Application of the AAK theory for sparse approximation of exponential sums

Gerlind Plonka* Vlada Pototskaia†

26th September 2016

In this paper, we derive a new method for optimal ℓ^1- and ℓ^2-approximation of discrete signals on $\mathbb{N}_0$ whose entries can be represented as an exponential sum of finite length. Our approach employs Prony's method in a first step to recover the exponential sum that is determined by the signal. In the second step we use the AAK-theory to derive an algorithm for computing a shorter exponential sum that approximates the original signal in the ℓ^p-norm well. AAK-theory originally determines best approximations of bounded periodic functions in Hardy-subspaces. We rewrite these ideas for our purposes and give a proof of the used AAK theorem based only on basic tools from linear algebra and Fourier analysis. The new algorithm is tested numerically in different examples.

Key words: Exponential sums, Prony method, AAK theory, infinite Hankel matrices, structured low-rank approximation, rational approximation.
Mathematics Subject Classification: 15A18, 41A30, 42A16, 65F15.

1 Introduction

In signal processing and system theory, we consider the problem of sparse approximation of structured signals. In many applications it can be assumed that the signal is either exactly or approximately a finite linear combination of non-increasing exponentials with complex exponents, i.e., $\mathbf{f} := (f_k)_{k=0}^{\infty}$ satisfies

$$f_k := f(k) = \sum_{j=1}^{N} a_j \, z_j^k \,, \tag{1.1}$$

where $a_j \in \mathbb{C} \setminus \{0\}$ and with pairwise different $z_j \in \mathbb{D} := \{z \in \mathbb{C} : \, 0 < |z| < 1\}$. The problem of recovering $\mathbf{f}$ from a suitable number of signal values $f(k)$, $k = 0, 1, \ldots, M$ with $M \geq 2N - 1$ is a well-studied parameter estimation problem. It can be solved using a stabilized Prony-like method as e.g. ESPRIT [19] or the approximative Prony method (APM), [18]. For a recent generalization and a review on Prony methods for recovery of structured functions we refer to [15, 17]. Beside system theory, the

*Institute for Numerical and Applied Mathematics, Göttingen University, Lotzestr. 16-18, 37083 Göttingen, Germany. Email: plonka@math.uni-goettingen.de
†Institute for Numerical and Applied Mathematics, Göttingen University, Lotzestr. 16-18, 37083 Göttingen, Germany. Email: v.pototskaia@math.uni-goettingen.de

approximation of special functions (as e.g. $1/x$) by finite exponential sums is also of high interest for solving higher dimensional integral equations as e.g. in coupled cluster analysis in quantum chemistry, see [5,7].

For sparse signal approximation we are interested in solving the following problem. For a given $\mathbf{f}$ of the form (1.1), we want to find a new signal $\tilde{\mathbf{f}} := (\tilde{f}_k)_{k=0}^{\infty}$ of the form

$$\tilde{f}_k := \tilde{f}(k) = \sum_{j=1}^{K} \tilde{a}_j \, \tilde{z}_j^k \tag{1.2}$$

with $\tilde{a}_j \in \mathbb{C}$ and $\tilde{z}_j \in \mathbb{D}$ such that $K < N$ and $\|\mathbf{f} - \tilde{\mathbf{f}}\|_p \le \epsilon$. We will focus on the two cases $p = 1$ and $p = 2$.

More precisely, we have to consider the following two questions. For a given accuracy level $\epsilon > 0$, what is the smallest $K \in \mathbb{N}$ such that $\tilde{\mathbf{f}}$ in (1.2) satisfies $\|\mathbf{f} - \tilde{\mathbf{f}}\|_p \le \epsilon$, and how to compute $\tilde{z}_j \in \mathbb{D}$ and $\tilde{a}_j \in \mathbb{C}$, $j = 1, \ldots, K$? Vice versa, for a given "storage budget", i.e., a given $K \in \mathbb{N}$, how do we have to choose the parameters in (1.2) in order to achieve the smallest possible error $\|\mathbf{f} - \tilde{\mathbf{f}}\|_p$?

In this paper, we will approach this problem using the theory of Adamjan, Arov and Krein (AAK-theory) [1]. Consider the infinite Hankel matrix generated by $\mathbf{f} = (f_k)_{k=0}^{\infty}$ of the form $\mathbf{\Gamma_f} = (f_{j+k})_{j,k=0}^{\infty}$ being bounded on ℓ^p, and let $\sigma_0 \ge \sigma_1 \ge \ldots$ denote its singular values, ordered by size and repeated according to multiplicities. Then the AAK-theory states that $\mathbf{\Gamma_f}$ can be approximated by an infinite Hankel matrix $\mathbf{\Gamma}_{\tilde{\mathbf{f}}}$ of finite K such that

$$\|\mathbf{\Gamma_f} - \mathbf{\Gamma}_{\tilde{\mathbf{f}}}\|_{\ell^p \to \ell^p} = \min_{\mathrm{rank} \mathbf{\Gamma_g} \le K} \|\mathbf{\Gamma_f} - \mathbf{\Gamma_g}\|_{\ell^p \to \ell^p} = \sigma_K.$$

This result ist non-trival since a usual spectral decomposition of $\mathbf{\Gamma_f}$ does not preserve the Hankel structure. Thus the result is strongly related to the problem of structured low rank approximation for Hankel matrices, see e.g. [9]. As presented in [1], the problem is equivalent to showing that for an arbitrary bounded function $f \in L^{\infty}([0, 2\pi))$, the best approximation by a function $\tilde{f}$ from the subspace $H^{\infty,[K]}$ ($K \in \mathbb{N}$) of the Hardy space H^{∞} exists and satisfies

$$\|f - \tilde{f}\|_{\infty} = \min_{g \in H^{\infty,[K]}} \|f - g\|_{\infty} = \sigma_K(\mathbf{\Gamma_f}),$$

where $\mathbf{\Gamma_f}$ is generated by $\mathbf{f} = (f(k))_{k=0}^{\infty}$. Here $H^{\infty,[K]}$ denotes the space of functions in H^{∞} whose extension to the the unit disc possess at most K poles in $\mathbb{D}$. For more details we refer e.g. to [1,6,10,11,13,14,20]. The proofs of these assertions are substantially based on analysis of bounded functions in Hardy spaces and operator theory, employing the Nehari's Theorem [12], Kronecker's Theorem (see e.g. [20], Theorem 16.3) as well as Beurling's Theorem [3,6].

For earlier approaches to the application of the AAK theory in order to solve sparse approximation problems using exponential sums we refer to [4] and [2]. In [4], a finite-dimensional approximation problem is considered using $2N + 1$ equidistant samples of a continuous function f. The goal is to approximate f by an exponential sum as accurate as possible with $K < N$ terms. Using $N \times N$ Hankel matrices H_N, the authors derived a reduction procedure such that the spectral norm of the difference between H_N and the obtained rank K Hankel approximation $\tilde{H}_N$ is approximately σ_K, where σ_K is the K-th singular value of H_N. In [2], relations between the AAK-theory and related discrete and continuous approximation problems on $\mathbb{R}^+$ and on the interval have been studied. However, the connection between AAK-theory and corresponding finite-dimensional approximation problems is still not completely understood.

Contribution of our paper. In this paper, we present a new algorithm for solving the sparse approximation problem (1.2) and give an explicit procedure to compute the nodes $\tilde{z}_j$ as well as the coefficients $\tilde{a}_j$, $j = 1, \ldots, K$ such that $\|\mathbf{f} - \tilde{\mathbf{f}}\|_p \leq \sigma_K$ is satisfied, where σ_K denotes the K-th singular value of $\mathbf{\Gamma_f}$. The procedure also includes an explicit computation of all non-zero singular values of $\mathbf{\Gamma_f}$. For this purpose, we investigate the structure of con-eigenvectors of the infinite Hankel matrix $\mathbf{\Gamma_f}$ with finite rank N and reduce the problem of characterizing the nonzero singular values of $\mathbf{\Gamma_f}$ to the problem of finding the singular values of an $N \times N$ kernel matrix. Further, we provide a new proof of the AAK Theorem in our context using only concepts from linear algebra and Fourier analysis. We show how these results intimately relate to the Prony method for recovering exponential sums.

The numerical application of our approach usually employs in the first step a stabilized Prony method, like APM [18] to recover the parameters of $\mathbf{f}$ in (1.1). In a second step, the new reduction method is used to compute all parameters of $\tilde{\mathbf{f}}$. For a fixed target error $\epsilon > 0$ we choose K as the smallest index such that $\sigma_K \leq \epsilon$.

This paper is organized as follows. In Section 2 we summarize some basic notations and properties of infinite Hankel matrices. Section 3 is devoted to the derivation of the new algorithm for sparse approximation by short exponential sums. In Section 4.1 we provide some important properties of infinite Toeplitz and Hankel matrices. The assertions of the AAK theory, stated already in Section 3, are proven in Subsection 4.2. Moreover, we give more insights into the structure of singular values, (con)-eigenvectors as well as the kernel of finite rank Hankel matrices and its rank K Hankel approximations. These insights provide the close relation of the theory to Prony's method. Finally, in Section 5 we present some numerical examples and applications of the new sparse approximation algorithm and comment on stability issues.

2 Preliminaries

In the following we denote by $\ell^p := \ell^p(\mathbb{N}_0)$ for $p = 1, 2$ the space of p-summable sequences $\mathbf{v} = (v_k)_{k=0}^{\infty}$ with the norm $\|v\|_p := (\sum_{k=0}^{\infty} |v_p|^p)^{1/p}$ and by $\mathbb{D}$ the open unit disc without zero $\{z \in \mathbb{C} : 0 < |z| < 1\}$. For a sequence $\mathbf{v} = (v_k)_{k=0}^{\infty} \in \ell^p$ and $z \in \mathbb{D}$ we call

$$P_{\mathbf{v}}(z) := \sum_{k=0}^{\infty} v_k z^k$$

its corresponding Laurent polynomial and $P_{\mathbf{v}}(e^{i\omega})$, $\omega \in \mathbb{R}$, its Fourier series. Further, for $\mathbf{f} \in \ell^1$ we define the infinite Hankel matrix

$$\mathbf{\Gamma_f} := \begin{pmatrix} f_0 & f_1 & f_2 & \cdots \\ f_1 & f_2 & f_3 & \cdots \\ f_2 & f_3 & f_4 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} = (f_{k+j})_{k,j=0}^{\infty}\,. \tag{2.1}$$

Then $\mathbf{\Gamma_f}$ determines an operator $\mathbf{\Gamma_f} : \ell^p \to \ell^p$ given by

$$\mathbf{\Gamma_f}\,\mathbf{v} = \left(\sum_{j=0}^{\infty} f_{k+j} v_j\right)_{k=0}^{\infty} \qquad \text{for } \mathbf{v} := (v_k)_{k=0}^{\infty} \in \ell^p.$$

This is a direct consequence of Youngs inequality, since $\mathbf{\Gamma_f}\,\mathbf{v}$ can be easily reinterpreted as a convolution by extending the sequence spaces to $\ell^p(\mathbb{Z})$.

Con-diagonalization of infinite Hankel matrices. Observe that $\mathbf{\Gamma_f}$ is symmetric. Generalizing the idea of unitary diagonalization of Hermitian matrices resp. compact selfadjoint operators, we will apply the concept of con-similarity and con-diagonalization, see e.g. [8] for the finite-dimensional case.

For an infinite Hankel matrix $\mathbf{\Gamma_f}$ we call $\lambda \in \mathbb{C}$ a *con-eigenvalue* with the corresponding *con-eigenvector* $\mathbf{v} \in \ell^p$ if it satisfies

$$\mathbf{\Gamma_f}\overline{\mathbf{v}} = \lambda\,\mathbf{v}.$$

Observe that for $\mathbf{\Gamma_f}\overline{\mathbf{v}} = \lambda\,\mathbf{v}$ also

$$\mathbf{\Gamma_f}(\overline{\mathrm{e}^{\mathrm{i}\alpha}\mathbf{v}}) = \mathrm{e}^{-\mathrm{i}\alpha}\mathbf{\Gamma_f}\overline{\mathbf{v}} = (\mathrm{e}^{-\mathrm{i}\alpha}\lambda)\mathbf{v} = (\mathrm{e}^{-2\mathrm{i}\alpha}\lambda)(\mathrm{e}^{\mathrm{i}\alpha}\mathbf{v})$$

for all $\alpha \in \mathbb{R}$. Thus, for each con-eigenvalue λ of $\mathbf{\Gamma_f}$ we can find a corresponding real non-negative con-eigenvalue $\sigma = |\lambda|$ by this rotation trick. In the following, we will restrict the con-eigenvalues to their unique nonnegative representatives. Now, it can be simply observed that a symmetric infinite Hankel matrix $\mathbf{\Gamma_f}$ with finite rank is compact and unitarily con-diagonalizable, see [8]. Since $\mathbf{\Gamma_f}\overline{\mathbf{v}} = \lambda\,\mathbf{v}$ implies

$$(\mathbf{\Gamma_f^*}\mathbf{\Gamma_f})\mathbf{v} = \mathbf{\Gamma_f^*}\lambda\overline{\mathbf{v}} = \lambda\overline{\mathbf{\Gamma}}_\mathbf{f}\overline{\mathbf{v}} = \lambda\overline{\mathbf{\Gamma_f}\mathbf{v}} = |\lambda|^2\mathbf{v},$$

we directly observe that the nonnegative con-eigenvalues and con-eigenvectors of $\mathbf{\Gamma_f}$ are also singular values and corresponding singular vectors of $\mathbf{\Gamma_f}$. Conversely, for symmetric matrices a singular pair $(\sigma, \mathbf{v})$ is also a con-eigenpair of $\mathbf{\Gamma_f}$ if the geometric multiplicity of σ is 1.

