-L)$ of $\tilde{\mathbf{V}}$ and $\tilde{\mathbf{u}}_k \in \mathbb{C}^{L+1}$ $(k = 0, \ldots, L)$ of $\tilde{\mathbf{U}}$ fulfill the conditions

$$\tilde{\mathbf{H}}\, \tilde{\mathbf{u}}_k = \tilde{\sigma}_k\, \tilde{\mathbf{v}}_k, \quad \tilde{\mathbf{H}}^{\mathrm{H}}\, \tilde{\mathbf{v}}_k = \tilde{\sigma}_k\, \tilde{\mathbf{u}}_k \quad (k = 0, \ldots, L), \tag{3.4}$$

i.e., $\tilde{\mathbf{u}}_k$ is a *right singular vector* and $\tilde{\mathbf{v}}_k$ is a *left singular vector* of $\tilde{\mathbf{H}}$ related to the *singular value* $\tilde{\sigma}_k \ge 0$ (see [6, p. 415]).
Note that $\sigma \ge 0$ is a singular value of the exact rectangular Hankel matrix $\mathbf{H}$ if and only if σ^2 is an eigenvalue of the Hermitian and positive semidefinite matrix $\mathbf{H}^{\mathrm{H}}\, \mathbf{H}$ (see [6, p. 414]). All eigenvalues of $\mathbf{H}^{\mathrm{H}}\, \mathbf{H}$ are nonnegative. Let $\sigma_0 \ge \sigma_1 \ge \ldots \ge \sigma_L \ge 0$ be the ordered singular values of the exact Hankel matrix $\mathbf{H}$. Note that $\ker \mathbf{H} = \ker \mathbf{H}^{\mathrm{H}}\, \mathbf{H}$, since obviously $\ker \mathbf{H} \subseteq \ker \mathbf{H}^{\mathrm{H}}\, \mathbf{H}$ and since from $\mathbf{u} \in \ker \mathbf{H}^{\mathrm{H}}\, \mathbf{H}$ it follows that

$$0 = (\mathbf{H}^{\mathrm{H}}\, \mathbf{H}\, \mathbf{u},\, \mathbf{u}) = \|\mathbf{H}\, \mathbf{u}\|^2\,,$$

i.e., $\mathbf{u} \in \ker \mathbf{H}$. Then by Lemma 2.1, we know that

$$\dim (\ker \mathbf{H}^{\rm H}\, \mathbf{H}) = L - M + 1,$$

and hence $\sigma_{M-1} > 0$ and $\sigma_k = 0$ $(k = M, \ldots, L)$. Then the basic perturbation bound for the singular values reads as follows (see [6, p. 419])

$$|\tilde{\sigma}_k - \sigma_k| \le \|\mathbf{E}\|_2 \quad (k = 0, \ldots, L)\,.$$

Thus $L - M + 1$ singular values of $\tilde{\mathbf{H}}$ are contained in $[0,\, \|\mathbf{E}\|_2]$, if the positive singular value σ_{M-1} of $\mathbf{H}$ is larger than $2\,\|\mathbf{E}\|_2$. In the following we use this property and evaluate a small singular value $\tilde{\sigma}$ $(0 \le \tilde{\sigma} \le \|\mathbf{E}\|_2)$ and a corresponding singular vector of the matrix $\tilde{\mathbf{H}}$.

The classical Prony method is known to perform poorly when noisy data are given. Therefore numerous modifications were attempted to improve the numerical behavior of the classical Prony method. Recently, a very interesting approach is described by G. Beylkin and L. Monzón [2]. Here we generalize a result of [2] to a perturbed rectangular Hankel matrix:

Theorem 3.1 (*cf.* [2, 11]) *Let $\tilde{\sigma} \in (0,\, \varepsilon_2]$ $(0 < \varepsilon_2 \ll 1)$ be a small singular value of the perturbed rectangular Hankel matrix (3.2) with a right singular vector $\tilde{\mathbf{u}} = \big(\tilde{u}_k\big)_{k=0}^{L} \in \mathbb{C}^{L+1}$ and a left singular vector $\tilde{\mathbf{v}} = \big(\tilde{v}_k\big)_{k=0}^{2N-L} \in \mathbb{C}^{2N-L+1}$. Assume that the polynomial*

$$\tilde{P}(z) = \sum_{k=0}^{L} \tilde{u}_k\, z^k \quad (z \in \mathbb{C}) \tag{3.5}$$

has L pairwise distinct zeros $\tilde{z}_n \in \mathbb{C}$ $(n = 1, \ldots, L)$. Further let $K > 2N$.
Then there exists a unique vector $(a_n)_{n=1}^{L} \in \mathbb{C}^L$ such that

$$\tilde{h}_k = \sum_{n=1}^{L} a_n\, \tilde{z}_n^k + \tilde{\sigma}\, d_k \quad (k = 0, \ldots, 2N)\,,$$

where the vector $\big(d_k\big)_{k=0}^{K-1} \in \mathbb{C}^K$ is defined by

$$d_k := \frac{1}{K} \sum_{l=0}^{K-1} \hat{d}_l\, {\rm e}^{2\pi {\rm i} kl/K} \quad (k = 0,\, \ldots,\, K-1)$$

with

$$\hat{d}_l := \begin{cases} \hat{v}_l / \bar{\hat{u}}_l & \text{if } \hat{u}_l \ne 0, \\ 1 & \text{if } \hat{u}_l = 0, \end{cases}$$

where

$$\hat{u}_l := \sum_{k=0}^{K-1} \bar{\tilde{u}}_k\, {\rm e}^{-2\pi {\rm i} kl/K} \quad (l = 0,\, \ldots,\, K-1)\,,$$

$$\hat{v}_l := \sum_{k=0}^{K-1} \tilde{v}_k\, {\rm e}^{-2\pi {\rm i} kl/K} \quad (l = 0,\, \ldots,\, K-1)$$

with the complex conjugate $\bar{\tilde{u}}_k$ *of* $\tilde{u}_k$*. The vector* $(a_n)_{n=1}^L$ *can be computed as solution of the linear Vandermonde system*

$$\sum_{n=1}^{L} a_n \, \tilde{z}_n^k = \tilde{h}_k - \tilde{\sigma}\, d_k \quad (k = 0, \ldots, L-1).$$

