Parameter estimation for nonincreasing exponential sums by Prony–like methods

Daniel Potts[*] Manfred Tasche[‡]

Let $z_j := \mathrm{e}^{f_j}$ with $f_j \in (-\infty, 0] + \mathrm{i}\,[-\pi, \pi)$ be distinct nodes for $j = 1, \ldots, M$. With complex coefficients $c_j \neq 0$, we consider a nonincreasing exponential sum $h(x) := c_1\,\mathrm{e}^{f_1\,x} + \ldots + c_M\,\mathrm{e}^{f_M\,x}$ $(x \geq 0)$. Many applications in electrical engineering, signal processing, and mathematical physics lead to the following problem: Determine all parameters of h, if $2\,N$ sampled values $h(k)$ $(k = 0, \ldots, 2N-1;\ N \geq M)$ are given. This parameter estimation problem is a nonlinear inverse problem. For noiseless sampled data, we describe the close connections between Prony–like methods, namely the classical Prony method, the matrix pencil method, and the ESPRIT method. Further we present a new efficient algorithm of matrix pencil factorization based on QR decomposition of a rectangular Hankel matrix. The algorithms of parameter estimation are also applied to sparse Fourier approximation and nonlinear approximation.

Key words and phrases: Parameter estimation, nonincreasing exponential sum, Prony–like method, exponential fitting problem, ESPRIT, matrix pencil factorization, companion matrix, Prony polynomial, eigenvalue problem, rectangular Hankel matrix, nonlinear approximation, sparse trigonometric polynomial, sparse Fourier approximation.

AMS *Subject Classifications*: 65D10, 41A45, 65F15, 65F20, 94A12.

1 Introduction

Let $M \geq 1$ be an integer. Let $f_j \in (-\infty, 0] + \mathrm{i}\,[-\pi, \pi)$ $(j = 1, \ldots, M)$ be distinct complex numbers. Further let $c_j \in \mathbb{C} \setminus \{0\}$ $(j = 1, \ldots, M)$. Assume that $|c_j| > \varepsilon$ for

[*]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 of Mathematics, D–18051 Rostock, Germany

convenient bound $0 < \varepsilon \ll 1$. In the following, we consider a *nonincreasing exponential sum of order* M

$$h(x) := \sum_{j=1}^{M} c_j \,\mathrm{e}^{f_j x} \quad (x \ge 0). \tag{1.1}$$

The real part $\mathrm{Re}\, f_j \le 0$ is the *damping factor* of the exponential $\mathrm{e}^{f_j x}$ such that $|\mathrm{e}^{f_j x}|$ is nonincreasing for $x \ge 0$. If $\mathrm{Re}\, f_j < 0$, then $\mathrm{e}^{f_j x}$ $(x \ge 0)$ is a damped exponential. If $\mathrm{Re}\, f_j = 0$, then $\mathrm{e}^{f_j x}$ $(x \ge 0)$ is an undamped exponential. The imaginary part $\mathrm{Im}\, f_j \in [-\pi,\, \pi)$ is the *angular frequency* of the exponential $\mathrm{e}^{f_j x}$. The nodes $z_j := \mathrm{e}^{f_j}$ $(j = 1, \ldots, M)$ are distinct values in the unit disk $\mathbb{D} := \{z \in \mathbb{C} :\, 0 < |z| \le 1\}$ without zero.

In the following, we recover all parameters of a nonincreasing exponential sum (1.1), if noiseless sampled data

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

with $N \ge M$ are given. This problem is known as frequency analysis problem, which is important within many disciplines in sciences and engineering (see [22]). For a survey of the most successful methods for the data fitting problem with linear combinations of complex exponentials, we refer to [21].

The aim of this paper is to present a unified approach to Prony–like methods for parameter estimation of (1.1), namely the classical Prony method, the matrix pencil method [16, 13], and the ESPRIT method (ESPRIT = Estimation of Signal Parameters via Rotational Invariance Techniques) [28, 29].

The basic idea of Prony–like methods can be easily explained in the case of known order M. We introduce the *Prony polynomial*

$$p(z) := \prod_{j=1}^{M} (z - z_j) = \sum_{r=0}^{M-1} p_r \, z^r + z^M \tag{1.3}$$

with distinct $z_j \in \mathbb{D}$ and the corresponding coefficient vector $\boldsymbol{p} := (p_r)_{r=0}^{M-1}$. Further we define the *companion matrix* $\boldsymbol{C}_M(p) \in \mathbb{C}^{M \times M}$ of the Prony polynomial (1.3) by

$$\boldsymbol{C}_M(p) := \begin{pmatrix} 0 & 0 & \ldots & 0 & -p_0 \\ 1 & 0 & \ldots & 0 & -p_1 \\ 0 & 1 & \ldots & 0 & -p_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -p_{M-1} \end{pmatrix}. \tag{1.4}$$

It is known that the companion matrix (1.4) has the property

$$\det \big(z\, \boldsymbol{I}_M - \boldsymbol{C}_M(p) \big) = p(z) \quad (z \in \mathbb{C})\,,$$

where $\boldsymbol{I}_M \in \mathbb{R}^{M\times M}$ denotes the identity matrix. Hence the eigenvalues of the companion matrix (1.4) coincide with the zeros of the Prony polynomial (1.3). Further we establish the Hankel matrices $\boldsymbol{H}_M(s) := \big(h(s+r+m)\big)_{m,r=0}^{M-1}$ $(s = 0,\,1)$ which are known by the sampled data. Then we see by Lemma 2.2 (with $N = L = M$) that $\boldsymbol{p}$ is a solution of the square Yule–Walker system (see e.g. [11])

$$\boldsymbol{H}_M(0)\,\boldsymbol{p} = -\big(h(k)\big)_{k=M}^{2M-1}$$

and that the Hankel matrices $\boldsymbol{H}_M(s)$ $(s = 0,\,1)$ and the companion matrix $\boldsymbol{C}_M(p)$ fulfil the useful relation

$$\boldsymbol{H}_M(0)\,\boldsymbol{C}_M(p) = \boldsymbol{H}_M(1)\,. \tag{1.5}$$

In the case $N = L = M$, the Algorithm 2.4 is the classical Prony method, which is mainly based on the solution of a square Yule–Walker system. From the property (1.5), it follows immediately that the eigenvalues of the square matrix pencil (see e.g. [14, p. 375])

$$z\,\boldsymbol{H}_M(0) - \boldsymbol{H}_M(1) \quad (z \in \mathbb{C})$$

are exactly the nodes z_j $(j = 1,\,\ldots,\,M)$. The advantage of the matrix pencil method is the fact that there is no need to compute the coefficients of the Prony polynomial (1.1). Using the QR decomposition of the rectangular Hankel matrix $\boldsymbol{H}_{M,M+1} := \big(h(l+m)\big)_{l,m=0}^{M-1,M}$, we obtain the new Algorithm 3.1 (with $N = L = M$) of the matrix pencil factorization. Using singular value decomposition (SVD) of $\boldsymbol{H}_{M,M+1}$, we obtain a short approach to the ESPRIT Algorithm 3.2 (with $N = L = M$). Finally one has to determine the coefficients c_j $(j = 1,\,\ldots,\,M)$ via solving a Vandermonde system.

In this paper, we firstly present the Prony method in the general case of unknown order M for the exponential sum (1.1), where L with $M \le L \le N$ is a given upper bound on M. If the order M is known, then one can choose $N = L = M$. In many practical applications one has to deal with the ill-conditioning of the Hankel and Vandermonde matrices. We show that one can attenuate this problem by using more sampled data (1.2) with $N \gg M$, see e.g. [3]. But then one has to deal with two rectangular Hankel matrices $\boldsymbol{H}_{2N-L,L}(s) := \big(h(s+r+m)\big)_{m,r=0}^{2N-L-1,L-1}$ for $s = 0,\,1$. Based on the fact that the relation

$$\boldsymbol{H}_{2N-L,L}(0)\,\boldsymbol{C}_L(q) = \boldsymbol{H}_{2N-L,L}(1)$$

between the rectangular Hankel matrices $\boldsymbol{H}_{2N-L,L}(s)$ $(s = 0,\,1)$ and the companion matrix $\boldsymbol{C}_L(q)$ is still valid, where q is a modified Prony polynomial of degree L with $q(z_j) = 0$ $(j = 1, \ldots, M)$, we use standard methods from numerical linear algebra (such as QR decomposition, SVD, and least squares problems) in order to compute the nodes $z_j \in \mathbb{D}$ $(j = 1, \ldots, M)$. Furthermore we are in position to determine the order M of the exponential sum (1.1). The Algorithm 2.4 is a slight generalization of the classical Prony method, which is now based on the least squares solution of a rectangular Yule–Walker system. We remark that the Prony method is equivalent to the annihilating filter method (see e.g. [31]).

