Computation of adaptive Fourier series by sparse approximation of exponential sums

Gerlind Plonka[*]   Vlada Pototskaia[†]

3rd January 2018

Abstract. In this paper, we study the convergence of adaptive Fourier sums for real-valued 2π-periodic functions. For this purpose, we approximate the sequence of classical Fourier coefficients by a short exponential sum with a pre-defined number of $N+1$ terms. The obtained approximation can be interpreted as an adaptive N-th Fourier sum with respect to the orthogonal Takenaka-Malmquist basis. Using the theoretical results on rational approximation in Hardy spaces and on the decay of singular values of special infinite Hankel matrices, we show that adaptive Fourier sums can converge essentially faster than the classical Fourier sums for a large class of functions. Further, we derive a method to compute almost optimal adaptive Fourier sums. We also present a much simpler greedy algorithm that possesses similar convergence behavior as the classical Fourier sums. Our numerical results show, that the significantly better convergence behavior of adaptive Fourier sums for optimally chosen basis elements can also be achieved in practice.

Key words: Adaptive Fourier sum, Prony method, AAK theory, infinite Hankel matrices, adaptive Fouries series, Takenaka-Malmquist basis, convergence rate.
Mathematics Subject Classification: 15A18, 41A30, 42A16, 65F15.

1 Introduction

In this paper, we want to study adaptive Fourier expansions of real-valued functions f in $L^2(0, 2\pi])$ using a generalized Fourier basis. In different papers, see e.g. [7, 18, 24, 25], generalized Fourier bases have been considered with particular applications in signal processing and system identification, see [7, 12].

Before we can state the problem and our results in detail, we need to introduce some notations. Let $\mathbb{T}$ be the unit circle in the complex domain, $\mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$. We identify $\mathbb{T}$ with the interval $[0, 2\pi)$ using the map $z = e^{it}$. Let $L^2([0, 2\pi))$ be the space of square integrable 2π-periodic functions. The Fourier expansion of $f \in L^2([0, 2\pi))$ reads

$$f(t) = \sum_{k=-\infty}^{\infty} c_k(f) \, e^{ikt},$$

---

[*]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

where $c_k(f) := \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-ikt} dt$ denote the Fourier coefficients of f. Further, let

$$\langle f, g \rangle_{L^2} := \frac{1}{2\pi} \int_0^{2\pi} f(t) \overline{g(t)} dt \tag{1.1}$$

be the inner product in $L^2([0, 2\pi))$, and $\|\cdot\|_{L^2}$ denotes the corresponding norm. We assume that f is real-valued, then $c_{-k}(f) = \overline{c_k(f)}$, and hence

$$f(t) = c_0(f) + \sum_{k=1}^{\infty} \left(c_k(f) e^{ikt} + \overline{c_k(f)} e^{-ikt} \right),$$

i.e., $f(t) = 2\mathrm{Re}\, f^+(t)$, where

$$f^+(t) := \frac{c_0(f)}{2} + \sum_{k=1}^{\infty} c_k(f) e^{ikt} = \sum_{k=0}^{\infty} c_k(f^+) e^{ikt}. \tag{1.2}$$

Let $H^2 \subset L^2([0, 2\pi))$ be the Hardy space of all functions f in $L^2([0, 2\pi))$ for which all negative Fourier coefficients vanish, i.e., $c_{-k}(f) = 0$ for $k \in \mathbb{N}$. Obviously, $f^+ \in H^2$. We introduce the generalized Fourier system, the so-called Takenaka-Malmquist system,

$$B_0 = \frac{\sqrt{1 - |z_0|^2}}{1 - z_0\, e^{it}}, \qquad B_\ell(t) = \frac{\sqrt{1 - |z_\ell|^2}}{1 - z_\ell\, e^{it}} \prod_{k=0}^{\ell-1} \frac{e^{it} - \overline{z}_k}{1 - z_k\, e^{it}}, \quad \ell = 1, 2, \ldots. \tag{1.3}$$

The function system $\{B_\ell\}_{\ell=0}^{\infty}$ is completely determined by the sequence of "zeros" $\{z_\ell\}_{\ell=0}^{\infty}$ where we assume that $|z_\ell| \leq c < 1$ for all $\ell \geq 0$, see e.g. [27]. This function system is an orthonormal system with respect to the inner product in (1.1), i.e., we have $\langle B_\ell, B_k \rangle_{L^2} = \delta_{\ell,k}$ with $\delta_{\ell,k}$ being the Kronecker symbol. Moreover, this system is complete in H^2 if and only if $\sum_{\ell \geq 0}(1 - |z_\ell|) = \infty$. For a proof we refer to [2], p. 244. Note that for $z_\ell = 0$, $\ell \geq 0$, we obtain the classical Fourier basis $B_\ell(t) = e^{i\ell t}$. The expansion of $f^+ \in H^2$ in this orthogonal system reads $f^+(t) = \sum_{k=0}^{\infty}\langle f^+, B_k \rangle_{L^2} B_k$ with adaptive Fourier coefficients $\langle f^+, B_k \rangle_{L^2}$. Further, let

$$f_N^+ := \sum_{k=0}^{N} \langle f^+, B_k \rangle_{L^2} B_k \tag{1.4}$$

denote the N-th adaptive partial Fourier sum. Bultheel and Carrette [7] have shown that similarly as for the classical Fourier basis, the generalized Fourier coefficients of a 2π-periodic function with continuous q-th derivative with $q > 2$ in the Takenaka-Malmquist basis decay with at least rate N^{-q}. In this case the approximation error $e_N(f^+) = \|f^+ - f_N^+\|_\infty$ decays with rate N^{2-q}. This result involves only the assumption that $z_0 = 0$ and $|z_\ell| \leq c < 1$.

