Error estimates for the ESPRIT algorithm

Daniel Potts* Manfred Tasche‡

Let $z_j := \mathrm{e}^{f_j}$ $(j = 1, \ldots, M)$ with $f_j \in [-\varphi, 0] + \mathrm{i}\,[-\pi, \pi)$ and small $\varphi \ge 0$ be distinct nodes. With complex coefficients $c_j \neq 0$, we consider an exponential sum $h(x) := c_1\,\mathrm{e}^{f_1\,x} + \ldots + c_M\,\mathrm{e}^{f_M\,x}$ $(x \ge 0)$. Many applications in electrical engineering, signal processing, and mathematical physics lead to the following problem: Determine all parameters of h, if N noisy sampled values $\tilde{h}_k := h(k) + e_k$ $(k = 0, \ldots, N-1)$ with $N \gg 2\,M$ are given, where e_k are small error terms. This parameter identification problem is a nonlinear inverse problem which can be efficiently solved by the ESPRIT algorithm. In this paper, we present mainly corresponding error estimates for the nodes z_j $(j = 1, \ldots, M)$. We show that under appropriate conditions, the results of the ESPRIT algorithm are relatively insensitive to small perturbations on the sampled data.

Key words and phrases: ESPRIT algorithm, exponential sum, frequency analysis problem, parameter identification problem, rectangular Hankel matrix, companion matrix, error estimates.

AMS *Subject Classifications*: 65F15, 65F20, 65F35, 94A12.

1 Introduction

The following *frequency analysis problem* resp. *parameter identification problem* arises in electrical engineering, signal processing, and mathematical physics (see e.g. [15]):
Recover the positive integer M, distinct numbers $f_j \in [-\varphi, 0] + \mathrm{i}\,[-\pi, \pi)$ with small $\varphi \ge 0$, and complex coefficients $c_j \neq 0$ $(j = 1, \ldots, M)$ in the *exponential sum of order* M

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

*potts@mathematik.tu-chemnitz.de, Technische Universität Chemnitz, Faculty of Mathematics, D–09107 Chemnitz, Germany

‡manfred.tasche@uni-rostock.de, University of Rostock, Institute of Mathematics, D–18051 Rostock, Germany

if noisy sampled data $\tilde{h}_k := h(k) + e_k$ $(k = 0, \ldots, N-1)$ with $N > 2\,M$ are given, where $e_k \in \mathbb{C}$ are small error terms with $|e_k| \le \varepsilon_1$ and $0 \le \varepsilon_1 \ll \min\{|c_j|;\, j = 1, \ldots, M\}$.
Note that $-\mathrm{Re}\, f_j \in [0,\, \varphi]$ is the *damping factor* and that $\mathrm{Im}\, f_j \in [-\pi,\, \pi)$ 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 annulus $\mathbb{D} := \{z \in \mathbb{C} :\, \mathrm{e}^{-\varphi} \le |z| \le 1\}$.
This frequency analysis problem can be seen as a nonlinear approximation problem to recover the best M-term approximation of h in the ∞-dimensional linear space $\{\mathrm{e}^{fx};\, f \in [-\varphi,\, 0] + \mathrm{i}\,[-\pi,\, \pi),\, x \ge 0\}$. One known method to solve the frequency analysis problem is the Prony method (see e.g. [17]). But the main drawback of the Prony method is that it may be unstable in some cases. By oversampling of the exponential sum (1.1), i.e. using N sampled data of (1.1) with $N \gg 2\,M$, and applying stable numerical methods, one can obtain efficient algorithms for the parameter identification [6]. A frequently used stable method is the so-called ESPRIT method (ESPRIT = Estimation of Signal Parameters via Rotational Invariance Techniques) [19]. If all nodes z_j $(j = 1, \ldots, M)$ are lying on the unit circle, then an alternative to ESPRIT is the so-called MUSIC algorithm (MUSIC = MUltiple SIgnal Classification). For a systematic numerical study of the MUSIC algorithm see [13].
The aim of this paper is to present error estimates for the ESPRIT algorithm. We show that under appropriate conditions, the results of the ESPRIT algorithm is relatively insensitive to small perturbations on the sampled data. Our study is mainly based on the research of F.S.V. Bazán [1, 3] and on perturbation theory for eigenvalues of a nonnormal $M \times M$ matrix. We extend these results and apply this method to the very popular ESPRIT algorithm. We emphasize that our numerical study is restricted to a moderate size of M.
The outline of this paper is as follows. Section 2 has preliminary character and summarizes known properties (see [1]) of the rectangular Vandermonde matrix

$$\boldsymbol{V}_{P,M}(\boldsymbol{z}) := (z_j^{k-1})_{k,j=1}^{P,M} \tag{1.2}$$

with $\boldsymbol{z} := (z_j)_{j=1}^{M}$. For well-separated nodes $z_j \in \mathbb{D}$ $(j = 1, \ldots, M)$, the rectangular Vandermonde matrix $\boldsymbol{V}_{P,M}(\boldsymbol{z})$ is well conditioned for sufficiently large $P > M$. In Section 3, we solve the frequency analysis problem by the ESPRIT method which is mainly based on singular value decomposition (SVD) of a rectangular Hankel matrix. Instead of SVD, one can use a QR factorization of this rectangular Hankel matrix too (see [17]). In Section 4, we consider the orthogonal projection onto the signal space and the projected companion matrix. In Section 5, we present error estimates of the nodes z_j for the ESPRIT method. Finally, some numerical examples are given in Section 6.
In the following we use standard notations. By $\mathbb{C}$ we denote the set of all complex numbers. The set of all positive integers is $\mathbb{N}$. The linear space of all column vectors with n complex components is denoted by $\mathbb{C}^n$, where $\boldsymbol{0}$ is the corresponding zero vector. By $\boldsymbol{e}_j := (\delta_{j,k})_{k=1}^{n}$ $(j = 1, \ldots, n)$ we denote the canonical basis vectors of $\mathbb{C}^n$. The 2-norm in $\mathbb{C}^n$ is $\|\cdot\|_2$. The linear space of all complex $m \times n$ 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 $\boldsymbol{A}_{m,n}^{\mathrm{T}}$, its conjugate–transpose is $\boldsymbol{A}_{m,n}^{*}$, and its Moore–Penrose pseudoinverse is $\boldsymbol{A}_{m,n}^{\dagger}$.

A square $m \times m$ matrix is abbreviated to $\boldsymbol{A}_m$. By $\boldsymbol{I}_m$ we denote the $m \times m$ identity matrix. The spectral norm in $\mathbb{C}^{m\times n}$ is denoted by $\|\cdot\|_2$, and the Frobenius norm is $\|\cdot\|_{\mathrm{F}}$. Further we use the known submatrix notation. Thus $\boldsymbol{A}_{m,m+1}(1:m,\, 2:m+1)$ is the $m \times m$ submatrix of $\boldsymbol{A}_{m,m+1}$ obtained by extracting rows 1 through m and columns 2 through $m+1$. Note that the first row or column of a matrix can be indexed by zero. Other notations are introduced when needed.

2 Rectangular Vandermonde matrices

Let M and P be given positive integers with $P > M$. Further let $z_j \in \mathbb{D}$ $(j = 1, \ldots, M)$ be given distinct nodes and $\boldsymbol{z} := (z_j)_{j=1}^M$. Under mild assumptions, it was shown in [1] that the $P \times M$ Vandermonde matrix (1.2) is well conditioned for sufficiently large $P > M$. Let

$$\mathrm{cond}_2\, \boldsymbol{V}_{P,M}(\boldsymbol{z}) := \|\boldsymbol{V}_{P,M}(\boldsymbol{z})\|_2\, \|\boldsymbol{V}_{P,M}(\boldsymbol{z})^{\dagger}\|_2$$

be the *spectral norm condition number* of (1.2), where $\boldsymbol{V}_{P,M}(\boldsymbol{z})^{\dagger}$ denotes the Moore–Penrose pseudoinverse of (1.2) (see [9, p. 382]). Since $\boldsymbol{V}_{P,M}(\boldsymbol{z})$ has full rank, we have

$$\boldsymbol{V}_{P,M}(\boldsymbol{z})^{\dagger} = \big(\boldsymbol{V}_{P,M}(\boldsymbol{z})^{*}\, \boldsymbol{V}_{P,M}(\boldsymbol{z})\big)^{-1}\, \boldsymbol{V}_{P,M}(\boldsymbol{z})^{*}\,.$$

Analogously,

$$\mathrm{cond}_{\mathrm{F}}\, \boldsymbol{V}_{P,M}(\boldsymbol{z}) := \|\boldsymbol{V}_{P,M}(\boldsymbol{z})\|_{\mathrm{F}}\, \|\boldsymbol{V}_{P,M}(\boldsymbol{z})^{\dagger}\|_{\mathrm{F}}$$

is the *Frobenius norm condition number* of (1.2). Since the spectral norm and the Frobenius norm are unitarily invariant norms, we conclude that

$$\|\boldsymbol{V}_{P,M}(\boldsymbol{z})\|_2 = \|\boldsymbol{V}_{P,M}(\boldsymbol{z})^{\mathrm{T}}\|_2\,, \quad \|\boldsymbol{V}_{P,M}(\boldsymbol{z})\|_{\mathrm{F}} = \|\boldsymbol{V}_{P,M}(\boldsymbol{z})^{\mathrm{T}}\|_{\mathrm{F}}\,.$$

Further we remark that $\big(\boldsymbol{V}_{P,M}(\boldsymbol{z})^{\dagger}\big)^{\mathrm{T}} = \big(\boldsymbol{V}_{P,M}(\boldsymbol{z})^{\mathrm{T}}\big)^{\dagger}$. Then we introduce the following entries

$$\begin{aligned}
\alpha &:= \max\{|z_j|;\, j = 1, \ldots, M\} \in [\mathrm{e}^{-\varphi},\, 1]\,, && (2.1)\\
\beta &:= \min\{|z_j|;\, j = 1, \ldots, M\} \in [\mathrm{e}^{-\varphi},\, 1]\,, && (2.2)\\
\mu &:= \sum_{j=1}^{M} |z_j|^2 \in [M\,\mathrm{e}^{-2\varphi},\, M]\,, && (2.3)\\
\nu &:= \prod_{j=1}^{M} |z_j|^2 \in [\mathrm{e}^{-2M\varphi},\, 1]\,.
\end{aligned}$$

The *separation distance* of all nodes $z_j \in \mathbb{D}$ $(j = 1, \ldots, M)$ is explained by

$$\delta := \min\{|z_j - z_k|;\, j,\, k = 1, \ldots, M,\, j \neq k\} > 0\,. \tag{2.4}$$

For $P > M$, the spectral norm of $\big(\boldsymbol{V}_{P,M}(\boldsymbol{z})^{\mathrm{T}}\big)^{\dagger}$ decreases monotonously with respect to P. If $|z_j| = 1$ for all $j = 1, \ldots, M$, then by [1, Theorem 1]

$$\lim_{P\to\infty} \|\big(\boldsymbol{V}_{P,M}(\boldsymbol{z})^{\mathrm{T}}\big)^{\dagger}\|_2 = 0\,.$$

Let $\boldsymbol{q}_P \in \mathbb{C}^P$ with $P > M$ be the minimum 2-norm solution of the underdetermined linear system

$$\boldsymbol{V}_{P,M}(\boldsymbol{z})^{\mathrm{T}}\, \boldsymbol{q}_P = -\big(z_j^P\big)_{j=1}^M \tag{2.5}$$

such that

$$\boldsymbol{q}_P = -\big(\boldsymbol{V}_{P,M}(\boldsymbol{z})^{\mathrm{T}}\big)^{\dagger}\, \big(z_j^P\big)_{j=1}^M .$$

Then by [1, Theorem 2], the norms $\|\boldsymbol{q}_P\|_2$ are bounded with respect to P. If either $|z_j| = 1$ $(j = 1, \ldots, M)$ or $|z_j| < 1$ $(j = 1, \ldots, M)$, then

$$\lim_{P\to\infty} \|\boldsymbol{q}_P\|_2 = 0 .$$

Let $P > M \ge 2$ and let $\boldsymbol{V}_{P,M}(\boldsymbol{z})$ be the $P \times M$ Vandermonde matrix with distinct nodes $z_j \in \mathbb{D}$ $(j = 1, \ldots, M)$. Then by [1, Theorem 6 and Lemma 7], the Frobenius norm condition number of $\boldsymbol{V}_{P,M}(\boldsymbol{z})$ can be estimated by

$$\mathrm{cond}_{\mathrm{F}}\, \boldsymbol{V}_{P,M}(\boldsymbol{z}) \le M\, \Big(1 + \frac{\|\boldsymbol{q}_P\|_2^2 + M + \nu - \mu - 1}{(M-1)\,\delta^2}\Big)^{(M-1)/2}\, \Phi_P(\alpha,\beta) \tag{2.6}$$

with

$$\Phi_P(\alpha,\beta) := \Big(\frac{1 + \alpha^2 + \alpha^4 + \ldots + \alpha^{2(P-1)}}{1 + \beta^2 + \beta^4 + \ldots + \beta^{2(P-1)}}\Big)^{1/2} ,$$

where α, β, μ, ν, and δ are defined by (2.1) – (2.4).