Then we present two matrix pencil methods. We observe that the rectangular matrix pencil

$$z\,\boldsymbol{H}_{2N-L,L}(0) - \boldsymbol{H}_{2N-L,L}(1) \quad (z \in \mathbb{C})$$

has the nodes z_j ($j = 1, \ldots, M$) as eigenvalues. The new Algorithm 3.1 is based on a simultaneous QR decomposition of the rectangular Hankel matrices $\boldsymbol{H}_{2N-L,L}(s)$ ($s = 0, 1$), which can be realized by one QR decomposition of the augmented Hankel matrix $\boldsymbol{H}_{2N-L,L+1} := \big(h(r+m)\big)_{m,r=0}^{2N-L-1,L}$. With Algorithm 3.1, we simplify a matrix pencil method proposed in [13]. The Algorithm 3.2 is based on a simultaneous SVD of the Hankel matrix $\boldsymbol{H}_{2N-L,L+1}$, which follows the same ideas as the Algorithm 3.1, but leads to the known ESPRIT method [28, 29]. In contrast to [28, 29], our approach is only based on simple properties of matrix computation without use of the rotational invariance property (see Remark 3.7). Note that a variety of papers compare the statistical properties of the different algorithms, see e.g. [16, 1, 2, 12]. We stress again that our aim is to propose a simple unified approach to Prony–like methods, such that the algorithms can be easily implemented, if routines for the SVD, QR decomposition, least squares problems, and computation of eigenvalues of a square matrix are available. Furthermore we mention that the Prony–like methods can be generalized to nonequispaced sampled data of (1.1) (see [10] and [23, Section 6]) as well as to multivariate exponential sums (see [26, 8]). But all this extensions are based on parameter estimation of the univariate exponential sum (1.1) as described in this paper.

The outline of this paper is as follows. In Section 2, we present results on Prony method for unknown order M and given upper bound L with $M \le L \le N$. Here we consider rectangular Vandermonde and Hankel matrices as well as companion matrices of modified Prony polynomials. In Section 3, we present a unified approach to two matrix pencil methods. Finally we present some numerical experiments in Section 4, where we apply our methods to the parameter estimation problems, to sparse Fourier approximation as well as to nonlinear approximation by exponential sums.

In the following we use standard notations. By $\mathbb{C}$, we denote the set of all complex numbers, and $\mathbb{N}$ is the set of all positive integers. The complex unit disk is denoted by $\mathbb{D}$. The Kronecker delta is δ_k. The linear space of all column vectors with N complex components is denoted by $\mathbb{C}^N$, where $\boldsymbol{o}$ is the corresponding zero vector. The linear space of all complex M-by-N matrices is denoted by $\mathbb{C}^{M\times N}$, where $\boldsymbol{O}_{M,N}$ is the corresponding zero matrix. For a matrix $\boldsymbol{A}_{M,N} \in \mathbb{C}^{M\times N}$, its transpose is denoted by $\boldsymbol{A}_{M,N}^{\mathrm{T}}$, its conjugate–transpose by $\boldsymbol{A}_{M,N}^{*}$, and its Moore–Penrose pseudoinverse by $\boldsymbol{A}_{M,N}^{\dagger}$. A square matrix $\boldsymbol{A}_{M,M}$ is abbreviated to $\boldsymbol{A}_M$. By $\boldsymbol{I}_M$ we denote the M-by-M identity matrix. By $\mathrm{null}\,\boldsymbol{A}_{M,N}$ we denote the null space of a matrix $\boldsymbol{A}_{M,N}$. Further we use the known submatrix notation. Thus $\boldsymbol{A}_{M,M+1}(1:M, 2:M+1)$ is the submatrix of $\boldsymbol{A}_{M,M+1}$ obtained by extracting rows 1 through M and columns 2 through $M+1$, and $\boldsymbol{A}_{M,M+1}(1:M, M+1)$ means the last column vector of $\boldsymbol{A}_{M,M+1}$. Note that the first row or column of $\boldsymbol{A}_{M,N}$ can be indexed by zero. Definitions are indicated by the symbol $:=$. Other notations are introduced when needed.

2 Prony method

We consider the case of unknown order M for the exponential sum (1.1) and given noiseless sampled data $h(k)$ $(k = 0, \ldots, 2N-1)$. 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 $L = N$. With the $2N$ sampled data $h(k) \in \mathbb{C}$ $(k = 0, \ldots, 2N-1)$ we form the *rectangular Hankel matrices*

$$\begin{aligned} \boldsymbol{H}_{2N-L,L+1} &:= \big(h(l+m)\big)_{l,m=0}^{2N-L-1,L}, &\quad (2.1)\\ \boldsymbol{H}_{2N-L,L}(s) &:= \big(h(s+l+m)\big)_{l,m=0}^{2N-L-1,L-1}, \quad (s=0,\,1)\,. &\quad (2.2)\end{aligned}$$

Then $\boldsymbol{H}_{2N-L,L}(1)$ is the “shifted” matrix of $\boldsymbol{H}_{2N-L,L}(0)$ and

$$\begin{aligned} \boldsymbol{H}_{2N-L,L+1} &= \big(\, \boldsymbol{H}_{2N-L,L}(0) \quad \boldsymbol{H}_{2N-L,L}(1)(1:2N-L,\,L)\,\big)\,,\\ \boldsymbol{H}_{2N-L,L}(s) &= \boldsymbol{H}_{2N-L,L+1}(1:2N-L,\,1+s:L+s) \quad (s=0,\,1)\,. \qquad (2.3)\end{aligned}$$

Using the coefficients p_k $(k = 0, \ldots, M-1)$ of the Prony polynomial (1.3), we form the vector $\boldsymbol{p}_L := (p_k)_{k=0}^{L-1}$ with $p_M := 1$, $p_{M+1} = \ldots = p_{L-1} := 0$. By $\boldsymbol{\Delta}_L := \big(\delta_{k-l-1}\big)_{k,l=0}^{L-1}$ we denote the forward shift matrix, where δ_k is the Kronecker delta. Analogously, we introduce $\boldsymbol{p}_{L+1} := (p_k)_{k=0}^{L}$ with $p_L := 0$ and $\boldsymbol{\Delta}_{L+1} := \big(\delta_{k-l-1}\big)_{k,l=0}^{L}$.

Lemma 2.1 *Let L, M, $N \in \mathbb{N}$ with $M \le L \le N$ be given. Furthermore, let* (1.2) *be noiseless sampled data of the exponential sum* (1.1) *with $c_j \in \mathbb{C} \setminus \{0\}$ and distinct nodes $z_j = \mathrm{e}^{f_j} \in \mathbb{D}$ $(j = 1, \ldots, M)$. Then*

$$\operatorname{rank} \boldsymbol{H}_{2N-L,L+1} = \operatorname{rank} \boldsymbol{H}_{2N-L,L}(s) = M \quad (s=0,\,1)\,. \qquad (2.4)$$

If $L = M$, then $\operatorname{null} \boldsymbol{H}_{2N-M,M+1} = \operatorname{span}\{\boldsymbol{p}_{M+1}\}$ *and* $\operatorname{null} \boldsymbol{H}_{2N-M,M}(s) = \{\boldsymbol{o}\}$ *for $s = 0,\,1$. If $L > M$, then*

$$\begin{aligned} \operatorname{null} \boldsymbol{H}_{2N-L,L+1} &= \operatorname{span}\{\boldsymbol{p}_{L+1},\, \boldsymbol{\Delta}_{L+1}\boldsymbol{p}_{L+1},\, \ldots,\, \boldsymbol{\Delta}_{L+1}^{L-M}\boldsymbol{p}_{L+1}\}\,,\\ \operatorname{null} \boldsymbol{H}_{2N-L,L}(s) &= \operatorname{span}\{\boldsymbol{p}_{L},\, \boldsymbol{\Delta}_{L}\boldsymbol{p}_{L},\, \ldots,\, \boldsymbol{\Delta}_{L}^{L-M-1}\boldsymbol{p}_{L}\} \quad (s=0,\,1)\end{aligned}$$

and

$$\begin{aligned} \dim(\operatorname{null} \boldsymbol{H}_{2N-L,L+1}) &= L-M+1\,,\\ \dim(\operatorname{null} \boldsymbol{H}_{2N-L,L}(s)) &= L-M \quad (s=0,\,1)\,.\end{aligned}$$

The proof follows similar lines as the proof of [25, Lemma 2.1] and is omitted here.