Infinite Hankel matrices of finite rank. In the following, we consider special sequences $\mathbf{f} = (f_k)_{k=0}^{\infty}$. Let $\mathbf{f}$ be of the form (1.1) where $N \in \mathbb{N}$, $a_j \in \mathbb{C} \neq \{0\}$ and with pairwise different nodes $0 < |z_N| \le \ldots \le |z_1| < 1$. Then $\mathbf{f} \in \ell^1$ since

$$\|\mathbf{f}\|_1 = \sum_{k=0}^{\infty} |f_k| = \sum_{k=0}^{\infty} \left|\sum_{j=1}^{N} a_j z_j^k\right| \le \sum_{j=1}^{N}\left(\sum_{k=0}^{\infty} |a_j z_j^k|\right) = \sum_{j=1}^{N} \frac{|a_j|}{1-|z_j|} < \infty.$$

First, we recall the following property of the corresponding infinite Hankel matrix $\mathbf{\Gamma_f}$, see e.g. [20], Theorem 16.13.

Theorem 2.1. (Kronecker's Theorem). *The Hankel operator $\mathbf{\Gamma_f} : \ell^p \to \ell^p$ generated by $\mathbf{f} = (f_k)_{k=0}^{\infty} \in \ell^1$ of the form (1.1) has finite rank N.*

Proof. For reader's convenience we provide a short proof that also gives some insight into the connection to difference equations. If $\mathbf{f}$ can be written in the form (1.1), we define the characteristic polynomial (Prony polynomial)

$$P(z) := \prod_{j=1}^{N}(z - z_j) = \sum_{k=0}^{N} b_k z^k. \tag{2.2}$$

Then

$$\sum_{l=0}^{N} b_l f_{k+l} = \sum_{l=0}^{N} b_l \sum_{j=1}^{N} a_j z_j^{k+l} = \sum_{j=1}^{N} a_j z_j^k (\sum_{l=0}^{N} p_l z_j l) = 0 \tag{2.3}$$

for all $k \in \mathbb{N}_0$, i.e., the $(N+k)$-th column of $\mathbf{\Gamma_f}$ is a linear combination of the N preceding columns. Thus rank $\mathbf{\Gamma_f} \leq N$. Since $P(z)$ has exact degree N it follows that rank $\mathbf{\Gamma_f} = N$. □

Remark 2.2. Conversely, if the infinite Hankel matrix $\mathbf{\Gamma_f}$ possesses rank N then $\mathbf{f}$ satisfies a difference equation of order N. Thus, there exist coefficients $b_0, \ldots, b_N$ such that

$$\sum_{l=0}^{N} b_l f_{k+l} = 0 \quad \forall k \in \mathbb{N}_0.$$

Assuming that the zeros z_j, $j = 1, \ldots, N$ of the characteristic polynomial $\sum_{l=0}^{N} b_l z^l$ are pairwise different, $\mathbf{f}$ can be written in the form (1.1). The zeros have modulus smaller than 1 since $\mathbf{f}$ has been assumed to be in ℓ^1.

3 Algorithm for sparse approximation of exponential sums

Let us assume now that the signal $\mathbf{f} = (f_k)_{k=0}^{\infty}$ is of the form (1.1) with pairwise different nodes z_j. Further, let a suitable number of it's samples f_k, $k = 0, 1, \ldots, M$ with $M \geq 2N - 1$ be given. Our approach for solving the sparse approximation problem (2) consists of two steps. In the first step we reconstruct the parameters a_j and z_j using Prony's method, see e.g. [17–19]. Once the representation (1.1) of $\mathbf{f}$ is known, in the second step we apply the AAK-theory [1] to compute a new signal $\tilde{\mathbf{f}}$ that can be represented by shorter exponential sum (1.2) with $K < N$ and satisfies $\|\mathbf{f} - \tilde{\mathbf{f}}\|_p < \epsilon$.

Classical Prony method. Let us shortly summarize Prony's method for the recovery of exponential sums. As seen already in the proof of Theorem 2.1, a signal $\mathbf{f}$ of the form (1.1) satisfies a homogeneous difference equation (2.3) of order N with constant coefficients b_l being determined by the coefficients of the Prony polynomial $P(z)$ in (2.2). For given samples f_k, $k = 0, \ldots, 2N$, (2.3) leads to the homogeneous system of equations

$$\sum_{l=0}^{N} b_l f_{l+k} = 0, \qquad k = 0, \ldots, N,$$

whose coefficient matrix $H_{N+1} \in \mathbb{C}^{(N+1)\times(N+1)}$ is the leading principal minor of $\mathbf{\Gamma_f}$ with rank $H_{N+1} = N$ by Theorem 2.1. Recalling that the leading coefficient of the Prony polynomial satisfies $b_N = 1$, there exists a unique solution of the homogeneous equation system, namely the eigenvector $\mathbf{b} = (b_l)_{l=0}^{N}$ of H_{N+1} to the single eigenvalue 0 normalized by $b_N = 1$. This observation leads to the following naive algorithm to recover the parameters z_j, a_j in (1.1) by the signal samples f_k, $k = 0, \ldots, 2N$.

1. Compute the eigenvector $\mathbf{b}$ of H_{N+1} with $b_N = 1$ corresponding to the zero-eigenvalue.
2. Determine the Prony polynomial $P(z)$ in (2.2) with coefficient vector $\mathbf{b}$ and compute its N zeros z_j, $j = 1, \ldots, N$.
3. Solve the overdetermined linear system (1.1) with $k = 0, \ldots, 2N$ to determine the weights a_j, $j = 1, \ldots, N$.

Remarks 3.1. 1. We observe that the zero-eigenvector of H_{N+1} also produces a zero-(con)-eigenvector of the infinite Hankel matrix $\boldsymbol{\Gamma}_{\mathbf{f}}$ by taking $\mathbf{v}^{(N)} = (v_l^{(N)})_{l=0}^{\infty}$ with $v_l^{(N)} = b_l$ for $l = 0, \ldots, N$ and $b_l = 0$ for $l > N$. Then the Laurent polynomial corresponding to $\mathbf{v}^{(N)}$ given by $P_{\mathbf{v}^{(N)}}(z) := \sum_{l=0}^{\infty} v_l^{(N)} z^l$ equals to the Prony polynomial $P(z)$ in (2.2).
2. Numerically stable algorithms for Prony's method usually employ more than $2N$ samples to achieve stability and use e.g. singular value decompositions of (rectangular) Hankel matrices as well as matrix pencil methods for evaluating the nodes z_j, $j = 1, \ldots, N$. In our numerical tests, we particularly use the APM method for stable evaluation of the parameters, see [18].

Sparse approximation based on AAK-theory. In order to compute an optimal approximation $\tilde{\mathbf{f}}$ of $\mathbf{f}$ using a shorter exponential sum, we want to apply the following theorem based on AAK theory, [1].

Theorem 3.2. *Let the Hankel matrix $\boldsymbol{\Gamma}_{\mathbf{f}}$ of rank N be generated by the the sequence $\mathbf{f}$ of the form (1.1) with $1 > |z_1| \geq \cdots \geq |z_N| > 0$. Let the N nonzero singular values of $\boldsymbol{\Gamma}_{\mathbf{f}}$ be ordered by size $\sigma_0 \geq \sigma_1 \ldots \geq \sigma_{N-1} > 0$. Then, for each $K \in \{0, \ldots, N-1\}$ satisfying $\sigma_K \neq \sigma_k$ for $K \neq k$ the Laurent polynomial of the corresponding con-eigenvector $\mathbf{v}^{(K)} = (v_l^{(K)})_{l=0}^{\infty}$,*

$$P_{\mathbf{v}^{(K)}}(z) := \sum_{l=0}^{\infty} v_l^{(K)}\, z^l,$$

has exactly K zeros $z_1^{(K)}, \ldots, z_K^{(K)}$ in $\mathbb{D}$, repeated according to their multiplicity. Furthermore, if $z_1^{(K)}, \ldots, z_K^{(K)}$ are pairwise different, then there exist coefficients $\tilde{a}_1, \ldots, \tilde{a}_K \in \mathbb{C}$ such that for

$$\tilde{\mathbf{f}}^{(K)} = \left(\tilde{f}_l^{(K)}\right)_{l=0}^{\infty} = \left(\sum_{j=1}^{K} \tilde{a}_j (z_j^{(K)})^l\right)_{l=0}^{\infty} \tag{3.1}$$

we have for $p = 1$ and $p = 2$

$$\|\mathbf{f} - \tilde{\mathbf{f}}^{(K)}\|_p \leq \sigma_K.$$

We will give a proof of this theorem in Section 4, where we also provide more insights into the structure of infinite Hankel matrices with finite rank and its spectral properties.

In order to apply this theorem to the sparse approximation problem (1.2) we need to find a numerical procedure to compute the singular pairs $(\sigma_n, \mathbf{v}^{(n)})$ of $\boldsymbol{\Gamma}_{\mathbf{f}}$ for $n = 0, \ldots, N-1$ and to find all zeros of the expansion $P_{\mathbf{v}^{(n)}}(z)$ lying inside $\mathbb{D}$. In a final step we have to compute the optimal coefficients $\tilde{a}_j$.

Investigating the special structure of the con-eigenvectors of $\boldsymbol{\Gamma}_{\mathbf{f}}$ corresponding to the non-zero con-eigenvalues (resp. singular values) we show the following result that provides us with an algorithm to compute all non-zero singular values of $\boldsymbol{\Gamma}_{\mathbf{f}}$ and the corresponding con-eigenvectors exactly.

Theorem 3.3. *Let* $\mathbf{f}$ *be of the form* (1.1)*. Then the con-eigenvector* $\mathbf{v}^{(l)} = (v_k^{(l)})_{k=0}^{\infty}$ *of* $\mathbf{\Gamma_f}$ *corresponding to a single nonzero singular value* σ_l *of* $\mathbf{\Gamma_f}$*,* $l \in \{0, 1, \ldots, N-1\}$ *is of the form*

$$v_k^{(l)} = \frac{1}{\sigma_l} \sum_{j=1}^{N} a_j P_{\overline{\mathbf{v}}^{(l)}}(z_j) z_j^k, \qquad k \in \mathbb{N}_0, \tag{3.2}$$

where the vector $(P_{\overline{\mathbf{v}}^{(l)}}(z_j))_{j=1}^N = (\overline{P_{\mathbf{v}^{(l)}}(\overline{z}_j)})_{j=1}^N$ *is determined by the con-eigenvector of the finite con-eigenvalue problem*

$$\mathbf{A}_N \mathbf{Z}_N (\overline{P_{\mathbf{v}^{(l)}}(\overline{z}_j)})_{j=1}^N = \sigma_l (P_{\mathbf{v}^{(l)}}(\overline{z}_j))_{j=1}^N \tag{3.3}$$

with

$$\mathbf{A}_N := \begin{pmatrix} a_1 & & & 0 \\ & a_2 & & \\ & & \ddots & \\ 0 & & & a_N \end{pmatrix}, \quad \mathbf{Z}_N := \begin{pmatrix} \frac{1}{1-|z_1|^2} & \frac{1}{1-z_1\bar{z}_2} & \cdots & \frac{1}{1-z_1\bar{z}_N} \\ \frac{1}{1-\bar{z}_1 z_2} & \frac{1}{1-|z_2|^2} & \cdots & \frac{1}{1-z_2\bar{z}_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{1-\bar{z}_1 z_N} & \frac{1}{1-\bar{z}_2 z_N} & \cdots & \frac{1}{1-|z_N|^2} \end{pmatrix}.$$

Proof. Let $(\sigma_l, \mathbf{v}^{(l)})$ with $\sigma_l \neq 0$ be a con-eigenpair of $\mathbf{\Gamma_f}$, i.e. $\mathbf{\Gamma_f} \overline{\mathbf{v}}^{(l)} = \sigma \, \mathbf{v}^{(l)}$. With the notation $P_{\overline{\mathbf{v}}^{(l)}}(z) := \sum_{k=0}^{\infty} \overline{v}_k^{(l)} z^k$ it follows by (1.1) that

$$\sigma_l v_k^{(l)} = (\mathbf{\Gamma_f} \overline{\mathbf{v}}^{(l)})_k = \sum_{r=0}^{\infty} f_{k+r} \overline{v}_r^{(l)} = \sum_{r=0}^{\infty} \sum_{j=1}^{N} a_j \, z_j^{k+r} \overline{v}_r^{(l)} = \sum_{j=1}^{N} a_j \, P_{\overline{\mathbf{v}}^{(l)}}(z_j) \, z_j^k$$

for all $k \in \mathbb{N}_0$, and hence (3.2) is true. The relation (3.3) is now a consequence of (3.2) observing that

$$\sigma_l P_{\mathbf{v}^{(l)}}(z) = \sigma_l \sum_{k=0}^{\infty} v_k^{(l)} z^k = \sum_{k=0}^{\infty} \sum_{j=1}^{N} a_j P_{\overline{\mathbf{v}}^{(l)}}(z_j) z_j^k z^k = \sum_{j=1}^{N} \frac{a_j P_{\overline{\mathbf{v}}^{(l)}}(z_j)}{1 - z_j z} \tag{3.4}$$

for $z \in \mathbb{D}$ by inserting $z = \bar{z}_r$, $r = 1, \ldots, N$. □

Equation (3.4) also shows that that $P_{\mathbf{v}^{(l)}}(z)$ is a rational function whose numerator is a polynomial of degree at most $N-1$. Thus, in order to find the zeros of $P_{\mathbf{v}^{(l)}}(z)$ we only need to compute the $N-1$ zeros of the numerator in this rational representation. Note that a similar idea of dimension reduction has been used by Beylkin and Monzón in [4]. But in contrast to the above approach, they considered the rank reduction of a finite Hankel matrix. We combine our observations with Theorem 3.2 and obtain the following procedure to compute the desired approximation $\tilde{\mathbf{f}}$ in the ℓ^2 norm.