The Vandermonde matrix $\boldsymbol{V}_{P,M}(\boldsymbol{z})$ with $P > M$ satisfies the inequality

$$M - 2 + \mathrm{cond}_2\, \boldsymbol{V}_{P,M}(\boldsymbol{z}) + \big(\mathrm{cond}_2\, \boldsymbol{V}_{P,M}(\boldsymbol{z})\big)^{-1} \le \mathrm{cond}_{\mathrm{F}}\, \boldsymbol{V}_{P,M}(\boldsymbol{z}) . \tag{2.7}$$

This inequality (2.7) follows directly from a corresponding result in [20] for an invertible, square matrix. Since the rectangular Vandermonde matrix $\boldsymbol{V}_{P,M}(\boldsymbol{z})$ possesses full rank M, the square matrix $\boldsymbol{V}_{P,M}(\boldsymbol{z})^*\, \boldsymbol{V}_{P,M}(\boldsymbol{z})$ is positive definite. Hence its square root $\big(\boldsymbol{V}_{P,M}(\boldsymbol{z})^*\, \boldsymbol{V}_{P,M}(\boldsymbol{z})\big)^{1/2}$ is defined. The eigenvalues of $\big(\boldsymbol{V}_{P,M}(\boldsymbol{z})^*\, \boldsymbol{V}_{P,M}(\boldsymbol{z})\big)^{1/2}$ coincide with the singular values of $\boldsymbol{V}_{P,M}(\boldsymbol{z})$. Thus one obtains for the spectral resp. Frobenius norm

$$\begin{aligned} \mathrm{cond}_2\, \big(\boldsymbol{V}_{P,M}(\boldsymbol{z})^*\, \boldsymbol{V}_{P,M}(\boldsymbol{z})\big)^{1/2} &= \mathrm{cond}_2\, \boldsymbol{V}_{P,M}(\boldsymbol{z}) , \\ \mathrm{cond}_{\mathrm{F}}\, \big(\boldsymbol{V}_{P,M}(\boldsymbol{z})^*\, \boldsymbol{V}_{P,M}(\boldsymbol{z})\big)^{1/2} &= \mathrm{cond}_{\mathrm{F}}\, \boldsymbol{V}_{P,M}(\boldsymbol{z}) . \end{aligned}$$

From (2.7) it follows that

$$\begin{aligned} \mathrm{cond}_2\, \boldsymbol{V}_{P,M}(\boldsymbol{z}) \;\le\; & \frac{1}{2}\, \Big(\mathrm{cond}_{\mathrm{F}}\, \boldsymbol{V}_{P,M}(\boldsymbol{z}) - M + 2\Big) \\ & + \frac{1}{2}\, \Big(\big(\mathrm{cond}_{\mathrm{F}}\, \boldsymbol{V}_{P,M}(\boldsymbol{z}) - M + 2\big)^2 - 4\Big)^{1/2} . \end{aligned}$$

If all nodes z_j are lying on the unit circle, i.e. $|z_j| = 1$ $(j = 1, \ldots, M)$, then $\alpha = \beta = \nu = \Phi_P(1,1) = 1$, $\mu = M$, and hence by (2.6)

$$\mathrm{cond}_{\mathrm{F}}\, \boldsymbol{V}_{P,M}(\boldsymbol{z}) \le M\, \Big(1 + \frac{\|\boldsymbol{q}_P\|_2^2}{(M-1)\,\delta^2}\Big)^{(M-1)/2} .$$

In the case $|z_j| = 1$ $(j = 1, \ldots, M)$, better estimates for the spectral norm condition number of $\boldsymbol{V}_{P,M}(\boldsymbol{z})$ are possible (see [16, 13]).

3 ESPRIT algorithm

In practice, the order M of the exponential sum (1.1) is often unknown. Assume that $L \in \mathbb{N}$ is a convenient upper bound of M with $M \le L \le \lceil \frac{N}{2} \rceil$, where N is a sufficiently large integer with $N \gg 2M$. In applications, such an upper bound L is mostly known *a priori*. If this is not the case, then one can choose $L = \lceil \frac{N}{2} \rceil$. Suppose that N noisy sampled data $\tilde{h}_k := h(k) + e_k \in \mathbb{C}$ $(k = 0, \ldots, N-1)$ of (1.1) are given, where $e_k \in \mathbb{C}$ are small error terms with $|e_k| \le \varepsilon_1$ and $0 \le \varepsilon_1 \ll 1$. Often the sequence $\{\tilde{h}_0, \tilde{h}_1, \ldots, \tilde{h}_{N-1}\}$ of sampled data is called as a *time series of length* N. Then we form the *L-trajectory matrix* of this time series

$$\tilde{\boldsymbol{H}}_{L,N-L+1} := \big(\tilde{h}_{\ell+m}\big)_{\ell,m=0}^{L-1,N-L} \tag{3.1}$$

with the *window length* $L \in \{M, \ldots, \lceil \frac{N}{2} \rceil\}$. Obviously, (3.1) is an $L \times (N-L+1)$ Hankel matrix.
The nonincreasingly ordered singular values $\tilde{\sigma}_k(\tilde{\boldsymbol{H}}_{L,N-L+1})$ of the L-trajectory matrix (3.1) possess the following property:

Lemma 3.1 *For fixed $N \gg 2\,M$, the singular values of* (3.1) *increase almost monotonously with respect to $L = M, \ldots, \lceil \frac{N}{2} \rceil$, i.e., they fulfil the inequalities*

$$\tilde{\sigma}_k(\tilde{\boldsymbol{H}}_{L,N-L+1})^2 \le \tilde{\sigma}_k(\tilde{\boldsymbol{H}}_{L+1,N-L})^2 + \|\tilde{\boldsymbol{h}}_L\|_2^2 \quad (k = 1, \ldots, M)\,, \tag{3.2}$$

where $\tilde{\boldsymbol{h}}_L := (\tilde{h}_k)_{k=N-L}^{N-1}$.

Proof. For $M \le L < \lceil \frac{N}{2} \rceil$, we represent the Hankel matrices $\tilde{\boldsymbol{H}}_{L,N-L+1}$ and $\tilde{\boldsymbol{H}}_{L+1,N-L}$ as block matrices

$$\tilde{\boldsymbol{H}}_{L,N-L+1} = \big(\tilde{\boldsymbol{H}}_{L,N-L}\ \tilde{\boldsymbol{h}}_L\big)\,, \quad \tilde{\boldsymbol{H}}_{L+1,N-L} = \begin{pmatrix} \tilde{\boldsymbol{H}}_{L,N-L} \\ \tilde{\boldsymbol{h}}_{N-L}^{\mathrm{T}} \end{pmatrix}$$

with $\tilde{\boldsymbol{h}}_{N-L} := (\tilde{h}_k)_{k=L}^{N-1}$. Setting $\tilde{\boldsymbol{B}}_L := \tilde{\boldsymbol{H}}_{L,N-L}\,\tilde{\boldsymbol{H}}_{L,N-L}^*$, we obtain that

$$\tilde{\boldsymbol{A}}_L := \tilde{\boldsymbol{H}}_{L,N-L+1}\,\tilde{\boldsymbol{H}}_{L,N-L+1}^* = \tilde{\boldsymbol{B}}_L + \tilde{\boldsymbol{h}}_L\,\tilde{\boldsymbol{h}}_L^*\,, \tag{3.3}$$

which is a rank-one Hermitian perturbation of the Hermitian matrix $\tilde{\boldsymbol{B}}_L$, and

$$\tilde{\boldsymbol{A}}_{L+1} := \tilde{\boldsymbol{H}}_{L+1,N-L}\,\tilde{\boldsymbol{H}}_{L+1,N-L}^* = \begin{pmatrix} \tilde{\boldsymbol{B}}_L & \tilde{\boldsymbol{y}}_L \\ \tilde{\boldsymbol{y}}_L^* & \|\tilde{\boldsymbol{h}}_{N-L}\|_2^2 \end{pmatrix}$$

with $\tilde{\boldsymbol{y}}_L := \tilde{\boldsymbol{H}}_{L,N-L}\,\overline{\tilde{\boldsymbol{h}}}_{N-L}$. Using Cauchy's Interlacing Theorem (see [10, p. 242]) for the bordered Hermitian matrix $\tilde{\boldsymbol{A}}_{L+1}$, the corresponding nondecreasingly ordered eigenvalues of $\tilde{\boldsymbol{A}}_{L+1}$ and $\tilde{\boldsymbol{B}}_L$ fulfil the inequalities

$$\tilde{\lambda}_j(\tilde{\boldsymbol{B}}_L) \le \tilde{\lambda}_{j+1}(\tilde{\boldsymbol{A}}_{L+1}) \quad (j = 1, \ldots, L)\,.$$

By (3.3) and by Weyl's Theorem (see [10, p. 239]), we obtain that

$$\tilde{\lambda}_1(-\tilde{\boldsymbol{h}}_L\,\tilde{\boldsymbol{h}}_L^*) + \tilde{\lambda}_j(\tilde{\boldsymbol{A}}_L) \le \tilde{\lambda}_j(\tilde{\boldsymbol{B}}_L) \le \tilde{\lambda}_{j+1}(\tilde{\boldsymbol{A}}_{L+1})\,.$$

Since the first eigenvalue of the rank-one matrix $-\tilde{\boldsymbol{h}}_L\,\tilde{\boldsymbol{h}}_L^*$ is equal to $-\|\tilde{\boldsymbol{h}}_L\|_2^2$, we obtain that

$$\tilde{\lambda}_j(\tilde{\boldsymbol{A}}_L) \le \tilde{\lambda}_{j+1}(\tilde{\boldsymbol{A}}_{L+1}) + \|\tilde{\boldsymbol{h}}_L\|_2^2\,.$$

The non-zero eigenvalues of $\tilde{\boldsymbol{A}}_L$ resp. $\tilde{\boldsymbol{A}}_{L+1}$ are the squares of the positive singular values of $\tilde{\boldsymbol{H}}_{L,N-L+1}$ resp. $\tilde{\boldsymbol{H}}_{L+1,N-L}$. This completes the proof of (3.2). ■

The convenient choice of the window length L is essential for the following ESPRIT method. By Lemma 3.1, a sufficiently large integer $L \approx \lceil \frac{N}{2} \rceil$ is a good choice. Then the L-trajectory matrix (3.1) with $L = \lceil \frac{N}{2} \rceil$ is almost square. Several numerical experiments in [8] confirm that the optimal window length L lies in the near of $\lceil \frac{N}{2} \rceil$.

The main step in the solution of the frequency analysis problem is the determination of the order M and the computation of the exponents f_j or alternatively of the nodes $z_j = \mathrm{e}^{f_j} \in \mathbb{D}$ $(j = 1, \ldots, M)$. Afterwards one can calculate the coefficient vector $\boldsymbol{c} := (c_j)_{j=1}^M \in \mathbb{C}^M$ as least squares solution of the overdetermined linear system

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

with the rectangular Vandermonde matrix (1.2), i.e., the coefficient vector $\boldsymbol{c}$ is the solution of the least squares problem

$$\|\boldsymbol{V}_{N,M}(\boldsymbol{z})\,\boldsymbol{c} - \big(\tilde{h}_k\big)_{k=0}^{N-1}\|_2 = \min\,.$$

As known, the square Vandermonde matrix $\boldsymbol{V}_M(\boldsymbol{z})$ is invertible and the matrix $\boldsymbol{V}_{N,M}(\boldsymbol{z})$ has full column rank. Additionally we introduce the rectangular Hankel matrices

$$\tilde{\boldsymbol{H}}_{L,N-L}(s) := \tilde{\boldsymbol{H}}_{L,N-L+1}(1:L,\, 1+s:N-L+s) \quad (s = 0,\, 1)\,. \tag{3.4}$$

In the case of exactly sampled data $\tilde{h}_k = h(k)$ $(k = 0, \ldots, N-1)$, the Hankel matrix (3.1) is denoted by $\boldsymbol{H}_{L,N-L+1}$ and the related Hankel matrices (3.4) are denoted by $\boldsymbol{H}_{L,N-L}(s)$ $(s = 0,\, 1)$.

Remark 3.2 The Hankel matrices $\boldsymbol{H}_{L,N-L+1}$ and $\boldsymbol{H}_{L,N-L}(s)$ $(s = 0,\, 1)$ have the same rank M for each window length $L \in \{M, \ldots, \lceil \frac{N}{2} \rceil\}$ (see [17, Lemma 2.1]). Consequently, the order M of the exponential sum (1.1) coincides with the rank of these Hankel matrices. □

First we assume that exactly sampled data $\tilde{h}_k = h(k)$ $(k = 0, \ldots, N-1)$ of (1.1) are given. We choose $L \approx \lceil \frac{N}{2} \rceil$. Then the matrix pencil

$$z\,\boldsymbol{H}_{L,N-L}(0) - \boldsymbol{H}_{L,N-L}(1) \quad (z \in \mathbb{C}) \tag{3.5}$$

has the nodes $z_j \in \mathbb{D}$ $(j = 1, \ldots, M)$ as eigenvalues (see e.g. [11, 17]). We start the ESPRIT method by the SVD of the exact L-trajectory matrix $\boldsymbol{H}_{L,N-L+1}$, i.e.

$$\boldsymbol{H}_{L,N-L+1} = \boldsymbol{U}_L\, \boldsymbol{D}_{L,N-L+1}\, \boldsymbol{W}^*_{N-L+1}\,,$$

where $\boldsymbol{U}_L \in \mathbb{C}^{L\times L}$ and $\boldsymbol{W}_{N-L+1} \in \mathbb{C}^{(N-L+1)\times(N-L+1)}$ are unitary matrices and where $\boldsymbol{D}_{L,N-L+1} \in \mathbb{R}^{L\times(N-L+1)}$ is a rectangular diagonal matrix. The diagonal entries of $\boldsymbol{D}_{L,N-L+1}$ are the singular values σ_j of the L-trajectory matrix arranged in nonincreasing order $\sigma_1 \geq \ldots \geq \sigma_M > \sigma_{M+1} = \ldots = \sigma_L = 0$. Thus we can determine the order M of the exponential sum (1.1) by the number of positive singular values σ_j.

Remark 3.3 For fixed N, the size of the lowest positive singular value σ_M of $\boldsymbol{H}_{L,N-L+1}$ with $M \leq L \leq \lceil \frac{N}{2} \rceil$ depends on the choice of L by Lemma 3.1. The M positive singular values of $\boldsymbol{H}_{L,N-L+1}$ are the square roots of the M positive eigenvalues of

$$\boldsymbol{A}_L := \boldsymbol{H}_{L,N-L+1}\, \boldsymbol{H}^*_{L,N-L+1}\,.$$

All the other singular values of $\boldsymbol{H}_{L,N-L+1}$ resp. eigenvalues of $\boldsymbol{A}_L$ are zero. The trace of $\boldsymbol{A}_L$ is equal to

$$\mathrm{tr}\,(\boldsymbol{A}_L) = \sum_{\ell=0}^{L-1} \sum_{j=\ell}^{N-L+\ell} |h(j)|^2$$

and the sum of all the principal minors of size 2 amounts

$$\begin{aligned} s_2(\boldsymbol{A}_L) \;&=\; \sum_{\ell=0}^{L-2} \sum_{k=1}^{L-\ell-1} \Big[\Big(\sum_{j=\ell}^{N-L+\ell} |h(j)|^2 \Big) \Big(\sum_{j=\ell}^{N-L+\ell} |h(j+k)|^2 \Big) \\ &\quad - \Big| \sum_{j=\ell}^{N-L+\ell} h(j)\, \overline{h(j+k)} \,\Big|^2 \Big]\,. \end{aligned}$$

In the case $M = 1$, i.e. $h(j) = c_1\, z_1^j$ with $c_1 \neq 0$ and $0 < |z_1| \leq 1$, the only positive eigenvalue λ_1 of $\boldsymbol{A}_L$ reads as follows

$$\lambda_1 = \mathrm{tr}\,(\boldsymbol{A}_L) = |c_1|^2 \Big(\sum_{j=0}^{N-L} |z_1|^{2j} \Big) \Big(\sum_{\ell=0}^{L-1} |z_1|^{2\ell} \Big)\,,$$

so that the only positive singular value σ_1 of $\boldsymbol{H}_{L,N-L+1}$ fulfils the estimate

$$\sigma_1 = |c_1| \sqrt{\Big(\sum_{j=0}^{N-L} |z_1|^{2j} \Big) \Big(\sum_{\ell=0}^{L-1} |z_1|^{2\ell} \Big)} \leq |c_1|\, \sqrt{(N-L+1)\,L} \qquad (3.6)$$

with equality for $|z_1| = 1$. Thus σ_1 is maximal for $L = \lceil \frac{N}{2} \rceil$.