In this paper we want to study the question, how well a given function $f^+ \in H^2$ can be approximated by its generalized partial Fourier sum f_N^+ if the zeros z_ℓ determining the Takenaka-Malmquist system are taken in an (almost) optimal way. Our considerations are inspired by the study of adaptive Fourier series in [24] on the one hand and by our own results on approximation by exponential sums on the other hand, [20, 21].

Our approach relies on the following idea. For a given classical Fourier expansion (1.2) of $f^+ \in H^2$ we try to approximate the sequence of Fourier coefficients $\mathbf{c}(f^+) =$

$(c_k(f^+))_{k=0}^\infty$ by a sequence $\tilde{\mathbf{c}} = \tilde{\mathbf{c}}^{(N)} = (\tilde{c}_k)_{k=0}^\infty$ being a finite linear combination of non-increasing exponentials, i.e.,

$$\tilde{c}_k = \tilde{c}_k^{(N)} = \sum_{\ell=0}^{N} a_\ell z_\ell^k, \qquad k \in \mathbb{N}_0, \tag{1.5}$$

with coefficients $\tilde{a}_\ell \in \mathbb{C}$ and $\tilde{z}_\ell \in \mathbb{D}_0$, where $\mathbb{D}_0 := \{z \in \mathbb{C} : |z| < 1\}$ denotes the open unit disk. The goal is now to choose a_ℓ and z_ℓ such that $\|\mathbf{c}(f^+) - \tilde{\mathbf{c}}\|_{\ell^2}^2 := \sum_{k=0}^\infty |c_k(f^+) - \tilde{c}_k|^2 < \epsilon$ with ϵ as small es possible. Then

$$\tilde{f}^+(t) := \sum_{k=0}^\infty \tilde{c}_k \, \mathrm{e}^{ikt}$$

satisfies $\|f^+ - \tilde{f}\|_{L^2}^2 < \epsilon$ by Parseval identity.

Using the special representation (1.5) for $\tilde{c}_k$ we will show that $\tilde{f}^+$ can be written as an N-th adaptive partial Fourier sum

$$\tilde{f}^+(t) = \sum_{\ell=0}^{N} b_\ell \, B_\ell(t)$$

in the Takenaka-Malmquist basis, where the basis elements $\{B_\ell\}_{\ell=0}^N$ are determined by the zeros z_ℓ in (1.5) and the coefficients b_ℓ can be obtained from a_ℓ in (1.5) by a bijective linear map. Therefore, the problem to find an appropriate sequence of zeros $\{z_\ell\}$ to determine the adaptive Fourier basis $\{B_\ell\}_{\ell=0}^N$ can be reformulated as a problem of sparse approximation with exponential sums:

For a given decaying sequence $\mathbf{c}(f^+)$ we want to find a new sequence $\tilde{\mathbf{c}}^{(N)} := (\tilde{c}_k)_{k=0}^\infty$ of the form (1.5) such that $\|\mathbf{c}(f^+) - \tilde{\mathbf{c}}^{(N)}\|_{\ell^2}^2 \le \epsilon$.

Practically, we have to solve the following problems. For a given accuracy level $\epsilon > 0$, find the smallest $N \in \mathbb{N}$ such that $\tilde{\mathbf{c}}^{(N)}$ in (1.5) satisfies $\|\mathbf{c}(f^+) - \tilde{\mathbf{c}}^{(N)}\|_{\ell^2} \le \epsilon$, and compute $z_\ell \in \mathbb{D}$ and $a_\ell \in \mathbb{C}$, $\ell = 0, \ldots, N$, determining $\tilde{\mathbf{c}}^{(N)}$. Vice versa, for a given "storage budget", i.e., a given $N \in \mathbb{N}$, find the parameters in (1.5) in order to achieve the smallest possible error $\|\mathbf{c}(f^+) - \tilde{\mathbf{c}}^{(N)}\|_{\ell^2}^2$.

Our paper has two goals. First, we want to study the following question. What is the asymptotic error behavior of $\|f^+ - f_N^+\|_{L^2}$ for $N \to \infty$ compared to the decay of the classical Fourier sums that can be achieved? Second, we want to present a new computational method to explicitly compute the adaptive partial Fourier sum f_N^+ with a presumed small error ϵ such that N is as small as possible by employing a (suitably large) finite number of classical Fourier coefficients $(c_k(f^+))_{k=0}^L$. This algorithm particularly includes the computation of the zeros $\{z_\ell\}_{\ell=0}^N$ needed to determine the Takenaka-Malmquist basis for the adaptive partial Fourier sum f_N^+. We will also observe that optimal zeros z_ℓ will depend on the considered length of the sum N.

We will apply the theory of Adamjan, Arov and Krein (AAK theory) [1] to solve the problem of sparse approximation of the sequence of classical Fourier coefficients by exponential sums. This theory is strongly related to the problem of rational approximation of functions in Hardy spaces and has been studied also with respect to its applications in signal processing and system identification, see e.g. [1, 8, 14, 15, 17, 19, 29]. For earlier approaches on the application of the AAK theory in order to solve sparse approximation problems using exponential sums we refer to [6] and [3].

The AAK theory gives us a tool to relate the approximation error $\|\mathbf{c}(f^+) - \tilde{\mathbf{c}}^{(N)}\|_{\ell^2}$ to singular values of special infinite Hankel matrices that are generated by the sequence $\mathbf{c}(f^+)$. Using this relation, we will be able to estimate the decay of the error $\|f^+ - f_N^+\|_{L^2}$ by inspecting the decay of the singular values of these infinite Hankel matrices. At this point, we can use recent results by Pushnitski and Yafaev [23], who derived sharp estimates for singular values of Hankel operators. Roughly speaking, we can conclude that the decay $\mathcal{O}(N^{-1} \log N)^{-\alpha}$ of the sequence $\mathbf{c}(f^+)$ of classical Fourier coefficients can lead to the decay $\mathcal{O}(N^{-\alpha})$ for the error $\|f^+ - f_N^+\|_{L^2}$ of the N-th adaptive partial Fourier sum. Thus, under certain side conditions, we will show that one can even achieve exponential decay of the error for the adaptive Fourier sum while the classical Fourier sum may not even possesses quadratic decay, see Theorem 3.6.