Algorithm for sparse approximation of exponential sums.

Input: samples f_k, $k = 0, \ldots, M$ for sufficiently large $M \geq 2N-1$,
target approximation error ϵ.

1. Find the parameters $z_j \in \mathbb{D}$ and $a_j \in \mathbb{C} \setminus \{0\}$, $j = 1, \ldots, N$ of the exponential representation of $\mathbf{f}$ in (1.1) using a Prony-like method.
2. Solve the con-eigenproblem for the matrix $\mathbf{A}_N \mathbf{Z}_N$ and determine the largest singular value σ_K with $\sigma_K < \epsilon$.
3. Compute the K zeros $z_j^{(K)} \in \mathbb{D}$, $j = 1, \ldots, K$, of the con-eigenpolynomial $P_{\mathbf{v}^{(K)}}(z)$ of $\mathbf{\Gamma_f}$ using its rational representation (3.4).

4. Compute the coefficients $\tilde{a}_j$ by solving the minimization problem

$$\min_{\tilde{a}_1,\ldots,\tilde{a}_K} \|\mathbf{f} - \tilde{\mathbf{f}}\|_{\ell^2}^2 = \min_{\tilde{a}_1,\ldots,\tilde{a}_K} \sum_{k=0}^{\infty} |f_k - \sum_{j=1}^{K} \tilde{a}_j (z_j^{(K)})^k|^2.$$

Output: sequence $\tilde{\mathbf{f}}$ of the form (1.2) such that $\|\mathbf{f} - \tilde{\mathbf{f}}\|_{\ell^2} \le \sigma_n < \epsilon$.

Remarks 3.4.
1. Since $\mathbf{A}_N\mathbf{Z}_N$ is con-diagonalizable by Theorem 3.3, it follows from Theorem 4.6.6 in [8] that $\overline{\mathbf{A}_N\mathbf{Z}_N}\mathbf{A}_N\mathbf{Z}_N$ has only real nonnegative eigenvalues λ_j. Now, if $(\lambda_j, \mathbf{w}^{(j)})$ is an eigenpair of $\overline{\mathbf{A}_N\mathbf{Z}_N}\mathbf{A}_N\mathbf{Z}_N$, then $\mathbf{v}^{(j)} := \mathbf{A}_N\mathbf{Z}_N\mathbf{w}^{(j)} + \sigma_j\overline{\mathbf{w}}^{(j)}$ is a con-eigenvector of $\mathbf{A}_N\mathbf{Z}_N$ to the con-eigenvalue $\sigma_j = \sqrt{\lambda_j}$, since

$$\begin{aligned}
\mathbf{A}_N\mathbf{Z}_N\overline{\mathbf{v}}^{(j)} &= \mathbf{A}_N\mathbf{Z}_N(\overline{\mathbf{A}}_N\overline{\mathbf{Z}}_N\overline{\mathbf{w}}^{(j)} + \sigma_j\mathbf{w}^{(j)}) \\
&= \overline{\overline{\mathbf{A}_N\mathbf{Z}_N}\mathbf{A}_N\mathbf{Z}_N\mathbf{w}^{(j)}} + \sigma_j\mathbf{A}_N\mathbf{Z}_N\mathbf{w}^{(j)} \\
&= \sigma_j^2\overline{\mathbf{w}}^{(j)} + \sigma_j\mathbf{A}_N\mathbf{Z}_N\mathbf{w}^{(j)} = \sigma_j\mathbf{v}^{(j)}.
\end{aligned}$$

Thus, to solve the con-eigenvalue problem in step 2 of the algorithm, we have to consider a usual eigenvalue-decomposition of $\overline{\mathbf{A}_N\mathbf{Z}_N}\mathbf{A}_N\mathbf{Z}_N$.
2. Observing that the components f_k have the form (1.1), the ℓ^2-minimization problem in step 4 of the algorithm breaks down to a least squares problem with complex coefficients of the form

$$\min_{\tilde{a}_1,\ldots,\tilde{a}_K} \left(2\mathrm{Re} \sum_{\nu=1}^{N}\sum_{j=1}^{K} \frac{\overline{a}_\nu \tilde{a}_j}{1 - \overline{z}_\nu z_j^{(K)}} + \sum_{j=1}^{K}\sum_{j'=1}^{K} \frac{\tilde{a}_j \overline{\tilde{a}}_{j'}}{1 - z_j^{(K)}\overline{z_j^{(K)}}} \right).$$

If we are interested to find an optimal sequence $\tilde{\mathbf{f}}$ in the ℓ^1-norm instead of the ℓ^2-norm, then we have to replace the minimization problem in step 4 accordingly. This problem can be reformulated as a linear program but its solution is more expensive than solving a least squares problem.
3. For a short summary of this section we also refer to [16].

4 The AAK theory revisited

In order to prove Theorem 3.2, we need some more inside information on infinite Hankel and Toeplitz matrices. We aim at showing the needed assertion using only tools from linear algebra and Fourier analysis. We hope that these insights will help to derive an analogous result on structured low-rank approximation of finite Hankel matrices in the future. The original result as well as all further presentations of the AAK theory that we are aware of require fundamental theorems in complex analysis for approximation of meromorphic functions, as the Nehari theorem and the Beurling theorem, see e.g. [1, 6, 14, 20].

4.1 Basic properties of infinite Hankel and Toeplitz matrices

Let us start with summarizing some important properties of special infinite Hankel and Toeplitz matrices. As shown in Section 2, we consider a Hankel matrix $\mathbf{\Gamma}_\mathbf{f}$ as in (2.1) generated by $\mathbf{f} \in \ell^1$ such that $\mathbf{\Gamma}_\mathbf{f} : \ell^p \to \ell^p$ is bounded for $p = 1$ and $p = 2$.

Characterization of the kernel of an infinite Hankel matrix. Let $\mathbf{v} := (v_k)_{k=0}^{\infty}$ be a sequence in ℓ^p and $p \in \{1, 2\}$. We define the *(forward) shift operator* $\mathbf{S} : \ell^p \to \ell^p$ by

$$\mathbf{S}\mathbf{v} := (0, v_0, v_1, v_2, \dots)$$

and the *backward shift operator* $\mathbf{S}^* : \ell^p \to \ell^p$ by

$$\mathbf{S}^*\mathbf{v} := (v_1, v_2, v_3, \dots).$$

The *shift invariant subspace* of ℓ^p generated by the sequence $\mathbf{v} \in \ell^2$ (resp. $\mathbf{v} \in \ell^1 \subset \ell^2$) is denoted by

$$\mathcal{S}_{\mathbf{v}} := \mathrm{clos}_{\ell^2}\mathrm{span}\,\{\mathbf{S}^k\mathbf{v} : k \in \mathbb{N}_0\}.$$

Note that the k-th row of $\mathbf{\Gamma_f}$ is the backward shift $(\mathbf{S}^*)^k\mathbf{f}$ of the first row. Due to the symmetry, the same holds for the columns. Thus for $\mathbf{v} \in \ell^2$ we have

$$\mathbf{\Gamma_f}\mathbf{S}\mathbf{v} = \left(\sum_{j=1}^{\infty} f_{k+j}\, v_{j-1}\right)_{k=0}^{\infty} = \left(\sum_{j=0}^{\infty} f_{k+1+j} v_j\right)_{k=0}^{\infty} = \mathbf{S}^*\mathbf{\Gamma_f}\mathbf{v}. \tag{4.1}$$

This commutator relation determines the structure of a Hankel operator and can be even used as a formal definition of $\mathbf{\Gamma_f}$, see [1].

The following lemma gives a useful characterization of the kernel of $\mathbf{\Gamma_f}$.

Lemma 4.1. *Let $\mathbf{f} := (f_k)_{k=0}^{\infty}$ be a sequence in ℓ^1 and $\mathbf{\Gamma_f}$ the corresponding infinite Hankel matrix as above. Then the following assertions hold.*

(i) *The kernel space $\ker(\mathbf{\Gamma_f}) := \{\mathbf{v} \in \ell^2 : \mathbf{\Gamma_f}\mathbf{v} = \mathbf{0}\}$ is $\mathbf{S}$-invariant, i.e., for $\mathbf{v} \in \ker(\mathbf{\Gamma_f})$ we have $\mathcal{S}_{\mathbf{v}} \subset \ker(\mathbf{\Gamma_f})$.*

(ii) *A vector $\mathbf{v} \in \ell^2$ is in $\ker(\mathbf{\Gamma_f})$ if and only if $\mathbf{f} \in (\mathcal{S}_{\overline{\mathbf{v}}})^{\perp}$.*

Proof. 1. Let $\mathbf{v} \in \ker(\mathbf{\Gamma_f})$. Then (4.1) implies

$$\mathbf{\Gamma_f}\mathbf{S}\mathbf{v} = \mathbf{S}^*\mathbf{\Gamma_f}\mathbf{v} = \mathbf{S}^*\mathbf{0} = \mathbf{0},$$

thus $\mathbf{S}\mathbf{v}$ is also in $\ker(\mathbf{\Gamma_f})$.

2. Using the definition of $\mathcal{S}_{\mathbf{v}}$ we obtain

$$\begin{aligned}
\mathbf{\Gamma_f}\mathbf{v} = 0 \;&\Leftrightarrow\; \sum_{k=0}^{\infty} f_{k+j}\, v_k = 0 \qquad \forall\, j \in \mathbb{N}_0\\
&\Leftrightarrow\; \sum_{k=0}^{\infty} (S^j\mathbf{v})_k f_k = 0 \quad \forall\, j \in \mathbb{N}_0\\
&\Leftrightarrow\; \langle \mathbf{f}, \mathbf{S}^j\bar{\mathbf{v}}\rangle_{\ell^2} = 0 \qquad \forall\, j \in \mathbb{N}_0\\
&\Leftrightarrow\; \mathbf{f} \perp \mathcal{S}_{\overline{\mathbf{v}}}
\end{aligned}$$

for every $\mathbf{v} \in \ell^2$. □

Let us now come back to the sequence $\mathbf{f}$ of the special form (1.1) with $z_j \in \mathbb{D}$. Then the structure of (con)-eigenvectors corresponding to the zero-con-eigenvalues of $\mathbf{\Gamma_f}$ can be described as follows.

Theorem 4.2. *Let $\mathbf{f}$ be a vector of the form (1.1). Then, $\mathbf{v} \in \ell^2$ satisfies $\mathbf{\Gamma_f}\overline{\mathbf{v}} = \mathbf{0}$ if and only if the corresponding Laurent polynomial satisfies $P_{\overline{\mathbf{v}}}(z_j) = 0$, for $j = 1, \dots, N$, where the z_j are given in (1.1).*

Proof. Observe first that $P_{\overline{\mathbf{v}}}(z)$ is well-defined for each $z \in \mathbb{D}$. The assertion $\mathbf{\Gamma_f} \overline{\mathbf{v}} = \mathbf{0}$ implies

$$0 = (\mathbf{\Gamma_f}\overline{\mathbf{v}})_k = \sum_{r=0}^{\infty} f_{k+r}\overline{v}_r = \sum_{r=0}^{\infty}\sum_{j=1}^{N} a_j z_j^{k+r}\overline{v}_r = \sum_{j=1}^{N} a_j z_j^k \sum_{r=0}^{\infty} \overline{v}_r z_j^r = \sum_{j=1}^{N} a_j P_{\overline{\mathbf{v}}}(z_j) z_j^k,$$

for all $k \in \mathbb{N}_0$ and thus

$$0 = \sum_{r=0}^{\infty}\sum_{j=1}^{N} a_j P_{\overline{\mathbf{v}}}(z_j) z_j^r z^r = \sum_{j=1}^{N} \frac{a_j P_{\overline{\mathbf{v}}}(z_j)}{1 - z_j z}$$

for all $z \in \mathbb{D}$. Hence, $P_{\overline{\mathbf{v}}}(z_j) = 0$ for $j = 1, \ldots, N$. Conversely, $P_{\overline{\mathbf{v}}}(z_j) = 0$ obviously implies that this equation is satisfied. □

Remark 4.3. A result similar to the assertion of Theorem 4.2 can be found e.g. in [20], see Lemma 16.11 in the context of of the Adamyan-Arov-Krein-Theory [1] for approximation of meromorphic functions in Hardy spaces.