In the case $M = 2$, i.e. $h(j) = c_1\, z_1^j + c_2\, z_2^j$ with $c_k \neq 0$ and $0 < |z_k| \le 1$ $(k = 1,\, 2)$, there exist only two positive eigenvalues λ_1, λ_2 of $\boldsymbol{A}_L$, all the other eigenvalues of $\boldsymbol{A}_L$ vanish. Then λ_1 and λ_2 are the solutions of the quadratic equation

$$\lambda^2 - \lambda\, \mathrm{tr}\,(\boldsymbol{A}_L) + s_2(\boldsymbol{A}_L) = 0$$

(see [10, p. 54]) so that

$$\lambda_{1,2} = \frac{1}{2}\, \mathrm{tr}\,(\boldsymbol{A}_L) \pm \frac{1}{2}\, \sqrt{(\mathrm{tr}\,(\boldsymbol{A}_L))^2 - 4\, s_2(\boldsymbol{A}_L)}\,.$$

Hence the two positive singular values of $\boldsymbol{H}_{L,N-L+1}$ are $\sigma_{1,2} = \sqrt{\lambda_{1,2}}$.
In the case $M > 1$, one can estimate the positive singular values of $\boldsymbol{H}_{L,N-L+1}$ by Weyl's Theorem (see [21, p. 68]). Since

$$\boldsymbol{H}_{L,N-L+1} = \sum_{k=1}^{M} \big(c_k\, z_k^{\ell+m}\big)_{\ell,m=0}^{L-1,N-L}$$

with $c_k \neq 0$ and $0 < |z_k| \le 1$ $(k = 1, \ldots, M)$, one obtains by (3.6) that

$$0 < \sigma_M \le \sigma_1 \le \sum_{k=1}^{M} |c_k|\, \sqrt{(N-L+1)\, L}\,. \tag{3.7}$$

A lower estimate of σ_M was presented in [4].
A good criterion for the choice of optimal window length L is to maximize the lowest positive singular value σ_M of $\boldsymbol{H}_{L,N-L+1}$. By Lemma 3.1, (3.6) and (3.7), one can see that $L = \lceil \frac{N}{2} \rceil$ is a good choice (see also [4, 8] and Example 6.4). For the ESPRIT Algorithm 3.4 (i.e. the determination of the numerical rank in step 1) and the corresponding error estimates (see Theorem 5.4), it is very important that σ_M is not too small. □

Introducing the matrices $\boldsymbol{U}_{L,M} := \boldsymbol{U}_M(1:L,\, 1:M)$ and $\boldsymbol{W}_{N-L+1,M} := \boldsymbol{W}_{N-L+1}(1:N-L+1,\, 1:M)$ with orthonormal columns as well as the diagonal matrix $\boldsymbol{D}_M := \mathrm{diag}\,(\sigma_j)_{j=1}^M$, we obtain the partial SVD of the matrix (3.1) with exact entries, i.e.

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

Setting

$$\boldsymbol{W}_{N-L,M}(s) := \boldsymbol{W}_{N-L+1,M}(1+s : N-L+s, 1:M) \quad (s = 0,\, 1)\,, \tag{3.8}$$

it follows by (3.8) and (3.4) that both Hankel matrices (3.4) can be simultaneously factorized in the form

$$\boldsymbol{H}_{L,N-L}(s) = \boldsymbol{U}_{L,M}\, \boldsymbol{D}_M\, \boldsymbol{W}_{N-L,M}(s)^* \quad (s = 0,\, 1)\,. \tag{3.9}$$

Since $\boldsymbol{U}_{L,M}$ has orthonormal columns and since $\boldsymbol{D}_M$ is invertible, the generalized eigenvalue problem of the matrix pencil

$$z\, \boldsymbol{W}_{N-L,M}(0)^* - \boldsymbol{W}_{N-L,M}(1)^* \quad (z \in \mathbb{C})$$

has the same non-zero eigenvalues $z_j \in \mathbb{D}$ $(j = 1, \ldots, M)$ as the matrix pencil (3.5) except for additional zero eigenvalues. Finally we determine the nodes $z_j \in \mathbb{D}$ $(j = 1, \ldots, M)$ as eigenvalues of the $M \times M$ matrix

$$\boldsymbol{F}_M^{\mathrm{SVD}} := \boldsymbol{W}_{N-L,M}(1)^* \left(\boldsymbol{W}_{N-L,M}(0)^*\right)^\dagger \tag{3.10}$$

Analogously, we can handle the general case of noisy data $\tilde{h}_k = h(k) + e_k \in \mathbb{C}$ $(k = 0, \ldots, N-1)$ with small error terms $e_k \in \mathbb{C}$, where $|e_k| \le \varepsilon_1$ and $0 < \varepsilon \ll 1$. For the Hankel matrix (3.1) with the singular values $\tilde{\sigma}_1 \ge \ldots \ge \tilde{\sigma}_L \ge 0$, we can calculate the numerical rank M of (3.1) by the property $\tilde{\sigma}_M \ge \varepsilon\, \tilde{\sigma}_1$ and $\tilde{\sigma}_{M+1} < \varepsilon\, \tilde{\sigma}_1$ with convenient chosen tolerance ε. Using the IEEE double precision arithmetic, one can choose $\varepsilon = 10^{-10}$ for given exact data. In the case of noisy data, one has to use a larger tolerance ε. Let

$$\boldsymbol{E}_{L,N-L+1} := \tilde{\boldsymbol{H}}_{L,N-L+1} - \boldsymbol{H}_{L,N-L+1}$$

be the *error matrix of given data*. Assuming $2\,\|\boldsymbol{E}_{L,N-L+1}\|_2 \ll \sigma_M$ and choosing $\varepsilon \approx 2\,\|\boldsymbol{E}_{L,N-L+1}\|_2/\tilde{\sigma}_1$, we find by Weyl's Theorem (see [21, p. 70]) that

$$|\tilde{\sigma}_j - \sigma_j| \le \|\boldsymbol{E}_{L,N-L+1}\|_2 \quad (j = 1, \ldots, L)\,.$$

Thus one obtains that $\tilde{\sigma}_M \ge \sigma_M - \|\boldsymbol{E}_{L,N-L+1}\|_2 \gg \|\boldsymbol{E}_{L,N-L+1}\|_2 \approx \varepsilon\, \tilde{\sigma}_1$ and $\tilde{\sigma}_{M+1} \le \|\boldsymbol{E}_{L,N-L+1}\|_2 \approx \frac{\varepsilon}{2}\, \tilde{\sigma}_1$, i.e. $\tilde{\sigma}_M/\tilde{\sigma}_1 \ge \varepsilon$ and $\tilde{\sigma}_{M+1}/\tilde{\sigma}_1 < \varepsilon$.
For the Hankel matrix (3.1) with noisy entries, we use its SVD

$$\tilde{\boldsymbol{H}}_{L,N-L+1} = \tilde{\boldsymbol{U}}_L\, \tilde{\boldsymbol{D}}_{L,N-L+1}\, \tilde{\boldsymbol{W}}^*_{N-L+1}$$

and define as above the matrices $\tilde{\boldsymbol{U}}_{L,M}$, $\tilde{\boldsymbol{D}}_M := \mathrm{diag}\left(\tilde{\sigma}_j\right)_{j=1}^M$, and $\tilde{\boldsymbol{W}}_{N-L+1,M}$. Then

$$\tilde{\boldsymbol{U}}_{L,M}\, \tilde{\boldsymbol{D}}_M\, \tilde{\boldsymbol{W}}^*_{N-L+1,M}$$

is a low-rank approximation of (3.1). Analogously to (3.8) and (3.10), we introduce corresponding matrices $\tilde{\boldsymbol{W}}_{N-L,M}(s)$ $(s = 0,\, 1)$ and $\tilde{\boldsymbol{F}}_M^{\mathrm{SVD}}$. Note that

$$\tilde{\boldsymbol{K}}_{L,N-L}(s) := \tilde{\boldsymbol{U}}_{L,M}\, \tilde{\boldsymbol{D}}_M\, \tilde{\boldsymbol{W}}_{N-L,M}(s)^* \quad (s = 0,\, 1) \tag{3.11}$$

is a low-rank approximation of $\tilde{\boldsymbol{H}}_{L,N-L}(s)$. Thus the SVD-based ESPRIT algorithm reads as follows:

Algorithm 3.4 (ESPRIT via SVD)
Input: $N \in \mathbb{N}$ $(N \gg 2\,M)$, M unknown order of (1.1), $L \approx \lceil \frac{N}{2} \rceil$ window length with $M \le L \le \lceil \frac{N}{2} \rceil$, $\tilde{h}_k = h(k) + e_k \in \mathbb{C}$ $(k = 0, \ldots, N-1)$ noisy sampled values of (1.1), $0 < \varepsilon \ll 1$ tolerance.

1. Compute the SVD of the rectangular Hankel matrix (3.1). Determine the numerical rank M of (3.1) such that $\tilde{\sigma}_M \ge \varepsilon\, \tilde{\sigma}_1$ and $\tilde{\sigma}_{M+1} < \varepsilon \tilde{\sigma}_1$. Form the matrices $\tilde{\boldsymbol{W}}_{N-L,M}(s)$ $(s = 0,\, 1)$ as in (3.8).

2. Calculate the $M \times M$ matrix $\tilde{\boldsymbol{F}}_M^{\mathrm{SVD}}$ as in (3.10) and compute all eigenvalues $\tilde{z}_j \in \mathbb{D}$ $(j = 1, \ldots, M)$ of $\tilde{\boldsymbol{F}}_M^{\mathrm{SVD}}$. Set $\tilde{f}_j := \log \tilde{z}_j$ $(j = 1, \ldots, M)$, where log denotes the principal value of the complex logarithm.
3. Compute the coefficient vector $\tilde{\boldsymbol{c}} := (\tilde{c}_j)_{j=1}^M \in \mathbb{C}^M$ as solution of the least squares problem

$$\|\boldsymbol{V}_{N,M}(\tilde{\boldsymbol{z}})\,\tilde{\boldsymbol{c}} - \big(\tilde{h}_k\big)_{k=0}^{N-1}\|_2 = \min\,,$$

where $\tilde{\boldsymbol{z}} := \big(\tilde{z}_j\big)_{j=1}^M$ denotes the vector of computed nodes.

Output: $M \in \mathbb{N}$, $\tilde{f}_j \in [-\varphi,\, 0] + \mathrm{i}\,[-\pi, \pi)$, $\tilde{c}_j \in \mathbb{C}$ $(j = 1, \ldots, M)$.

Remark 3.5 One can pass on the computation of the Moore–Penrose pseudoinverse in (3.10). Then the second step of Algorithm 3.4 reads as follows (see [17, Algorithm 3.1]):

2′. Calculate the matrix products

$$\tilde{\boldsymbol{A}}_M := \tilde{\boldsymbol{W}}_{N-L,M}(0)^*\,\tilde{\boldsymbol{W}}_{N-L,M}(0)\,, \quad \tilde{\boldsymbol{B}}_M := \tilde{\boldsymbol{W}}_{N-L,M}(1)^*\,\tilde{\boldsymbol{W}}_{N-L,M}(0)$$

and compute all eigenvalues $\tilde{z}_j \in \mathbb{D}$ $(j = 1, \ldots, M)$ of the square matrix pencil $z\,\tilde{\boldsymbol{A}}_M - \tilde{\boldsymbol{B}}_M$ $(z \in \mathbb{C})$ by the QZ–Algorithm (see [7, pp. 384 – 385]). Set $\tilde{f}_j := \log \tilde{z}_j$ $(j = 1, \ldots, M)$. □

In the second step of Algorithm 3.4, the matrix (3.10) can be replaced by the matrix

$$\boldsymbol{F}_M := \boldsymbol{X}_{N-L,M}(1)^*\,\boldsymbol{X}_{N-L,M}(0)\,, \tag{3.12}$$

where

$$\boldsymbol{X}_{N-L,M}(s) := \boldsymbol{W}_{N-L,M}(s)\big(\boldsymbol{W}_{N-L,M}(0)^*\,\boldsymbol{W}_{N-L,M}(0)\big)^{-1/2} \quad (s = 0,\, 1)\,. \tag{3.13}$$

Since $\boldsymbol{W}_{N-L,M}(0)^*\,\boldsymbol{W}_{N-L,M}(0)$ is positive definite, the above matrix (3.13) is well-defined. Obviously, we have

$$\boldsymbol{X}_{N-L,M}(0)^*\,\boldsymbol{X}_{N-L,M}(0) = \boldsymbol{I}_M\,, \tag{3.14}$$

i.e., the columns of $\boldsymbol{X}_{N-L,M}(0)$ are orthonormal. As later will be shown in Lemma 4.2, the new matrix (3.12) has the same eigenvalues $z_j \in \mathbb{D}$ $(j = 1, \ldots, M)$ as (3.10).

4 Orthogonal projection onto the signal space

First we consider the ESPRIT method for exact sampled data $\tilde{h}_k = h(k)$ $(k = 0, \ldots, N-1)$ of the exponential sum (1.1) of order M. We choose a convenient window length $L \approx \lceil \frac{N}{2} \rceil$ such that $M \le L \le \lceil \frac{N}{2} \rceil$.
Analogously to (2.5), the vector $\boldsymbol{q}_{N-L} := (q_k)_{k=0}^{N-L-1} \in \mathbb{C}^{N-L}$ is defined as the minimum 2-norm solution of the (underdetermined) linear system

$$\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\,\boldsymbol{q}_{N-L} = -(z_j^{N-L})_{j=1}^M\,. \tag{4.1}$$

Forming the corresponding monic polynomial q_{N-L} of degree $N-L$

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

then by (4.1) this polynomial has all nodes $z_j \in \mathbb{D}$ $(j = 1, \ldots, M)$ as roots. By (4.1) and by the factorization

$$\boldsymbol{H}_{L,N-L}(0) = \boldsymbol{V}_{L,M}(\boldsymbol{z})\,(\mathrm{diag}\,\boldsymbol{c})\,\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\,,$$

the vector $\boldsymbol{q}_{N-L}$ is also the minimum 2-norm solution of the *Yule–Walker system*

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

Now we introduce the *companion matrix* of the monic polynomial q_{N-L} resp. of the vector $\boldsymbol{q}_{N-L}$

$$\boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L}) := \big(\boldsymbol{e}_2\,|\,\boldsymbol{e}_3\,|\,\ldots\,|\,\boldsymbol{e}_{N-L}\,|\,-\boldsymbol{q}_{N-L}\big)\,, \tag{4.2}$$

where $\boldsymbol{e}_j \in \mathbb{C}^{N-L}$ are the canonical basis vectors.