In the second part of the paper, we present a new algorithm for solving the sparse approximation problem (1.5). The procedure consists of two steps. In the first step we employ a stabilized Prony method, like APM [22], to approximate $\mathbf{c}(f^+)$ by a sequence $\check{\mathbf{c}} = (\check{c}_k)_{k=0}^{\infty}$ of the form

$$\check{c}_k := \sum_{\ell=0}^{M} \check{a}_\ell \, \check{z}_\ell^k$$

with possibly large M where $\check{a}_\ell \in \mathbb{C} \setminus \{0\}$ and with pairwise different $\check{z}_\ell \in \mathbb{D} := \{z \in \mathbb{C} : 0 < |z| < 1\}$. In a second step, we present our new algorithm to explicitly compute the non-zero singular values of the infinite Hankel matrix generated by $\check{\mathbf{c}}$. The decay of these singular values will give us a criteria, how to choose the number N such that a sequence $\tilde{\mathbf{c}}^{(N)}$ in (1.5) built by an exponential sum of length $N+1$ provides an error $\|\check{\mathbf{c}} - \tilde{\mathbf{c}}\|_{\ell^2}^2$ being smaller than a predefined ϵ. Furthermore, our algorithm will provide the zeros z_ℓ and coefficients a_ℓ to compute $\tilde{\mathbf{c}}^{(N)}$ in (1.5).

We emphasize that the sequences $\tilde{\mathbf{c}}^{(N)}$ in (1.5) obtained by our algorithm lead to adaptive partial Fourier sums that are optimal in the sense of the AAK theory if the sequence $\mathbf{c}(f^+)$ of classical Fourier coefficients of f^+ coincides with $\check{\mathbf{c}}$ for some M. This is the case if f^+ is a rational function.

For comparison, we also provide a greedy algorithm that searches for appropriate zeros z_ℓ for the Takenaka-Malmquist basis step by step. Here we apply a simple adaption of Prony's method for exponential sums of length 1.

Our numerical experiments impressively support our theoretical result that the appropriate choice of zeros z_ℓ in the adaptive Fourier basis leads to a significantly faster decay of the adaptive partial Fourier sums. In contrast, the simple greedy approach leads to an error that is comparable to the error of the classical partial Fourier sums.

This paper is organized as follows. In Section 2 we present a further basis $\{P_\ell\}_{\ell=0}^{\infty}$ such that $\{B_\ell\}_{\ell=0}^{N}$ and $\{P_\ell\}_{\ell=0}^{N}$ span the same subspaces for each N, and we describe the corresponding basis transform. We prefer the basis $\{P_\ell\}_{\ell=0}^{\infty}$ in the theoretical observation since it simplifies our representations. Section 3 is devoted to the study of the decay of adaptive partial Fourier series that can be achieved. In Section 4 we derive a method to compute (almost) optimal zeros $z_\ell = z_\ell^{(N)}$, $\ell = 0, \ldots, N$, such that the corresponding N-th adaptive partial Fourier sum f_N^+ satisfies a small predetermined error bound. In Section 5 we also present a much simpler greedy algorithm to compute the zeros z_ℓ determining the Takenaka-Malmquist basis consecutively. Finally, in Section 6 we present numerical results for the decay of Fourier sums for

two different examples. In the first example, we consider the Fourier expansion of a rational function, where our algorithm gives the optimal results according to AAK theory. The second example considers a non-rational function. Our numerical results impressively show that a good choice of zeros in the adaptive Fourier basis leads to a significantly faster decay of adaptive partial Fourier sums compared to the classical Fourier sums.

2 Adaptive Fourier bases

Preliminary, we want to present a different basis that is closely related to the Takenaka-Malmquist basis and simplifies our representations using approximations with exponential sums.

For a given vector of pairwise different values $(z_0, z_1, \ldots, z_N)$ with $0 < |z_\ell| < 1$ let

$$P_\ell(t) := \frac{1}{(1 - z_\ell e^{it})}, \quad \ell = 0, \ldots, N. \tag{2.1}$$

More generally, if 0 or multiple values occur in the vector $(z_0, \ldots, z_N)$, let $M_\ell + 1$ be the number of different values in $(z_0, \ldots, z_\ell)$ that are denoted by $\tilde{z}_0, \ldots, \tilde{z}_{M_\ell}$. Let $K_{j,\ell}$ be the multiplicity of $\tilde{z}_j$ in $(z_0, \ldots, z_\ell)$ such that $K_{0,\ell} + K_{1,\ell} + \ldots + K_{M_\ell,\ell} = \ell + 1$. Further, let μ_ℓ be the multiplicity of 0 in the vector $(z_0, \ldots, z_{\ell-1})$. If $z_\ell = \tilde{z}_j$, we define P_ℓ by

$$P_\ell(t) := \frac{e^{it\mu_\ell}}{(1 - \tilde{z}_j e^{it})^{K_{j,\ell}}}. \tag{2.2}$$

Theorem 2.1. *The set of functions $\{B_\ell : \ell = 0, \ldots, N\}$ determined by (1.3) with the zeros $z_0, z_1, \ldots z_N$ in $\mathbb{D}_0$ spans the same subspace as $\{P_\ell : \ell = 0, \ldots, N\}$ given in (2.1) and (2.2), respectively.*