Trianguar Toeplitz matrices. For $\mathbf{g} = (g_k)_{k=0}^{\infty} \in \ell^p$ $p \in \{1, 2\}$, we define the infinite triangular Toeplitz matrix $\mathbf{T_g}$ by

$$\mathbf{T_g} := \begin{pmatrix} g_0 & & & \\ g_1 & g_0 & & \\ g_2 & g_1 & g_0 & \\ \vdots & \vdots & & \ddots \end{pmatrix}.$$

For $\mathbf{g} \in \ell^1$, $\mathbf{T_g}$ determines a bounded operator $\mathbf{T_g} : \ell^\nu \to \ell^\nu$ for $\nu \geq 1$ given by

$$\mathbf{T_g v} := \left(\sum_{j=0}^{k} g_{k-j} v_j \right)_{k=0}^{\infty} = \mathbf{g} * \mathbf{v}, \qquad \mathbf{v} \in \ell^\nu,$$

since

$$\|\mathbf{T_g v}\|_\nu = \|\mathbf{g} * \mathbf{v}\|_\nu \leq \|\mathbf{g}\|_1 \|\mathbf{v}\|_\nu$$

by Young's inequality. Similarly, for $\mathbf{g} \in \ell^2$, $\mathbf{T_g} : \ell^1 \to \ell^1$ is bounded. We summarize some important properties of $\mathbf{T_g}$ in the following two lemmas.

Lemma 4.4. *For two sequences* $\mathbf{f} \in \ell^1$ *and* $\mathbf{g} \in \ell^p$ $p \in \{1, 2\}$ *we have*

(1) *The convolution* $\mathbf{f} * \mathbf{g}$ *is a sequence in* ℓ^p *and*

$$\mathbf{T_{f*g}} = \mathbf{T_f} \cdot \mathbf{T_g} = \mathbf{T_g} \cdot \mathbf{T_f}.$$

The corresponding Fourier series satisfy $P_{\mathbf{f*g}}(e^{i\omega}) = P_{\mathbf{f}}(e^{i\omega}) \cdot P_{\mathbf{g}}(e^{i\omega})$ *for all* $\omega \in \mathbb{R}$.

(2) *For* $\mathbf{g} \in \ell^1$, $\mathbf{\Gamma_f T_g}$ *is a bounded Hankel operator on* ℓ^ν *for* $\nu \geq 1$. *For* $\mathbf{g} \in \ell^2$, $\mathbf{\Gamma_f T_g}$ *is a bounded Hankel operator on* ℓ^1.

(3) *We have* $\mathbf{\Gamma_f T_g} = \mathbf{T_g}^T \mathbf{\Gamma_f}$.

Proof. 1. We observe that for $l \geq k$

$$(\mathbf{T_f T_g})_{l,k} = \sum_{r=0}^{l-k} f_{(l-k)-r} g_r = \sum_{r=0}^{l-k} f_r g_{(l-k)-r} = (\mathbf{T_f T_g})_{l,k} = (\mathbf{f} * \mathbf{g})_{l-k}$$

while $(\mathbf{T_f T_g})_{l,k} = 0$ for $l < k$. Young's inequality ensures that $(\mathbf{f} * \mathbf{g}) \in \ell^p$ and thus the product of Toeplitz operators is a bounded operator on ℓ^ν, $\nu \geq 1$ for $p = 1$ and on ℓ^1 for $p = 2$. The relation for the corresponding Fourier series follows by the convolution theorem.

2. Since the j-th row of $\mathbf{\Gamma_f}$ is $(\mathbf{S}^*)^j \mathbf{f}$ and the k-th column of $\mathbf{T_g}$ is $\mathbf{S}^k \mathbf{g}$, it follows that

$$(\mathbf{\Gamma_f T_g})_{j,k} = ((\mathbf{S}^*)^j \mathbf{f})^T (\mathbf{S}^k \mathbf{g}) = \mathbf{f}^T (\mathbf{S}^{j+k} \mathbf{g}),$$

thus the entries of $\mathbf{\Gamma_f T_g}$ only depend on the sum of its indices. Therefore, $\mathbf{\Gamma_f T_g}$ has again Hankel structure. For $p = 1$, the obtained Hankel matrix is generated by $\mathbf{\Gamma_f g} \in \ell^1$, and for $p = 2$ we get $\mathbf{\Gamma_f g} \in \ell^2$ by Young's inequality.

3. Similarly, since the j-th row of $\mathbf{T_g}$ is $\mathbf{S}^j \mathbf{g}$ and the k-th column of $\mathbf{\Gamma_f}$ is $(\mathbf{S}^*)^k \mathbf{f}$ we obtain

$$(\mathbf{T_g}^T \mathbf{\Gamma_f})_{j,k} = (\mathbf{S}^j \mathbf{g})^T ((\mathbf{S}^*)^k \mathbf{f}) = (\mathbf{S}^{j+k} \mathbf{g})^T \mathbf{f} = \mathbf{f}^T (\mathbf{S}^{j+k} \mathbf{g}) = (\mathbf{\Gamma_f T_g})_{j,k}.$$

□

Lemma 4.5. *For some $K \in \mathbb{N}_0$ let $\mathbf{b} = (b_k)_{k=0}^\infty$ be given by the Blaschke product*

$$B(e^{i\omega}) = \sum_{k=0}^{\infty} b_k e^{i\omega k} := \begin{cases} \prod_{j=1}^K \frac{e^{i\omega} - \alpha_j}{1 - \overline{\alpha}_j e^{i\omega}} & K > 0, \\ 1 & K = 0, \end{cases} \tag{4.2}$$

where $\alpha_1, \ldots, \alpha_K \in \mathbb{D}$. Then $\mathbf{b} \in \ell^1$ and the infinite triangular Toeplitz matrix $\mathbf{T_b}$ generated by $\mathbf{b}$ satisfies the following properties.

(1) $\mathbf{T_b^* T_b} = \mathbf{I}$, *i.e.* $\mathbf{T_b^*}$ *is the left inverse of* $\mathbf{T_b}$.

(2) *The operator $\mathbf{T_b} : \ell^p \to \ell^p$ has the norm $\|\mathbf{T_b}\|_{\ell^p \to \ell^p} = 1$ for $p \in \{1, 2\}$.*

(3) *Let $\mathbf{\Gamma_f}$ be an infinite Hankel matrix being generated by $\mathbf{f} \in \ell^1$. Let $\sigma_n(\mathbf{\Gamma_f})$ and $\sigma_n(\mathbf{\Gamma_f T_b})$ be the n-th singular values of $\mathbf{\Gamma_f}$ and $\mathbf{\Gamma_f T_b}$ being ordered decreasingly as $\sigma_0(\mathbf{\Gamma_f}) \geq \sigma_1(\mathbf{\Gamma_f}) \geq \ldots$ and $\sigma_0(\mathbf{\Gamma_f T_b}) \geq \sigma_1(\mathbf{\Gamma_f T_b}) \geq \ldots$. Then, for all $n \in \mathbb{N}_0$, we have*

$$\sigma_n(\mathbf{\Gamma_f T_b}) \leq \sigma_n(\mathbf{\Gamma_f}).$$

Proof. 1. Obviously, $\mathbf{T_b^* T_b}$ is hermitian. For the (l, k)-th entry of $\mathbf{T_b^* T_b}$ we obtain for $l \geq k$

$$(\mathbf{T_b^* T_b})_{l,k} = \sum_{j=l}^{\infty} \overline{b}_{j-l} b_{j-k} = \sum_{j=0}^{\infty} \overline{b}_j b_{j+(l-k)}.$$

The coefficients b_k are the Fourier coefficients of $B(e^{i\omega})$,

$$b_k = \frac{1}{2\pi} \int_0^{2\pi} B(e^{i\omega})\, e^{-i\omega k} d\omega, \qquad k = 0, 1, 2, \ldots.$$

Thus,

$$\sum_{j=0}^{\infty} \overline{b}_j b_{j+(l-k)} = \frac{1}{2\pi} \sum_{j=0}^{\infty} \overline{b}_j \int_0^{2\pi} B(e^{i\omega}) e^{-i\omega(j+l-k)} d\omega$$

$$= \frac{1}{2\pi}\int_0^{2\pi} B(e^{i\omega})e^{-i\omega(l-k)}\sum_{j=0}^{\infty}\overline{b}_j e^{-i\omega j}d\omega$$
$$= \frac{1}{2\pi}\int_0^{2\pi} |B(e^{i\omega})|^2 e^{-i\omega(l-k)}d\omega = \delta_{l,k}$$

since
$$|B(e^{i\omega})|^2 = \prod_{j=1}^{K}\left(\frac{e^{i\omega}-\alpha_j}{1-\overline{\alpha_j}e^{i\omega}}\right)\left(\frac{e^{-i\omega}-\overline{\alpha_j}}{1-\alpha_j e^{-i\omega}}\right) = 1.$$

2. Now the second assertion follows directly from the first.

3. Using the definition of the singular value and the properties of $\mathbf{T_b}$ we obtain for $p \in \{1,2\}$

$$\begin{aligned}
\sigma_n(\mathbf{\Gamma_f}) &= \min\{\|\mathbf{\Gamma_f}-\mathbf{R}\| \ : \ \mathbf{R}\in\mathcal{L}(\ell^p),\ \mathrm{rank}(\mathbf{R})\le n\}\\
&= \min\{\|\mathbf{\Gamma_f}-\mathbf{R}\|\|\mathbf{T_b}\| \ : \ \mathbf{R}\in\mathcal{L}(\ell^p),\ \mathrm{rank}(\mathbf{R})\le n\}\\
&\ge \min\{\|(\mathbf{\Gamma_f}-\mathbf{R})\mathbf{T_b}\| \ : \ \mathbf{R}\in\mathcal{L}(\ell^p),\ \mathrm{rank}(\mathbf{R})\le n\}\\
&= \min\{\|\mathbf{\Gamma_f}\mathbf{T_b}-\mathbf{R}\mathbf{T_b}\| \ : \ \mathbf{R}\in\mathcal{L}(\ell^p),\ \mathrm{rank}(\mathbf{R})\le n\}\\
&= \min\{\|\mathbf{\Gamma_f}\mathbf{T_b}-\tilde{\mathbf{R}}\| \ : \ \tilde{\mathbf{R}}\in\mathcal{L}(\ell^p),\ \mathrm{rank}(\tilde{\mathbf{R}})\le n\}\\
&= \sigma_n(\mathbf{\Gamma_f}\mathbf{T_b}),
\end{aligned}$$

since $\mathrm{rank}(\mathbf{RT_b})$ is still at most n. Here $\mathcal{L}(\ell^p)$ denotes the set of all linear operators from ℓ^p to ℓ^p. □

Construction of infinite Hankel matrices with special properties. Next, we will construct an infinite Hankel matrix with operator norm 1 that possesses a predetermined con-eigenvector $\mathbf{v}\in\ell^1$ to the con-eigenvalue 1. For that purpose, we first need to understand the image of an infinite Hankel matrix.

Lemma 4.6. *For given sequences* $\mathbf{f}\in\ell^1$ *and* $\mathbf{v}\in\ell^p$, $p\in\{1,2\}$ *with corresponding Fourier series* $P_{\mathbf{f}}(e^{i\omega})$ *and* $P_{\mathbf{v}}(e^{i\omega})$ *the vector* $\mathbf{w}=(w_k)_{k=0}^{\infty}$ *obtained by*

$$\mathbf{w} = \mathbf{\Gamma_f}\mathbf{v}$$

satisfies
$$w_k = \frac{1}{2\pi}\int_0^{2\pi} P_{\mathbf{f}}(\mathrm{e}^{\mathrm{i}t})P_{\mathbf{v}}(\mathrm{e}^{-\mathrm{i}t})\mathrm{e}^{-\mathrm{i}tk}\mathrm{d}t, \qquad k\in\mathbb{N}_0.$$

Proof. Let $P_{\mathbf{w}}(e^{i\omega}) := \sum_{k=0}^{\infty} w_k e^{i\omega k}$. Then, on the one hand, we find

$$P_{\mathbf{w}}(e^{i\omega}) = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} f_{k+j}v_j e^{i\omega k} = \sum_{j=0}^{\infty}\sum_{k=j}^{\infty} f_k v_j e^{i\omega(k-j)}.$$

On the other hand,

$$P_{\mathbf{f}}(e^{i\omega})P_{\mathbf{v}}(e^{-i\omega}) = \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} f_k v_j e^{i\omega(k-j)} = P_{\mathbf{w}}(e^{i\omega}) + \sum_{j=0}^{\infty}\sum_{k=0}^{j-1} f_k v_j e^{i\omega(k-j)},$$

where in the second sum occur only negative powers of $e^{i\omega}$. Hence, $P_{\mathbf{w}}(\mathrm{e}^{\mathrm{i}t})$ possesses the Fourier coefficients

$$w_k = \frac{1}{2\pi}\int_0^{2\pi} P_{\mathbf{f}}(\mathrm{e}^{\mathrm{i}t})P_{\mathbf{v}}(\mathrm{e}^{-\mathrm{i}t})\mathrm{e}^{-\mathrm{i}tk}\mathrm{d}t$$

for $k \in \mathbb{N}_0$. □

Now we consider the construction of a special infinite Hankel matrix with operator norm 1.