For $\boldsymbol{z} = (z_j)_{j=1}^{M}$ with distinct $z_j \in \mathbb{D}$, we introduce the *rectangular Vandermonde matrix*

$$\boldsymbol{V}_{2N-L,M}(\boldsymbol{z}) := \Big(z_j^{k-1}\Big)_{k,j=1}^{2N-L,M}. \qquad (2.5)$$

Then the rectangular Hankel matrices (2.1) and (2.2) can be factorized in the following form

$$\begin{aligned} \boldsymbol{H}_{2N-L,L+1} &= \boldsymbol{V}_{2N-L,M}(\boldsymbol{z})\,(\operatorname{diag}\boldsymbol{c})\,\boldsymbol{V}_{L+1,M}(\boldsymbol{z})^{\mathrm{T}}\,, \\ \boldsymbol{H}_{2N-L,L}(s) &= \boldsymbol{V}_{2N-L,M}(\boldsymbol{z})\,(\operatorname{diag}\boldsymbol{c})\,(\operatorname{diag}\boldsymbol{z})^{s}\,\boldsymbol{V}_{L,M}(\boldsymbol{z})^{\mathrm{T}} \quad (s=0,\,1)\,. \end{aligned}$$

The classical Prony method (for unknown order M) is based on the following result.

Lemma 2.2 *Let $L,\ M,\ N \in \mathbb{N}$ with $M \le L \le N$ be given. Let (1.2) be noiseless sampled data of the exponential polynomial (1.1) with $c_j \in \mathbb{C}\setminus\{0\}$ and distinct nodes $z_j \in \mathbb{D}$ $(j=1,\ldots,M)$. Then following assertions are equivalent:*
(i) *The polynomial*

$$q(z) = \sum_{k=0}^{L-1} q_k\, z^k + z^L \quad (z \in \mathbb{C}) \tag{2.6}$$

with complex coefficients q_k has all $z_j \in \mathbb{D}$ $(j=1,\ldots,M)$ as roots. Such a monic polynomial q of degree L is called modified Prony polynomial.
(ii) *The vector $\boldsymbol{q} = (q_k)_{k=0}^{L-1}$ is a solution of the linear system*

$$\boldsymbol{H}_{2N-L,L}(0)\,\boldsymbol{q} = -\big(h(k)\big)_{k=L}^{2N-1}\,.$$

(iii) *The companion matrix $\boldsymbol{C}_L(q) \in \mathbb{C}^{L\times L}$, cf. (1.4), has the property*

$$\boldsymbol{H}_{2N-L,L}(0)\,\boldsymbol{C}_L(q) = \boldsymbol{H}_{2N-L,L}(1)\,. \tag{2.7}$$

Proof. The equivalence of (i) and (ii) can be shown in a similar manner as in [25, Lemma 2.2]. Therefore we prove only that (ii) and (iii) are equivalent. From (2.7) it follows immediately that

$$-\boldsymbol{H}_{2N-L,L}(0)\,\boldsymbol{q} = \big(h(k)\big)_{k=L}^{2N-1}\,,$$

since the last column of $\boldsymbol{C}_L(q)$ reads as $-\boldsymbol{q}$ and since the last column of $\boldsymbol{H}_{2N-L}(1)$ is equal to $\big(h(k)\big)_{k=L}^{2N-1}$. Conversely, by

$$\begin{aligned} \boldsymbol{H}_{2N-L,L}(0)\,(\delta_{k-j})_{k=0}^{L-1} &= \big(h(k+j)\big)_{k=0}^{2N-L-1} \quad (j=1,\ldots,L-1)\,, \\ -\boldsymbol{H}_{2N-L,L}(0)\,\boldsymbol{q} &= \big(h(k)\big)_{k=L}^{2N-1} \end{aligned}$$

we obtain (2.7) column by column. This completes the proof. ■

Thus (2.6) is a modified Prony polynomial of degree L if and only if $\boldsymbol{q} = (q_k)_{k=0}^{L-1}$ is a solution of the linear system

$$\boldsymbol{H}_{2N-L,L}(0)\,\boldsymbol{q} = -\big(h(k)\big)_{k=L}^{2N-1}\,.$$

Then (2.6) has the same zeros $z_j \in \mathbb{D}$ $(j=1,\ldots,M)$ as the Prony polynomial (1.3), but (2.6) has $L-M$ additional zeros, if $L > M$. For example,

$$q(z) = z^{L-M}\,p(z)$$

is the simplest modified Prony polynomial of degree L. If r is an arbitrary monic polynomial of degree $L-M$, then $q(z) = r(z)\,p(z)$ is a modified Prony polynomial of degree L too. A modified Prony polynomial is not uniquely determined in the case $L > M$. If M is known and if $L = M$, then the modified Prony polynomial (2.6) and the Prony polynomial (1.3) coincide.

Remark 2.3 Previously, modified Prony polynomials of *moderate* degree L ($M \le L \le N$) were considered. A modified Prony polynomial q of *higher* degree $2N-L$ ($M \le L \le N$) has the form

$$q(z) = \sum_{k=0}^{2N-L-1} q_k\, z^k + z^{2N-L} \quad (z \in \mathbb{C}),$$

where the coefficient vector $\boldsymbol{q} = (q_k)_{k=0}^{2N-L-1}$ is now a solution of the underdetermined linear system

$$\boldsymbol{H}_{L,2N-L}(0)\,\boldsymbol{q} = -\big(h(k)\big)_{k=2N-L}^{2N-1}.$$

with $\boldsymbol{H}_{L,2N-L}(0) = \boldsymbol{H}_{2N-L,L}(0)^{\mathrm{T}}$. The proof follows similar lines as the proof of Lemma 2.2 (see [25, Lemma 2.2]). □

Now we formulate Lemma 2.2 as an algorithm. Since the unknown coefficients c_j ($j = 1,\ldots,M$) do not vanish, we can assume that $|c_j| > \varepsilon$ for convenient bound ε ($0 < \varepsilon \ll 1$).

Algorithm 2.4 (Prony method)
Input: $L,\ N \in \mathbb{N}$ ($N \gg 1$, $3 \le L \le N$, L is upper bound of the order M of (1.1)), $h(k) \in \mathbb{C}$ ($k = 0,\ldots,2N-1$), $0 < \varepsilon \ll 1$.

1. Compute the least squares solution of the rectangular Yule–Walker system

$$\boldsymbol{H}_{2N-L,L}(0)\,\boldsymbol{q} = -\big(h(m+L)\big)_{m=0}^{2N-L-1}.$$

2. Determine the simple roots $\tilde{z}_j \in \mathbb{D}$ ($j = 1,\ldots,\tilde{M}$), $\tilde{M} \le L$, of the modified Prony polynomial (2.6), i.e., compute all eigenvalues $\tilde{z}_j \in \mathbb{D}$ ($j = 1,\ldots,\tilde{M}$) of the companion matrix $\boldsymbol{C}_L(q)$. Note that rank $\boldsymbol{H}_{2N-L,L}(0) = M \le \tilde{M}$.
3. Compute $\tilde{c}_j \in \mathbb{C}$ ($j = 1,\ldots,\tilde{M}$) as least squares solution of the overdetermined linear Vandermonde–type system

$$\boldsymbol{V}_{2N,\tilde{M}}(\tilde{\boldsymbol{z}})\,(\tilde{c}_j)_{j=1}^{\tilde{M}} = \big(h(k)\big)_{k=0}^{2N-1}$$

with $\tilde{\boldsymbol{z}} := (\tilde{z}_j)_{j=1}^{\tilde{M}}$.
4. Delete all the $\tilde{z}_l$ ($l \in \{1,\ldots,\tilde{M}\}$ with $|\tilde{c}_l| \le \varepsilon$ and denote the remaining values by z_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.
5. Repeat step 3 and compute the coefficients $c_j \in \mathbb{C}$ ($j = 1,\ldots,M$) as least squares solution of the overdetermined linear Vandermonde–type system

$$\boldsymbol{V}_{2N,M}(\boldsymbol{z})\,\boldsymbol{c} = (h(k))_{k=0}^{2N-1}$$

with $\boldsymbol{z} := (z_j)_{j=1}^M$ and $\boldsymbol{c} := (c_j)_{j=1}^M$.