Proof. Let us first assume that $z_0, \ldots, z_N$ are pairwise different and nonzero. Then the functions $B_\ell(t)$ in (1.3) can be written in the form

$$B_\ell(t) = \sqrt{1 - |z_\ell|^2} \sum_{j=0}^{\ell} \frac{p_{\ell,j}}{1 - z_j e^{it}}, \qquad \ell = 0, \ldots, N, \tag{2.3}$$

where the coefficients $p_{\ell,j}$ are uniquely determined by employing a partial fraction decomposition. More exactly, determining the functions

$$\Lambda_{\ell,j}(t) := \frac{\prod_{k=0}^{\ell}(1 - z_k e^{it})}{(1 - z_j e^{it})} \quad j = 0, \ldots, \ell,$$

we find from (1.3) and (2.3) that

$$\sum_{j=0}^{\ell} p_{\ell,j} \Lambda_{\ell,j}(t) = \prod_{k=0}^{\ell-1} (e^{it} - \overline{z}_k), \qquad \ell = 0, \ldots, N,$$

and $p_{\ell,j}$ are uniquely determined by comparing the coefficients of the trigonometric polynomials of degree ℓ with respect to the powers e^{itk}, $k = 0, \ldots, \ell$. In particular, it follows that $p_{\ell,\ell} \neq 0$. Thus, comparison with (2.1) yields

$$B_\ell(t) = \sqrt{1 - |z_\ell|^2} \sum_{j=0}^{\ell} p_{\ell,j} P_j(t), \qquad \ell = 0, \ldots, N. \tag{2.4}$$

The system (2.4) implies

$$
\begin{pmatrix} (1-|z_0|^2)^{-1/2}B_0(t) \\ (1-|z_1|^2)^{-1/2}B_1(t) \\ \vdots \\ (1-|z_N|^2)^{-1/2}B_N(t) \end{pmatrix} = \begin{pmatrix} p_{0,0} & 0 & \cdots & & 0 \\ p_{1,0} & p_{1,1} & & & \vdots \\ \vdots & & \ddots & & 0 \\ p_{N,0} & \cdots & p_{N,N-1} & p_{N,N} \end{pmatrix} \begin{pmatrix} P_0(t) \\ P_1(t) \\ \vdots \\ P_N(t) \end{pmatrix},
$$

showing that the transform is invertible.

If the vector $(z_0, \ldots, z_N)$ contains the value $z_j = 0$ or multiple values, then we find with the same notations as for the general definition of P_ℓ in (2.2) the generalized partial fraction decomposition of the form

$$
B_\ell(t) = \sqrt{1-|z_\ell|^2}\, \mathrm{e}^{\mu_\ell it} \sum_{j=0}^{M_\ell} \sum_{r=1}^{K_{j,\ell}} \frac{p_{\ell,j,r}}{(1-\tilde{z}_j\, \mathrm{e}^{it})^{K_{j,\ell}}}.
$$

Now, similarly as before we observe that $B_\ell(t)$ can be written as a linear combination of the functions P_j, $j = 0, \ldots, \ell$ obtained from these $\ell + 1$ zeros, and in particular the coefficients $p_{\ell,j,K_{j,\ell}}$ do not vanish. Therefore we get again an invertible linear system relating $P_\ell(t)$ and $B_\ell(t)$. $\qquad\square$

3 Adaptive Fourier sums and its convergence

Let us assume that $f(t) = 2\mathrm{Re}\, f^+(t)$ is a real-valued function in $L^2(\mathbb{R})$ with

$$
f^+(t) = \sum_{k=0}^{\infty} c_k(f^+)\, \mathrm{e}^{ikt}.
$$

Our main goal is to generate a new adaptive Fourier basis $\{B_\ell\}$ as given in (1.3) such that the adaptive N-th Fourier sums f_N^+ of f^+ in (1.4) have significantly faster decay for $N \to \infty$ than the classical Fourier sum $S_N(f) := \sum_{k=0}^N c_k(f^+)\mathrm{e}^{ik\cdot}$. For this purpose we need to choose the sequence of zeros $\{z_\ell\}_{\ell=0}^N$ defining the adaptive Fourier basis $\{B_\ell\}_{\ell=0}^N$ in an appropriate way. In this section we will show that indeed a strong decay of the adaptive Fourier sums can be achieved. The crucial idea to find the adaptive basis is the approach to approximate the sequence of classical Fourier coefficients by finite exponential sums.

For a given function $f^+ \in H^2$ we will approximate $\mathbf{c}(f^+) = (c_k(f^+))_{k=0}^{\infty}$ by a new sequence $\widetilde{\mathbf{c}}^{(N)} = \widetilde{\mathbf{c}} = (\widetilde{c}_k)_{k=0}^{\infty}$ with coefficients being given by an exponential sum of length $N + 1$,

$$
\widetilde{c}_k = \sum_{\ell=0}^{N} a_\ell z_\ell^k, \tag{3.1}
$$

where the parameters $a_\ell \in \mathbb{C}$ and $z_\ell \in \mathbb{D}$ have to be suitably chosen such that $\|\widetilde{\mathbf{c}}^{(N)} - \mathbf{c}(f^+)\|_{\ell^2}^2 := \sum_{k=0}^{\infty} |\tilde{c}_k - c_k(f^+)|^2 < \epsilon$ for some small ϵ. Taking now the approximation $f_N^+(t) := \sum_{k=0}^{\infty} \widetilde{c}_k\, \mathrm{e}^{ikt}$ we immediately obtain the error estimate

$$
\|f^+ - f_N^+\|_{L^2} = \|\mathbf{c}(f^+) - \widetilde{\mathbf{c}}^{(N)}\|_{\ell^2} < \epsilon
$$

by Parseval identity. Approximating the sequence of Fourier coefficients of f^+ as given in (3.1), we find

$$
f_N^+(t) \;=\; \sum_{k=0}^{\infty} \widetilde{c}_k\, \mathrm{e}^{ikt} = \sum_{k=0}^{\infty} \Big(\sum_{\ell=0}^{N} a_\ell z_\ell^k \Big) \mathrm{e}^{ikt}
$$

$$= \sum_{\ell=0}^{N} a_\ell \sum_{k=0}^{\infty} (z_\ell \mathrm{e}^{it})^k = \sum_{\ell=0}^{N} a_\ell \frac{1}{(1 - z_\ell \mathrm{e}^{it})} = \sum_{\ell=0}^{N} a_\ell \, P_\ell(t), \qquad (3.2)$$

i.e., f_N^+ can be written with only $N + 1$ terms in the basis $\{P_\ell\}_{\ell=0}^{N}$. This observation justifies that we call this approximation f_N^+. By Theorem 2.1, f_N^+ can also be written as an expansion of $N + 1$ terms in the orthogonal Takenaka-Malmquist basis determined by the zeros $(z_0, \ldots, z_N)$ in (3.1).