Lemma 4.7. *Let* $\mathbf{v} \in \ell^1$ *be given with the corresponding Fourier series* $P_{\mathbf{v}}(e^{i\omega})$. *Assume that* $P_{\mathbf{v}}(e^{i\omega}) \neq 0$ *for all* $\omega \in [0, 2\pi)$. *Further, let* $\mathbf{w} = (w_k)_{k=0}^{\infty}$ *be given by*

$$w_k := \frac{1}{2\pi} \int_0^{2\pi} \frac{P_{\mathbf{v}}(\mathrm{e}^{\mathrm{i}t})}{P_{\overline{\mathbf{v}}}(\mathrm{e}^{-\mathrm{i}t})} \mathrm{e}^{-\mathrm{i}tk} \mathrm{d}t, \qquad k \in \mathbb{N}_0.$$

Then $\mathbf{w} \in \ell^2$, *and* $\|\mathbf{w}\|_2 = 1$. *Further, the Hankel operator* $\mathbf{\Gamma_w}$ *satisfies* $\mathbf{\Gamma_w} \overline{\mathbf{v}} = \mathbf{v}$ *and*

$$\|\mathbf{\Gamma_w}\|_{\ell^2 \to \ell^2} := \sup_{\mathbf{u} \in \ell^2 \setminus \{\mathbf{0}\}} \frac{\|\mathbf{\Gamma_w u}\|_2}{\|\mathbf{u}\|_2} = \frac{\|\mathbf{\Gamma_w} \overline{\mathbf{v}}\|_2}{\|\mathbf{v}\|_2} = 1.$$

Proof. First we observe that by Parseval's identity

$$\|\mathbf{w}\|_2^2 = \sum_{k=0}^{\infty} |w_k|^2 \leq \left\| \frac{P_{\mathbf{v}}(\mathrm{e}^{\mathrm{i}\cdot})}{P_{\overline{\mathbf{v}}}(\mathrm{e}^{-\mathrm{i}\cdot})} \right\|_{L^2([0,2\pi))}^2 = \frac{1}{2\pi} \int_0^{2\pi} \left| \frac{P_{\mathbf{v}}(\mathrm{e}^{\mathrm{i}t})}{P_{\overline{\mathbf{v}}}(\mathrm{e}^{-\mathrm{i}t})} \right|^2 \mathrm{d}t = 1$$

and thus $\mathbf{w} \in \ell^2$. Further, we obtain

$$\begin{aligned} (\mathbf{\Gamma_w} \overline{\mathbf{v}})_k &= \sum_{j=0}^{\infty} w_{k+j} \bar{v}_j = \sum_{j=0}^{\infty} \frac{1}{2\pi} \int_0^{2\pi} \frac{P_{\mathbf{v}}(\mathrm{e}^{\mathrm{i}t})}{P_{\overline{\mathbf{v}}}(\mathrm{e}^{-\mathrm{i}t})} \mathrm{e}^{-\mathrm{i}t(k+j)} \bar{v}_j \mathrm{d}t \\ &= \frac{1}{2\pi} \int_0^{2\pi} \frac{P_{\mathbf{v}}(\mathrm{e}^{\mathrm{i}t})}{P_{\overline{\mathbf{v}}}(\mathrm{e}^{-\mathrm{i}t})} \mathrm{e}^{-\mathrm{i}tk} \sum_{j=0}^{\infty} \bar{v}_j \mathrm{e}^{-\mathrm{i}tj} \mathrm{d}t = \frac{1}{2\pi} \int_0^{2\pi} P_{\mathbf{v}}(\mathrm{e}^{\mathrm{i}t}) \mathrm{e}^{-\mathrm{i}tk} \mathrm{d}t = v_k. \end{aligned}$$

for all $k \in \mathbb{N}_0$ and thus $\mathbf{\Gamma_w} \overline{\mathbf{v}} = \mathbf{v}$. The norm of $\mathbf{\Gamma_w}$ is indeed equal to 1 since for arbitrary $\mathbf{u} \in \ell^2$ it follows by Lemma 4.6 and Parseval's identity

$$\begin{aligned} \|\mathbf{\Gamma_w} \overline{\mathbf{u}}\|_2^2 &= \sum_{k=0}^{\infty} \left| \frac{1}{2\pi} \int_0^{2\pi} P_{\mathbf{w}}(\mathrm{e}^{\mathrm{i}t}) P_{\overline{\mathbf{u}}}(\mathrm{e}^{-\mathrm{i}t})\, \mathrm{e}^{-\mathrm{i}kt} \mathrm{d}t \right|^2 \\ &\leq \sum_{k=-\infty}^{\infty} \left| \frac{1}{2\pi} \int_0^{2\pi} \frac{P_{\mathbf{v}}(\mathrm{e}^{\mathrm{i}t})}{P_{\overline{\mathbf{v}}}(\mathrm{e}^{-\mathrm{i}t})} P_{\overline{\mathbf{u}}}(\mathrm{e}^{-\mathrm{i}t})\, \mathrm{e}^{-\mathrm{i}kt} \mathrm{d}t \right|^2 \\ &= \frac{1}{2\pi} \int_0^{2\pi} \left| \frac{P_{\mathbf{v}}(\mathrm{e}^{\mathrm{i}t})}{P_{\overline{\mathbf{v}}}(\mathrm{e}^{-\mathrm{i}t})} \right|^2 |P_{\overline{\mathbf{u}}}(\mathrm{e}^{-\mathrm{i}t})|^2 \, \mathrm{d}t \\ &\leq \sup_{t \in [0,2\pi)} \left| \frac{P_{\mathbf{v}}(\mathrm{e}^{\mathrm{i}t})}{P_{\overline{\mathbf{v}}}(\mathrm{e}^{-\mathrm{i}t})} \right|^2 \frac{1}{2\pi} \int_0^{2\pi} |P_{\overline{\mathbf{u}}}(\mathrm{e}^{-\mathrm{i}t})|^2 \, \mathrm{d}t = \sum_{k=-\infty}^{\infty} |\bar{u}_k|^2 = \|\mathbf{u}\|_2^2, \end{aligned}$$

and thus the assertion holds. □

This result immediately implies also $\|\mathbf{\Gamma_w u}\|_2 \leq \|\mathbf{w}\|_2 \|\mathbf{u}\|_1 = \|\mathbf{u}\|_1$ for all $\mathbf{u} \in \ell^1$ by Young's inequality.

4.2 Proof of the AAK-Theorem for Hankel matrices with finite rank

Let us come back to the Hankel matrix $\mathbf{\Gamma_f}$ of rank N generated by the the sequence $\mathbf{f}$ of the form (1.1) with $1 > |z_1| \geq \cdots \geq |z_N| > 0$, and with con-eigenvectors $\mathbf{v}^{(l)}$, $l = 0, \ldots, N-1$ corresponding to the nonzero con-eigenvalues (resp. singular values)

$\sigma_0 \geq \sigma_1 \ldots \geq \sigma_{N-1} > 0$. As shown in (3.4), the Laurent polynomial corresponding to $\mathbf{v}^{(l)}$ has the form

$$P_{\mathbf{v}^{(l)}}(z) = \sum_{j=0}^{\infty} v_j^{(l)} z^j = \sum_{j=1}^{N} \frac{a_j P_{\overline{\mathbf{v}}^{(l)}}(z_j)}{1 - z_j z} = \frac{q^{(l)}(z)}{z^N P(z^{-1})}$$

with $q^{(l)}(z)$ being a polynomial of degree $N-1$ and with the Prony polynomial $P(z)$ in (2.2).

We want to show now that for each single nonzero singular value σ_K of $\mathbf{\Gamma_f}$ the Laurent series of the corresponding con-eigenvector $\mathbf{v}^{(K)}$ possesses exactly K zeros in $\mathbb{D}$, and moreover, that these zeros $z_1^{(K)}, \ldots, z_K^{(K)}$ can be used to construct a new Hankel matrix $\mathbf{\Gamma}_{\tilde{\mathbf{f}}}$ of rank K with $\tilde{\mathbf{f}}$ of the form (3.1) and $\|\mathbf{\Gamma_f} - \mathbf{\Gamma}_{\tilde{\mathbf{f}}}\| = \sigma_K$.

The above relation implies that the zeros of $P_{\mathbf{v}^{(K)}}(z)$ are the $N-1$ zeros of $q^{(K)}(z)$. Let n_K denote the number of zeros of $q^{(K)}$ in $\mathbb{D}$, where $0 \leq n_K \leq N-1$. We show first that $n_K \leq K$.

We can write

$$q^{(K)}(z) = \prod_{j=1}^{n_K} (z - \alpha_j) \prod_{j=n_K+1}^{N-1} (z - \beta_j)$$

with $|\alpha_j| < 1$ and $|\beta_j| \geq 1$. Now, let

$$P_{\mathbf{u}^{(K)}}(z) := \frac{1}{\sigma_K} \frac{\prod\limits_{j=1}^{n_K} (1 - \overline{\alpha}_j z) \prod\limits_{j=n_K+1}^{N-1} (z - \beta_j)}{z^N p(z^{-1})}. \tag{4.3}$$

Then $P_{\mathbf{u}^{(K)}}(z)$ has no zeros in $\mathbb{D}$ and defines a sequence $\mathbf{u}^{(K)} = (u_r^{(l)})_{r=0}^{\infty} \in \ell^1$ via

$$P_{\mathbf{u}^{(K)}}(z) = \sum_{r=0}^{\infty} u_r^{(K)} z^r, \quad z \in \mathbb{D}.$$

Denoting by

$$B^{(K)}(e^{i\omega}) := \prod_{j=1}^{n_K} \frac{e^{i\omega} - \alpha_j}{1 - \overline{\alpha}_j e^{i\omega}} = \sum_{k=0}^{\infty} b_k^{(K)} e^{i\omega k} \tag{4.4}$$

the Blaschke product on the unit circle, it follows that

$$P_{\mathbf{v}^{(K)}}(e^{i\omega}) = B^{(K)}(e^{i\omega}) \, P_{\mathbf{u}^{(K)}}(e^{i\omega}),$$

or equivalently,

$$\mathbf{v}^{(K)} = \mathbf{T}_{\mathbf{b}^{(K)}} \mathbf{u}^{(K)}, \tag{4.5}$$

where $\mathbf{T}_{\mathbf{b}^{(K)}}$ denotes the triangular Toeplitz matrix corresponding to $B^{(K)}(e^{i\omega})$. Now we can show

Theorem 4.8. *Let $\mathbf{\Gamma_f}$ be the infinite Hankel matrix of finite rank N generated by $\mathbf{f} = (f_k)_{k=0}^{\infty}$ of the form (1.1) and with nonzero singular values $\sigma_0 \geq \sigma_1 \geq \ldots \geq \sigma_{N-1} > 0$. Further, let $(\sigma_K, \mathbf{v}^{(K)})$ be the K-th con-eigenpair of $\mathbf{\Gamma_f}$ with $\sigma_K \neq \sigma_k$ for $K \neq k$. Let $\mathbf{T}_{\mathbf{b}^{(K)}}$ be the Toeplitz matrix corresponding to $\mathbf{v}^{(K)}$ being defined by $B^{(K)}(e^{i\omega})$ as in (4.4). Then $\mathbf{\Gamma_f} \mathbf{T}_{\mathbf{b}^{(K)}}$ possesses the singular value σ_K with multiplicity at least $n_K + 1$, where n_K denotes the number of zeros of $P_{\mathbf{v}^{(K)}}$ in $\mathbb{D}$. In particular, we have $n_K \leq K$.*

Proof. 1. Considering the Blaschke product in (4.4), we define its partial products by

$$B_j^{(K)}(e^{i\omega}) := \sum_{r=0}^{\infty} (b_j^{(K)})_r e^{i\omega r} = \prod_{k=1}^{j} \frac{e^{i\omega} - \alpha_k}{1 - \overline{\alpha}_k e^{i\omega}}, \qquad j = 1, \ldots, n_K,$$

where $\alpha_1, \ldots, \alpha_{n_K}$ are the zeros of $P_{\mathbf{v}^{(K)}}(z)$ inside $\mathbb{D}$. We employ the notation $\mathbf{T}_{\mathbf{b}^{(K)}/\mathbf{b}_j^{(K)}}$ for the triangular Toeplitz matrix generated by the sequence of Fourier coefficients of $\prod_{k=j+1}^{n_K} \frac{e^{i\omega}-\alpha_k}{1-\overline{\alpha}_k e^{i\omega}}$ such that

$$\mathbf{T}_{\mathbf{b}^{(K)}} = \mathbf{T}_{\mathbf{b}^{(K)}/\mathbf{b}_j^{(K)}} \cdot \mathbf{T}_{\mathbf{b}_j^{(K)}}.$$