Furthermore, if

$$|\hat{v}_l| \le \gamma \, |\hat{u}_l| \quad (l = 0, \ldots, K-1)$$

for certain constant $\gamma > 0$*, then*

$$\sum_{k=0}^{K-1} |d_k|^2 \le \gamma^2 , \quad |h(k) - \sum_{n=1}^{L} a_n \, \tilde{z}_n^k| \le \varepsilon_1 + \gamma \, \varepsilon_2 \quad (k = 0, \ldots, 2N) \, .$$

For shortness, the proof is omitted here (see [11]).

Remark 3.2 *This Theorem* 3.1 *yields a different representation for each* $K > 2N$ *even though* $\tilde{z}_n$ *and* $\tilde{\sigma}$ *remain the same. If* K *is chosen as power of 2, then the entries* $\hat{u}_l$, $\hat{v}_l$ *and* d_k $(l,\, k = 0, \ldots, K-1)$ *can be computed by fast Fourier transforms. Note that the least squares solution* $\left(\tilde{b}_n\right)_{n=1}^L$ *of the overdetermined linear system*

$$\sum_{n=1}^{L} \tilde{b}_n \, \tilde{z}_n^k = \tilde{h}_k \quad (k = 0, \ldots, 2N) \tag{3.6}$$

has an error with Euclidean norm less than $\gamma \, \varepsilon_2$*, since*

$$\begin{aligned}\sum_{k=0}^{2N} |\tilde{h}_k - \sum_{n=1}^{L} \tilde{b}_n \, \tilde{z}_n^k|^2 \;&\le\; \sum_{k=0}^{2N} |\tilde{h}_k - \sum_{n=1}^{L} a_n \, \tilde{z}_n^k|^2 \\ &=\; \sum_{k=0}^{2N} |\tilde{\sigma}\, d_k|^2 \le \tilde{\sigma}^2 \sum_{k=0}^{L-1} |d_k|^2 \le (\gamma \, \varepsilon_2)^2.\end{aligned}$$

□

Thus a first algorithm of the APM reads as follows:

Algorithm 3.3 (APM 1)

Input: $L,\, N \in \mathbb{N}$ $(3 \le L \le N$, *L is upper bound of the number of exponentials*$)$, $\tilde{h}_k \in \mathbb{C}$ $(k = 0, \ldots, 2N)$, *accuracies* $\varepsilon_1,\, \varepsilon_2 > 0$.

1. *Compute a small singular value* $\tilde{\sigma} \in (0,\, \varepsilon_2]$ *and a corresponding singular vector* $\tilde{\mathbf{u}} = \left(\tilde{u}_l\right)_{l=0}^L \in \mathbb{C}^{L+1}$ *of the perturbed rectangular Hankel matrix* (3.2).
2. *Determine all zeros* $\tilde{z}_n \in \mathbb{C}$ $(n = 1, \ldots, L)$ *of the corresponding polynomial* (3.5). *Assume that all zeros of* (3.5) *are simple.*

3. *Determine the least squares solution* $\left(\tilde{b}_n\right)_{n=1}^{L} \in \mathbb{C}^L$ *of the overdetermined linear Vandermonde–type system*

$$\sum_{n=1}^{L} \tilde{b}_n\, \tilde{z}_n^k = \tilde{h}_k \quad (k = 0, \ldots, 2N)\,.$$

4. *Denote by* $(z_j,\, b_j)$ $(j = 1, \ldots, M)$ *all the pairs* $(\tilde{z}_k,\, \tilde{b}_k)$ $(k = 1, \ldots, L)$ *with the properties* $\tilde{w}_k \in \mathbb{D}$ *and* $|\tilde{b}_k| \geq \varepsilon_1$. *Set* $f_j = \log z_j \in \mathbb{F}$ $(j = 1, \ldots, M)$.

Output: $M \in \mathbb{N}$, $b_j \in \mathbb{C}$, $f_j \in \mathbb{F}$ $(j = 1, \ldots, M)$.

Similarly as in [2], we are not interested in exact representations of the sampled values

$$\tilde{h}_k = \sum_{n=1}^{L} \tilde{b}_n\, \tilde{z}_n^k \quad (k = 0, \ldots, 2N)$$

but rather in *approximate representations*

$$\big|\tilde{h}_k - \sum_{j=1}^{M} b_j\, z_j^k\big| \leq \varepsilon \quad (k = 0, \ldots, 2N)$$

for very small accuracy $\varepsilon > 0$ and minimal number M of nontrivial summands.

Now we present a second APM by means of matrix perturbation theory. First we introduce the *rectangular Vandermonde–type matrix*

$$\mathbf{V} := \big(z_j^k\big)_{k=0,j=1}^{2N-L,M} \in \mathbb{C}^{(2N-L+1)\times M}\,. \tag{3.7}$$

Note that in the special case $|z_j| = 1$ $(j = 1, \ldots, M)$, $\mathbf{V}$ is a *nonequispaced Fourier matrix* (see [10]). We discuss the properties of $\mathbf{V}$. Especially, we show that $\mathbf{V}$ is left–invertible and estimate the spectral norm of its left inverse.