Using this approach we can reformulate the problem of finding an (almost) optimal adaptive Fourier basis as the problem to find an optimal approximation of the sequence of classical Fourier coefficients by a sequence obtained from a short exponential sum. Therefore, we consider now the question how to obtain for a fixed N a sequence $\tilde{\mathbf{c}}^{(N)} = (\widetilde{c}_k)_{k=0}^{\infty}$ of the form (3.1) such that $\|\widetilde{\mathbf{c}}^{(N)} - \mathbf{c}(f^+)\|_{\ell^2}$ is as small as possible. In this section we first consider the theoretic problem. Using the Theorem of Adamjan, Arov and Krein and the properties of infinite Hankel matrices, we will be able to show that one can achieve very fast convergence rates for adaptive Fourier sums for a large class of functions.

3.1 Infinite Hankel matrices and the AAK theorem

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

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

its corresponding Laurent polynomial and $P_{\mathbf{v}}(\mathrm{e}^{it})$, $t \in \mathbb{R}$, its Fourier series. Further, for $\mathbf{c} \in \ell^2$ we define the infinite Hankel matrix

$$\mathbf{\Gamma_c} := \begin{pmatrix} c_0 & c_1 & c_2 & \cdots \\ c_1 & c_2 & c_3 & \cdots \\ c_2 & c_3 & c_4 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} = (c_{k+j})_{k,j=0}^{\infty}. \qquad (3.3)$$

If the sequence $\mathbf{c}$ can be obtained as the (sub)sequence of Fourier coefficients, i.e., $\mathbf{c} = (c_k(f))_{k=0}^{\infty}$ for a function $f \in L^{\infty}([0, 2\pi))$, then $\mathbf{\Gamma_c}$ is a bounded operator on ℓ^2, see the Theorem of Nehari [16] or [29], Theorem 15.18. Moreover, $\mathbf{\Gamma_c}$ is compact on ℓ^2 if $\mathbf{c} = (c_k(f))_{k=0}^{\infty}$ for a continuous function $f \in C([0, 2\pi))$, see [10].

We assume that $\mathbf{c}$ is contained in the weighted Hilbert space

$$\ell_w^2 := \{\mathbf{c} \in \ell^2 : \sum_{k=0}^{\infty} (k + 1)|c_k|^2 < \infty\} \subset \ell^2.$$

Then $\mathbf{\Gamma_c}$ maps a sequence $\mathbf{v} = (v_j)_{j=0}^{\infty} \in \ell^2$ into ℓ^2, and with

$$\mathbf{\Gamma_c}\, \mathbf{v} = \left(\sum_{j=0}^{\infty} c_{k+j}\, v_j \right)_{k=0}^{\infty} \quad \text{for } \mathbf{v} := (v_k)_{k=0}^{\infty}$$

it follows by Cauchy Schwarz inequality that

$$\|\mathbf{\Gamma_c v}\|_{\ell^2}^2 = \sum_{k=0}^{\infty} \Big| \sum_{j=0}^{\infty} c_{j+k}\, v_j \Big|^2 \le \sum_{k=0}^{\infty} \|(c_{k+j})_{j=0}^{\infty}\|_{\ell_2}^2 \, \|\mathbf{v}\|_{\ell^2}^2$$

$$= \Big(\sum_{k=0}^{\infty} (k+1)\,|c_k|^2 \Big) \|\mathbf{v}\|_{\ell^2}^2 = \|\mathbf{c}\|_{\ell_w^2}^2 \, \|\mathbf{v}\|_{\ell^2}^2 .$$

Observe that $\ell^1 \subset \ell_w^2$ since $\mathbf{c} \in \ell^1$ implies that

$$\sum_{k=0}^{\infty} (k+1)|c_k|^2 \le \sup_{k \in \mathbb{N}_0} |(k+1)c_k| \, \|\mathbf{c}\|_{\ell^1} < \infty .$$

We want to apply now the Theorem of Adamjan, Arov and Krein [1] together with some new results on the decay of singular values of Hankel operators by Pushnitski and Yafarev [23] in order to show the strong convergence properties of adaptive N-th Fourier sums in (1.4) if we take appropriate zeros $\{z_\ell\}_{\ell=0}^N$ to determine $B_\ell(t)$.

Theorem 3.1. (*Adamjan, Arov and Krein*). *Let $\mathbf{c} \in \ell_w^2$ be a given sequence and let $\mathbf{\Gamma_c}$ be the corresponding infinite Hankel matrix on ℓ^2. Further, let $\sigma_0(\mathbf{\Gamma_c}) \ge \sigma_1(\mathbf{\Gamma_c}) \ge \sigma_2(\mathbf{\Gamma_c}) \ge \dots$ denote the singular values of $\mathbf{\Gamma_c}$ in decreasing order. Then, for each $N \in \mathbb{N}_0$ there exists an infinite Hankel matrix $\mathbf{\Gamma_{\tilde{c}}}$ of rank $N+1$ such that*

$$\|\mathbf{\Gamma_c} - \mathbf{\Gamma_{\tilde{c}}}\|_{\ell^2 \to \ell^2} = \sigma_{N+1}(\mathbf{\Gamma_c}).$$

For the proof of this theorem we refer to [1] or to [19].