Remark 4.1 The companion matrix (4.2) has the known property

$$\det\big(z\,\boldsymbol{I}_{N-L} - \boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L})\big) = q_{N-L}(z) \quad (z \in \mathbb{C})\,,$$

where $\boldsymbol{I}_{N-L}$ denotes the identity matrix. All singular values of (4.2) can be explicitly determined (see [12] or [10, p. 197]). □

By [17, Lemma 2.2], the companion matrix (4.2) has the property

$$\boldsymbol{H}_{L,N-L}(0)\,\boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L}) = \boldsymbol{H}_{L,N-L}(1)\,. \tag{4.3}$$

Now we show interesting relations between the $M \times M$ matrix (3.10) resp. (3.12) and the $(N-L) \times (N-L)$ companion matrix (4.2).

Lemma 4.2 *Between the matrices* (3.10)*,* (3.12)*, and* (4.2) *there consist the following relations*

$$\boldsymbol{F}_M^{\mathrm{SVD}} = \boldsymbol{W}_{N-L,M}(0)^*\,\boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L})\,\big(\boldsymbol{W}_{N-L,M}(0)^*\big)^{\dagger}\,, \tag{4.4}$$

$$\boldsymbol{F}_M = \boldsymbol{X}_{N-L,M}(0)^*\,\boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L})\,\boldsymbol{X}_{N-L,M}(0)\,, \tag{4.5}$$

where $\boldsymbol{W}_{N-L,M}(0)^*$ *is the third factor in the factorization* (3.9) *of the Hankel matrix* $\boldsymbol{H}_{L,N-L}(0)$ *and where* $\boldsymbol{X}_{N-L,M}(0)$ *is defined by* (3.13)*. Further the matrix* (3.10) *is similar to* (3.12) *by*

$$\boldsymbol{F}_M^{\mathrm{SVD}} = \big(\boldsymbol{W}_{N-L,M}(0)^*\,\boldsymbol{W}_{N-L,M}(0)\big)^{1/2}\,\boldsymbol{F}_M\,\big(\boldsymbol{W}_{N-L,M}(0)^*\,\boldsymbol{W}_{N-L,M}(0)\big)^{-1/2} \tag{4.6}$$

so that both matrices have the same eigenvalues.

Proof. 1) By (4.3) and (3.9) we obtain that

$$\boldsymbol{D}_M\, \boldsymbol{W}_{N-L,M}(0)^*\, \boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L}) = \boldsymbol{D}_M\, \boldsymbol{W}_{N-L,M}(1)^*\,,$$

since $\boldsymbol{U}_{L,M}^*\, \boldsymbol{U}_{L,M} = \boldsymbol{I}_M$. Multiplying the above equation with

$$\boldsymbol{D}_M^{-1} = \mathrm{diag}\,(\sigma_j^{-1})_{j=1}^M\,,$$

where $\sigma_1 \ge \ldots \ge \sigma_M > 0$ denote all positive singular values of $\boldsymbol{H}_{L,N-L+1}$, it follows that

$$\boldsymbol{W}_{N-L,M}(0)^*\, \boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L}) = \boldsymbol{W}_{N-L,M}(1)^*\,. \tag{4.7}$$

Thus we receive by (3.10) that

$$\begin{aligned} \boldsymbol{W}_{N-L,M}(0)^*\, \boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L})\, \big(\boldsymbol{W}_{N-L,M}(0)^*\big)^\dagger &= \boldsymbol{W}_{N-L,M}(1)^*\, \big(\boldsymbol{W}_{N-L,M}(0)^*\big)^\dagger \\ &= \boldsymbol{F}_M^{\mathrm{SVD}}\,. \end{aligned} \tag{4.8}$$

2) Formula (4.5) is an immediate consequence of (4.7) and (3.13), if we multiply (4.7) by $\boldsymbol{X}_{N-L,M}(0)$ from the right and by $\big(\boldsymbol{W}_{N-L,M}(0)^*\, \boldsymbol{W}_{N-L,M}(0)\big)^{-1/2}$ from the left.
3) Using the representations (4.4) – (4.5) as well as formula (3.13), we obtain (4.6), since

$$\big(\boldsymbol{W}_{N-L,M}(0)^*\big)^\dagger = \boldsymbol{W}_{N-L,M}(0)\, \big(\boldsymbol{W}_{N-L,M}(0)^*\, \boldsymbol{W}_{N-L,M}(0)\big)^{-1}$$

(cf. (4.18)). As known, similar matrices possess the same eigenvalues. This completes the proof. ■

The *signal space* $\mathcal{S}_{N-L} \subset \mathbb{C}^{N-L}$ is defined as the range of the matrix $\boldsymbol{V}_{N-L,M}(\bar{\boldsymbol{z}})$, where $\bar{\boldsymbol{z}} := (\bar{z}_j)_{j=1}^M$, i.e., the signal space $\mathcal{S}_{N-L}$ is spanned by the M linearly independent vectors $(\bar{z}_j^k)_{k=0}^{N-L-1}$ $(j = 1, \ldots, M)$. Thus the M-dimensional signal space is fully characterized by the distinct nodes $z_j \in \mathbb{D}$ $(j = 1, \ldots, M)$. By the properties of the Moore–Penrose pseudoinverse $\big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^\dagger$ it follows that

$$\boldsymbol{P}_{N-L} := \big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^\dagger\, \boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}} \tag{4.9}$$

is the *orthogonal projection onto the signal space* $\mathcal{S}_{N-L}$. Further, we remark that

$$\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\, \big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^\dagger = \boldsymbol{I}_M\,, \tag{4.10}$$

since $\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}$ has full row rank. Note that $\boldsymbol{q}_{N-L} \in \mathcal{S}_{N-L}$, since by (4.1) and (4.9)

$$\begin{aligned} \boldsymbol{P}_{N-L}\, \boldsymbol{q}_{N-L} &= -\big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^\dagger\, \boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\, \big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^\dagger\, \big(z_j^{N-L}\big)_{j=1}^M \\ &= -\big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^\dagger\, \big(z_j^{N-L}\big)_{j=1}^M = \boldsymbol{q}_{N-L}\,. \end{aligned}$$

Now we describe the properties of the *projected companion matrix*

$$\boldsymbol{P}_{N-L}\, \boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L}) \in \mathbb{C}^{(N-L)\times(N-L)}\,. \tag{4.11}$$

Note that in [3] the matrix $\boldsymbol{F}_M$ with the representation (4.5) is called “projected companion matrix”.

Theorem 4.3 *The projected companion matrix* (4.11) *can be represented in the following forms*

$$\begin{aligned} \boldsymbol{P}_{N-L}\,\boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L}) &= \big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^{\dagger}\,(\operatorname{diag}\boldsymbol{z})\,\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}} && (4.12)\\ &= \big(\boldsymbol{W}_{N-L,M}(0)^{*}\big)^{\dagger}\,\boldsymbol{W}_{N-L,M}(1)^{*} && (4.13)\\ &= \boldsymbol{X}_{N-L,M}(0)\,\boldsymbol{X}_{N-L,M}(1)^{*}\,. && (4.14) \end{aligned}$$

The signal space $\mathcal{S}_{N-L}$ *coincides with the range of* $\boldsymbol{H}_{L,N-L}(0)^*$ *and also with the range of* $\boldsymbol{W}_{N-L,M}(0)$. *The columns of* $\boldsymbol{X}_{N-L,M}(0)$ *form an orthonormal basis of the* M-*dimensional signal space* $\mathcal{S}_{N-L}$. *Moreover, the signal space* $\mathcal{S}_{N-L}$ *is an invariant subspace for* $\boldsymbol{X}_{N-L,M}(0)$. *Further,* $\boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L})^*$ *maps the signal space* $\mathcal{S}_{N-L}$ *into itself. The orthogonal projection* $\boldsymbol{P}_{N-L}$ *onto* $\mathcal{S}_{N-L}$ *can be represented as follows*

$$\boldsymbol{P}_{N-L} = \big(\boldsymbol{W}_{N-L,M}(0)^{*}\big)^{\dagger}\,\boldsymbol{W}_{N-L,M}(0)^{*} = \boldsymbol{X}_{N-L,M}(0)\,\boldsymbol{X}_{N-L,M}(0)^{*}\,. \qquad (4.15)$$

The nonvanishing eigenvalues of the projected companion matrix (4.11) *coincide with the eigenvalues of* (3.10) *resp.* (3.12).

Proof. 1) By (4.1) and (4.2) we obtain that

$$\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\,\boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L}) = (\operatorname{diag}\boldsymbol{z})\,\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\,. \qquad (4.16)$$

Note that (4.16) shows a close relationship between the Vandermonde matrix $\boldsymbol{V}_{N-L,M}(\boldsymbol{z})$ and the companion matrix $\boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L})$. From (4.16) it follows immediately that

$$\boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L})^{*}\,\boldsymbol{V}_{N-L,M}(\bar{\boldsymbol{z}}) = \boldsymbol{V}_{N-L,M}(\bar{\boldsymbol{z}})\,(\operatorname{diag}\bar{\boldsymbol{z}})\,,$$

i.e., $\boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L})^*$ maps the signal space $\mathcal{S}_{N-L}$ into itself. Multiplying (4.16) with $\big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^{\dagger}$, we receive the factorization (4.12) of the projected companion matrix (4.11).
2) Using the factorization

$$\boldsymbol{H}_{L,N-L}(0) = \boldsymbol{V}_{L,M}(\boldsymbol{z})\,(\operatorname{diag}\boldsymbol{c})\,\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\,,$$

we obtain that

$$\boldsymbol{H}_{L,N-L}(0)^{*} = \boldsymbol{V}_{N-L,M}(\bar{\boldsymbol{z}})\,(\operatorname{diag}\bar{\boldsymbol{c}})\,\boldsymbol{V}_{L,M}(\boldsymbol{z})^{*}\,.$$

Consequently, $\mathcal{S}_{N-L}$ coincides with the range of $\boldsymbol{H}_{L,N-L}(0)^*$. By (3.9) for $s = 0$ it follows that

$$\boldsymbol{H}_{L,N-L}(0)^{*} = \boldsymbol{W}_{N-L,M}(0)\,\boldsymbol{D}_{M}\,\boldsymbol{U}_{L,M}^{*}\,.$$

Hence $\mathcal{S}_{N-L}$ coincides with the range of $\boldsymbol{W}_{N-L,M}(0)$ too. Further

$$\big(\boldsymbol{W}_{N-L,M}(0)^{*}\big)^{\dagger}\,\boldsymbol{W}_{N-L,M}(0)^{*}$$

is the orthogonal projection onto the range of $\big(\boldsymbol{W}_{N-L,M}(0)^{*}\big)^{\dagger}$ which coincides with the range of $\boldsymbol{W}_{N-L,M}(0)$. Since the range of $\boldsymbol{W}_{N-L,M}(0)$ is equal to the signal space $\mathcal{S}_{N-L}$, we conclude that

$$\boldsymbol{P}_{N-L} = \big(\boldsymbol{W}_{N-L,M}(0)^{*}\big)^{\dagger}\,\boldsymbol{W}_{N-L,M}(0)^{*}\,. \qquad (4.17)$$

Multiplying (4.7) with $\big(\boldsymbol{W}_{N-L,M}(0)^*\big)^\dagger$ from the left, by (4.17) we receive the factorization (4.13) of the projected companion matrix (4.11). Formula (4.14) follows immediately from (4.13) and (3.13), since the Moore–Penrose pseudoinverse of the full rank matrix $\boldsymbol{W}_{N-L,M}(0)^*$ reads as follows

$$\big(\boldsymbol{W}_{N-L,M}(0)^*\big)^\dagger = \boldsymbol{W}_{N-L,M}(0)\,\big(\boldsymbol{W}_{N-L,M}(0)^*\,\boldsymbol{W}_{N-L,M}(0)\big)^{-1} \tag{4.18}$$

and since the inverse square root of the positive definite matrix $\boldsymbol{W}_{N-L,M}(0)^*\,\boldsymbol{W}_{N-L,M}(0)$ is well defined.

By the properties of the Moore–Penrose pseudoinverse $\big(\boldsymbol{W}_{N-L,M}(0)^*\big)^\dagger$, the matrix $\big(\boldsymbol{W}_{N-L,M}(0)^*\big)^\dagger\,\boldsymbol{W}_{N-L,M}(0)^*$ is the orthogonal projection onto the range of $\boldsymbol{W}_{N-L,M}(0)$ which coincides with the signal space $\mathcal{S}_{N-L}$. Hence we conclude that

$$\boldsymbol{P}_{N-L} = \big(\boldsymbol{W}_{N-L,M}(0)^*\big)^\dagger\,\boldsymbol{W}_{N-L,M}(0)^*\,.$$

Using (4.18) and (3.13), we obtain (4.15).

3) By the property (3.14), the M columns of $\boldsymbol{X}_{N-L,M}(0)$ are orthonormal and are contained in the M-dimensional signal space $\mathcal{S}_{N-L}$, because $\mathcal{S}_{N-L}$ coincides with the range of $\boldsymbol{W}_{N-L,M}(0)$. Hence the M columns of $\boldsymbol{X}_{N-L,M}(0)$ form an orthonormal basis of $\mathcal{S}_{N-L}$. Using (3.12) and (4.14), we obtain the relation

$$\boldsymbol{P}_{N-L}\,\boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L})\,\boldsymbol{X}_{N-L,M}(0) = \boldsymbol{X}_{N-L,M}(0)\,\boldsymbol{F}_M\,.$$

From this it follows that the signal space $\mathcal{S}_{N-L}$ is an invariant subspace for $\boldsymbol{X}_{N-L,M}(0)$.