Theorem 3.4 *Let* L, M, $N \in \mathbb{N}$ *with* $M \leq L \leq N$ *be given. Let* h *be an exponential sum* (1.1) *with* $c_j \in \mathbb{C} \setminus \{0\}$ *and pairwise different* $z_j = \mathrm{e}^{f_j} \in \mathbb{D}$ $(j = 1, \ldots, M)$. *Assume that the perturbed Hankel matrix* (3.2) *has* $\tilde{\sigma} \in (0, \varepsilon_2]$ $(0 < \varepsilon_2 \leq \|\mathbf{E}\|_2)$ *as singular value with the corresponding right/left singular vectors* $\tilde{\mathbf{u}} = (\tilde{u}_n)_{n=0}^{L} \in \mathbb{C}^{L+1}$ *and* $\tilde{\mathbf{v}} = (\tilde{v}_n)_{n=0}^{2N-L} \in \mathbb{C}^{2N-L+1}$, *respectively. Let*

$$\tilde{P}(z) := \sum_{k=0}^{L} \tilde{u}_k\, z^k \quad (z \in \mathbb{C})$$

be the polynomial related to $\tilde{\mathbf{u}}$.
Then the rectangular Vandermonde–type matrix (3.7) *has a left inverse* $\mathbf{L} = (\mathbf{V}^{\mathrm{H}}\mathbf{V})^{-1}\mathbf{V}^{\mathrm{H}}$. *Further the values* $\tilde{P}(z_j)$ $(j = 1, \ldots, M)$ *fulfill the estimate*

$$\sum_{j=1}^{M} |c_j|^2 |\tilde{P}(z_j)|^2 \leq (\varepsilon_2 + \|\mathbf{E}\|_2)^2\, \|\mathbf{L}\|_2^2\,.$$

Proof. 1. By assumption we have $\tilde{\mathbf{H}}\,\tilde{\mathbf{u}} = \tilde{\sigma}\,\tilde{\mathbf{v}}$, i.e.,

$$\sum_{l=0}^{L} \tilde{h}_{l+m}\,\tilde{u}_l = \tilde{\sigma}\,\tilde{v}_m \quad (m = 0,\,\ldots,\,2N-L)\,.$$

Using (1.2) and $\tilde{h}_k = h(k) + e_k$ $(k = 0, \ldots, 2N)$, we receive the $2N - L + 1$ equations

$$\sum_{j=1}^{M} c_j\,\tilde{P}(z_j)\,z_j^k = \tilde{\sigma}\,\tilde{v}_k - \sum_{l=0}^{L} e_{l+k}\,\tilde{u}_l \quad (k = 0, \ldots, 2N-L). \tag{3.8}$$

2. Using the matrix–vector notation of (3.8) with the rectangular Vandermonde–type matrix $\mathbf{V}$, we obtain

$$\mathbf{V}\left(c_j\,\tilde{P}(z_j)\right)_{j=1}^{M} = \tilde{\sigma}\,\tilde{\mathbf{v}} - \mathbf{E}\,\tilde{\mathbf{u}}.$$

By $N \ge L \ge M$ the matrix $\mathbf{V}$ has full rank M, since the quadratic submatrix $\left(z_j^k\right)_{k=0,j=1}^{M-1,M}$ of order M is a regular Vandermonde matrix. Hence $\mathbf{V}^{\mathrm{H}}\mathbf{V}$ is Hermitian and positive definite. Thus the matrix $\mathbf{V}$ has a left inverse $\mathbf{L} = (\mathbf{V}^{\mathrm{H}}\mathbf{V})^{-1}\mathbf{V}^{\mathrm{H}}$ such that

$$\left(c_j\,\tilde{P}(z_j^k)\right)_{j=1}^{M} = \tilde{\sigma}\,\mathbf{L}\,\tilde{\mathbf{v}} - \mathbf{L}\,\mathbf{E}\,\tilde{\mathbf{u}}\,.$$

Then by $\|\tilde{\mathbf{u}}\| = \|\tilde{\mathbf{v}}\| = 1$ it follows that

$$\sum_{j=1}^{M} |c_j|^2\,|\tilde{P}(z_j)|^2 \le (\tilde{\sigma} + \|\mathbf{E}\|_2)^2\,\|\mathbf{L}\|_2^2 \le (\varepsilon_2 + \|\mathbf{E}\|_2)^2\,\|\mathbf{L}\|_2^2\,.$$

This completes the proof. ■

In [1], the condition number of the rectangular Vandermonde–type matrix $\mathbf{V}$ is estimated. It is shown that this matrix is well conditioned, provided the nodes z_j are close to the unit circle but not extremely close to each other and provided N is large enough. Now we formulate an analog of the Rayleigh–Ritz Theorem (see [6, pp. 176–178]) for singular values of the symmetric Hankel matrix $\mathbf{H}$:

Lemma 3.5 *Let L, M, $N \in \mathbb{N}$ with $M \le L \le N$ be given. Assume that the ordered singular values of the exact rectangular Hankel matrix $\mathbf{H} \in \mathbb{C}^{(2N-L+1)\times(L+1)}$ are $\sigma_0 \ge \ldots \ge \sigma_{M-1} > 0$ and $\sigma_M = \ldots = \sigma_L = 0$. Then for all $\mathbf{x} \in \mathbb{C}^{L+1}$ with $\mathbf{x} \perp \ker \mathbf{H}$*

$$\sigma_{M-1}\,\|\mathbf{x}\| \le \|\mathbf{H}\,\mathbf{x}\| \le \sigma_0\,\|\mathbf{x}\|\,.$$