We consider now Hankel operators with finite rank $N+1$ in more detail. Let $\tilde{\mathbf{c}} = (\tilde{c}_k)_{k=0}^{\infty}$ be a special sequence of the form

$$\tilde{c}_k = \sum_{j=0}^{N} a_j z_j^k, \qquad k \in \mathbb{N}_0 \tag{3.4}$$

where $N \in \mathbb{N}_0$, $a_j \in \mathbb{C} \ne \{0\}$ and with pairwise different values $0 < |z_N| \le \dots \le |z_0| < 1$. Then $\tilde{\mathbf{c}} \in \ell^1$ and thus in ℓ_w^2 since

$$\|\tilde{\mathbf{c}}\|_{\ell^1} = \sum_{k=0}^{\infty} |\tilde{c}_k| = \sum_{k=0}^{\infty} \Big| \sum_{j=0}^{N} a_j z_j^k \Big| \le \sum_{j=0}^{N} \Big(\sum_{k=0}^{\infty} |a_j z_j^k| \Big) = \sum_{j=0}^{N} \frac{|a_j|}{1 - |z_j|} < \infty .$$

We recall the following property of the corresponding infinite Hankel matrix $\mathbf{\Gamma_{\tilde{c}}}$, see e.g. [29], Theorem 16.13.

Theorem 3.2. (*Kronecker*). *The Hankel operator $\mathbf{\Gamma_{\tilde{c}}} : \ell^2 \to \ell^2$ generated by $\tilde{\mathbf{c}} = (\tilde{c}_k)_{k=0}^{\infty} \in \ell^1$ of the form (3.4) has finite rank $N+1$.*

Proof. For reader's convenience we provide a short proof that also gives some insight into the connection to difference equations. If $\tilde{\mathbf{c}}$ can be written in the form (3.4), we define the characteristic polynomial $P(z)$ by the values $z_0, \dots, z_N$,

$$P(z) := \prod_{j=0}^{N} (z - z_j) = \sum_{\ell=0}^{N+1} b_\ell z^\ell, \tag{3.5}$$

where the b_ℓ are the coefficients in the monomial representation of $P(z)$. Then (3.4) yields

$$\sum_{\ell=0}^{N+1} b_\ell \tilde{c}_{k+\ell} = \sum_{\ell=0}^{N+1} b_\ell \sum_{j=0}^{N} a_j z_j^{k+\ell} = \sum_{j=0}^{N} a_j z_j^k \left(\sum_{\ell=0}^{N} b_\ell z_j^\ell\right) = \sum_{j=0}^{N} a_j z_j^k P(z_j) = 0 \quad (3.6)$$

for all $k \in \mathbb{N}_0$, i.e., $(\tilde{c}_k)_{k=0}^\infty$ satisfies a difference equation of order $N+1$. Thus, the $(N+1+k)$-th column of $\boldsymbol{\Gamma}_{\tilde{\mathbf{c}}}$ is a linear combination of the $N+1$ preceding columns, and we conclude rank $\boldsymbol{\Gamma}_{\tilde{\mathbf{c}}} \leq N+1$. Since $P(z)$ has exact degree $N+1$ it follows that rank $\boldsymbol{\Gamma}_{\tilde{\mathbf{c}}} = N+1$. $\qquad\square$

Remark 3.3. Conversely, if the infinite Hankel matrix $\boldsymbol{\Gamma}_{\tilde{\mathbf{c}}}$ possesses rank $N+1$, then $\tilde{\mathbf{c}}$ satisfies a difference equation of order $N+1$. Thus, there exist coefficients $b_0, \ldots, b_{N+1}$ such that

$$\sum_{l=0}^{N+1} b_l \tilde{c}_{k+l} = 0 \quad \forall k \in \mathbb{N}_0,$$

and $b_0 \neq 0$, $b_{N+1} \neq 0$. Assuming that the zeros z_j, $j = 0, \ldots, N$, of the characteristic polynomial $\sum_{l=0}^{N+1} b_l z^l$ are pairwise different, $\tilde{\mathbf{c}}$ can be written in the form (3.4). The zeros z_j have modulus smaller than 1 since $\mathbf{f}$ has been assumed to be in ℓ^1.

In the general case, if the characteristic polynomial possesses m pairwise different zeros $\tilde{z}_1, \ldots, \tilde{z}_m$ with multiplicities $K_1, \ldots, K_m$ and $K_1 + \ldots + K_m = N+1$, then the $\tilde{c}_k$ take the form

$$\tilde{c}_k = \sum_{j=1}^{m} \sum_{r=1}^{K_j} a_{j,r} k^{r-1} z_j^k, \qquad k \in \mathbb{N}_0, \tag{3.7}$$

see [4].

3.2 Decay of adaptive Fourier sums

Let us assume now that for a given decaying sequence $\mathbf{c} = \mathbf{c}(f^+) \in \ell_w^2$ of classical Fourier coefficients of $f^+ \in H^2$ we have found the sequence $\tilde{\mathbf{c}}$ of the form (3.4) (or (3.7)) that generates an infinite Hankel matrix of rank $N+1$ as given in Theorem 3.2 satisfying the relation $\|\boldsymbol{\Gamma}_{\mathbf{c}} - \boldsymbol{\Gamma}_{\tilde{\mathbf{c}}}\|_{\ell^2 \to \ell^2} = \sigma_{N+1}(\boldsymbol{\Gamma}_{\mathbf{c}})$ in Theorem 3.1. Then we can conclude that

$$\|\mathbf{c} - \tilde{\mathbf{c}}\|_{\ell^2} \leq \sigma_{N+1}(\boldsymbol{\Gamma}_{\mathbf{c}}). \tag{3.8}$$

This follows easily from the observation

$$\|\mathbf{c} - \tilde{\mathbf{c}}\|_{\ell^2} = \|\boldsymbol{\Gamma}_{\mathbf{c}-\tilde{\mathbf{c}}}\, \mathbf{e}_0\|_{\ell^2} \leq \sup_{\|\mathbf{v}\|_{\ell^2}=1} \|\boldsymbol{\Gamma}_{\mathbf{c}-\tilde{\mathbf{c}}}\, \mathbf{v}\|_{\ell^2} = \|\boldsymbol{\Gamma}_{\mathbf{c}} - \boldsymbol{\Gamma}_{\tilde{\mathbf{c}}}\|_{\ell^2 \to \ell^2},$$

where $\mathbf{e}_0 := (1, 0, 0, \ldots) \in \ell^2$.