We show now that the $n_K + 1$ vectors $\mathbf{v}^{(K)}$, $\mathbf{T}^*_{\mathbf{b}_1^{(K)}} \mathbf{v}^{(K)}, \ldots, \mathbf{T}^*_{\mathbf{b}_{n_K}^{(K)}} \mathbf{v}^{(K)}$ are linearly independent singular vectors of $\mathbf{\Gamma_f T_{b^{(K)}}}$ to the singular value σ_K. For $j = 0, \ldots, n_K$ (with $\mathbf{T}_{\mathbf{b}_0^{(K)}} := \mathbf{I}$) we obtain by Lemma 4.4 and Lemma 4.5

$$\begin{aligned}
& (\mathbf{\Gamma_f T}_{\mathbf{b}^{(K)}})^* \mathbf{\Gamma_f T}_{\mathbf{b}^{(K)}} \mathbf{T}^*_{\mathbf{b}_j^{(K)}} \mathbf{v}^{(K)} \\
= \; & \mathbf{T}^*_{\mathbf{b}^{(K)}} \mathbf{\Gamma_f^* \Gamma_f T}_{\mathbf{b}^{(K)}/\mathbf{b}_j^{(K)}} \mathbf{T}_{\mathbf{b}_j^{(K)}} \mathbf{T}^*_{\mathbf{b}_j^{(K)}} \mathbf{T}_{\mathbf{b}_j^{(K)}} \mathbf{T}_{\mathbf{b}^{(K)}/\mathbf{b}_j^{(K)}} \mathbf{u}^{(K)} \\
= \; & \mathbf{T}^*_{\mathbf{b}^{(K)}} \mathbf{\Gamma_f^* T}^T_{\mathbf{b}^{(K)}/\mathbf{b}_j^{(K)}} \mathbf{\Gamma_f T}_{\mathbf{b}_j^{(K)}} \mathbf{T}_{\mathbf{b}^{(K)}/\mathbf{b}_j^{(K)}} \mathbf{u}^{(K)} \\
= \; & \mathbf{T}^*_{\mathbf{b}^{(K)}} \mathbf{\Gamma_f^* T}^T_{\mathbf{b}^{(K)}/\mathbf{b}_j^{(K)}} \mathbf{\Gamma_f} \mathbf{v}^{(K)} \\
= \; & \sigma_K \mathbf{T}^*_{\mathbf{b}^{(K)}} \mathbf{\Gamma_f^* T}^T_{\mathbf{b}^{(K)}/\mathbf{b}_j^{(K)}} \overline{\mathbf{v}}^{(K)} \\
= \; & \sigma_K \mathbf{T}^*_{\mathbf{b}^{(K)}} \mathbf{\Gamma_f^*} \overline{\mathbf{T}^*_{\mathbf{b}^{(K)}/\mathbf{b}_j^{(K)}} \mathbf{T}_{\mathbf{b}^{(K)}/\mathbf{b}_j^{(K)}} \mathbf{T}_{\mathbf{b}_j^{(K)}}} \overline{\mathbf{u}}^{(K)} \\
= \; & \sigma_K \mathbf{T}^*_{\mathbf{b}^{(K)}} \mathbf{\Gamma_f^*} \overline{\mathbf{T}_{\mathbf{b}_j^{(K)}}} \overline{\mathbf{u}}^{(K)} \\
= \; & \sigma_K \mathbf{\Gamma_f^*} \overline{\mathbf{T}_{\mathbf{b}_j^{(K)}} \mathbf{T}_{\mathbf{b}^{(K)}/\mathbf{b}_j^{(K)}} \mathbf{T}_{\mathbf{b}_j^{(K)}}} \overline{\mathbf{u}}^{(K)} \\
= \; & \sigma_K \mathbf{T}^*_{\mathbf{b}_j^{(K)}} \mathbf{\Gamma_f^*} \overline{\mathbf{v}}^{(K)} = \sigma_K^2 \mathbf{T}^*_{\mathbf{b}_j^{(K)}} \mathbf{v}^{(K)}.
\end{aligned}$$

Moreover, the vectors $\mathbf{v}^{(K)}$, $\mathbf{T}^*_{\mathbf{b}_1^{(K)}} \mathbf{v}^{(K)}, \ldots, \mathbf{T}^*_{\mathbf{b}_{n_K}^{(K)}} \mathbf{v}^{(K)}$ are linearly independent, since

$$\sum_{j=0}^{n_K} \gamma_j \mathbf{T}^*_{\mathbf{b}_j^{(K)}} \mathbf{v}^{(K)} = \mathbf{0}$$

is by Lemma 4.4(3) equivalent with

$$\mathbf{\Gamma}_{\mathbf{v}^{(K)}} \left(\sum_{j=0}^{n_K} \gamma_j \mathbf{b}_j^{(K)} \right) = \mathbf{0},$$

i.e., $\left(\sum_{j=0}^{n_K} \gamma_j \mathbf{b}_j^{(K)}\right)$ is a zero-(con)-eigenvector of $\mathbf{\Gamma}_{\mathbf{v}^{(K)}}$. Thus, by Theorem 4.2 and (3.2), the Laurent polynomial

$$\sum_{r=0}^{\infty} \sum_{j=0}^{n_K} \gamma_j (\mathbf{b}_j^{(K)})_r z^r = \gamma_0 + \sum_{j=1}^{n_K} \gamma_j \prod_{k=1}^{j} \frac{z - \alpha_k}{1 - \overline{\alpha}_k z}$$

possesses all zeros $z_1, \ldots, z_N$. Since $n_K \leq N-1$, we conclude that $\gamma_0 = \ldots = \gamma_{n_K} = 0$. Therefore, $\mathbf{\Gamma_f T_{b^{(K)}}}$ possesses the singular value σ_K with multiplicity at least $n_K + 1$. On the other hand, since $\sigma_K(\mathbf{\Gamma_f T_{b^{(K)}}}) \leq \sigma_K(\mathbf{\Gamma_f})$ by Lemma 4.5, it follows that $n_K \leq K$. □

In the next step, we construct a sequence $\tilde{\mathbf{f}}^{(K)} = \mathbf{f} - \mathbf{g}^{(K)}$ such that $\mathrm{rank}(\mathbf{\Gamma}_{\mathbf{f}-\mathbf{g}^{(K)}}) = K$ and $\|\mathbf{\Gamma}_{\mathbf{g}^{(K)}}\| = \|\mathbf{\Gamma}_{\mathbf{f}-\tilde{\mathbf{f}}^{(K)}}\| = \sigma_K$.

Let $\mathbf{\Gamma}_{\mathbf{g}^{(K)}}$ be the Hankel matrix generated by $\mathbf{g}^{(K)} = (g_l^{(K)})_{l=0}^{\infty}$ with

$$g_l^{(K)} := \frac{\sigma_K}{2\pi} \int_0^{2\pi} \frac{P_{\mathbf{v}^{(K)}}(\mathrm{e}^{\mathrm{i}t})}{P_{\overline{\mathbf{v}}^{(K)}}(\mathrm{e}^{-\mathrm{i}t})} \mathrm{e}^{-\mathrm{i}tl} \mathrm{d}t \tag{4.6}$$

for $l \in \mathbb{N}_0$. Then by Lemma 4.7 it follows that $\|\mathbf{\Gamma}_{\mathbf{g}^{(K)}}\|_{\ell^2 \to \ell^2} = \sigma_K$ and $\mathbf{\Gamma}_{\mathbf{g}^{(K)}} \overline{\mathbf{v}}^{(K)} = \sigma_K \mathbf{v}^{(K)}$.

Now we can show

Theorem 4.9. *Let $\mathbf{\Gamma_f}$ be a Hankel operator of finite rank N generated by $\mathbf{f} = (f_k)_{k=0}^{\infty}$ with f_k of the form (1.1) with the nonzero con-eigenvalues $\sigma_0 \geq \sigma_1 \geq \ldots \geq \sigma_{N-1} > 0$. Further, let $(\sigma_K, \mathbf{v}^{(K)})$ be the K-th con-eigenpair of $\mathbf{\Gamma_f}$ with $\sigma_K \neq \sigma_k$ for $K \neq k$. Then the shift-invariant space*

$$\mathcal{S}_{\overline{\mathbf{v}}^{(K)}} := \mathrm{clos}_{\ell^2} \mathrm{span}\, \{S^l \overline{\mathbf{v}}^{(K)} : l \in \mathbb{N}_0\}$$

has at least co-dimension K in ℓ_2, and the matrix $\mathbf{\Gamma}_{\mathbf{f}-\mathbf{g}^{(K)}}$ with $\mathbf{g}^{(K)}$ determined by (4.6) has at least rank K. Moreover, for the operator norm of $\mathbf{\Gamma}_{\mathbf{g}^{(K)}}$ we have

$$\|\mathbf{\Gamma}_{\mathbf{g}^{(K)}}\|_{\ell^2 \to \ell^2} = \|\mathbf{\Gamma_f} - \mathbf{\Gamma}_{\mathbf{f}-\mathbf{g}^{(K)}}\|_{\ell^2 \to \ell^2} = \sigma_K.$$

Proof. 1. Similarly as in the proof of Lemma 4.7 we observe that

$$\begin{aligned}
(\mathbf{\Gamma}_{\mathbf{g}^{(K)}} \overline{\mathbf{v}}^{(K)})_l &= \sum_{j=0}^{\infty} g_{l+j} \bar{v}_j^{(K)} = \sum_{j=0}^{\infty} \frac{\sigma_K}{2\pi} \int_0^{2\pi} \frac{P_{\mathbf{v}^{(K)}}(\mathrm{e}^{\mathrm{i}t})}{P_{\overline{\mathbf{v}}^{(K)}}(\mathrm{e}^{-\mathrm{i}t})} \mathrm{e}^{-\mathrm{i}t(l+j)} \bar{v}_j^{(K)} \mathrm{d}t \\
&= \frac{\sigma_K}{2\pi} \int_0^{2\pi} \frac{P_{\mathbf{v}^{(K)}}(\mathrm{e}^{\mathrm{i}t})}{P_{\overline{\mathbf{v}}^{(K)}}(\mathrm{e}^{-\mathrm{i}t})} \mathrm{e}^{-\mathrm{i}tl} \sum_{j=0}^{\infty} \bar{v}_j^{(K)} \mathrm{e}^{-\mathrm{i}tj} \mathrm{d}t \\
&= \frac{\sigma_K}{2\pi} \int_0^{2\pi} P_{\mathbf{v}}^{(K)}(\mathrm{e}^{\mathrm{i}t}) \mathrm{e}^{-\mathrm{i}tl} \mathrm{d}t = \sigma_K v_l^{(K)},
\end{aligned}$$

for all $k \in \mathbb{N}_0$ and thus $\mathbf{\Gamma}_{\mathbf{g}^{(K)}} \overline{\mathbf{v}}^{(K)} = \sigma \mathbf{v}^{(K)}$, resp. $\mathbf{\Gamma}_{\mathbf{f}-\mathbf{g}^{(K)}} \overline{\mathbf{v}}^{(K)} = \mathbf{0}$. Moreover, by Lemma 4.7 it follows that $\|\mathbf{\Gamma}_{\mathbf{g}^{(K)}}\|_{\ell^2 \to \ell^2} = \sigma_K$.
We consider now the operator $\mathbf{\Gamma}_{\mathbf{f}-\mathbf{g}^{(K)}}$. By Lemma 4.1, the shift-invariant space $\mathcal{S}_{\overline{\mathbf{v}}^{(K)}}$ is a subset of $\ker \mathbf{\Gamma}_{\mathbf{f}-\mathbf{g}^{(K)}}$. On the other hand, we observe that for $r = 0, \ldots, K-1$,

$$\begin{aligned}
\|\mathbf{\Gamma}_{\mathbf{f}-\mathbf{g}^{(K)}} \overline{\mathbf{v}}^{(r)}\|_2 &= \|\mathbf{\Gamma_f} \overline{\mathbf{v}}^{(r)} - \mathbf{\Gamma}_{\mathbf{g}^{(K)}} \overline{\mathbf{v}}^{(r)}\|_2 \\
&\geq |\|\mathbf{\Gamma_f} \overline{\mathbf{v}}^{(r)}\|_2 - \|\mathbf{\Gamma}_{\mathbf{g}^{(K)}} \overline{\mathbf{v}}^{(r)}\|_2| \geq (\sigma_r - \sigma_k) \|\mathbf{v}^{(r)}\|_2 > 0.
\end{aligned}$$

Thus, the K linearly independent con-eigenvectors $\overline{\mathbf{v}}^{(0)}, \ldots, \overline{\mathbf{v}}^{(K-1)}$ to the larger con-eigenvalues $\sigma_0 \geq \ldots \geq \sigma_{k-1}$ are not contained in the kernel of $\mathbf{\Gamma}_{\mathbf{f}-\mathbf{g}^{(K)}}$ and thus not in $\mathcal{S}_{\overline{\mathbf{v}}^{(K)}}$. Hence, codim $\mathcal{S}_{\overline{\mathbf{v}}^{(K)}} \geq K$, and $\mathbf{\Gamma}_{\mathbf{f}-\mathbf{g}^{(K)}}$ possesses at least rank K. □

Since $P_{\overline{\mathbf{v}}^{(K)}}$ has by construction no zeros on the unit circle, it follows that $\mathbf{g}^{(K)}$ is also in ℓ^1 and therefore also $\|\mathbf{\Gamma}_{\mathbf{g}^{(K)}}\|_{\ell^1 \to \ell^1} = \sigma_K$.