Proof. 1. By the assumptions, σ_k^2 $(k = 0, \ldots, L)$ are the eigenvalues of the Hermitian matrix $\mathbf{H}^{\mathrm{H}}\,\mathbf{H}$. As usual, we denote the ordered eigenvalues of $\mathbf{H}^{\mathrm{H}}\,\mathbf{H}$ by $\lambda_0 = \ldots = \lambda_{L-M} = 0$, $0 < \lambda_{L-M+1} \le \ldots \le \lambda_L$ such that $\sigma_{M-1}^2 = \lambda_{L-M+1}$. Here we have used Lemma 2.1 and $\ker \mathbf{H}^{\mathrm{H}}\,\mathbf{H} = \ker \mathbf{H}$. Now we apply the spectral theorem for the Hermitian matrix $\mathbf{H}^{\mathrm{H}}\,\mathbf{H}$ (see [6, p. 171]). Then there exists a unitary matrix $\mathbf{U} \in \mathbb{C}^{(L+1)\times(L+1)}$ (see [6, pp. 204–205]) such that

$$\mathbf{H}^{\mathrm{H}}\,\mathbf{H} = \mathbf{U}\left(\operatorname{diag}(\lambda_k)_{k=0}^{L}\right)\mathbf{U}^{\mathrm{H}}\,. \tag{3.9}$$

Let $\mathbf{u}_k \in \mathbb{C}^{L+1}$ $(k = 0, \ldots, L)$ be the k-th column vector of $\mathbf{U}$. Then by (3.9) we receive that

$$\mathbf{H}^{\mathrm{H}}\,\mathbf{H}\,\mathbf{u}_k = \lambda_k\,\mathbf{u}_k \quad (k = 0, \ldots, L)\,.$$

From $\lambda_0 = \ldots = \lambda_{L-M} = 0$ it follows that $\mathbf{u}_k \in \ker \mathbf{H}^{\mathrm{H}}\,\mathbf{H} = \ker \mathbf{H}$ $(k = 0, \ldots, L - M)$ which form an orthonormal basis of $\ker \mathbf{H}$.
2. For arbitrary $\mathbf{x} \in \mathbb{C}^{L+1}$, we obtain by (3.9) that

$$\|\mathbf{H}\,\mathbf{x}\|^2 = \mathbf{x}^{\mathrm{H}}\,\mathbf{H}^{\mathrm{H}}\,\mathbf{H}\,\mathbf{x} = \mathbf{x}^{\mathrm{H}}\,\mathbf{U}\,\left(\operatorname{diag}(\lambda_k)_{k=0}^{L}\right)\,\mathbf{U}^{\mathrm{H}}\,\mathbf{x} = \sum_{k=L-M+1}^{L} \lambda_k\,|(\mathbf{u}_k,\,\mathbf{x})|^2\,. \tag{3.10}$$

The condition $\mathbf{x} \perp \ker \mathbf{H}$ is equivalent to $(\mathbf{u}_k,\,\mathbf{x}) = 0$ $(k = 0, \ldots, L - M)$. Thus for all $\mathbf{x} \in \mathbb{C}^{L+1}$ with $\mathbf{x} \perp \ker \mathbf{H}$ we obtain that

$$\|\mathbf{x}\|^2 = \|\mathbf{U}\,\mathbf{x}\|^2 = \sum_{k=L-M+1}^{L} |(\mathbf{u}_k,\,\mathbf{x})|^2\,. \tag{3.11}$$

Thus from (3.10) and (3.11) it follows that for all $\mathbf{x} \in \mathbb{C}^{L+1}$ with $\mathbf{x} \perp \ker \mathbf{H}$

$$\lambda_{L-M+1}\,\|\mathbf{x}\|^2 \le \|\mathbf{H}\,\mathbf{x}\|^2 \le \lambda_L\,\|\mathbf{x}\|^2\,.$$

This completes the proof. ■

Lemma 3.6 *If the assumptions of Theorem* 3.4 *are fulfilled with sufficiently small accuracies* $\varepsilon_1,\, \varepsilon_2 > 0$ *and if* $\sigma_{M-1} > 0$ *is the smallest singular value* $\neq 0$ *of* $\mathbf{H}$*, then*

$$\|\tilde{\mathbf{u}} - \mathbf{P}\,\tilde{\mathbf{u}}\| \le \frac{\varepsilon_2 + (N+1)\,\varepsilon_1}{\sigma_{M-1}}\,, \tag{3.12}$$

where $\mathbf{P}$ *is the orthogonal projector of* $\mathbb{C}^{L+1}$ *onto* $\ker \mathbf{H}$*. Furthermore the polynomial* (3.5) *has zeros close to* $z_j \in \mathbb{D}$ $(j = 1, \ldots, M)$*, where*

$$\sum_{j=1}^{M} |\tilde{P}(z_j)|^2 \le (2N - L + 1)M\left(\frac{\varepsilon_2 + (N+1)\varepsilon_1}{\sigma_{M-1}}\right)^2\,.$$

Proof. 1. Let $\tilde{\mathbf{u}}$ be a right singular vector of $\tilde{\mathbf{H}}$ with respect to the singular value $\tilde{\sigma} \in [0,\, \varepsilon_2]$. Using Lemma 3.5, we receive for all $\mathbf{u} \in \mathbb{C}^{L+1}$ with $\mathbf{u} \perp \ker \mathbf{H}$ that

$$\sigma_{M-1}\,\|\mathbf{u}\| \le \|\mathbf{H}\,\mathbf{u}\|$$

i.e., the following estimate

$$\sigma_{M-1}\,\|\tilde{\mathbf{u}}-\mathbf{u}\| \le \|\mathbf{H}(\tilde{\mathbf{u}}-\mathbf{u})\|$$

is valid for all $\mathbf{u} \in \mathbb{C}^{L+1}$ with $\tilde{\mathbf{u}} - \mathbf{u} \perp \ker \mathbf{H}$. Especially for $\mathbf{u} = \mathbf{P}\,\tilde{\mathbf{u}}$, we see that $\tilde{\mathbf{u}} - \mathbf{P}\,\tilde{\mathbf{u}} \perp \ker \mathbf{H}$ and hence by (3.1)