4) By simple calculations, one can see that the nonvanishing eigenvalues of the projected companion matrix (4.11) coincide with the eigenvalues of (3.10) resp. (3.12). Let $(z, \boldsymbol{y})$ with $z \neq 0$ and $\boldsymbol{y} \in \mathbb{C}^{N-L}$ ($\boldsymbol{y} \neq \boldsymbol{0}$) be a right eigenpair of (4.11), i.e.

$$\boldsymbol{P}_{N-L}\,\boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L})\,\boldsymbol{y} = z\,\boldsymbol{y}\,.$$

Hence $z\,\boldsymbol{P}_{N-L}\boldsymbol{y} = z\boldsymbol{y}$ and thus $\boldsymbol{P}_{N-L}\boldsymbol{y} = \boldsymbol{y}$ by $z \neq 0$. For $\boldsymbol{x} := \boldsymbol{X}_{N-L,M}(0)^*\boldsymbol{y}$ we obtain $\boldsymbol{X}_{N-L,M}(0)\,\boldsymbol{x} = \boldsymbol{P}_{N-L}\,\boldsymbol{y} = \boldsymbol{y}$ by (4.15) so that $\boldsymbol{x} \neq \boldsymbol{0}$. Further by (3.14) and (4.5) it follows that

$$\begin{aligned}\boldsymbol{X}_{N-L,M}(0)^*\,\boldsymbol{P}_{N-L}\,\boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L})\,\boldsymbol{y} &= \boldsymbol{X}_{N-L,M}(0)^*\,\boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L})\,\boldsymbol{y}\\ &= \boldsymbol{X}_{N-L,M}(0)^*\,\boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L})\,\boldsymbol{X}_{N-L,M}(0)\,\boldsymbol{x}\\ &= \boldsymbol{F}_M\,\boldsymbol{x} = z\,\boldsymbol{x}\,,\end{aligned}$$

i.e., $(z,\,\boldsymbol{X}_{N-L,M}(0)^*\boldsymbol{y})$ is a right eigenpair of (3.12). Analogously, one can show that each eigenvalue of (3.12) is an eigenvalue of (4.11) too. By Lemma 4.2, the eigenvalues of (3.12) coincide with the eigenvalues of (3.10). This completes the proof. ■

Remark 4.4 The singular values of the matrix $\boldsymbol{F}_M$ can be characterized by [1, Theorem 4]. Assume that $2 \le M \le L \le \lceil \frac{N}{2} \rceil$. Let $\boldsymbol{p}_{N-L} := \boldsymbol{P}_{N-L}\, \boldsymbol{e}_1$ be the first column of $\boldsymbol{P}_{N-L}$. Then the singular values of the matrix (3.12) are $\omega_2 = \ldots = \omega_{M-1} = 1$ and

$$\begin{aligned}\omega_1 &= \frac{1}{\sqrt{2}} \Big(2 + \|\boldsymbol{q}_{N-L}\|_2^2 - \|\boldsymbol{p}_{N-L}\|_2^2 + \sqrt{(\|\boldsymbol{q}_{N-L}\|_2^2 + \|\boldsymbol{p}_{N-L}\|_2^2)^2 - 4\,|q_0|^2}\Big)^{1/2}, \\ \omega_M &= \frac{1}{\sqrt{2}} \Big(2 + \|\boldsymbol{q}_{N-L}\|_2^2 - \|\boldsymbol{p}_{N-L}\|_2^2 - \sqrt{(\|\boldsymbol{q}_{N-L}\|_2^2 + \|\boldsymbol{p}_{N-L}\|_2^2)^2 - 4\,|q_0|^2}\Big)^{1/2},\end{aligned}$$

where q_0 is the first component of $\boldsymbol{q}_{N-L}$. Further, the spectral resp. Frobenius norm of the matrix (3.12) is equal to

$$\|\boldsymbol{F}_M\|_2 = \omega_1\,, \quad \|\boldsymbol{F}_M\|_{\mathrm{F}} = \sqrt{M + \|\boldsymbol{q}_{N-L}\|_2^2 - \|\boldsymbol{p}_{N-L}\|_2^2}\,. \tag{4.19}$$

□

5 Error estimates of the nodes

From the matrix perturbation theory, the following results on the perturbation of eigenvalues are known. Let $\boldsymbol{A}_P \in \mathbb{C}^{P\times P}$ be a square nonnormal matrix with the eigenvalues $y_k \in \mathbb{C}$ $(k = 1, \ldots, P)$ and let $\tilde{\boldsymbol{A}}_P \in \mathbb{C}^{P\times P}$ be a perturbation of $\boldsymbol{A}_P$. If y_j be a simple eigenvalue of $\boldsymbol{A}_P$ with right resp. left eigenvectors $\boldsymbol{u}_j$ resp. $\boldsymbol{v}_j$, then there exists a unique eigenvalue $\tilde{y}_j$ of $\tilde{\boldsymbol{A}}_P$ (see e.g. [22, pp. 183 – 184]) such that

$$\tilde{y}_j = y_j + \frac{\boldsymbol{v}_j^*\,(\tilde{\boldsymbol{A}}_P - \boldsymbol{A}_P)\,\boldsymbol{u}_j}{\boldsymbol{v}_j^*\,\boldsymbol{u}_j} + \mathcal{O}(\|\tilde{\boldsymbol{A}}_P - \boldsymbol{A}_P\|_2^2)$$

Note that the right and left eigenvectors of a simple eigenvalue y_j cannot be orthogonal, i.e. $\boldsymbol{v}_j^*\,\boldsymbol{u}_j \neq 0$. If the left and right eigenvectors of an eigenvalue of $\boldsymbol{A}_P$ are nearly orthogonal, then $\boldsymbol{A}_P$ must be near one with multiple eigenvalue (see [23]). For sufficiently small spectral norm $\|\tilde{\boldsymbol{A}}_P - \boldsymbol{A}_P\|_2$, we obtain the first–order estimate

$$|\tilde{y}_j - y_j| \le \frac{|\boldsymbol{v}_j^*\,(\tilde{\boldsymbol{A}}_P - \boldsymbol{A}_P)\,\boldsymbol{u}_j|}{|\boldsymbol{v}_j^*\,\boldsymbol{u}_j|} \le \frac{\|\boldsymbol{v}_j\|_2\,\|\boldsymbol{u}_j\|_2}{|\boldsymbol{v}_j^*\,\boldsymbol{u}_j|}\,\|\tilde{\boldsymbol{A}}_P - \boldsymbol{A}_P\|_2\,. \tag{5.1}$$

Thus the quantity

$$\kappa_j(\boldsymbol{A}_P) := \frac{\|\boldsymbol{v}_j\|_2\,\|\boldsymbol{u}_j\|_2}{|\boldsymbol{v}_j^*\,\boldsymbol{u}_j|} \ge 1$$

measures the sensitivity of the eigenvalue y_j to perturbations on $\boldsymbol{A}_P$, see also [5]. Therefore this number is called the *condition number of the eigenvalue* y_j. By definition we have $\kappa_j(\boldsymbol{A}_P) = \kappa_j(\boldsymbol{A}_P^*)$. By [20], the condition number of a simple eigenvalue y_j of $\boldsymbol{A}_P$ can be estimated by

$$\kappa_j(\boldsymbol{A}_P) \le \Big(1 + \frac{\|\boldsymbol{A}_P\|_{\mathrm{F}}^2 - \sum_{k=1}^P |y_k|^2}{(P-1)\,d_j^2}\Big)^{(P-1)/2}. \tag{5.2}$$

where $d_j := \min\{|y_j - y_k|;\, k = 1, \ldots, P,\, k \neq j\}$ denotes the *separation distance for the eigenvalue* y_j. Note that

$$\Delta(\boldsymbol{A}_P) := \sqrt{\|\boldsymbol{A}_P\|_{\mathrm{F}}^2 - \sum_{k=1}^{P} |y_k|^2}$$

is the so-called *departure from normality* of $\boldsymbol{A}_P$ (see [20]). If $\Delta(\boldsymbol{A}_P) = 0$, then $\boldsymbol{A}_P$ is normal.

Now we apply these results for the perturbation of eigenvalues concerning the projected companion matrix (4.11) and the $M \times M$ matrix (3.12).

Theorem 5.1 *Let $M, L, N \in \mathbb{N}$ with $M \leq L \leq \lceil \frac{N}{2} \rceil$ be given. Then the projected companion matrix* (4.11) *has $z_j \in \mathbb{D}$ $(j = 1, \ldots, M)$ as simple eigenvalues. Further, the matrix* (4.11) *has* 0 *as an eigenvalue with algebraic multiplicity $N - L - M$. Moreover, $\big(z_j, \big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^{\dagger}\, \boldsymbol{e}_j\big)$ $(j = 1, \ldots, M)$ is a right eigenpair and $(z_j, \boldsymbol{V}_{N-L,M}(\bar{\boldsymbol{z}})\, \boldsymbol{e}_j)$ $(j = 1, \ldots, M)$ is a left eigenpair of* (4.11)*, where $\boldsymbol{e}_j \in \mathbb{C}^M$ is the jth canonical basis vector. The condition number of the eigenvalue z_j of the matrix* (4.11) *fulfills*

$$\kappa_j(\boldsymbol{P}_{N-L}\, \boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L})) = \|\big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^{\dagger}\, \boldsymbol{e}_j\|_2\, \|\boldsymbol{V}_{N-L,M}(\bar{\boldsymbol{z}})\, \boldsymbol{e}_j\|_2\,. \tag{5.3}$$

Corresponding to each eigenvalue $z_j \in \mathbb{D}$ $(j = 1, \ldots, M)$ of $\boldsymbol{P}_{N-L}\, \boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L})$, there exists a unique eigenvalue $\tilde{z}_j$ of $\tilde{\boldsymbol{P}}_{N-L}\, \boldsymbol{C}_{N-L}(\tilde{\boldsymbol{q}}_{N-L})$ so that

$$|z_j - \tilde{z}_j| \leq \kappa_j(\boldsymbol{P}_{N-L}\, \boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L}))\Big(\|\boldsymbol{P}_{N-L} - \tilde{\boldsymbol{P}}_{N-L}\|_2 + \|\boldsymbol{q}_{N-L} - \tilde{\boldsymbol{q}}_{N-L}\|_2\Big)\,. \tag{5.4}$$

Here $\tilde{\boldsymbol{q}}_{N-L} \in \mathbb{C}^{N-L}$ denotes the minimum 2-norm solution of the linear system

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

where $\tilde{\boldsymbol{K}}_{L,N-L}(0)$ is the low-rank approximation (3.11) *of $\tilde{\boldsymbol{H}}_{L,N-L}(0)$.*

Proof. 1) By the representation (4.12) of the projected companion matrix (4.11) and the property (4.10) it follows immediately that $\big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^{\dagger}\, \boldsymbol{e}_j$ resp. $\big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^{*}\, \boldsymbol{e}_j = \boldsymbol{V}_{N-L,M}(\bar{\boldsymbol{z}})\, \boldsymbol{e}_j$ is a right resp. left eigenvector of (4.11) with respect to the eigenvalue z_j. By (4.10), these eigenvectors possess the property

$$\Big(\big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^{*}\, \boldsymbol{e}_j\Big)^{*}\, \big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^{\dagger}\, \boldsymbol{e}_j = \boldsymbol{e}_j^{\mathrm{T}}\, \boldsymbol{e}_j = 1\,.$$

Then the condition number of the projected companion matrix (4.11) with respect to the eigenvalue $z_j \in \mathbb{D}$ is given by (5.3).
By assumption, it holds $N - L > M$. Since $\boldsymbol{V}_{N-L,M}(\boldsymbol{z})$ has full rank, we see that by (4.12) the null space of $\boldsymbol{P}_{N-L}\, \boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L})$ coincides with the null space of $\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}$. Hence the null space of $\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}$ has the dimension $N - L - M$. This means that 0

is an eigenvalue with algebraic multiplicity $N-L-M$ and that $z_j \in \mathbb{D}$ ($j = 1, \ldots, M$) are simple eigenvalues of (4.11).
2) Let $\tilde{z}_j$ ($j = 1, \ldots, M$) denote an eigenvalue of (4.11) that is the closest to $z_j \in \mathbb{D}$. Our goal is to estimate the error $|\tilde{z}_j - z_j|$ ($j = 1, \ldots, M$). Let $\boldsymbol{P}_{N-L}$ resp. $\tilde{\boldsymbol{P}}_{N-L}$ denote the orthogonal projector onto the corresponding signal space $\mathcal{S}_{N-L}$ resp. $\tilde{\mathcal{S}}_{N-L}$. Setting

$$\boldsymbol{A}_{N-L} := \boldsymbol{P}_{N-L}\, \boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L}) - \tilde{\boldsymbol{P}}_{N-L}\, \boldsymbol{C}_{N-L}(\tilde{\boldsymbol{q}}_{N-L})\,,$$

by (5.1) the following first–order estimate holds

$$|\tilde{z}_j - z_j| \le \kappa_j(\boldsymbol{P}_{N-L}\, \boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L}))\, \|\boldsymbol{A}_{N-L}\|_2$$

for $j = 1, \ldots, M$. Using the special structure of the companion matrix (4.2), one can see that

$$\begin{aligned} \boldsymbol{A}_{N-L}\, \boldsymbol{A}_{N-L}^* &= (\boldsymbol{P}_{N-L} - \tilde{\boldsymbol{P}}_{N-L})\, (\boldsymbol{P}_{N-L} - \tilde{\boldsymbol{P}}_{N-L})^* + (\boldsymbol{q}_{N-L} - \tilde{\boldsymbol{q}}_{N-L})\, (\boldsymbol{q}_{N-L} - \tilde{\boldsymbol{q}}_{N-L})^* \\ &\quad - (\boldsymbol{p}_{N-L} - \tilde{\boldsymbol{p}}_{N-L})\, (\boldsymbol{p}_{N-L} - \tilde{\boldsymbol{p}}_{N-L})^*\,, \end{aligned}$$

where $\boldsymbol{p}_{N-L}$ and $\tilde{\boldsymbol{p}}_{N-L}$ are the first columns of $\boldsymbol{P}_{N-L}$ resp. $\tilde{\boldsymbol{P}}_{N-L}$. Then for each unit vector $\boldsymbol{x} \in \mathbb{C}^{N-L}$ we receive that

$$\begin{aligned} \boldsymbol{x}^*\, \boldsymbol{A}_{N-L}\, \boldsymbol{A}_{N-L}^*\, \boldsymbol{x} &= \|\boldsymbol{A}_{N-L}^* \boldsymbol{x}\|_2^2 \\ &= \|(\boldsymbol{P}_{N-L} - \tilde{\boldsymbol{P}}_{N-L})^* \boldsymbol{x}\|_2^2 + |\boldsymbol{x}^*\, (\boldsymbol{q}_{N-L} - \tilde{\boldsymbol{q}}_{N-L})|^2 \\ &\quad - |\boldsymbol{x}^*\, (\boldsymbol{p}_{N-L} - \tilde{\boldsymbol{p}}_{N-L})|^2 \\ &\le \|\boldsymbol{P}_{N-L} - \tilde{\boldsymbol{P}}_{N-L}\|_2^2 + \|\boldsymbol{q}_{N-L} - \tilde{\boldsymbol{q}}_{N-L}\|_2^2\,. \end{aligned}$$