Output: $M \in \mathbb{N}$, $f_j \in (-\infty, 0] + \mathrm{i}\,[-\pi, \pi)$, $c_j \in \mathbb{C}$ $(j = 1, \ldots, M)$.

Remark 2.5 In signal processing, the Prony method is also known as the *annihilating filter method* (see e.g. [31]). For distinct $z_j \in \mathbb{D}$ and for $c_j \in \mathbb{C} \setminus \{0\}$ $(j = 1, \ldots, M)$, we consider a discrete signal

$$h_n := \sum_{j=1}^{M} c_j\, z_j^n \quad (n \in \mathbb{Z})\,. \tag{2.8}$$

For simplicity, we assume that M is known. Then a discrete signal a_n $(n \in \mathbb{Z})$ is called an *annihilating filter* of (2.8), if the discrete convolution of the signals a_n and h_n vanishes, i.e.

$$(a * h)_n := \sum_{l=-\infty}^{\infty} a_l\, h_{n-l} = 0 \quad (n \in \mathbb{Z})\,.$$

If a_n is chosen in the following manner

$$a(z) := \prod_{j=1}^{M} (1 - z_j\, z^{-1}) = \sum_{n=0}^{M} a_n\, z^{-n} \quad (z \in \mathbb{C} \setminus \{0\})\,,$$

then a_n is an annihilating filter of (2.8), since for all $n \in \mathbb{Z}$

$$\begin{aligned} (a * h)_n &= \sum_{l=0}^{M} a_l\, h_{n-l} = \sum_{l=0}^{M} \sum_{j=1}^{M} c_j\, a_l\, z_j^{n-l} \\ &= \sum_{j=1}^{M} c_j\, a(z_j)\, z_j^n = 0\,. \end{aligned}$$

Note that $a(z)$ is the z–transform of the annihilating filter a_n. Further $a(z)$ and the Prony polynomial (1.3) have the same zeros $z_j \in \mathbb{D}$ $(j = 1, \ldots, M)$, since $z^M\, a(z) = p(z)$ for all $z \in \mathbb{C} \setminus \{0\}$. Hence the Prony method and the method of annihilating filter are equivalent. For details of the annihilating filter method see e.g. [31]. □

3 Matrix pencil method

Now we show that the Prony method can be simplified to a matrix pencil method. Note that a rectangular matrix pencil may not have eigenvalues in general. But this is not the case for our *rectangular matrix pencil*

$$z\, \boldsymbol{H}_{2N-L,L}(0) - \boldsymbol{H}_{2N-L,L}(1)\,, \tag{3.1}$$

which has $z_j \in \mathbb{D}$ $(j = 1, \ldots, M)$ as eigenvalues. This follows by using (2.7) from

$$\big(z\, \boldsymbol{H}_{2N-L,L}(0) - \boldsymbol{H}_{2N-L,L}(1)\big)\, \boldsymbol{v} = \boldsymbol{H}_{2N-L,L}(0)\, \big(z\, \boldsymbol{I}_L - \boldsymbol{C}_L(q)\big)\, \boldsymbol{v}$$

and

$$\det\big(z\,\boldsymbol{I}_L - \boldsymbol{C}_L(q)\big) = q(z)\,.$$

If $z = z_j$ $(j = 1, \ldots, M)$, then $\boldsymbol{v} = (v_k)_{k=0}^{L-1} \in \mathbb{C}^L$ is an eigenvector of the square eigenvalue problem $\boldsymbol{C}_L(q)\,\boldsymbol{v} = z_j\,\boldsymbol{v}$ with

$$v_k = -z_j^{L-1-k}\,\rho_k(z_j) \quad (k = 0, \ldots, L-1)$$

and $v_{L-1} = z_j^L$, where

$$\rho_k(z) := \sum_{r=0}^{k} q_r\, z^r \quad (z \in \mathbb{C};\ k = 0, \ldots, L-1)$$

is the truncated modified Prony polynomial of degree k and where (2.6) is a modified Prony polynomial of degree L. The generalized eigenvalue problem of the rectangular matrix pencil (3.1) can be reduced to a classical eigenvalue problem of a square matrix.

Therefore one can simultaneously factorize the rectangular Hankel matrices (2.2) under the assumption $2N \ge 3L$. Then there are at least twice as many rows as there are columns in the matrix pencil (3.1). Following [7, p. 598], one can apply a QR decomposition to the matrix pair

$$\begin{pmatrix} \boldsymbol{H}_{2N-L,L}(0) & \boldsymbol{H}_{2N-L,L}(1) \end{pmatrix} \in \mathbb{C}^{(2N-L)\times 2L}\,.$$

Here we simplify this idea. Without the additional assumption $2N \ge 3L$, we compute the QR decomposition of the rectangular Hankel matrix (2.1). By (2.4), the rank of the Hankel matrix $\boldsymbol{H}_{2N-L,L+1}$ is equal to M. Hence $\boldsymbol{H}_{2N-L,L+1}$ is rank deficient. Therefore we apply QR factorization with column pivoting (see [14, pp. 248 – 250]) and obtain

$$\boldsymbol{H}_{2N-L,L+1}\,\boldsymbol{\Pi}_{L+1} = \boldsymbol{Q}_{2N-L}\,\boldsymbol{R}_{2N-L,L+1} \tag{3.2}$$

with a unitary matrix $\boldsymbol{Q}_{2N-L}$, a permutation matrix $\boldsymbol{\Pi}_{L+1}$, and a trapezoidal matrix

$$\boldsymbol{R}_{2N-L,L+1} = \begin{pmatrix} \boldsymbol{R}_{2N-L,L+1}(1:M,\,1:L+1) \\ \boldsymbol{O}_{2N-L-M,L+1} \end{pmatrix},$$

where $\boldsymbol{R}_{2N-L,L+1}(1:M,\,1:M)$ is a nonsingular upper triangular matrix. By the QR decomposition we can determine the rank M of the Hankel matrix (2.1) and hence the order of the exponential sum (1.1). Note that the permutation matrix $\boldsymbol{\Pi}_{L+1}$ is chosen such that the diagonal entries of $\boldsymbol{R}_{2N-L,L+1}(1:M,\,1:M)$ have nonincreasing absolute values. We denote the diagonal matrix containing these diagonal entries by $\boldsymbol{D}_M$. With

$$\boldsymbol{S}_{2N-L,L+1} := \boldsymbol{R}_{2N-L,L+1}\,\boldsymbol{\Pi}_{L+1}^{\mathrm{T}} = \begin{pmatrix} \boldsymbol{S}_{2N-L,L+1}(1:M,\,1:L+1) \\ \boldsymbol{O}_{2N-L-M,L+1} \end{pmatrix}, \tag{3.3}$$

we infer that by (2.3)

$$\boldsymbol{H}_{2N-L,L}(s) = \boldsymbol{Q}_{2N-L}\,\boldsymbol{S}_{2N-L,L}(s) \quad (s = 0,\, 1)$$

with

$$\boldsymbol{S}_{2N-L,L}(s) := \boldsymbol{S}_{2N-L,L+1}(1:2N-L,\, 1+s:L+s) \quad (s=0,\,1)\,.$$

Since $\boldsymbol{Q}_{2N-L}$ is unitary, the generalized eigenvalue problem of the rectangular matrix pencil (3.1) is equivalent to the generalized eigenvalue problem of the matrix pencil

$$z\,\boldsymbol{S}_{2N-L,L}(0) - \boldsymbol{S}_{2N-L,L}(1) \quad (z \in \mathbb{C})\,.$$

Using the special structure of (3.3), we can simplify the matrix pencil to

$$z\,\boldsymbol{T}_{M,L}(0) - \boldsymbol{T}_{M,L}(1) \tag{3.4}$$

with

$$\boldsymbol{T}_{M,L}(s) := \boldsymbol{S}_{2N-L,L}(1:M,\, 1+s:L+s) \quad (s=0,\,1)\,.$$

Here one can use the matrix $\boldsymbol{D}_M$ as diagonal preconditioner and proceed with $\boldsymbol{T}'_{M,L}(s) := \boldsymbol{D}_M^{-1}\,\boldsymbol{T}_{M,L}(s)$. Then the generalized eigenvalue problem of the transposed matrix pencil

$$z\,\boldsymbol{T}'_{M,L}(0)^{\mathrm{T}} - \boldsymbol{T}'_{M,L}(1)^{\mathrm{T}}$$

has the same eigenvalues as the matrix pencil (3.4) except for the zero eigenvalues and it can be solved as eigenvalue problem of the M-by-M matrix

$$\boldsymbol{F}_M^{QR} := \big(\boldsymbol{T}'_{M,L}(0)^{\mathrm{T}}\big)^{\dagger}\,\boldsymbol{T}'_{M,L}(1)^{\mathrm{T}}\,. \tag{3.5}$$

Finally we obtain the nodes $z_j \in \mathbb{D}$ $(j = 1, \ldots, M)$ as the eigenvalues of (3.5).