$$\begin{aligned}\sigma_{M-1}\,\|\tilde{\mathbf{u}}-\mathbf{P}\,\tilde{\mathbf{u}}\| &\le \|\mathbf{H}\,\tilde{\mathbf{u}}\| = \|(\tilde{\mathbf{H}}-\mathbf{E})\,\tilde{\mathbf{u}}\| = \|\tilde{\sigma}\,\tilde{\mathbf{u}} - \mathbf{E}\,\tilde{\mathbf{u}}\| \\ &\le \tilde{\sigma} + \|\mathbf{E}\|_2 \le \varepsilon_2 + (N+1)\,\varepsilon_1\end{aligned}$$

such that (3.12) follows.
2. Thereby $\mathbf{u} = \mathbf{P}\,\tilde{\mathbf{u}}$ is a right singular vector of $\mathbf{H}$ with respect to the singular value 0. Thus the corresponding polynomial P has the values $z_j \in \mathbb{D}$ $(j = 1, \ldots, M)$ as zeros by Lemma 2.1. By (3.12), the coefficients of P differ only a little from the coefficients of $\tilde{P}$ with respect to $\tilde{\mathbf{u}}$. Consequently, the zeros of $\tilde{P}$ lie nearby the zeros of P, i.e., $\tilde{P}$ has zeros close to z_j $(j = 1, \ldots, M)$ (see [6, pp. 539–540]).
By $\|\mathbf{V}^{\mathrm{H}}\|_2^2 \le \sum_{k=0}^{2N-L} \sum_{j=1}^{M} |z_j^k|^2 \le (2N-L+1)M$ and (3.12), we obtain the estimate

$$\begin{aligned}\sum_{j=1}^{M} |\tilde{P}(z_j)|^2 &= \sum_{j=1}^{M} |P(z_j) - \tilde{P}(z_j)|^2 = \|\mathbf{V}^{\mathrm{H}}(\mathbf{u}-\tilde{\mathbf{u}})\|^2 \\ &\le \|\mathbf{V}^{\mathrm{H}}\|_2^2\,\|\mathbf{u}-\tilde{\mathbf{u}}\|^2 \le (2N-L+1)M \left(\frac{\varepsilon_2 + (N+1)\varepsilon_1}{\sigma_{M-1}}\right)^2 .\end{aligned}$$

This completes the proof. ■

For noisy data we can not assume that a reconstruction yields $z_j \in \mathbb{D}$. Therefore we introduce the disk $\tilde{\mathbb{D}}(r) := \{z \in \mathbb{C} : |z| \le r\}$ with radius r. Now we can formulate a second algorithm of the APM.

Algorithm 3.7 (APM 2)

Input: $L, N \in \mathbb{N}$ $(3 \le L \le N$, *L is upper bound of the number of exponentials*), $\tilde{h}_k = h(k) + e_k \in \mathbb{C}$ $(k = 0, \ldots, 2N)$ *with* $|e_k| \le \varepsilon_1$, *accuracy bounds* ε_1, ε_2, *radius* r.

1. *Compute a right singular vector* $\tilde{\mathbf{u}} = (\tilde{u}_k)_{k=0}^{L}$ *corresponding to a singular value* $\tilde{\sigma} \in (0, \varepsilon_2]$ *of the perturbed rectangular Hankel matrix* (3.2).
2. *Form the corresponding polynomial* (3.5) *and evaluate all zeros* $\tilde{z}_j \in \tilde{\mathbb{D}}(r)$ $(j = 1, \ldots, \tilde{M})$. *Note that* $L \ge \tilde{M}$.
3. *Compute* $\tilde{c}_j \in \mathbb{C}$ $(j = 1, \ldots, \tilde{M})$ *as least squares solution of the overdetermined linear Vandermonde–type system*

$$\sum_{j=1}^{\tilde{M}} \tilde{c}_j\,\tilde{z}_j^k = \tilde{h}_k \quad (k = 0, \ldots, 2N)\,. \tag{3.13}$$

4. *Delete all the* $\tilde{z}_l$ $(l \in \{1,\dots,\tilde{M}\})$ *with* $|\tilde{c}_l| \le \varepsilon_1$ *and denote the remaining set by* $\{z_j : j = 1,\dots,M\}$ *with* $M \le \tilde{M}$.
5. *Repeat step 3 and solve the overdetermined linear Vandermonde–type system*

$$\sum_{j=1}^{M} c_j\, z_j^k = \tilde{h}_k \quad (k = 0,\dots,2N)$$

with respect to the new set $\{z_j : j = 1,\dots,M\}$ *again. Set* $f_j := \log z_j \in \mathbb{F}$ $(j = 1,\dots,M)$.

Output: $M \in \mathbb{N}$, $c_j \in \mathbb{C}$, $f_j \in \mathbb{F}$ $(j = 1,\dots,M)$.

There exist a variety of algorithms to recover the complex exponents f_j like ESPRIT [12, 13]. Now we replace the steps 1 and 2 of Algorithm 3.7 by the ESPRIT method [7, p. 493]. Note that the following method needs a further parameter P as an estimation for upper bound of the number of exponentials. We choose this parameter $P = L - 1$

Algorithm 3.8 (APM via ESPRIT)

Input: L, N, $P \in \mathbb{N}$ $(3 \le L \le N$, *P is upper bound of the number of exponentials with* $P+1 \le L)$, $\tilde{h}_k = h(k) + e_k \in \mathbb{C}$ $(k = 0,\dots,2N)$ *with* $|e_k| \le \varepsilon_1$, *accuracy bounds* ε_1, *radius* r.