Finally, we conclude the following theorem.

Theorem 4.10. *Let $\mathbf{\Gamma_f}$ be the Hankel operator of finite rank N generated by $\mathbf{f} = (f_k)_{k=0}^{\infty}$ of the form (1.1) and with nonzero singular values $\sigma_0 \geq \sigma_1 \geq \ldots \geq \sigma_{N-1} > 0$. Further, let $(\sigma_K, \mathbf{v}^{(K)})$ be the K-th con-eigenpair of $\mathbf{\Gamma_f}$. Then for each $K \in \{0, \ldots, N-1\}$ with σ_K being a single singular value we have:*

(1) *The Laurent polynomial $P_{\mathbf{v}^{(K)}}(z)$ corresponding to the con-eigenvector $\mathbf{v}^{(K)}$ has exactly K zeros $z_1^{(K)}, \ldots, z_K^{(K)}$ in $\mathbb{D}$ repeated according to multiplicity.*

(2) *Considering the Hankel matrix $\mathbf{\Gamma}_{\mathbf{g}^{(K)}}$ given by the sequence $\mathbf{g}^{(K)} = (g_k)_{k=0}^{\infty}$ in (4.6), it follows that $\mathbf{\Gamma}_{\mathbf{f}-\mathbf{g}^{(K)}}$ possesses the rank K and*

$$\|\mathbf{\Gamma}_{\mathbf{g}^{(K)}}\|_{\ell^p \to \ell^p} = \|\mathbf{\Gamma_f} - \mathbf{\Gamma}_{\mathbf{f}-\mathbf{g}^{(K)}}\|_{\ell^p \to \ell^p} = \sigma_K, \qquad p \in \{1, 2\}.$$

(3) *The kernel of $\mathbf{\Gamma}_{\mathbf{f}-\mathbf{g}^{(K)}}$ has co-dimension K. If the zeros $z_1^{(K)}, \ldots, z_K^{(K)}$ are pairwise different, then it satisfies*

$$\ker(\mathbf{\Gamma}_{\mathbf{f}-\mathbf{g}^{(K)}}) \;=\; \mathcal{S}_{\overline{\mathbf{v}}^{(K)}} = (\text{clos}_{\ell^2}\text{span}\{((z_1^{(K)})^l)_{l=0}^{\infty}, \ldots, ((z_K^{(K)})^l)_{l=0}^{\infty}\})^{\perp},$$

where $\mathcal{S}_{\overline{\mathbf{v}}^{(K)}} := \text{clos}_{\ell^2}\text{span}\,\{S^l \overline{\mathbf{v}}^{(K)} : l \in \mathbb{N}_0\}$.

Proof. 1. First we show that $\mathcal{S}_{\overline{\mathbf{v}}^{(K)}} = (\text{clos}_{\ell^2}\text{span}\{((z_1^{(K)})^l)_{l=0}^{\infty}, \ldots, ((z_{n_K}^{(K)})^l)_{l=0}^{\infty}\})^{\perp}$, where $z_1^{(K)}, \ldots, z_{n_K}^{(K)}$ are all pairwise different zeros of $P_{\mathbf{v}^{(K)}}(z)$ inside $\mathbb{D}$. Indeed for all $l \in \mathbb{N}_0$,

$$\begin{aligned} \langle ((z_j^{(K)})^r)_{r=0}^{\infty}, S^l \overline{\mathbf{v}}^{(K)} \rangle_{\ell^2} &= \langle (S^*)^l ((z_j^{(K)})^r)_{r=0}^{\infty}, \overline{\mathbf{v}}^{(K)} \rangle_{\ell^2} \\ &= \sum_{r=0}^{\infty} (z_j^{(K)})^{r+l} \, v_r^{(K)} \\ &= (z_j^{(K)})^l \sum_{r=0}^{\infty} (z_j^{(K)})^r \, v_r^{(K)} = (z_j^{(K)})^l \, P_{\mathbf{v}^{(K)}}(z_j^{(K)}) = 0. \end{aligned}$$

Thus,

$$\mathcal{S}_{\overline{\mathbf{v}}^{(K)}} \perp \text{span}\{((z_1^{(K)})^l)_{l=0}^{\infty}, \ldots, ((z_{n_K}^{(K)})^l)_{l=0}^{\infty}\}).$$

Assume now, that $\mathbf{u} \in \ell^2(\mathbb{N}_0)$ satisfies $\mathbf{u} \perp \text{span}\{((z_1^{(K)})^l)_{l=0}^{\infty}, \ldots, ((z_{n_K}^{(K)})^l)_{l=0}^{\infty}\})$, i.e., that $P_{\overline{\mathbf{u}}}(z_j^{(K)}) = 0$ for $j = 1, \ldots, n_K$. We show that $\mathbf{u} \in \mathcal{S}_{\overline{\mathbf{v}}^{(K)}}$. We can rewrite

$$P_{\overline{\mathbf{u}}}(e^{i\omega}) = \prod_{j=1}^{n_K} \frac{(e^{i\omega} - z_j^{(K)})}{(1 - \overline{z}_j^{(K)} e^{i\omega})} \, P_{\mathbf{w}}(e^{i\omega}) = B^{(K)}(e^{i\omega}) P_{\mathbf{w}}(e^{i\omega})$$

with the same Blaschke product as in (4.4), where $P_{\mathbf{w}}(e^{i\omega})$ still corresponds to a sequence $\mathbf{w} = (w_l)_{l=0}^{\infty} \in \ell^1(\mathbb{N}_0)$. Equivalently, we have $\overline{\mathbf{u}} = \mathbf{T}_{\mathbf{b}^{(K)}} \mathbf{w}$. Since $\mathbf{T}_{\overline{\mathbf{v}}^{(K)}}$ contains the columns $\overline{\mathbf{v}}^{(K)}$, $S\overline{\mathbf{v}}^{(K)}, \ldots$, the assertion $\mathbf{u} \in \mathcal{S}_{\overline{\mathbf{v}}^{(K)}}$ is equivalent to the assertion that there exists a sequence $\mathbf{y} \in \ell^2(\mathbb{N}_0)$, such that

$$\overline{\mathbf{u}} = \mathbf{T}_{\mathbf{v}^{(K)}} \mathbf{y}.$$

By Lemma 4.4 and (4.5) this is equivalent to

$$\mathbf{T}_{\mathbf{b}^{(K)}} \mathbf{w} = \mathbf{T}_{\mathbf{b}^{(K)}} \mathbf{T}_{\mathbf{u}^{(K)}} \, \mathbf{y},$$

and thus to

$$\mathbf{w} = \mathbf{T}_{\mathbf{b}^{(K)}}^* \mathbf{T}_{\mathbf{b}^{(K)}} \mathbf{w} = \mathbf{T}_{\mathbf{u}^{(K)}} \, \mathbf{y}.$$

The assertion now follows since $\mathbf{T}_{\mathbf{u}^{(K)}}$ is invertible. Indeed, (4.3) implies

$$P_{\mathbf{u}^{(K)}}(z) = \frac{1}{\sigma_K} \frac{\prod_{j=1}^{n_K} (1 - \overline{z}_j^{(K)} z) \prod_{j=n_K+1}^{N-1} (-\beta_j^{(K)})(1 - (\beta_j^{(K)})^{(-1)} z)}{\prod_{j=1}^{N} (1 - z_j z)},$$

and thus

$$\mathbf{T}_{\mathbf{u}^{(K)}}^{-1} = \frac{\sigma_K}{\prod_{j=n_K+1}^{N-1} (-\beta_j^{(K)})} \left(\prod_{j=1}^{n_K} \mathbf{T}_{\overline{z}_j^{(K)}} \right) \left(\prod_{j=n_K+1}^{N-1} \mathbf{T}_{(\beta_j^{(K)})^{-1}} \right) \mathbf{T}_{\tilde{\mathbf{p}}},$$

where $\mathbf{T}_{\overline{z}_j^{(K)}}$, $\mathbf{T}_{(\beta_j^{(K)})^{-1}}$ and $\mathbf{T}_{\tilde{\mathbf{p}}}$ are the infinite Toeplitz matrices generated by the sequences $((\overline{z}_j^{(K)})^r)_{r=0}^{\infty}$, $((\beta_j^{(K)})^{-r})_{r=0}^{\infty}$ and by the finite sequence $\tilde{\mathbf{p}} = (1, p_{N-1}, \ldots, p_0)$ containing the coefficients of the Prony polynomial in (2.2).

2. By Lemma 4.8 we have $n_K \leq K$, i.e., $\mathcal{S}_{\overline{\mathbf{v}}^{(K)}}$ possesses at most co-dimension K. On the other hand, $\mathcal{S}_{\overline{\mathbf{v}}^{(K)}} \subseteq \ker(\mathbf{\Gamma}_{\mathbf{f}-\mathbf{g}^{(K)}})$ and $\ker(\mathbf{\Gamma}_{\mathbf{f}-\mathbf{g}^{(K)}})$ has at least co-dimension K by Theorem 4.9. Thus, $n_K = K$, i.e., $P_{\mathbf{v}^{(K)}}(z)$ possesses exactly K zeros in $\mathbb{D}$, and $\mathcal{S}_{\overline{\mathbf{v}}^{(K)}} = \ker(\mathbf{\Gamma}_{\mathbf{f}-\mathbf{g}^{(K)}})$ has co-dimension K. Assertion (2) follows directly from Theorem 4.9. □

Proof of Theorem 3.2. Theorem 3.2 is now a corollary of Theorem 4.10. Theorem 4.10 contains the explicit sequence $\tilde{\mathbf{f}} = \mathbf{g} - \mathbf{f}$. From Theorem 4.10(3) it follows that $\tilde{\mathbf{f}} \in \mathrm{clos}_{\ell^2}\mathrm{span}\{((z_1^{(K)})^l)_{l=0}^{\infty}, \ldots, ((z_K^{(K)})^l)_{l=0}^{\infty}\}$, i.e., it can be written as a finite linear combination of the form (3.1). Moreover, for $\mathbf{e}_0 := (1, 0, 0, \ldots) \in \ell^1 \subset \ell^2$ we have for $p \in \{1, 2\}$

$$\|\mathbf{f}\|_p = \frac{\|\mathbf{\Gamma}_{\mathbf{f}} \mathbf{e}_0\|_p}{\|\mathbf{e}_0\|_p} \leq \|\mathbf{\Gamma}_{\mathbf{f}}\|_{\ell^p \to \ell^p}.$$

Thus the assertion follows. □

Connection to Prony's method. There is now obviously a close connection to Prony's method. When taking a zero-(con)-eigenvector $\mathbf{v}$ of the Hankel operator $\mathbf{\Gamma}_{\mathbf{f}}$ generated by $\mathbf{f}$ in (1.1), then by Theorem 4.2, the Laurent polynomial corresponding to $\mathbf{v}$ satisfies $P_{\overline{\mathbf{v}}}(z_j) = 0$ for all $j = 1, \ldots, N$ and has therefore at least N zeros inside the unit disk $\mathbb{D}$. In particular we have

Corollary 4.11. Let $\mathbf{v} = \mathbf{v}^{(N)} \in \ell^1$ be a zero-(con)-eigenvector of $\mathbf{\Gamma}_{\mathbf{f}}$ with $\mathbf{f}$ in (1.1), such that

$$P_{\overline{\mathbf{v}}}(z) = \prod_{j=1}^{N} (z - z_j) = P(z)$$

is the Prony polynomial in (2.2). Then

$$\ker(\mathbf{\Gamma}_{\mathbf{f}}) = \mathcal{S}_{\overline{\mathbf{v}}} = \mathrm{clos}_{\ell^2}\mathrm{span}\left\{S^l \overline{\mathbf{v}}^{(N)} : l \in \mathbb{N}_0\right\} = (\mathrm{clos}_{\ell^2}\mathrm{span}\{(z_1^l)_{l=0}^{\infty}, \ldots, (z_N^l)_{l=0}^{\infty}\})^{\perp}.$$

Proof. We observe as before that $\ker(\mathbf{\Gamma}_{\mathbf{f}}) \perp \mathrm{clos}_{\ell^2}\mathrm{span}\{(z_1^l)_{l=0}^{\infty}, \ldots, (z_N^l)_{l=0}^{\infty}\}$. Since on the one hand $\mathcal{S}_{\overline{\mathbf{v}}} \subseteq \ker(\mathbf{\Gamma}_{\mathbf{f}})$ by Lemma 4.1, and both $\ker(\mathbf{\Gamma}_{\mathbf{f}})$ and $\mathcal{S}_{\overline{\mathbf{v}}}$ have co-dimension N, equality follows. □

Remark 4.12. The proof given in this subsection does not explicitly use the Theorems of Beurling and Nehari for Hankel operators. Nehari's result states that the norm of the operator $\mathbf{\Gamma_f}$ is equal to the infimum of the L^∞-norm over all bounded function 2π-periodic functions whose Fourier coefficients coincide with f_k for $k \in \mathbb{N}_0$, see e.g. [20]. This result is "hidden" in Lemma 4.7, where a sequence $\mathbf{w}$ is constructed by the Fourier coefficients of a special function with norm 1 in L^∞.