Algorithm 3.1 (Matrix pencil factorization based on QR decomposition)
Input: $L,\, N \in \mathbb{N}$ $(N \gg 1,\, 3 \le L < N,\, L$ is upper bound of the order M of (1.1)), $h(k) \in \mathbb{C}$ $(k = 0, \ldots, 2N-1)$, $0 < \varepsilon \ll 1$.

1. Compute QR factorization of the rectangular Hankel matrix (2.1). Determine the rank M of (3.2) such that the diagonal entries of $\boldsymbol{R}_{2N-L,L+1}(M+1, M+1) < \varepsilon\,\boldsymbol{R}_{2N-L,L+1}(1,1)$ and form the preconditioned matrices $\boldsymbol{T}'_{M,L}(s)$ $(s = 0,\, 1)$.
2. Determine the eigenvalues $z_j \in \mathbb{D}$ $(j = 1, \ldots, M)$ of the square matrix (3.5). Form $f_j := \log z_j$ $(j = 1, \ldots, M)$.
3. Compute the coefficients $c_j \in \mathbb{C}$ $(j = 1, \ldots, M)$ as least squares solution of the overdetermined linear Vandermonde system

$$\boldsymbol{V}_{2N,M}(\boldsymbol{z})\,\boldsymbol{c} = (h(k))_{k=0}^{2N-1}$$

with $\boldsymbol{z} := (z_j)_{j=1}^M$ and $\boldsymbol{c} := (c_j)_{j=1}^M$.

Output: $M \in \mathbb{N}$, $f_j \in (-\infty,\, 0] + \mathrm{i}\,[-\pi, \pi)$, $c_j \in \mathbb{C}$ $(j = 1, \ldots, M)$.

In the following we derive the ESPRIT method by similar ideas as above, but now we use the SVD of the Hankel matrix (2.1), which is rank deficient by (2.4). We obtain a method which is known as the ESPRIT method. In [16, 30] a relationship between the matrix pencil methods and several variants of the ESPRIT method [28, 29] is derived

showing comparable performance. The essence of ESPRIT method lies in the rotational property between staggered subspaces, see [20, Section 9.6.5].
Therefore we use the factorization

$$\boldsymbol{H}_{2N-L,L+1} = \boldsymbol{U}_{2N-L}\,\boldsymbol{D}_{2N-L,L+1}\,\boldsymbol{W}_{L+1}\,, \tag{3.6}$$

where $\boldsymbol{U}_{2N-L}$ and $\boldsymbol{W}_{L+1}$ are unitary matrices and where $\boldsymbol{D}_{2N-L,L+1}$ is a rectangular diagonal matrix. The diagonal entries of $\boldsymbol{D}_{2N-L,L+1}$ are the singular values of (2.1) arranged in nonincreasing order $\sigma_1 \ge \sigma_2 \ge \ldots \ge \sigma_M > \sigma_{M+1} = \ldots = \sigma_{L+1} = 0$. Thus we can determine the rank M of the Hankel matrix (2.1) which coincides with the order of the exponential sum (1.1). Introducing the matrices

$$\begin{aligned} \boldsymbol{D}_{2N-L,M} &:= \boldsymbol{D}_{2N-L,L+1}(1:2N-L,\,1:M) = \begin{pmatrix} \operatorname{diag}(\sigma_j)_{j=1}^M \\ \boldsymbol{O}_{2N-L-M,M} \end{pmatrix}, \\ \boldsymbol{W}_{M,L+1} &:= \boldsymbol{W}_{L+1}(1:M,\,1:L+1)\,, \end{aligned}$$

we can simplify the SVD of the Hankel matrix (2.1) as follows

$$\boldsymbol{H}_{2N-L,L+1} = \boldsymbol{U}_{2N-L}\,\boldsymbol{D}_{2N-L,M}\,\boldsymbol{W}_{M,L+1}\,.$$

Note that $\boldsymbol{W}_{M,L+1}\boldsymbol{W}^*_{M,L+1} = \boldsymbol{I}_M$. Setting

$$\boldsymbol{W}_{M,L}(s) = \boldsymbol{W}_{M,L+1}(1:M,\,1+s:L+s) \quad (s = 0,\,1)\,, \tag{3.7}$$

it follows from (2.3) that

$$\boldsymbol{H}_{2N-L,L}(s) = \boldsymbol{U}_{2N-L}\,\boldsymbol{D}_{2N-L,M}\,\boldsymbol{W}_{M,L}(s) \quad (s = 0,\,1)\,.$$

Since $\boldsymbol{U}_{2N-L}$ is unitary, the generalized eigenvalue problem of the rectangular matrix pencil (3.1) is equivalent to the generalized eigenvalue problem of the matrix pencil

$$z\,\boldsymbol{D}_{2N-L,M}\,\boldsymbol{W}_{M,L}(0) - \boldsymbol{D}_{2N-L,M}\,\boldsymbol{W}_{M,L}(1)\,. \tag{3.8}$$

If we multiply the transposed matrix pencil (3.8) from the right side with

$$\begin{pmatrix} \operatorname{diag}(\sigma_j^{-1})_{j=1}^M \\ \boldsymbol{O}_{2N-L-M,M} \end{pmatrix},$$

we obtain the generalized eigenvalue problem of the matrix pencil

$$z\,\boldsymbol{W}_{M,L}(0)^{\mathrm{T}} - \boldsymbol{W}_{M,L}(1)^{\mathrm{T}}\,,$$

which has the same eigenvalues as the matrix pencil (3.8) except for the additional zero eigenvalues. Finally we determine the nodes $z_j \in \mathbb{D}$ $(j = 1, \ldots, M)$ as eigenvalues of the matrix

$$\boldsymbol{F}_M^{SVD} := \big(\boldsymbol{W}_{M,L}(0)^{\mathrm{T}}\big)^{\dagger}\,\boldsymbol{W}_{M,L}(1)^{\mathrm{T}}\,. \tag{3.9}$$

Thus the ESPRIT algorithm reads as follows:

Algorithm 3.2 (ESPRIT method)
Input: $L,\, N \in \mathbb{N}$ ($N \gg 1$, $3 \le L \le N$, L is upper bound of the order M of (1.1)), $h(k) \in \mathbb{C}$ ($k = 0, \ldots, 2N-1$), $0 < \varepsilon \ll 1$.

1. Compute the SVD of the rectangular Hankel matrix (3.6). Determine the rank M of (2.1) such that $\sigma_{M+1} < \varepsilon \sigma_1$ and form the matrices (3.7).
2. Compute all eigenvalues $z_j \in \mathbb{D}$ ($j = 1, \ldots, M$) of the square matrix (3.9). Set $f_j := \log z_j$ ($j = 1, \ldots, M$).
3. Compute the coefficients $c_j \in \mathbb{C}$ ($j = 1, \ldots, M$) as least squares solution of the overdetermined linear Vandermonde–type system

$$\boldsymbol{V}_{2N,M}(\boldsymbol{z})\, \boldsymbol{c} = \big(h(k)\big)_{k=0}^{2N-1}$$

with $\boldsymbol{z} := (z_j)_{j=1}^{M}$ and $\boldsymbol{c} := (c_j)_{j=1}^{M}$

Output: $M \in \mathbb{N}$, $f_j \in (-\infty,\, 0] + \mathrm{i}\,[-\pi, \pi)$, $c_j \in \mathbb{C}$ ($j = 1, \ldots, M$).