1. *Compute the singular value decomposition of the perturbed rectangular Hankel matrix* (3.2) *with a diagonal matrix* $\mathbf{S} \in \mathbb{R}^{(2N-L+1)\times(L+1)}$ *with nonnegative diagonal elements in decreasing order, and unitary matrices* $\mathbf{L} \in \mathbb{C}^{(2N-L+1)\times(2N-L+1)}$ *and* $\mathbf{U} := \big(u_{k,l}\big)_{k,l=0}^{L} \in \mathbb{C}^{(L+1)\times(L+1)}$ *such that* $\tilde{\mathbf{H}} = \mathbf{L}\,\mathbf{S}\,\mathbf{U}^{\mathrm{H}}$.
2. *For* $\mathbf{U}_1 := (u_{k,l})_{k,l=0}^{L-1,P}$, $\mathbf{U}_2 := (u_{k+1,l})_{k,l=0}^{L-1,P}$ *compute the matrix* $\mathbf{P} := \mathbf{U}_1^{\dagger}\mathbf{U}_2 \in \mathbb{C}^{(P+1)\times(P+1)}$, *where* $\mathbf{U}_1^{\dagger} := (\mathbf{U}_1^{\mathrm{H}}\mathbf{U}_1)^{-1}\mathbf{U}_1^{\mathrm{H}}$ *is the Moore–Penrose pseudoinverse of* $\mathbf{U}_1$. *Compute all the eigenvalues* $\tilde{z}_j \in \mathbb{D}(r)$ $(j = 1,\dots,\tilde{M})$ *of the matrix* $\mathbf{P}$. *Note that* $\tilde{M} \le P + 1$.
3. *Continue with steps* $3-6$ *of Algorithm* 3.7.

Output: $M \in \mathbb{N}$, $c_j \in \mathbb{C}$, $f_j \in \mathbb{F}$ $(j = 1,\dots,M)$.

Further we can replace the steps 1 and 2 of Algorithm 3.7 by solving the overdetermined linear Hankel system (2.6) such that we dispense with the computation of singular values and right/left singular vectors of the perturbed rectangular Hankel matrix (3.2). This approach is justified by the Lemmas 3.5 and 3.6. We denote this modification of Algorithm 3.7 as least squares Prony method (LSPM).

Algorithm 3.9 (LSPM)

Input: L, $N \in \mathbb{N}$ $(3 \le L \le N$, *L is upper bound of the number of exponentials)*, $\tilde{h}_k = h(k) + e_k \in \mathbb{C}$ $(k = 0,\dots,2N)$ *with* $|e_k| \le \varepsilon_1$, *accuracy bounds* ε_1, *radius* r.

1. *Solve the overdetermined linear Hankel system*

$$\sum_{l=0}^{L-1} h(l+m)\, p_l = -h(L+m) \quad (m = 0, \ldots, 2N-L)\,.$$

2. *Compute all the zeros* $\tilde{z}_j \in \tilde{\mathbb{D}}(r)$ $(j = 1, \ldots, \tilde{M})$ *of the polynomial*

$$z^L + \sum_{k=0}^{L-1} p_k\, z^k\,.$$

Note that $\tilde{M} \le L$.

3. *Continue with the steps* $3-6$ *of Algorithm* 3.7.

Output: $M \in \mathbb{N}$, $c_j \in \mathbb{C}$, $f_j \in \mathbb{F}$ $(j = 1, \ldots, M)$.

Note that Algorithm 3.8 requires more arithmetical operations than Algorithm 3.7. The Algorithm 3.9 is the cheapest method.

4 Numerical examples

Now we illustrate the behavior of the suggested Algorithms 3.7 – 3.9. We have implemented our algorithms in MATLAB with IEEE double precision arithmetic. The relative error of the complex exponents is given by

$$e(\boldsymbol{f}) := \frac{\max\limits_{j=1,\ldots,M} |f_j - \tilde{f}_j|}{\max\limits_{j=1,\ldots,M} |f_j|}\,,$$

where $\tilde{f}_j$ are the exponents computed by our algorithms. Analogously, the relative error of the coefficients is defined by

$$e(\boldsymbol{c}) := \frac{\max\limits_{j=1,\ldots,M} |c_j - \tilde{c}_j|}{\max\limits_{j=1,\ldots,M} |c_j|}\,,$$

where $\tilde{c}_j$ are the coefficients computed by our algorithms. Further we determine the relative error of the exponential sum by

$$e(h) := \frac{\max |h(x) - \tilde{h}(x)|}{\max |h(x)|}\,,$$

where the maximum is built from 10000 equidistant points of $[0,\, 2N]$, and where

$$\tilde{h}(x) := \sum_{j=1}^{M} \tilde{c}_j\, \mathrm{e}^{\tilde{f}_j x}$$

is the exponential sum recovered by our algorithms. For increasing number of sampled data, we obtain a very precise parameter estimation. With other words, N acts as regularization parameter of this inverse problem. Algorithm 2.3 can not be applied, since 0 is not a singular value of the perturbed rectangular Hankel matrix (3.2).

Example 4.1 We simulate a typical damped five–peak NMR signal [9]. The signal is given in the form (1.1) with $M = 5$, the complex coefficients

$$(c_1, c_2, c_3, c_4, c_5) = \mathrm{e}^{15\mathrm{i}}\,(6.1, 9.9, 6.0, 2.8, 17)$$

and the complex exponents

$$\begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \end{pmatrix} = \frac{1}{50000} \begin{pmatrix} -208 - 2\pi\mathrm{i} \cdot 1379 \\ -256 - 2\pi\mathrm{i} \cdot 685 \\ -197 - 2\pi\mathrm{i} \cdot 271 \\ -117 + 2\pi\mathrm{i} \cdot 353 \\ -808 + 2\pi\mathrm{i} \cdot 478 \end{pmatrix}.$$