Now the question remains, how the singular values of the infinite Hankel matrix $\boldsymbol{\Gamma}_{\mathbf{c}}$ decay compared to the decay of the generating sequence $\mathbf{c}$. Here, the following recent theorem about the decay of singular values of Hankel operators comes to our help, see [23].

Theorem 3.4. (*Pushnitski and Yafaev*). *Let* $\mathbf{c} = (c_k)_{k=0}^\infty \in \ell_w^2$ *satisfy the decay condition*

$$c_k = \mathcal{O}(k^{-1}(\log k)^{-\alpha}), \quad k \to \infty \tag{3.9}$$

for some $\alpha > 0$.

1. If $0 < \alpha < 1/2$, then the singular values of $\Gamma_{\mathbf{c}}$ satisfy the estimate

$$\sigma_N(\Gamma_{\mathbf{c}}) = \mathcal{O}(N^{-\alpha}), \quad N \to \infty,$$

and moreover,

$$\sigma_N(\Gamma_{\mathbf{c}}) = C_\alpha N^{-\alpha} + o(N^{-\alpha})$$

with $C_\alpha = 2^{-\alpha}\pi^{1-\alpha}(B(\frac{1}{2\alpha}, \frac{1}{2}))^\alpha$, where $B(\cdot, \cdot)$ denotes the Beta function.

2. Let Δ denote the forward difference operator with $\Delta c_k := c_{k+1} - c_k$, and let recursively $\Delta^m c_k := \Delta^{m-1} c_{k+1} - \Delta^{m-1} c_k$ for $m > 1$. If $\alpha > 1/2$ and $\mathbf{c}$ satisfies (3.9) and

$$\Delta^m c_k = \mathcal{O}(k^{-1-m}(\log k)^{-\alpha}), \qquad k \to \infty \tag{3.10}$$

for $m = 0, \ldots, \lfloor \alpha \rfloor + 1$, then

$$\sigma_N(\Gamma_{\mathbf{c}}) = \mathcal{O}(N^{-\alpha}).$$

For a proof we refer to [23].

Remark 3.5. The estimates in Theorem 3.4 on the decay of the singular values of the operator $\Gamma_{\mathbf{c}}$ imply that $\Gamma_{\mathbf{c}}$ is in the Schatten class $S_{p,\infty}$ with the norm

$$\|\Gamma_{\mathbf{c}}\|_{S_{p,\infty}} := \sup_{N \in \mathbb{N}} N^{1/p}\sigma_N(\Gamma_{\mathbf{c}}) < \infty$$

for $p = 1/\alpha$. It has been already shown by Peller [19] that $\Gamma_{\mathbf{c}} \in S_{p,p}$ with

$$\|\Gamma_{\mathbf{c}}\|_{S_{p,p}}^p := \sum_{n=1}^{\infty} \sigma_n(\Gamma_{\mathbf{c}})^p < \infty,$$

if and only if $\mathbf{c}$ is a sequence of Fourier coefficients for a function f in the Besov space $B_p^{1/p}$. The above theorem also implies that we can even have a decay of singular values of $\Gamma_{\mathbf{c}}$ being faster than any polynomial. In [28] it has been shown that for $c_k = (1 + k)^{-\gamma}$, $\gamma > 1$, the singular values of $\Gamma_{\mathbf{c}}$ satisfy

$$\sigma_N(\Gamma_{\mathbf{c}}) = \exp\left(-\pi\sqrt{2\gamma N} + o(\sqrt{N})\right), \quad N \to \infty.$$

Using the considerations above, we arrive at the following result.

Theorem 3.6. *Let $f^+(t) = \sum_{k=0}^{\infty} c_k(f^+) \, e^{ikt}$ be a function in H^2. If $\mathbf{c} = (c_k(f^+))_{k=0}^{\infty}$ satisfies the decay conditions (3.9) and (3.10) for $\alpha > 1/2$ then for each $N \in \mathbb{N}$, there exists a vector of zeros $z_0^{(N)}, \ldots, z_N^{(N)}$ determining a Takenaka-Malmquist system $\{B_k\}_{k=0}^N$ such that the N-th adaptive partial Fourier sum $f_N^+ = \sum_{k=0}^N \langle f, B_k \rangle_{L^2} B_k$ satisfies the asymptotic estimate*

$$\|f^+ - f_N^+\|_{L^2} \le CN^{-\alpha}$$

where C does not depend on N.