Remark 3.3 The original approach to the ESPRIT method (see [28, 29]) is essentially based on the *rotational invariance property* of the Vandermonde matrix (2.5), i.e.

$$\boldsymbol{V}'_{2N-L,M}(\boldsymbol{z}) = \boldsymbol{V}_{2N-L,M}(\boldsymbol{z})\, (\mathrm{diag}\, \boldsymbol{z})\,,$$

with $\boldsymbol{V}'_{2N-L,M}(\boldsymbol{z}) := \big(z_j^k\big)_{k,j=1}^{2N-L,M}$. Note that there exists a close relationship between the Vandermonde matrix (2.5) and the transposed companion matrix $\boldsymbol{C}_{2N-L,M}(q)^{\mathrm{T}}$, namely

$$\boldsymbol{V}'_{2N-L,M}(\boldsymbol{z}) = \boldsymbol{V}_{2N-L,M}(\boldsymbol{z})\, (\mathrm{diag}\, \boldsymbol{z}) = \boldsymbol{C}_{2N-L}(q)^{\mathrm{T}}\, \boldsymbol{V}_{2N-L,M}(\boldsymbol{z})\,,$$

where q is any monic polynomial of degree $2N-L$ with $q(z_j) = 0$ ($j = 1, \ldots, M$).
In contrast to [28, 29], we mainly use the relation (2.7) between the given Hankel matrices (2.2) and the companion matrix $\boldsymbol{C}_L(q)$ of a modified Prony polynomial (2.6). In this sense, we simplify the approach to the ESPRIT method. □

In the case of parameter estimation of (1.1) with unknown order M, we have seen that Algorithm 2.4 computes the nodes z_j ($j = 1, \ldots, M$) as eigenvalues of the L-by-L companion matrix of a modified Prony polynomial (2.6). The Algorithms 3.1 and 3.2 determine exactly the nodes z_j ($j = 1, \ldots, M$) as eigenvalues of an M-by-M matrix (3.5) and (3.9), respectively, which is similar to the companion matrix (1.4) of the Prony polynomial (1.3).

4 Numerical examples

Now we illustrate the behavior of the suggested algorithms. Using IEEE standard floating point arithmetic with double precision, we have implemented our algorithms in MATLAB. The signal is measured in the form (1.1) with complex exponents $f_j \in$

$[-1,\, 0] + \mathrm{i}\,[-\pi,\, \pi)$ and complex coefficients $c_j \neq 0$. The relative error of the complex exponents is given by

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

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

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

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 taken over $10 \cdot (2N-1) + 1$ equispaced points from a grid of $[0, 2N-1]$, and where

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

is the exponential sum recovered by our algorithms.

Example 4.1 We start with an example which is often used in testing system identification algorithms (see [3]). With $M = 6$, $c_j = j$ $(j = 1, \ldots, 6)$, and

$$\big(z_j\big)_{j=1}^{6} = \begin{pmatrix} 0.9856 - 0.1628\,\mathrm{i} \\ 0.9856 + 0.1628\,\mathrm{i} \\ 0.8976 - 0.4305\,\mathrm{i} \\ 0.8976 + 0.4305\,\mathrm{i} \\ 0.8127 - 0.5690\,\mathrm{i} \\ 0.8127 + 0.5690\,\mathrm{i} \end{pmatrix},$$

we form the sampled data (1.2) at the nodes $k = 0, \ldots, 2N-1$. Then we apply our Algorithms 2.4, 3.1 and 3.2. For several parameters N and L, the resulting errors are presented in Table 4.1. As the bound ε in the algorithms we use 10^{-10}. We see that the classical Prony method does not work very well. Using matrix pencil factorization or ESPRIT method, we obtain excellent parameter reconstructions for only few sampled data. See [25] for further examples. □

N	L	$e(\boldsymbol{f})$	$e(\boldsymbol{c})$	$e(h)$
Algorithm 2.4 (Prony method)				
6	6	4.93e+00	1.89e–01	2.41e–04
7	6	1.65e–09	9.86e–10	7.12e–13
7	7	7.27e–10	4.89e–10	6.21e–13
Algorithm 3.1 (Matrix pencil)				
6	6	7.76e–09	4.44e–09	3.52e–14
7	6	2.23e–10	1.75e–10	5.92e–15
7	7	5.53e–10	3.62e–10	7.81e–14
Algorithm 3.2 (ESPRIT)				
6	6	7.44e–09	4.31e–09	6.52e–13
7	6	1.01e–10	7.73e–11	2.23e–13
7	7	5.69e–10	3.87e–10	823e–14

Table 4.1: Results of Example 4.1.

Example 4.2 Now we consider the exponential sum (1.1) of order $M = 6$ with the complex exponents $(f_j)_{j=1}^6 = \frac{\mathrm{i}}{1000}\,(7,\,21,\,200,\,201,\,53,\,1000)^{\mathrm{T}}$ and coefficients $(c_j)_{j=1}^6 = (6,\,5,\,4,\,3,\,2,\,1)^{\mathrm{T}}$. For the $2N$ sampled data (1.2), we apply the Algorithms 2.4, 3.1 and 3.2. As the bound ε in the algorithms we use again 10^{-10}. For several parameters N and L, the resulting errors are presented in Table 4.2. Introducing the separation distance $\delta := \min\{|f_j - f_k|;\, j \neq k\} = 0.001$, all parameters of (1.1) can be recovered by results of [24, Lemma 4.1] if N is sufficiently large with $N > \frac{\pi^2}{\delta\sqrt{3}} \approx 5698$. However we observe that much less sampled data are sufficient for a good parameter reconstruction. We present in Table 4.3 some condition numbers of the involved matrices. Note that the matrix $\boldsymbol{H}_{2N-L,L}$ is rank deficient for $L \geq M$. Furthermore the condition number for the the Vandermonde-type matrices are estimated in [24].
This and also the next example is related to *sparse Fourier approximation*, because $f(t) := h(1000\,t)$ is a sparse trigonometric polynomial

$$f(t) = \sum_{j=1}^{M} c_j\, \mathrm{e}^{\mathrm{i}\omega_j t} \quad (t \geq 0)$$

with distinct integer frequencies $\omega_j := -1000\, f_j\, \mathrm{i}$, see [27, 19, 9, 17]. The aim of sparse Fourier approximation is the efficient recovery of all parameters ω_j, c_j $(j = 1, \ldots, M)$ using as few sampled values of f as possible. Using the deterministic Algorithms 2.4, 3.1 or 3.2 for the sampled data $f(k/1000) = h(k)$ for $k = 0, \ldots, 2N-1$ with $N \geq 10$, we obtain the exact frequencies ω_j by rounding. □

N	L	$e(\boldsymbol{f})$	$e(\boldsymbol{c})$	$e(h)$
Algorithm 2.4 (Prony method)				
7	7	9.47e–01	1.73e–01	1.48e–12
8	8	9.47e–01	2.86e–01	8.55e–14
9	9	6.43e–05	6.96e–02	1.02e–13
10	10	1.21e–05	1.22e–02	7.82e–14
15	15	2.83e–07	2.88e–04	1.65e–13
20	20	6.26e–09	6.37e–06	2.87e–13
30	30	5.83e–10	5.91e–07	1.23e–13
30	20	9.73e–11	1.01e–07	4.59e–12
30	10	1.80e–08	1.83e–05	1.18e–10
Algorithm 3.1 (Matrix Pencil)				
7	7	9.47e–01	1.76e–01	1.59e–13
8	8	1.79e–01	1.67e–01	2.82e–13
9	9	5.47e–05	5.80e–02	2.32e–13
10	10	5.62e–06	5.68e–03	7.68e–14
15	15	6.48e–08	6.59e–05	5.51e–13
20	20	1.96e–09	1.99e–06	1.46e–13
30	30	1.08e–10	1.09e–07	1.19e–13
30	10	7.39e–09	7.44e–06	1.21e–10
Algorithm 3.2 (ESPRIT)				
7	7	9.47e–01	1.76e–01	1.58e–13
8	8	1.37e–04	1.33e–01	4.01e–13
9	9	1.20e–05	1.20e–02	2.81e–13
10	10	2.20e–05	2.20e–02	2.48e–13
15	15	5.72e–08	5.81e–05	1.81e–13
20	20	1.75e–09	1.78e–06	7.88e–14
30	30	2.51e–10	2.55e–07	2.88e–13
30	10	2.02e–08	2.04e–05	9.82e–11

Table 4.2: Results of Example 4.2.