We sample this exponential sum (1.1) at the equidistant nodes $x = k$ $(k = 0, \ldots, 2N)$. Then we apply our algorithms for exact sampled data $h_k = h(k)$, i.e., $e_k = 0$ $(k = 0, \ldots, 2N)$. Thus the accuracy ε_1 can be chosen as the unit roundoff $\varepsilon_1 = 2^{-53} \approx 1.11 \times 10^{-16}$ (see [5, p. 45]) and furthermore we choose the accuracies $\varepsilon_2 = 10^{-5}$, the radiuses $r = 1.1$ for $L = 5$ and $r = 1$ is $L = 100$. We consider noisy sampled data $\tilde{h}_k = h(k) + 10^{-s}\, e_k$ $(k = 0, \ldots, 2N)$, where $(e_k)_{k=0}^{2N}$ is a vector with samples from a normal distribution with mean 1 and standard deviation 2. The third column of the tables contains the signal–to–noise ratio SNR $:= 10 \log_{10} (\|(h(k))_{k=0}^{2N}\| \,/\, \|(e_k)_{k=0}^{2N}\|)$, where ∞ indicates that no noise is added. Note that SNR ≈ 66 for $s = 6$ and SNR ≈ 96 for $s = 9$. We repeat each experiment 100 times and present the averages of the errors in Table 4.1. It is remarkable that we obtain very precise results even in the case, where the unknown number $M = 5$ is estimated by $L = 100$. Furthermore we find out that the cheapest Algorithm 3.9 yields precise results too. □

Example 4.2 This example is often used in testing system identification algorithms (see [1]). We choose $M = 6$, $c_j = 1$ $(j = 1, \ldots, 6)$, and

$$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \end{pmatrix} = \begin{pmatrix} 0.9856 - 0.1628\mathrm{i} \\ 0.9856 + 0.1628\mathrm{i} \\ 0.8976 - 0.4305\mathrm{i} \\ 0.8976 + 0.4305\mathrm{i} \\ 0.8127 - 0.5690\mathrm{i} \\ 0.8127 + 0.5690\mathrm{i} \end{pmatrix}.$$

The results of Algorithms 3.7 – 3.9 are presented in Table 4.2. Note that this example was chosen in [1] in order to justify the relative small upper bounds for the condition number of the rectangular Vandermonde–type matrix $\mathbf{V}$. The upper bound depends on

N	L	SNR	$e(\boldsymbol{f})$	$e(\boldsymbol{c})$	$e(h)$
Algorithm 3.7					
6	5	∞	7.67e-05	5.44e-05	2.48e-14
250	5	∞	1.96e-09	1.52e-08	7.38e-09
250	5	95.7	3.98e-06	1.70e-05	7.34e-06
250	5	65.7	3.82e-03	1.55e-02	7.20e-03
250	100	∞	9.61e-15	2.73e-13	1.71e-13
250	100	95.7	7.30e-11	5.94e-10	1.71e-10
250	100	65.7	7.74e-08	5.28e-07	1.61e-07
Algorithm 3.8					
6	5	∞	7.67e-05	5.44e-05	1.98e-14
250	5	∞	1.25e-09	7.64e-09	3.64e-09
250	5	95.7	3.49e-06	1.60e-05	6.56e-06
250	5	65.7	3.79e-03	1.55e-02	7.02e-03
250	100	∞	1.52e-14	3.07e-13	7.15e-14
250	100	95.7	7.64e-11	6.80e-10	2.23e-10
250	100	65.7	7.92e-08	6.87e-07	1.82e-07
Algorithm 3.9					
6	5	∞	8.40e-05	6.16e-05	2.05e-14
250	5	∞	1.96e-09	1.40e-08	6.86e-09
250	5	95.7	4.00e-06	1.83e-05	7.52e-06
250	5	65.7	4.10e-01	2.71e-01	1.28e-01
250	100	∞	8.57e-15	1.72e-13	9.01e-14
250	100	95.7	2.82e-11	2.42e-10	6.79e-11
250	100	65.7	2.63e-08	2.23e-07	6.54e-08

Table 4.1: Results of Example 4.1.

the separation of the nodes z_j in the unit disk, the departure from normality, $\min_j |z_j|$, and $\max_j |z_j|$. We choose the accuracies $\varepsilon_1 = 10^{-10}$, $\varepsilon_2 = 10^{-5}$ and the radius $r = 1.5$. □

Example 4.3 We choose 30 equidistant nodes on each circle with radius 0.7, 0.8, and 0.9. Using random coefficients $c_j \in [0, 1]$ $(j = 0, \ldots, 89)$, we sample the exponential sum (1.1) at the equidistant nodes $x = k$ $(k = 0, \ldots, 2N)$ without noise. We choose the accuracies $\varepsilon_1 = 10^{-4}$, $\varepsilon_2 = 10^{-20}$ and the radius $r = 1$. Applying the Algorithms 3.7 – 3.9, we present the results in Table 4.3. Here a dash means that we could not found the values z_j. A typical example is presented in Figure 4.1. The given 90 nodes are shown as circles. In Figure 4.1 (left) we show the values $\tilde{z}_j$ after step 2 of Algorithm 3.8 with the

N	L	SNR	$e(\boldsymbol{f})$	$e(\boldsymbol{c})$	$e(h)$
Algorithm 3.7					
7	6	∞	9.78e-12	3.24e-11	5.74e-15
7	6	90.9	1.11e-04	3.48e-04	1.52e-09
7	6	60.7	9.15e-02	5.50e-01	1.63e-06
Algorithm 3.8					
7	6	∞	1.01e-11	3.51e-11	5.92e-15
7	6	90.8	1.22e-04	3.83e-04	1.57e-09
7	6	60.9	8.85e-02	4.52e-01	1.50e-06
Algorithm 3.9					
7	6	∞	1.00e-11	3.74e-11	2.00e-14
7	6	90.9	1.08e-04	3.39e-04	1.55e-09
7	6	60.8	9.68e-02	6.13e-01	1.53e-06