Thus it follows that for *all* unit vectors $\boldsymbol{x} \in \mathbb{C}^{N-L}$

$$\|\boldsymbol{A}_{N-L}^* \boldsymbol{x}\|_2 \le \|\boldsymbol{P}_{N-L} - \tilde{\boldsymbol{P}}_{N-L}\|_2 + \|\boldsymbol{q}_{N-L} - \tilde{\boldsymbol{q}}_{N-L}\|_2$$

and hence

$$\|\boldsymbol{A}_{N-L}\|_2 \le \|\boldsymbol{P}_{N-L} - \tilde{\boldsymbol{P}}_{N-L}\|_2 + \|\boldsymbol{q}_{N-L} - \tilde{\boldsymbol{q}}_{N-L}\|_2\,.$$

Thus we obtain the above estimate of $|\tilde{z}_j - z_j|$. This completes the proof. ■

In the next theorem, we show that $\kappa_j(\boldsymbol{P}_{N-L}\, \boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L})) = \kappa_j(\boldsymbol{F}_M)$. The matrix $\boldsymbol{F}_M$ is computed from exactly sampled data $h(k)$ ($k = 0, \ldots, N-1$). Analogously, the matrix $\tilde{\boldsymbol{F}}_M$ is obtained from noisy sampled data $\tilde{h}_k$ ($k = 0, \ldots, N-1$). Thus $\tilde{\boldsymbol{F}}_M$ has a similar form as $\boldsymbol{F}_M$, namely

$$\tilde{\boldsymbol{F}}_M = \tilde{\boldsymbol{X}}_{N-L,M}(1)^*\, \tilde{\boldsymbol{X}}_{N-L,M}(0)\,.$$

Theorem 5.2 *The matrix* (3.12) *has only* $z_j \in \mathbb{D}$ $(j = 1, \ldots, M)$ *as simple eigenvalues. Further,* $(z_j, \boldsymbol{X}_{N-L,M}(0)^* \big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^\dagger \boldsymbol{e}_j)$ $(j = 1, \ldots, M)$ *is a right eigenpair and* $(z_j, \boldsymbol{X}_{N-L,M}(0)^\dagger \boldsymbol{V}_{N-L,M}(\bar{\boldsymbol{z}})\, \boldsymbol{e}_j)$ $(j = 1, \ldots, M)$ *is a left eigenpair of* (3.12)*. The condition number of the eigenvalue* z_j *of the matrix* (3.12) *fulfils*

$$\begin{aligned}\kappa_j(\boldsymbol{F}_M) &= \|\boldsymbol{X}_{N-L,M}(0)^* \big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^\dagger \boldsymbol{e}_j\|_2 \, \|\boldsymbol{X}_{N-L,M}(0)^\dagger \boldsymbol{V}_{N-L,M}(\bar{\boldsymbol{z}})\, \boldsymbol{e}_j\|_2 \\ &= \kappa_j(\boldsymbol{P}_{N-L}\, \boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L})) \qquad (5.5)\\ &\le \Big(1 + \frac{M + \|\boldsymbol{q}_{N-L}\|_2^2 - \|\boldsymbol{p}_{N-L}\|_2^2 - \mu}{(M-1)\,\delta_j^2}\Big)^{(M-1)/2} \qquad (5.6)\end{aligned}$$

with the first column $\boldsymbol{p}_{N-L}$ *of* $\boldsymbol{P}_{N-L}$*, with* μ *defined by* (2.3)*, and with* $\delta_j := \min\{|z_j - z_k|;\, k = 1, \ldots, M,\, k \neq j\}$*.*

Proof. Using (4.5) and (4.12), we obtain the matrix factorizations

$$\begin{aligned}\boldsymbol{F}_M &= \boldsymbol{X}_{N-L,M}(0)^*\, \boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L})\, \boldsymbol{X}_{N-L,M}(0)\\ &= \boldsymbol{X}_{N-L,M}(0)^*\, \boldsymbol{P}_{N-L}\, \boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L})\, \boldsymbol{X}_{N-L,M}(0)\\ &= \boldsymbol{X}_{N-L,M}(0)^* \big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^\dagger (\operatorname{diag} \boldsymbol{z})\, \boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\, \boldsymbol{X}_{N-L,M}(0)\,.\end{aligned}$$

Consequently, $\boldsymbol{X}_{N-L,M}(0)^* \big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^\dagger \boldsymbol{e}_j$ resp. $\boldsymbol{X}_{N-L,M}(0)^\dagger \boldsymbol{V}_{N-L,M}(\bar{\boldsymbol{z}})\, \boldsymbol{e}_j$ is a right resp. left eigenvector of (3.12) with respect to z_j. Since these eigenvectors possess the property

$$\begin{aligned}&\boldsymbol{e}_j^{\mathrm{T}}\, \boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}} \big(\boldsymbol{X}_{N-L,M}(0)\, \boldsymbol{X}_{N-L,M}(0)^* \big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^\dagger \boldsymbol{e}_j\\ &= \boldsymbol{e}_j^{\mathrm{T}}\, \boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\, \boldsymbol{P}_{N-L} \big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^\dagger \boldsymbol{e}_j = \boldsymbol{e}_j^{\mathrm{T}}\, \boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}} \big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^\dagger \boldsymbol{e}_j\\ &= \boldsymbol{e}_j^{\mathrm{T}}\, \boldsymbol{e}_j = 1\,,\end{aligned}$$

the condition number of the simple eigenvalue z_j of the matrix (3.12) is given by

$$\kappa_j(\boldsymbol{F}_M) = \|\boldsymbol{X}_{N-L,M}(0)^* \big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^\dagger \boldsymbol{e}_j\|_2 \, \|\boldsymbol{X}_{N-L,M}(0)^* \boldsymbol{V}_{N-L,M}(\bar{\boldsymbol{z}})\, \boldsymbol{e}_j\|_2$$

From (5.2) it follows the estimate (5.6). Since both $\big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^\dagger \boldsymbol{e}_j$ and $\boldsymbol{V}_{N-L,M}(\bar{\boldsymbol{z}})\, \boldsymbol{e}_j$ belong to the signal space $\mathcal{S}_{N-L}$ and since the columns of $\boldsymbol{X}_{N-L,M}(0)$ form an orthonormal basis of $\mathcal{S}_{N-L}$, it is clear that

$$\begin{aligned}\|\boldsymbol{X}_{N-L,M}(0)^* \big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^\dagger \boldsymbol{e}_j\|_2 &= \|\big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^\dagger \boldsymbol{e}_j\|_2\,,\\ \|\boldsymbol{X}_{N-L,M}(0)^* \boldsymbol{V}_{N-L,M}(\bar{\boldsymbol{z}})\, \boldsymbol{e}_j\|_2 &= \|\boldsymbol{V}_{N-L,M}(\bar{\boldsymbol{z}})\, \boldsymbol{e}_j\|_2\,.\end{aligned}$$

Thus we receive (5.5). From (5.2) and (4.19) it follows immediately the nice upper bound (5.6) of $\kappa_j(\boldsymbol{F}_M)$. Using (5.1), we obtain the above estimate of $|z_j - \tilde{z}_j|$. Similarly to (5.6), a corresponding estimate was also presented in [3, Proposition 3]. This completes the proof. ■

Remark 5.3 The matrix (3.12) is not Hermitian in general. By balancing one can often improve the accuracy of the computed eigenvalues of (3.12). *Balancing* is a convenient diagonal scaling of (3.12), i.e., a diagonal matrix $\boldsymbol{\Delta}_M$ is computed in $\mathcal{O}(M^2)$ operations, so that the jth column and the jth row of $\boldsymbol{\Delta}_M^{-1}\,\boldsymbol{F}_M\,\boldsymbol{\Delta}_M$ for each $j = 1,\ldots,M$ have almost the same 1-norm. Since the diagonal entries of $\boldsymbol{\Delta}_M$ are chosen as powers of 2, the balanced matrix $\boldsymbol{\Delta}_M^{-1}\,\boldsymbol{F}_M\,\boldsymbol{\Delta}_M$ can be calculated without roundoff (see [14]). □

By construction, the columns of the matrices $\boldsymbol{X}_{N-L,M}(0)$ and $\tilde{\boldsymbol{X}}_{N-L,M}(0)$ form orthonormal bases for the M-dimensional signal spaces $\mathcal{S}_{N-L}$ and $\tilde{\mathcal{S}}_{N-L}$, respectively. Assume that $\boldsymbol{U}_M\,\boldsymbol{D}_M\,\boldsymbol{V}_M^*$ is the singular value decomposition of $\boldsymbol{X}_{N-L,M}(0)^*\,\tilde{\boldsymbol{X}}_{N-L,M}(0)$, where $\boldsymbol{U}_M$ and $\boldsymbol{V}_M$ are unitary matrices and $\boldsymbol{D}_M$ is a diagonal matrix with the diagonal entries d_j $(j = 1,\ldots,M)$ arranged in nonincreasing order $1 \ge d_1 \ge \ldots \ge d_M \ge 0$. Then

$$\theta_j := \arccos d_{M-j+1} \quad (j = 1,\ldots,M)$$

are the *canonical angles* between $\mathcal{S}_{N-L}$ and $\tilde{\mathcal{S}}_{N-L}$ (see [22, p. 43 and p. 45]). We remark that

$$\frac{\pi}{2} \ge \theta_1 \ge \ \ldots \ \ge \theta_M \ge 0$$

such that θ_1 is the largest canonical angle between $\mathcal{S}_{N-L}$ and $\tilde{\mathcal{S}}_{N-L}$.
Note that $\|\boldsymbol{P}_{N-L} - \tilde{\boldsymbol{P}}_{N-L}\|_2$ is the *distance* between the M-dimensional signal spaces $\mathcal{S}_{N-L}$ and $\tilde{\mathcal{S}}_{N-L}$ (cf. [7, p. 76]). Since $\boldsymbol{P}_{N-L}\,(\boldsymbol{P}_{N-L} - \tilde{\boldsymbol{P}}_{N-L}) = \boldsymbol{P}_{N-L}\,(\boldsymbol{I}_{N-L} - \tilde{\boldsymbol{P}}_{N-L})$, we see immediately that

$$\|\boldsymbol{P}_{N-L} - \tilde{\boldsymbol{P}}_{N-L}\|_2 \le \|\boldsymbol{P}_{N-L}\|_2\,\|\boldsymbol{I}_{N-L} - \tilde{\boldsymbol{P}}_{N-L}\|_2 \le 1\,.$$

As known (see [22, pp. 43 – 44]), the largest singular value of $\boldsymbol{P}_{N-L} - \tilde{\boldsymbol{P}}_{N-L}$ is equal to $\sin\theta_1$. Hence the distance between $\mathcal{S}_{N-L}$ and $\tilde{\mathcal{S}}_{N-L}$ amounts to

$$\|\boldsymbol{P}_{N-L} - \tilde{\boldsymbol{P}}_{N-L}\|_2 = \sin\theta_1\,.$$

Now we estimate $\|\boldsymbol{P}_{N-L} - \tilde{\boldsymbol{P}}_{N-L}\|_2 = \sin\theta_1$, where $\boldsymbol{P}_{N-L}$ is the orthogonal projection onto the signal space $\mathcal{S}_{N-L}$ which coincides with the range of $\boldsymbol{H}_{L,N-L}(0)^*$ by Theorem 4.3. After the construction (see step 1 of Algorithm 3.4), $\tilde{\boldsymbol{P}}_{N-L}$ is the orthogonal projection onto the signal space $\tilde{\mathcal{S}}_{N-L}$ which is the range of $\tilde{\boldsymbol{K}}_{L,N-L}(0)^*$, where $\tilde{\boldsymbol{K}}_{L,N-L}(0)$ defined by (3.11) is the rank-M approximation of the given noisy matrix $\tilde{\boldsymbol{H}}_{L,N-L}(0)$. Thus the *error matrix of low-rank approximation* can be estimated by

$$\|\tilde{\boldsymbol{H}}_{L,N-L}(0) - \tilde{\boldsymbol{K}}_{L,N-L}(0)\|_2 \le \tilde{\sigma}_{M+1} < \varepsilon\,\tilde{\sigma}_1\,, \tag{5.7}$$

where $\tilde{\sigma}_1$ is the largest singular value of (3.1) and $\varepsilon > 0$ is a convenient chosen tolerance. Let

$$\boldsymbol{E}_{L,N-L} = \tilde{\boldsymbol{H}}_{L,N-L}(0) - \boldsymbol{H}_{L,N-L}(0) = (e_{\ell+m})_{\ell,m=0}^{L-1,N-L-1}$$

be the error matrix of given data. Using the maximum column resp. row sum norm of $\boldsymbol{E}_{L,N-L}$, we obtain by $|e_k| \le \varepsilon_1$ $(k = 0,\ldots,N-1)$ that

$$\|\boldsymbol{E}_{L,N-L}\|_2 \le \sqrt{\|\boldsymbol{E}_{L,N-L}\|_1\,\|\boldsymbol{E}_{L,N-L}\|_\infty} \le \sqrt{L\,(N-L)}\,\varepsilon_1 \le \frac{N}{2}\,\varepsilon_1\,.$$

Theorem 5.4 *Let $N \in \mathbb{N}$ $(N \gg 1)$ be given. Assume that the order M of the exponential sum* (1.1) *fulfills $2\,M \ll N$ and that the coefficients c_j of* (1.1) *satisfy the condition $|c_j| \ge \rho > 0$ $(j = 1, \ldots, M)$. Let σ_M be the lowest positive singular value of $\boldsymbol{H}_{L,N-L+1}$ with $L \approx \lceil \frac{N}{2} \rceil$ $(M \le L \le \lceil \frac{N}{2} \rceil)$.*
If $2\,\|\boldsymbol{E}_{L,N-L}\|_2 \ll \sigma_M$, then the spectral norm $\|\boldsymbol{P}_{N-L} - \tilde{\boldsymbol{P}}_{N-L}\|_2 = \sin\theta_1$ can be estimated by

$$\|\boldsymbol{P}_{N-L} - \tilde{\boldsymbol{P}}_{N-L}\|_2 \le \frac{2}{\sigma_M}\,\|\boldsymbol{E}_{L,N-L}\|_2\,. \tag{5.8}$$

Further, it holds

$$\|\boldsymbol{H}_{L,N-L}(0)^\dagger\|_2 = \frac{1}{\sigma_M} \le \frac{1}{\rho}\,\|\boldsymbol{V}_{L,M}(\boldsymbol{z})^\dagger\|_2^2\,,$$

where

$$\|\boldsymbol{V}_{L,M}(\boldsymbol{z})^\dagger\|_2^2 \le \begin{cases} \frac{M\,(1-\beta^2)}{1-\beta^{2L}}\left(1 + \frac{M+\|\boldsymbol{q}_L\|_2^2-\|\boldsymbol{p}_L\|_2^2-\mu}{(M-1)\,\delta^2}\right)^{M-1} & \text{if } \beta < 1\,, \\ \frac{M}{L}\left(1 + \frac{\|\boldsymbol{q}_L\|_2^2-\|\boldsymbol{p}_L\|_2^2}{(M-1)\,\delta^2}\right)^{M-1} & \text{if } \beta = 1 \end{cases}$$

with the first column $\boldsymbol{p}_L$ of $\boldsymbol{P}_L$. Note that β, μ, and δ are defined by (2.2), (2.3) *resp.* (2.4).