Example 4.3 We consider the exponential sum (1.1) of order $M = 6$ with the complex exponents $(f_j)_{j=1}^{6} = \frac{\mathrm{i}}{1000}\,(200, 201, 202, 203, 204, 205)^{\mathrm{T}}$ and the coefficients $(c_j)_{j=1}^{6} = (6, 5, 4, 3, 2, 1)^{\mathrm{T}}$. For the $2N$ sampled data (1.2), we apply the Algorithms 2.4, 3.1 and 3.2. As the bound ε in the algorithms we use again 10^{-10}. The corresponding results are presented in Table 4.4. The marker $*$ in Table 4.4 means that we could not recover all complex exponents. However we approximate the signal very well with fewer exponentials. For the separation distance $\delta = 0.001$, we obtain the same theoretical

N	L	$\boldsymbol{H}^*_{2N-L,L}(0)\boldsymbol{H}_{2N-L,L}(0)$	$\boldsymbol{V}^*_{2N,M}(\boldsymbol{z})\boldsymbol{V}_{2N,M}(\boldsymbol{z})$	$\boldsymbol{F}_M^{\mathrm{QR}}$	$\boldsymbol{F}_M^{\mathrm{SVD}}$
7	6	4.26e+16	1.90e+09	1.19e+02	1.94e+02
10	6	3.81e+16	4.42e+08	3.27e+02	1.94e+02
10	7	6.95e+16	4.07e+08	8.99e+01	8.21e+01
20	7	7.21e+16	1.22e+06	9.83e+01	8.21e+01
20	20	4.21e+17	1.22e+06	6.19e+00	3.28e+00

Table 4.3: Condition numbers of Example 4.2.

bounds for N as in Example 4.2. But now we need much more sampled data than in Example 4.2, since all exponents f_j are clustered. This example shows that one can deal with the ill–conditioning of the matrices by choice of higher N and L.
We note that the reconstruct of the trigonometric polynomial $f(t) = h(1000\,t)$ is much simpler. □

N	L	$e(\boldsymbol{f})$	$e(\boldsymbol{c})$	$e(h)$
Algorithm 2.4 (Prony method)				
300	300	∗	∗	1.85e–11
400	400	∗	∗	3.01e–08
500	500	∗	∗	2.81e–08
600	600	∗	∗	2.76e–08
Algorithm 3.1 (Matrix pencil)				
300	300	5.94e–03	2.47e–01	5.26e–12
400	400	8.46e–05	6.87e–03	8.12e–12
500	500	2.68e–06	2.27e–04	1.25e–11
600	600	4.71e–07	4.20e–05	1.18e–11
Algorithm 3.2 (Matrix ESPRIT)				
300	300	∗	∗	3.22e–12
400	400	4.49e–05	3.60e–03	7.61e–12
500	500	4.53e–06	3.82e–04	7.17e–12
600	600	6.26e–07	5.74e–05	1.58e–11

Table 4.4: Results of Example 4.3.

Example 4.4 Exponential sums are very often studied in *nonlinear approximation*, see also [5, 6]. The starting point in the consideration of exponential sums is an approximation problem encountered for the analysis of decay processes in science and engineering.

A given function $g : [\alpha, \beta] \to \mathbb{C}$ with $0 \le \alpha < \beta < \infty$ is to be approximated by an exponential sum

$$g(t) \approx \sum_{j=1}^{M} \gamma_j \, \mathrm{e}^{\varphi_j t} \quad (t \in [\alpha, \beta]) \tag{4.1}$$

of fixed order M, where the parameters φ_j, γ_j $(j = 1, \dots, M)$ are to be determined. We set

$$\tilde{h}(x) := g(\alpha + \frac{\beta - \alpha}{2N} x) \quad (x \in [0, 2N]) \,.$$

For the equidistant sampled data $\tilde{h}(k)$ $(k = 0, \dots, 2N - 1)$, we apply the Algorithm 3.1, where M is now known. The result of the Algorithm 3.1 is an exponential sum (1.1) of the order M. Substituting $x := 2N\,(t - \alpha)/(\beta - \alpha)$ $(t \in [\alpha, \beta])$ in (1.1), we obtain an exponential sum (4.1) approximating the given function g on the interval $[\alpha, \beta]$.
First we approximate the function $g(t) = 1/t$ $(t \in [1, 10^6])$ by an exponential sum of order $M = 20$ on the interval $[1, 10^6]$. We choose $N = 500$ and $L = 250$. For the 10^3 sampled values $\tilde{h}(k) = g(1 + k\,(10^6 - 1)/1000)$, $(k = 0, \dots, 999)$, the Algorithm 3.1 provides an exponential sum (1.1) with negative damping factors f_j, where $\operatorname{Im} f_j = 0$, and coefficients c_j $(j = 1, \dots, 20)$. Finally, the substitution $x = 1000\,(t - 1)/(10^6 - 1)$ in (1.1) delivers the exponential sum (4.1) approximating the function $g(t) = 1/t$ on the interval $[1, 10^6]$, see Table 4.5. We plot the absolute error between the function $g(t) = 1/t$ and (4.1) in Figure 4.1, where the absolute error is computed on 10^7 equispaced points in $[1, 10^6]$. We remark that the method in [15], which is based on nonequispaced sampling and the Remez algorithm, leads to slightly better results. □

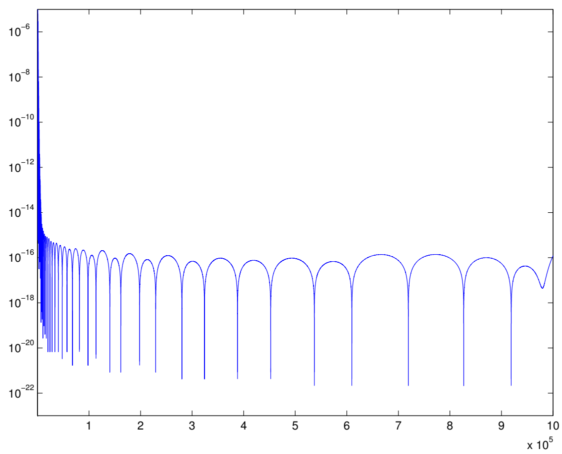

Figure 4.1: Absolute error between $1/t$ and the exponential sum (4.1) of order 20 on the interval $[1, 10^6]$.

j	φ_j	γ_j
1	–1.131477118248638e-02	1.007272522767242e+00
2	–3.135207069583116e-03	1.630338699249195e–03
3	–1.992050185224761e-03	8.190832864884086e–04
4	–1.346933589233963e-03	5.077389797304821e–04
5	–9.311468710314101e-04	3.383442795104141e–04
6	–6.495138717190783e-04	2.324544964268483e–04
7	–4.546022251286615e-04	1.619000817653079e–04
8	–3.184055272088167e-04	1.134465879412991e–04
9	–2.228354952207525e-04	7.970080578182607e–05
10	–1.556526041917767e-04	5.605845462805709e–05
11	–1.083789751443881e-04	3.946592554286590e–05
12	–7.507628403357719e-05	2.782732966146225e–05
13	–5.156651439477524e-05	1.967606126541986e–05
14	–3.491193584036177e-05	1.397100067893288e–05
15	–2.306126330251814e-05	9.962349111371446e–06
16	–1.460537606626776e-05	7.104216597682316e–06
17	–8.602447909553079e-06	5.001095886653157e–06
18	–3.291629375734129e-07	8.476195898630604e–07
19	–1.760073594077830e-06	2.032258842381446e–06
20	–4.445510593794362e-06	3.373949779113571e–06

Table 4.5: Damping factors φ_j and coefficients γ_j of the exponential sum (4.1) approximating $1/t$ on the interval $[1, 10^6]$.