Table 4.2: Results for h from Example 4.2.

parameters $N = 500$ and $L = 400$ as points. Since the Vandermonde–type system (3.13) is very ill–conditioned we cannot separate the 90 correct values. The ill–conditioning of (3.13) is due to the small distance of the points $\tilde{z}_j$ (see [1]). In Figure 4.1 (right) we show the values $\tilde{z}_j$ after step 2 of Algorithm 3.9 with the parameters $N = 500$ and $L = 200$ as points. The given values are indicated as small circles. We cannot recover the values z_j, because we do not find a point in all small circles. Since in the following steps of the algorithm we delete only the additional points, there is no possibility to solve the problem correctly. This is a consequence of the ill–conditioning of the Hankel system in step 1 of Algorithm 3.9. Fortunately our suggested Algorithm 3.7 works, since we separate the true values z_j from the additional values by using the advance information $z_j \in \tilde{\mathbb{D}}(r)$. □

Acknowledgment

The first named author gratefully acknowledges the support by the German Research Foundation within the project KU 2557/1-1.

References

[1] F. S. V. Bazán. Conditioning of rectangular Vandermonde matrices with nodes in the unit disk. *SIAM J. Matrix Anal. Appl.*, 21:679 – 693, 1999.

N	L	$e(\boldsymbol{f})$	$e(\boldsymbol{c})$	$e(h)$
Algorithm 3.7				
500	90	8.99e-06	2.00e-05	6.70e-07
500	200	1.46e-05	1.91e-05	1.09e-06
500	400	1.20e-05	1.10e-05	9.13e-07
Algorithm 3.8				
500	90	8.99e-06	2.00e-05	6.71e-07
500	200	3.73e-02	1.10e-01	1.41e-02
500	400	–	–	–
Algorithm 3.9				
500	90	1.48e-08	5.67e-07	1.27e-08
500	200	–	–	–
500	400	–	–	–

Table 4.3: Results of Example 4.3.

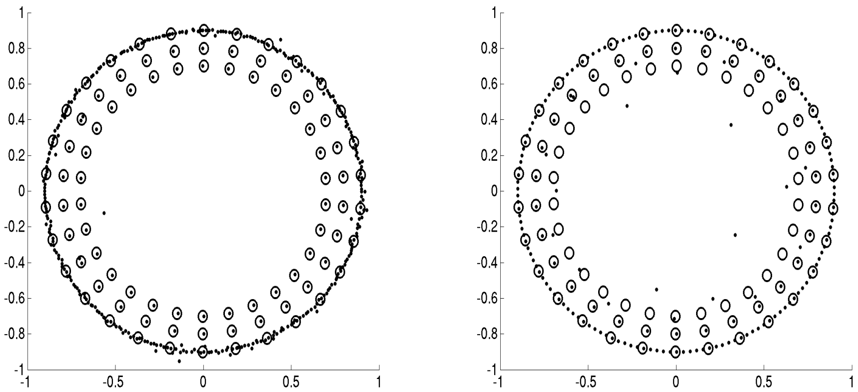

Figure 4.1: The polar grid is generated by 30 equidistant nodes on each circle with radius $0.7, 0.8$, and 0.9, respectively. The given nodes z_j are shown by circles. The reconstructed nodes $\tilde{z}_j$ are indicated as points. Left: The results after step 2 of Algorithm 3.8 with $N = 500$ and $L = 400$. Right: The results after step 2 of Algorithm 3.9 with $N = 500$ and $L = 200$.

[2] G. Beylkin and L. Monzón. On approximations of functions by exponential sums. *Appl. Comput. Harmon. Anal.*, 19:17 – 48, 2005.

[3] P. L. Dragotti, M. Vetterli, and T. Blu. Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang-Fix. *IEEE Trans. Signal*

Process., 55:1741 – 1757, 2007.

[4] G. H. Golub, P. Milanfar, and J. Varah. A stable numerical method for inverting shape from moments. *SIAM J. Sci. Comput.*, 21:1222 – 1243, 1999.

[5] N. J. Higham. *Accuracy and Stability of Numerical Algorithms.* SIAM, Philadelphia, 1996.

[6] R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambridge University Press, Cambridge, 1985.

[7] D. G. Manolakis, V. K. Ingle, and S. M. Kogon. *Statistical and Adaptive Signal Processing.* McGraw-Hill, Boston, 2005.

[8] S. L. Marple. *Digital Spectral Analysis with Applications.* Prentice Hall Inc., Englewood Cliffs, 1987.

[9] J. M. Papy, L. De Lathauwer, and S. Van Huffel. Exponential data fitting using multilinear algebra: the decimative case. *J. Chemometrics*, 23:341 – 351, 2004.

[10] D. Potts, G. Steidl, and M. Tasche. Fast Fourier transforms for nonequispaced data: A tutorial. In: J. J. Benedetto and P. J. S. G. Ferreira (Eds.), *Modern Sampling Theory: Mathematics and Applications*, pages 247 – 270, Birkhäuser, Boston, 2001.

[11] D. Potts and M. Tasche. Parameter estimation for exponential sums by approximate Prony method. *Signal Process., in print.*

[12] R. Roy and T. Kailath. ESPRIT—estimation of signal parameters via rotational invariance techniques. *IEEE Trans. Acoustic Speech Signal Process.*, 37:984 – 994, 1989.

[13] R. Roy and T. Kailath. ESPRIT—estimation of signal parameters via rotational invariance techniques. In: L. Auslander, F. A. Grünbaum, J. W. Helton, T. Kailath, P. Khargoneka, and S. Mitter (Eds.), *Signal Processing, Part II: Control Theory and Applications*, pages 369–411, Springer, New York, 1990.