Proof. 1) For the orthogonal projections $\boldsymbol{P}_{N-L}$ and $\tilde{\boldsymbol{P}}_{N-L}$ we obtain that

$$\begin{aligned}\|\boldsymbol{P}_{N-L} - \tilde{\boldsymbol{P}}_{N-L}\|_2 &= \|\boldsymbol{P}_{N-L}\,(\boldsymbol{P}_{N-L} - \tilde{\boldsymbol{P}}_{N-L})\|_2 = \|\boldsymbol{P}_{N-L} - \boldsymbol{P}_{N-L}\,\tilde{\boldsymbol{P}}_{N-L}\|_2 \\ &= \|\boldsymbol{P}_{N-L}\,(\boldsymbol{I}_{N-L} - \tilde{\boldsymbol{P}}_{N-L})\|_2\,.\end{aligned}$$

Since $\boldsymbol{P}_{N-L}$ is the orthogonal projection onto the range of $\boldsymbol{H}_{L,N-L}(0)^*$, this projection has the form $\boldsymbol{P}_{N-L} = \boldsymbol{H}_{L,N-L}(0)^\dagger\,\boldsymbol{H}_{L,N-L}(0)$. Analogously, the orthogonal projection $\tilde{\boldsymbol{P}}_{N-L}$ onto the range of $\tilde{\boldsymbol{K}}_{L,N-L}(0)^*$ is given by $\tilde{\boldsymbol{P}}_{N-L} = \tilde{\boldsymbol{K}}_{L,N-L}(0)^\dagger\,\tilde{\boldsymbol{K}}_{L,N-L}(0)$. Then it follows that

$$\boldsymbol{P}_{N-L}\,(\boldsymbol{I}_{N-L} - \tilde{\boldsymbol{P}}_{N-L}) = \boldsymbol{H}_{L,N-L}(0)^\dagger\,\boldsymbol{H}_{L,N-L}(0)\,(\boldsymbol{I}_{N-L} - \tilde{\boldsymbol{P}}_{N-L})\,,$$

where

$$\boldsymbol{H}_{L,N-L}(0) = \tilde{\boldsymbol{K}}_{L,N-L}(0) + (\tilde{\boldsymbol{H}}_{L,N-L}(0) - \tilde{\boldsymbol{K}}_{L,N-L}(0)) - \boldsymbol{E}_{L,N-L}\,.$$

Since $\tilde{\boldsymbol{K}}_{L,N-L}(0) = \tilde{\boldsymbol{K}}_{L,N-L}(0)\,\tilde{\boldsymbol{P}}_{N-L}$ and since $\boldsymbol{I}_{N-L} - \tilde{\boldsymbol{P}}_{N-L}$ is an orthogonal projection too, we obtain by $\|\boldsymbol{H}_{L,N-L}(0)^\dagger\|_2 = \frac{1}{\sigma_M}$ that

$$\|\boldsymbol{P}_{N-L} - \tilde{\boldsymbol{P}}_{N-L}\|_2 \le \frac{1}{\sigma_M}\left(\|\tilde{\boldsymbol{H}}_{L,N-L}(0) - \tilde{\boldsymbol{K}}_{L,N-L}(0)\|_2 + \|\boldsymbol{E}_{L,N-L}\|_2\right).$$

From (5.7) and $\tilde{\sigma}_{M+1} \le \|\boldsymbol{E}_{L,N-L}\|_2$ by Weyl's Theorem (see [21, p. 70]), it follows the inequality (5.8).

2) The exact Hankel matrix $\boldsymbol{H}_{L,N-L}(0)$ has the rank M and can be factorized into the following product of full rank matrices

$$\boldsymbol{H}_{L,N-L}(0) = \boldsymbol{V}_{L,M}(\boldsymbol{z})\,(\mathrm{diag}\,\boldsymbol{c})\,\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\,.$$

Thus the Moore–Penrose pseudoinverse of $\boldsymbol{H}_{L,N-L}(0)$ has the form

$$\boldsymbol{H}_{L,N-L}(0)^{\dagger} = \big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^{\dagger}\,(\mathrm{diag}\,\boldsymbol{c})^{-1}\,\boldsymbol{V}_{L,M}(\boldsymbol{z})^{\dagger}\,.$$

Hence its norm can be estimated as follows

$$\|\boldsymbol{H}_{L,N-L}(0)^{\dagger}\|_2 \le \frac{1}{\rho}\,\|\big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^{\dagger}\|_2\,\|\boldsymbol{V}_{L,M}(\boldsymbol{z})^{\dagger}\|_2 \le \frac{1}{\rho}\,\|\big(\boldsymbol{V}_{L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^{\dagger}\|_2^2\,,$$

since for $M \le L \le N-L$ it holds by [1, Theorem 1]

$$\|\big(\boldsymbol{V}_{N-L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^{\dagger}\|_2 \le \|\big(\boldsymbol{V}_{L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^{\dagger}\|_2 = \|\boldsymbol{V}_{L,M}(\boldsymbol{z})^{\dagger}\|_2\,.$$

3) Finally, we estimate $\|\big(\boldsymbol{V}_{L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^{\dagger}\|_2^2$ for $L \ge M$. We start with

$$\begin{aligned}\|\big(\boldsymbol{V}_{L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^{\dagger}\|_2^2 &\le \|\big(\boldsymbol{V}_{L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^{\dagger}\|_{\mathrm{F}}^2 = \sum_{j=1}^{M}\|\big(\boldsymbol{V}_{L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^{\dagger}\,\boldsymbol{e}_j\|_2^2\\ &= \sum_{j=1}^{M}\frac{\|\big(\boldsymbol{V}_{L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^{\dagger}\,\boldsymbol{e}_j\|_2^2\,\|\boldsymbol{V}_{L,M}(\bar{\boldsymbol{z}})\,\boldsymbol{e}_j\|_2^2}{\|\boldsymbol{V}_{L,M}(\bar{\boldsymbol{z}})\,\boldsymbol{e}_j\|_2^2}\end{aligned}$$

As shown in the proof of Theorem 5.2, we know that

$$\|\big(\boldsymbol{V}_{L,M}(\boldsymbol{z})^{\mathrm{T}}\big)^{\dagger}\,\boldsymbol{e}_j\|_2^2\,\|\boldsymbol{V}_{L,M}(\bar{\boldsymbol{z}})\,\boldsymbol{e}_j\|_2^2 \le \Big(1 + \frac{M + \|\boldsymbol{q}_L\|_2^2 - \|\boldsymbol{p}_L\|_2^2 - \mu}{(M-1)\,\delta^2}\Big)^{M-1}$$

with $\delta = \min\{|z_j - z_k|;\, j,k = 1,\ldots,M,\, j \ne k\}$. Finally we use the estimate

$$\|\boldsymbol{V}_{L,M}(\bar{\boldsymbol{z}})\,\boldsymbol{e}_j\|_2^2 = \sum_{k=0}^{L-1}|z_j|^{2k} \ge \sum_{k=0}^{L-1}\beta^{2k}\,.$$

This completes the proof. ■

We summarize: If the error bound ε_1 in Algorithm 3.4 are very small for sufficiently large integer N ($N \gg 2\,M$) so that $2\,\|\boldsymbol{E}_{L,N-L}\|_2 \ll \sigma_M$ for a window length $L \approx \lceil \frac{N}{2} \rceil$ ($M \le L \le \lceil \frac{N}{2} \rceil$), if all nodes z_j ($j = 1,\ldots,M$) are lying near to the unit circle with $\beta < 1$ (see (2.2)) but not extremely close to each other, and if all coefficients c_j ($j = 1,\ldots,M$) fulfill $|c_j| \ge \rho > 0$, then for each node z_j ($j = 1,\ldots,M$) there exists a unique node $\tilde{z}_j$ such that

$$\begin{aligned}|z_j - \tilde{z}_j| &\le \Big(1 + \frac{M + \|\boldsymbol{q}_{N-L}\|_2^2 - \|\boldsymbol{p}_{N-L}\|_2^2 - \mu}{(M-1)\,\delta_j^2}\Big)^{(M-1)/2}\\ &\quad \times \Big(\sin\theta_1 + \|\boldsymbol{q}_{N-L} - \tilde{\boldsymbol{q}}_{N-L}\|_2\Big)\end{aligned}$$

where μ and δ_j are defined by (2.3) and in Theorem 5.2, and where

$$\sin\theta_1 \le \frac{2}{\sigma_M}\,\|\boldsymbol{E}_{L,N-L}\|_2\,.$$

If the nodes $z_j \in \mathbb{D}$ $(j = 1, \ldots, M)$ are computed with low errors, then the nonvanishing coefficients $c_j \in \mathbb{C}$ $(j = 1, \ldots, M)$ can be determined as solution $\boldsymbol{c} = (c_j)_{j=1}^M$ of the least squares problem

$$\|\boldsymbol{V}_{N,M}(\boldsymbol{z})\,\boldsymbol{c} - \boldsymbol{h}\|_2 = \min$$

with the vector $\boldsymbol{h} = \big(h(k)\big)_{k=0}^{N-1}$ of exact data and $N > 2\,M$ (see Algorithm 3.4). Note that $\boldsymbol{V}_{N,M}(\boldsymbol{z})$ has full rank. Let $\tilde{\boldsymbol{h}} = (\tilde{h}_k)_{k=0}^{N-1}$ be the vector of noisy data and let $\tilde{\boldsymbol{z}} = (\tilde{z}_j)_{j=1}^M$ be the vector of computed nodes $\tilde{z}_j \in \mathbb{D}$ with $\tilde{z}_j \approx z_j$. Let $\tilde{\boldsymbol{c}} = (\tilde{c}_j)_{j=1}^M$ be the solution of the least squares problem

$$\|\boldsymbol{V}_{N,M}(\tilde{\boldsymbol{z}})\,\tilde{\boldsymbol{c}} - \tilde{\boldsymbol{h}}\|_2 = \min\,.$$

For large $N \gg 2\,M$, the Vandermonde matrix $\boldsymbol{V}_{N,M}(\tilde{\boldsymbol{z}})$ has full rank and is well conditioned with respect to the spectral norm. Assume that $\varepsilon_2 > 0$ fulfils the inequalities

$$\begin{aligned} \|\boldsymbol{V}_{N,M}(\boldsymbol{z}) - \boldsymbol{V}_{N,M}(\tilde{\boldsymbol{z}})\|_2 &\le \varepsilon_2\,\|\boldsymbol{V}_{N,M}(\boldsymbol{z})\|_2\,, \quad \|\boldsymbol{h} - \tilde{\boldsymbol{h}}\|_2 \le \varepsilon_2\,\|\boldsymbol{h}\|_2\,, \\ \varepsilon_2\,\mathrm{cond}_2\,\boldsymbol{V}_{N,M}(\boldsymbol{z}) &< 1\,. \end{aligned}$$

By the perturbation theory of the least squares problem one obtains the normwise estimate (see [9, p. 382 and pp. 400 – 402])

$$\begin{aligned} \frac{\|\boldsymbol{c} - \tilde{\boldsymbol{c}}\|_2}{\|\boldsymbol{c}\|_2} &\le \frac{\varepsilon_2\,\mathrm{cond}_2\,\boldsymbol{V}_{N,M}(\boldsymbol{z})}{1 - \varepsilon_2\,\mathrm{cond}_2\,\boldsymbol{V}_{N,M}(\boldsymbol{z})} \\ &\quad \times \Big[2 + (\mathrm{cond}_2\,\boldsymbol{V}_{N,M}(\boldsymbol{z}) + 1)\,\frac{\|\boldsymbol{V}_{N,M}(\boldsymbol{z})\,\boldsymbol{c} - \boldsymbol{h}\|_2}{\|\boldsymbol{V}_{N,M}(\boldsymbol{z})\|_2\,\|\boldsymbol{c}\|_2}\Big]\,. \end{aligned}$$

Consequently, the sensitivity of the least squares problem can be measured by the spectral norm condition number $\mathrm{cond}_2\,\boldsymbol{V}_{N,M}(\boldsymbol{z})$ when $\|\boldsymbol{V}_{N,M}(\boldsymbol{z})\,\boldsymbol{c} - \boldsymbol{h}\|_2$ is small and by the square $\big(\mathrm{cond}_2\,\boldsymbol{V}_{N,M}(\boldsymbol{z})\big)^2$ otherwise.

6 Numerical examples

Finally we illustrate the results by some numerical experiments. All computations are performed in MATLAB with IEEE double–precision arithmetic.

First we summarize the corresponding assumptions of our study:
(A1) The number N of noisy sampled data $\tilde{h}_k = h(k) + e_k$ $(k = 0, \ldots, N-1)$ satisfies the condition $N \gg 2\,M$. In other words, we use oversampling of the exponential sum (1.1). The order M of the exponential sum (1.1) is only of moderate size.
(A2) The coefficients $c_j \in \mathbb{C}$ $(j = 1, \ldots, M)$ of the exponential sum (1.1) fulfill the condition $|c_j| \ge \rho > 0$, where ρ is not too small.

(A3) The distinct nodes $z_j = \mathrm{e}^{f_j} \in \mathbb{D}$ $(j = 1, \ldots, M)$ are lying in the near of the unit circle.
(A4) The error terms e_k $(k = 0, \ldots, N-1)$ are relatively small so that $|e_k| \le \varepsilon_1$ with $0 < \varepsilon_1 \ll \rho$ and $2\,\|\boldsymbol{E}_{L,N-L}\|_2 \ll \sigma_M$, where σ_M is the lowest positive singular value of the L-trajectory matrix $\boldsymbol{H}_{L,N-L+1}$ with the window length $L \approx \lceil \frac{N}{2} \rceil$, where $M \le L \le \lceil \frac{N}{2} \rceil$.