Beurling's theorem essentially says that the linear span of all shifts of a given sequence $\mathbf{v}$ in ℓ^2 is characterized by the inner factor of its corresponding Laurent polynomial $P_{\mathbf{v}}(z)$. Thus assertion (3) of Theorem 4.10 is a direct consequence of Beurling's theorem. We have proven it directly by showing invertibility of the Toeplitz matrix $\mathbf{T}_{\mathbf{u}^{(K)}}$.

5 Numerical Examples

In this section we present some numerical examples demonstrating the performance of our algorithm. In all examples the approximate Prony method APM2 provided in [18] was used in the first step of the algorithm in Section 3.

Example 1. We approximate the function $f : \mathbb{R}^+ \to \mathbb{C}$ of the form

$$f(x) = \sum_{j=1}^{N} a_j\, z_j^x\,,$$

with $N = 10$ using $M = 50$ samples. We denote by $\mathbf{f} := (f_k)_{k=0}^M := (f(k))_{k=0}^M$ the vector of samples of the function f and by $\tilde{\mathbf{f}}^{(n)} := (\tilde{f}_k^{(n)})_{k=0}^M$ the output vector of the n-term approximation $\tilde{f}$ of f. Both, the nodes z_j and the coefficients $a_j,\ j = 1, \ldots, 10$ were chosen randomly in $\mathbb{D}$ and in the interval $(0,1)$ respectively and are given as follows

$$\begin{pmatrix} z_1\\ z_2\\ z_3\\ z_4\\ z_5\\ z_6\\ z_7\\ z_8\\ z_9\\ z_{10} \end{pmatrix} = \begin{pmatrix} -0.0159 + 0.3739i\\ -0.0770 + 0.0394i\\ -0.0639 - 0.1791i\\ -0.2324 + 0.5268i\\ 0.0102 + 0.4511i\\ 0.0129 + 0.0602i\\ 0.3812 + 0.1470i\\ 0.3538 + 0.1045i\\ 0.1732 - 0.3507i\\ -0.1457 - 0.2385i \end{pmatrix}, \qquad \begin{pmatrix} a_1\\ a_2\\ a_3\\ a_4\\ a_5\\ a_6\\ a_7\\ a_8\\ a_9\\ a_{10} \end{pmatrix} = \begin{pmatrix} 0.4709 + 0.4302i\\ 0.2305 + 0.1848i\\ 0.8443 + 0.9049i\\ 0.1948 + 0.9797i\\ 0.2259 + 0.4389i\\ 0.1707 + 0.1111i\\ 0.2277 + 0.2581i\\ 0.4357 + 0.4087i\\ 0.3111 + 0.5949i\\ 0.9234 + 0.2622i \end{pmatrix}.$$

The minimization problem from step 4 was performed by the least squares method using the given M samples. In the first step of the algorithm the accuracies $\varepsilon_1 = \varepsilon_2 = 10^{-15}$, the radius $r = 1$ and the upper bound for the number of exponentials $L = 20$ were chosen for APM2. The following table shows the coneigenvalues σ_n of the matrix $\mathbf{A}_N\mathbf{Z}_N$ and the corresponding approximation errors for different values

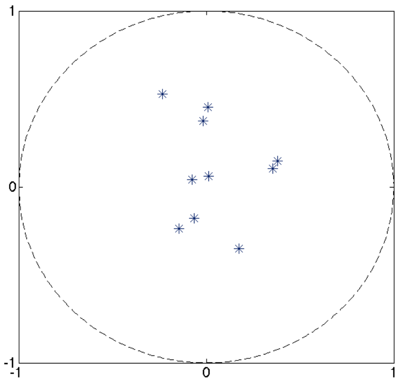

Figure 1 *The nodes $z_1, \ldots, z_{10}$ in the unit circle from Example 1.*

of n.

n	σ_n	$\|\mathbf{f} - \tilde{\mathbf{f}}^{(n)}\|_2$
1	4.4340e-01	4.4142e-01
2	5.5171e-02	5.3850e-02
3	1.8185e-02	1.8096e-02
4	8.1149e-03	8.1145e-03
5	7.8571e-05	7.8571e-05
6	4.3647e-06	4.3647e-06
7	2.6711e-07	2.6711e-07
8	6.2531e-08	6.2531e-08
9	1.4512e-10	1.4512e-10

Example 2. In this example we approximate the function $f(x) = 1/x$ using $M = 100$ samples in the interval $[1, 50]$. Let $\mathbf{f} := (f_k)_{k=0}^M$ be the vector of samples of f and $\tilde{\mathbf{f}}^{(n)} := (\tilde{f}_k^{(n)})_{k=0}^M$ the output vector of its n-term approximation by above algorithm. The initial $N = 11$ nodes z_j and weights a_j were obtained with parameters $\varepsilon_1 = \varepsilon_2 = 10^{-10}, r = 1,1$ and $L = 23$ by applying APM2 and are given as follows:

$$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \\ z_7 \\ z_8 \\ z_9 \\ z_{10} \\ z_{11} \end{pmatrix} = \begin{pmatrix} 0.9959 \\ 0.9781 \\ 0.9443 \\ 0.8919 \\ 0.8178 \\ 0.7198 \\ 0.5981 \\ 0.4568 \\ 0.3060 \\ 0.1634 \\ 0.0533 \end{pmatrix}, \qquad \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \\ a_8 \\ a_9 \\ a_{10} \\ a_{11} \end{pmatrix} = \begin{pmatrix} 0.0214 \\ 0.0507 \\ 0.0818 \\ 0.1137 \\ 0.1422 \\ 0.1597 \\ 0.1585 \\ 0.1339 \\ 0.0895 \\ 0.0405 \\ 0.0082 \end{pmatrix}.$$

In the following table we compare the ℓ^2-error of the above algorithm with the n-term exponential sum approximation obtained by W. Hackbusch in [7]. Let $\mathbf{f}_H^{(n)}$ be a vector of samples of the n-term approximation of f from [7], then we obtain the following errors using the same sampling for both approximations.

n	σ_n	$\|\mathbf{f}-\tilde{\mathbf{f}}^{(n)}\|_2$	$\|\mathbf{f}-\mathbf{f}_H^{(n)}\|_2$
1	1.5789e-00	1.0479e-00	–
2	4.3137e-01	3.7340e-01	1.4145e-01
3	9.9203e-02	9.4372e-02	2.4771e-02
4	1.9627e-02	1.9207e-02	4.4988e-03
5	3.3233e-03	3.2870e-03	7.8479e-04
6	4.7360e-04	4.6840e-04	1.3138e-04
7	5.5123e-05	5.4309e-05	2.2138e-05
8	4.9665e-06	4.8884e-06	3.6552e-06
9	3.1299e-07	3.1581e-07	5.9684e-07
10	1.0840e-08	4.5328e-08	9.8033e-08

Note that in [7] an algorithm for function approximation based on the Remez algorithm was used, whereas we approximate sequences of samples. The nodes $\tilde{z}_j$ and weights $\tilde{a}_j$ for $j = 1, \ldots, 10$ obtained by Hackbusch are very different from the ones computed by our algorithm, especially for small n. Even though, for the most values of n we obtain the same order of the ℓ^2-error. For $n \geq 9$ the error of our algorithm is even slightly better. The nodes and weights obtained by our algorithm are given below.

n	1	2	3	4	5	6	7	8	9	10
$\tilde{z}_1$	0.9804	0.8725	0.6982	0.5254	0.3856	0.2816	0.2063	0.1516	0.1112	0.0802
$\tilde{z}_2$		0.9933	0.9545	0.8706	0.7544	0.6279	0.5079	0.4022	0.3123	0.2355
$\tilde{z}_3$			0.9953	0.9710	0.9187	0.8386	0.7391	0.6309	0.5220	0.4157
$\tilde{z}_4$				0.9958	0.9760	0.9358	0.8731	0.7901	0.6917	0.5814
$\tilde{z}_5$					0.9959	0.9776	0.9421	0.8869	0.8113	0.7154
$\tilde{z}_6$						0.9959	0.9780	0.9439	0.8911	0.8171
$\tilde{z}_7$							0.9959	0.9780	0.9443	0.8919
$\tilde{z}_8$								0.9959	0.9781	0.9443
$\tilde{z}_9$									0.9959	0.9781
$\tilde{z}_{10}$										0.9959

n	1	2	3	4	5	6	7	8	9	10
$\tilde{a}_1$	0.2419	0.6728	0.7437	0.5561	0.3499	0.2045	0.1158	0.0645	0.0354	0.0186
$\tilde{a}_2$		0.0434	0.1717	0.3228	0.3886	0.3592	0.2832	0.2011	0.1320	0.0797
$\tilde{a}_3$			0.0290	0.0921	0.1756	0.2434	0.2669	0.2474	0.2014	0.1455
$\tilde{a}_4$				0.0225	0.0637	0.1173	0.1685	0.1988	0.2001	0.1753
$\tilde{a}_5$					0.0217	0.0542	0.0929	0.1318	0.1597	0.1686
$\tilde{a}_6$						0.0215	0.0514	0.0841	0.1172	0.1444
$\tilde{a}_7$							0.0214	0.0508	0.0821	0.1140
$\tilde{a}_8$								0.0214	0.0507	0.0818
$\tilde{a}_9$									0.0214	0.0507
$\tilde{a}_{10}$										0.0214

Acknowledgement

The authors gratefully acknowledge the funding of this work by the DFG in the framework of the GRK 2088 and the project PL 170/16-1.

References

[1] V.M. Adamjan, D.Z. Arov, and M.G. Krein, *Analytic properties of the Schmidt pairs of a Hankel operator and the generalized Schur-Takagi problem*, Mat. Sb. **86** (1971), 34–75 (in Russian).

[2] F. Andersson, M. Carlsson, and M.V. de Hoop, *Sparse approximation of functions using sums of exponentials and AAK theory*, J. Approx. Theory **163** (2011), 213–248.

[3] A. Beurling, *On two problems concerning linear transformations in Hilbert space*, Acta Math. **81** (1949), 239-255.

[4] G. Beylkin and L. Monzón, *On approximation of functions by exponential sums*, Appl. Comput. Harmon. Anal. **19** (2005), 17–48.

[5] D. Braess and W. Hackbusch, *Approximation of* $1/x$ *by exponential sums in* $[1, \infty)$, IMA J. Numer. Anal. **25** (2005), 685–697.

[6] C.K. Chui and G. Chen, *Discrete* H^∞ *Optimization*, Springer, Berlin, 1997.

[7] W. Hackbusch, *Entwicklungen nach Exponentialsummen*, Techn. Bericht 4 (2005).

[8] R.A. Horn and C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1990.

[9] I. Markovsky, Low Rank Approximation: Algorithms, Implementation, Applications, Springer, London, 2012.

[10] R.A. Martínez-Avendaño and Peter Rosenthal, An Introduction to Operators on the Hardy-Hilbert Space, Graduate Texts im Mathematics, Springer, 2007.

[11] J. Meinguet, *A simplified presentation of the Adamjan-Arov-Krein approximation theory*, In: Computational Aspects of Complex Analysis, edited by H. Werner, L. Wuytack, E. Ng, and H.J. Bünger, Volume 102 of the series NATO Advanced Study Institutes Series, D. Reidel, Dordrecht, 1983, pp. 217–248.

[12] Z. Nehari, *On bounded linear forms*, Annals of Mathematics **65** (1957), 153–162.

[13] N.K. Nikolski, Operators, Functions, and Systems: An Easy Reading (Mathematical Surveys and Monographs), Vol. **92**, AMS, 2002.

[14] V.V. Peller, Hankel Operators and Their Applications, In: Springer Monographs in Mathematics, Springer, New York, 2003.

[15] T. Peter and G. Plonka, *A generalized Prony method for reconstruction of sparse sums of eigenfunctions of linear operators*, Inverse Problems **29** (2013), 025001.

[16] G. Plonka and V. Pototskaia, Sparse approximation by Prony's method and AAK theory. Oberwolfach Reports, Volume 33/2016, 16-19, DOI: 10.4171/OWR/2016/1627b

[17] G. Plonka and M. Tasche, *Prony methods for recovery of structured functions*, GAMM-Mitt. **37**(2) (2014), 239–258.

[18] D. Potts, M. Tasche, *Parameter estimation for nonincreasing exponential sums by Prony-like methods*, Linear Algebra Appl. 439 (2013), 1024-1039.

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

[20] N. Young, An Introduction to Hilbert Space, Cambridge University Press, Cambridge, 1988.