Example 4.5 Finally we consider the function

$$g(t) = J_0(t) := \sum_{k=0}^{\infty} \frac{(-1)^k}{(k!)^2} \left(\frac{t}{2}\right)^{2k} \quad (t \in [0, 1000])\,,$$

where J_0 denotes the Bessel function of first kind of order 0, see Figure 4.3 (left). We approximate this function by an exponential sum (4.1) of order $M = 20$ on the interval $[0, 1000]$, see [5]. Choosing $N = 500$ and $L = 250$, the linear substitution reads $x = t$ and we apply the Algorithm 3.1, where $M = 20$ is now known. For the sampled values $\tilde{h}(k) = J_0(k)$ $(k = 0, \ldots, 999)$, we obtain the exponential sum (4.1) of order 20 with the complex exponents φ_j (shown in Figure 4.2 (left)) and the complex coefficients γ_j (shown in Figure 4.2 (right)). The absolute error between J_0 and the exponential sum (4.1) of order 20 is shown in Figure 4.3 (right), where the absolute error is computed on 10^7 equispaced points in $[0, 1000]$. □

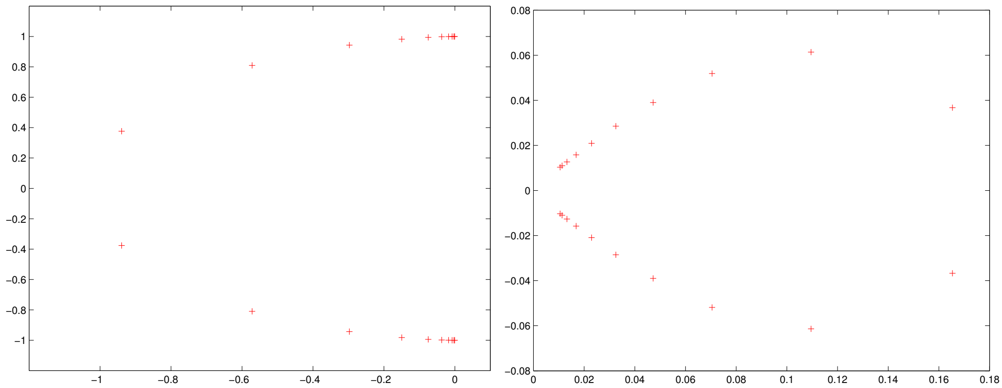

Figure 4.2: Complex exponents φ_j (left) and complex coefficients γ_j for the exponential sum (4.1) of order 20 approximating the Bessel function J_0 on the interval [0, 1000].

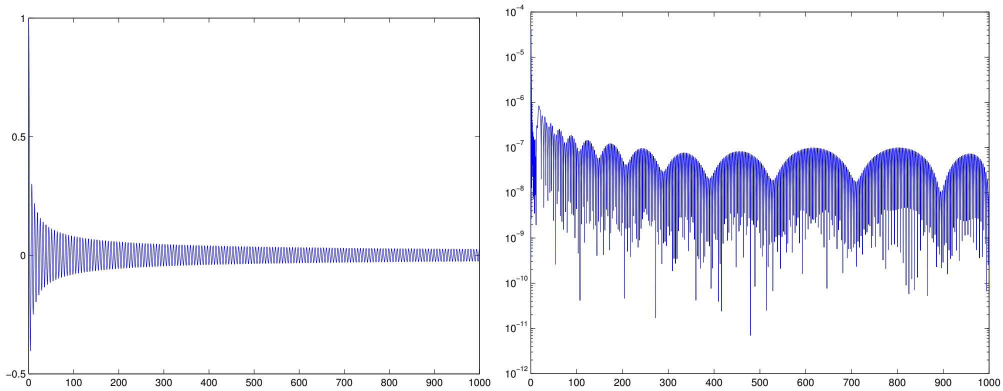

Figure 4.3: Bessel function J_0 (left) and absolute error between J_0 and the exponential sum (4.1) of order 20 (right) on the interval [0, 1000].

Acknowledgment

The first named author gratefully acknowledges the support by the German Research Foundation within the project PO 711/10–1. Moreover, the authors would like to thank the referees for their valuable suggestions.

References

[1] R. Badeau, B. David, and G. Richard. A new perturbation analysis for signal enumeration in rotational invariance techniques. *IEEE Trans. Signal Process.*, 54:450 – 458, 2006.

[2] R. Badeau, B. David, and G. Richard. Cramér–Rao bounds for multiple poles

and coefficients of quasi-polynomials in colored noise. *IEEE Trans. Signal Process.*, 56:3458 – 3467, 2008.

[3] F.S.V. Bazán. Conditioning of rectangular Vandermonde matrices with nodes in the unit disk. *SIAM J. Matrix Anal. Appl.*, 21:679 – 693, 2000.

[4] B. Beckermann, G.H. Golub, and G. Labahn. On the numerical condition of a generalized Hankel eigenvalue problem. *Numer. Math.*, 106:41 – 68, 2007.

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

[6] G. Beylkin and L. Monzón. Approximation by exponential sums revisited. *Appl. Comput. Harmon. Anal.*, 28:131 – 149, 2010.

[7] G. Boutry, M. Elad, G.H. Golub, and P. Milanfar. The generalized eigenvalue problem for nonsquare pencils using a minimal perturbation approach. *SIAM J. Matrix Anal. Appl.*, 27:582 – 601, 2005.

[8] E.J. Candès and C. Fernandez–Granda. Towards a mathematical theory of super-resolution. *arXiv, abs/1203.5871*, 2012.

[9] E.J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE *Trans. Inform. Theory*, 52:489 – 509, 2006.

[10] L. Coluccio, A. Eisinberg, and G. Fedele. A Prony–like method for non–uniform sampling. *Signal Process.*, 87:2484 – 2490, 2007.

[11] 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.

[12] F. Filbir, H.N. Mhaskar, and J. Prestin. On the problem of parameter estimation in exponential sums. *Constr. Approx.*, 35:323 – 343, 2012.

[13] 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.

[14] G.H. Golub and C.F. Van Loan. *Matrix Computations.* Johns Hopkins Univ. Press, Baltimore, 1983.

[15] W. Hackbusch. Entwicklungen nach Exponentialsummen. Technischer Bericht 4, Max–Planck–Institute für Mathematik, Leipzig, 2005.

[16] Y. Hua and T.K. Sarkar. Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. *IEEE Trans. Acoust. Speech Signal Process.*, 38:814 – 824, 1990.

[17] M.A. Iwen. Combinatorial sublinear–time Fourier algorithms. *Found. Comput. Math.*, 10:303–338, 2010.

[18] F. Kittaneh. Singular values of companion matrices and bounds on zeros of polynomials. *SIAM J. Matrix Anal. Appl.*, 16:333 – 340, 1995.

[19] S. Kunis and H. Rauhut. Random sampling of sparse trigonometric polynomials II, Orthogonal matching pursuit versus basis pursuit. *Found. Comput. Math.*, 8:737 – 763, 2008.

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

[21] V. Pereyra and G. Scherer. Exponential data fitting. In V. Pereyra and G. Scherer, editors, *Exponential Data Fitting and its Applications*, pages 1 – 26. Bentham Sci. Publ., Sharjah, 2010.

[22] V. Pereyra and G. Scherer. *Exponential Data Fitting and its Applications.* Bentham Sci. Publ., Sharjah, 2010.

[23] T. Peter, D. Potts, and M. Tasche. Nonlinear approximation by sums of exponentials and translates. *SIAM J. Sci. Comput.*, 33:314 – 334, 2011.

[24] D. Potts and M. Tasche. Parameter estimation for exponential sums by approximate Prony method. *Signal Process.*, 90:1631 – 1642, 2010.

[25] D. Potts and M. Tasche. Nonlinear approximation by sums of nonincreasing exponentials. *Appl. Anal.*, 90:609 – 626, 2011.

[26] D. Potts and M. Tasche. Parameter estimation for multivariate exponential sums. *Preprint, Chemnitz University of Technology*, 2011.

[27] H. Rauhut. Random sampling of sparse trigonometric polynomials. *Appl. Comput. Harmon. Anal.*, 22:16 – 42, 2007.

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

[29] R. Roy and T. Kailath. ESPRIT – estimation of signal parameters via rotational invariance techniques. In L. Auslander, F. Grünbaum, J. Helton, T. Kailath, P. Khargoneka, and S. Mitter, editors, *Signal Processing, Part II*, volume 23 of *IMA Vol. Math. Appl.*, pages 369 – 411. Springer, New York, 1990.

[30] T.K. Sarkar and O. Pereira. Using the matrix pencil method to estimate the parameters of a sum of complex exponentials. *IEEE Antennas Propagation Magazine*, 37:48 – 55, 1995.

[31] M. Vetterli, P. Marziliano, and T. Blu. Sampling signals with finite rate of innovation. *IEEE Trans. Signal Process.*, 50:1417 – 1428, 2002.