We start with the estimate from Theorem 5.1 and show that the estimates (5.4) are sharp for some parameters, but also useless in other cases. To this end we compute $\kappa_j(\boldsymbol{P}_{N-L}\boldsymbol{C}_{N-L}(\boldsymbol{q}_{N-L}))$ and the estimate by the RHS of (5.6) as well as the values $\|\boldsymbol{P}_{N-L} - \tilde{\boldsymbol{P}}_{N-L}\|_2$ and $\|\boldsymbol{q}_{N-L} - \tilde{\boldsymbol{q}}_{N-L}\|_2$.

Example 6.1 We choose M equispaced nodes $z_j = \exp(2\pi \mathrm{i} j/M)$ $(j = 1, \ldots, M)$ on the unit circle and set the coefficients $c_j = 1$ $(j = 1, \ldots, M)$. We form the exponential sum (1.1) so that

$$h(k) = \sum_{j=1}^{M} z_j^k \quad (k = 0, \ldots, N-1)\,. \tag{6.1}$$

We use noisy sampled data $\tilde{h}_k := h(k) + e_k \in \mathbb{C}$ $(k = 0, \ldots, N-1)$ of (1.1), where $e_k \in [-10^{-s},\, 10^{-s}] + \mathrm{i}\,[-10^{-s},\, 10^{-s}]$ $(s = 4,\, 6,\, 8)$ are uniformly random error terms. The corresponding results are shown in Table 6.1, where we have chosen $L = N - L + 1$. We observe that $\kappa_j(\boldsymbol{F}_M) = 1$ and furthermore that the RHS of (5.6) is also one, i.e. the estimate is sharp. The condition number $\kappa_j(\boldsymbol{F}_M^{\mathrm{SV}D})$ is only slightly larger. □

Further examples are given in [17, Examples 4.1 – 4.2] and [18, Example 6.1].

M	$N-L$	s	$\max\|z_j - \tilde{z}_j\|$	$\|\boldsymbol{P}_{N-L} - \tilde{\boldsymbol{P}}_{N-L}\|_2$	$\|\boldsymbol{q}_{N-L} - \tilde{\boldsymbol{q}}_{N-L}\|_2$
10	10	4	4.733e-06	2.023e-15	3.340e-05
10	20	4	2.029e-06	1.014e-05	1.453e-05
10	30	4	1.305e-06	1.066e-05	1.039e-05
10	10	6	4.002e-08	5.793e-15	2.764e-07
10	20	6	1.587e-08	7.933e-08	1.336e-07
10	30	6	1.259e-08	1.028e-07	9.445e-08
10	100	6	1.623e-09	4.663e-08	2.177e-08
50	100	6	2.803e-09	7.009e-08	6.918e-08
50	100	8	2.562e-11	6.406e-10	7.030e-10
50	200	8	1.081e-11	6.042e-10	3.315e-10

Table 6.1: Maximum errors of the nodes and the related estimates for noisy sampled data in Example 6.1.

Example 6.2 Now we choose M nodes on an Archimedean spiral given in the form

$$z_j = \sqrt{\frac{j+M}{2M}}\, \exp \frac{8\pi \mathrm{i}\, \sqrt{j+M}}{5} \quad (j = 1, \ldots, M)$$

and the coefficients $c_j = 1$ $(j = 1, \ldots, M)$. The exact data of (1.1) are denoted by (6.1). As in Example 6.1, we use noisy sampled data $\tilde{h}_k := h(k) + e_k \in \mathbb{C}$ $(k = 0, \ldots, N-1)$ of (1.1), where $e_k \in [-10^{-s},\, 10^{-s}] + \mathrm{i}\, [-10^{-s},\, 10^{-s}]$ $(s = 4, 6, 8)$ are uniformly random error terms. In Table 6.2, we present maximum errors of the nodes, where we have chosen $L = N - L + 1$. The condition numbers $\kappa_j(\boldsymbol{F}_M)$ and $\kappa_j(\boldsymbol{F}_M^{\mathrm{SVD}})$ of the eigenvalues are very similar. □

M	$N-L$	s	$\max \lvert z_j - \tilde{z}_j \rvert$	$\max \kappa_j$	RHS of (5.6)	$\lVert \boldsymbol{P}_{N-L} - \tilde{\boldsymbol{P}}_{N-L} \rVert_2$	$\lVert \boldsymbol{q}_{N-L} - \tilde{\boldsymbol{q}}_{N-L} \rVert_2$
10	10	6	9.746e-07	2.143e+00	3.330e+01	4.145e-15	2.869e-06
10	20	6	6.977e-07	1.749e+00	1.112e+01	1.879e-06	1.044e-06
10	30	6	4.991e-07	1.731e+00	9.864e+00	1.486e-06	1.905e-06
10	100	6	9.097e-07	1.718e+00	9.206e+00	2.703e-06	1.584e-06
30	100	6	2.415e-04	3.658e+01	1.036e+14	3.445e-04	3.385e-04
30	100	4	5.758e-04	3.658e+01	1.028e+14	2.476e-03	2.968e-04

Table 6.2: Maximum errors of the nodes and the related estimates for noisy sampled data in Example 6.2, where $\max \kappa_j$ denotes the maximal condition number of $\kappa_j(\boldsymbol{F}_M)$.

The Examples 6.1 – 6.2 show that our estimates for $|z_j - \tilde{z}_j|$ $(j = 1, \ldots, M)$ based on Theorem 5.1 are very precise. The estimations of the condition numbers are sharp and cannot be improved in some cases. However we observe also that the estimates of the condition numbers of the eigenvalues based on RHS of (5.6) are useless for higher order M.

In the following Example 6.3 we show that the orthogonal projection onto the signal space is essential for good error estimates for the ESPRIT Algorithm 3.4. Applying (5.1) to the matrices $\boldsymbol{F}_M$ and $\tilde{\boldsymbol{F}}_M$, we obtain the first-order error estimate

$$|z_j - \tilde{z}_j| \le \kappa_j(\boldsymbol{F}_M)\, \lVert \boldsymbol{F}_M - \tilde{\boldsymbol{F}}_M \rVert_2 \tag{6.2}$$

for $j = 1, \ldots, M$. In Example 6.3, one can see that the norm $\lVert \boldsymbol{F}_M - \tilde{\boldsymbol{F}}_M \rVert_2$ is not small also for large $N - L$. In other words, one cannot explain the good error behavior of the ESPRIT Algorithm 3.4 by the estimate (6.2). If we replace (3.12) by (3.10), then the same statement is true.

Example 6.3 As in Example 6.1, we choose the M equispaced nodes $z_j = \exp(2\pi \mathrm{i} j/M)$ $(j = 1, \ldots, M)$ on the unit circle and the coefficients $c_j = 1$ $(j = 1, \ldots, M)$. The corresponding results are shown in Table 6.3. If we use (3.10) instead of (3.12), then we obtain similar results. □

M	$N-L$	s	$\max \lvert z_j - \tilde{z}_j \rvert$	$\max \kappa_j(\boldsymbol{F}_M)$	$\lVert \boldsymbol{F}_M - \tilde{\boldsymbol{F}}_M \rVert_2$
10	20	4	2.145e-06	1.007e+00	2.014e+00
10	30	4	1.354e-06	1.004e+00	2.069e+00
10	100	4	2.317e-07	1.000e+00	1.999e+00
50	100	4	2.719e-07	1.002e+00	2.161e+00
50	100	8	2.772e-11	1.002e+00	2.101e+00

Table 6.3: Maximum errors of the nodes and the related estimates for noisy sampled data in Example 6.3.

Example 6.4 Finally we use the same parameters of a nuclear magnetic resonance (NMR) signal as in [2, Table 1], i.e., $M = 5$ with the nodes $z_1 = 0.6342 - 0.7463\,\mathrm{i}$, $z_2 = 0.8858 - 0.4067\,\mathrm{i}$, $z_3 = 0.9663 - 0.1661\,\mathrm{i}$, $z_4 = 0.9642 + 0.2174\,\mathrm{i}$, $z_5 = 0.8811 + 0.2729\,\mathrm{i}$ and the coefficients $c_1 = 5.8921 + 1.5788\,\mathrm{i}$, $c_2 = 9.5627 + 2.5623\,\mathrm{i}$, $c_3 = 5.7956 + 1.5529\,\mathrm{i}$, $c_4 = 2.7046 + 0.7247\,\mathrm{i}$, $c_5 = 16.4207 + 4.3999\,\mathrm{i}$. For $N = 160$, the left Figure 6.1 shows the 5 positive singular values of the exact Hankel matrix $\boldsymbol{H}_{L,160-L+1}$ for different window lengths $L = 5, \ldots, 80$. As expected (see Lemma 3.1 and Remark 3.3), the positive singular values of $\boldsymbol{H}_{L,160-L+1}$ increase for increasing window length $L = 5, \ldots, 80$. Thus $L = N/2$ is an optimal window length, where in practical applications it may be enough to choose $L = 2M$ or $L = 4M$. Note that the computational cost of the ESPRIT Algorithm 3.4 may be better for $L = 2M$ or $L = 4M$. For exactly sampled data, the right Figure 6.1 shows the errors $|z_j - \tilde{z}_j|$ between the given nodes z_j and the reconstructed nodes $\tilde{z}_j$ for $j = 1, \ldots, 5$. Note that both axes of ordinates in Figures 6.1 and 6.2 have a logarithmic scale. In Figure 6.2 we show the corresponding results for noisy sampled data $\tilde{h}_k = h(k) + e_k$ $(k = 0, \ldots, 159)$, where $\mathrm{Re}\, e_k$ and $\mathrm{Im}\, e_k$ zero-mean Gaussian random numbers with standard deviation 1. □

Acknowledgment

The first named author gratefully acknowledges the support by the German Research Foundation within the project PO 711/10–2. Special thanks are given to our friend Albrecht Böttcher for providing very valuable suggestions.

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, 2000.

[2] F.S.V. Bazán. Sensitivity eigenanalysis for single shift-invariant subspace-based methods. *Signal Process.*, 80:89 – 100, 2000.

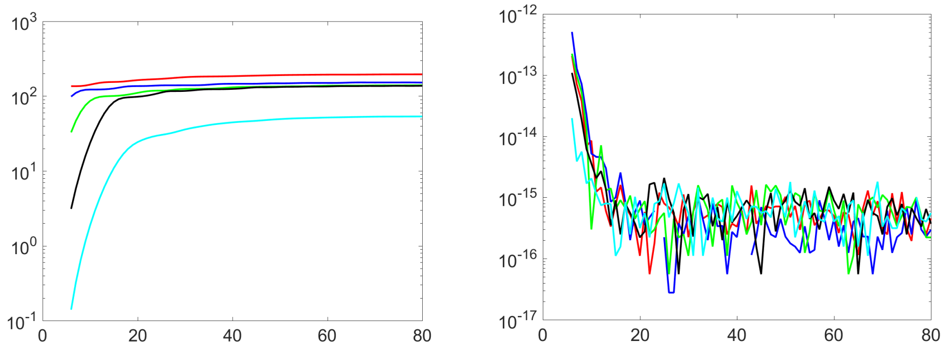

Figure 6.1: Singular values of $\boldsymbol{H}_{L,160-L+1}$ for different window lengths $L = 5, \ldots, 80$ (left) and the errors $|z_j - \tilde{z}_j|$ for $j = 1, \ldots, 5$ (right) in the case of exactly sampled data.

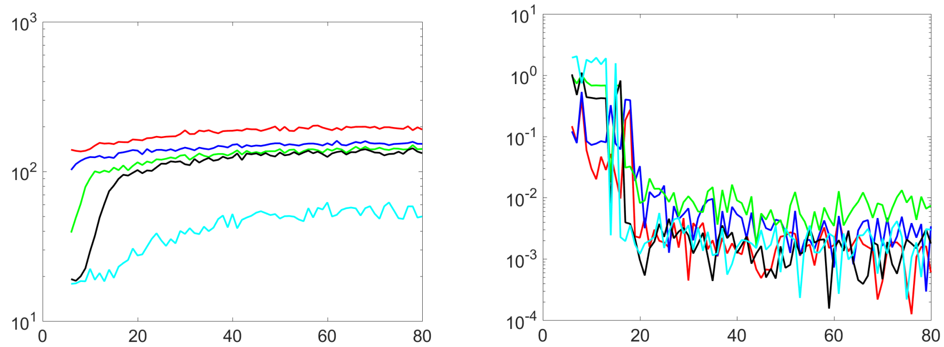

Figure 6.2: Singular values of $\tilde{\boldsymbol{H}}_{L,160-L+1}$ for different window lengths $L = 5, \ldots, 80$ (left) and the errors $|z_j - \tilde{z}_j|$ for $j = 1, \ldots, 5$ (right) in the case of noisy sampled data.

[3] F.S.V. Bazán. Error analysis of signal zeros: a projected companion matrix approach. *Linear Algebra Appl.*, 369:153 – 167, 2003.

[4] F.S.V. Bazán and P.L. Toint. Conditioning of infinite Hankel matrices of finite rank. *Systems Control Letters*, 41:347 – 359, 2000.

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

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

[7] G.H. Golub and C.F. Van Loan. *Matrix Computations*. Third edn. Johns Hopkins Univ. Press, Baltimore, 1996.

[8] N. Golyandina. On the choice of parameters in singular spectrum analysis and related subspace–based methods. *Stat. Interface*, 3:259 – 279, 2010.

[9] N.J. Higham. *Accuracy and Stability of Numerical Algorithms.* Second edn. SIAM, Philadelphia, 2002.

[10] R.A. Horn and C.R. Johnson. *Matrix Analysis.* Second edn. Cambridge Univ. Press, Cambridge, 2013.

[11] 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.

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

[13] W. Liao and A. Fannjiang. MUSIC for single-snapshot spectral estimation: Stability and super-resolution. *Appl. Comput. Harmon. Anal.*, in press.

[14] B.N. Parlett and C. Reinsch. Balancing a matrix for calculation of eigenvalues and eigenvectors. *Numer. Math.*, 13:293 – 304, 1969.

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

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

[17] D. Potts and M. Tasche. Parameter estimation for nonincreasing exponential sums by Prony-like methods. *Linear Algebra Appl.*, 439:1024 – 1039, 2013.

[18] D. Potts and M. Tasche. Fast ESPRIT algorithms based on partial singular value decompositions. *Appl. Numer. Math.*, 88:31 – 45, 2015.

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

[20] R.A. Smith. The condition numbers of the matrix eigenvalue problem. *Numer. Math.*, 10:232 – 240, 1967.

[21] G.W. Stewart. *Matrix Algorithms.* Vol. I: Basic Decompositions. SIAM, Philadelphia, 1998.

[22] G.W. Stewart and J.-G. Sun. *Matrix Perturbation Theory.* Academic Press, Boston, 1990.

[23] J.H. Wilkinson. Note on matrices with a very ill-conditioned eigenproblem. *Numer. Math.*, 19:176 – 178, 1972.