Proof. From Theorem 2.1 and (3.2) it follows that f_N^+ can be written in the basis using $\{P_k : k = 0, \ldots, N\}$, such that $f_N^+(t) = \sum_{k=0}^{\infty} \tilde{c}_k e^{itk}$, where the coefficients $\tilde{c}_k$ are of the form

$$\tilde{c}_k = \sum_{j=0}^N a_j^{(N)}(z_j^{(N)})^k.$$

Here we have for simplicity assumed that the zeros $z_j^{(N)}$ are pairwise different. The sequence $\tilde{\mathbf{c}} = (\tilde{c}_k)_{k=0}^\infty$ determines a Hankel operator of rank $N + 1$. Thus, if $\tilde{\mathbf{c}}$, (i.e., the zeros $z_j^{(N)}$ and the coefficients $a_j^{(N)}$) are taken such that

$$\|\mathbf{\Gamma_c} - \mathbf{\Gamma_{\tilde{c}}}\|_{\ell^2 \to \ell^2} = \sigma_{N+1}(\mathbf{\Gamma_c}),$$

then it follows from (3.8) that $\|\mathbf{c} - \tilde{\mathbf{c}}\|_{\ell^2} \le \sigma_{N+1}(\mathbf{\Gamma_c})$. By Theorem 3.1, this sequence $\tilde{\mathbf{c}}$ can always be found, and by Theorem 3.4 we have $\sigma_{N+1}(\mathbf{\Gamma_c}) \le \mathcal{O}(N^{-\alpha})$. The assertion finally follows from Parseval identity,

$$\|f^+ - f_N^+\|_{L^2} = \|\mathbf{c} - \tilde{\mathbf{c}}\|_{\ell^2} \le \|\mathbf{\Gamma_c} - \mathbf{\Gamma_{\tilde{c}}}\|_{\ell^2 \to \ell^2} = \sigma_{N+1}(\mathbf{\Gamma_c}) \le C\,N^{-\alpha}$$

with some C being independent from N. $\qquad\square$

4 Computation of adaptive Fourier sums

In this section we derive an algorithm to compute the partial adaptive Fourier sums of f^+ in H^2. For this purpose , we suppose that we know the vector of Fourier coefficients $(c_k(f^+))_{k=0}^L$ in the classical Fourier expansion for a sufficiently large L.

If f^+ is a rational function in H^2 of the form $q(\mathrm{e}^{-it})/p(\mathrm{e}^{-it})$ with $\deg q \le \deg p = M$ and p possesses pairwise different zeros, then our algorithm provides optimal adaptive partial sums for any length $N + 1 < M + 1$ to approximate f. Note that for $N = M$ we can represent f^+ exactly, i.e., $f_M^+ = f^+$. If f is not rational, then the theoretical results in the last subsection imply that we still find a very good approximations of f^+ by adaptive Fourier sums.

First, let us assume that the sequence of Fourier coefficients $\mathbf{c}(f) = \mathbf{c} = (c_k)_{k=0}^\infty$ of some $f^+ \in H^2$ possesses the form

$$c_k = c_k(f^+) = \sum_{j=0}^M \breve{a}_j\, \breve{z}_j^k, \qquad k \in \mathbb{N}_0, \tag{4.1}$$

with possibly large M and with pairwise different zeros $\breve{z}_j \in \mathbb{D}$, and $\breve{a}_j \in \mathbb{C}$. However, we assume that we are given only the classical Fourier coefficients $c_k = c_k(f^+)$ for $k = 0, \ldots, L$ with $L \ge 2M + 3$ and will compute the parameters $\breve{a}_j\ \breve{z}_j$, $j = 0, \ldots, M$ in a first step by Prony's method. In this section, we present an algorithm to solve the problem of finding an optimal sequence $\tilde{\mathbf{c}} = \tilde{\mathbf{c}}^{(N)} = (\tilde{c}_k)_{k=0}^\infty$ of the form

$$\tilde{c}_k = \sum_{j=0}^N a_j^{(N)}\, (z_j^{(N)})^k \tag{4.2}$$

for a pre-defined $N < M$ such that

$$\|\mathbf{c} - \tilde{\mathbf{c}}\|_{\ell^2} \le \|\mathbf{\Gamma_c} - \mathbf{\Gamma_{\tilde{c}}}\|_{\ell^2 \to \ell^2} = \sigma_{N+1}(\mathbf{\Gamma_c})$$

holds, i.e., $\tilde{\mathbf{c}}$ is the optimal sequence for approximation by Theorem 3.1.

Once we have computed the sequence $\tilde{\mathbf{c}}$ in (4.2) by determining the appropriate parameters $a_j^{(N)}$, $z_j^{(N)}$, $j = 0, \ldots, N$, we can find according to (3.2) the adaptive N-th Fourier sum

$$f_N^+ = \sum_{j=0}^N a_j^{(N)} P_j$$

where the coefficients $a_j^{(N)}$, $j = 0, \ldots, N$ appear in (4.2) and where the P_j are defined by the $z_j^{(N)}$ in (4.2). If we are interested in representing the approximation f_N^+ of f^+ in terms of the Takenaka-Malmquist basis, then we need to use $z_0^{(N)}, \ldots z_N^{(N)}$ from (4.2) to define the basis $\{B_\ell\}_{\ell=0}^N$ and the corresponding coefficients $b_\ell^{(N)}$ in the representation $f_N^+ = \sum_{\ell=0}^N b_\ell^{(N)} B_\ell$ can be obtained by employing the basis transform as given in (2.4).

Our approach to find appropriate parameters $z_j^{(N)}$ and $a_j^{(N)}$ in (4.2) consists of two steps. In the first step we reconstruct the parameters $\breve{a}_j$ and $\breve{z}_j$ in (4.1) using a Prony-like method, see e.g. [21, 22, 26].

Remark 4.1. If f^+ is rational then the representation (4.1) of $c_k(f^+)$ is exact. However, we emphasize that the idea can be applied in practice to all sequences $c_k(f^+)$ with sufficient decay. In the general case we obtain in the first step a very good approximation of the form (4.1) if M is taken large enough. In our practical experiments it has been usually sufficient to take e.g. M between 10 and 20 to get approximations with accuracy 10^{-8} or better.

Once the representation (4.1) of $\mathbf{c}$ is known we will proceed in the second step with a new algorithm to compute the sequence $\tilde{\mathbf{c}}$ and particularly the parameters $a_j^{(N)}$ and $z_j^{(N)}$, $j = 0, \ldots, N$, in the shorter exponential sum (4.2) with $N < M$ such that $\|\mathbf{c} - \tilde{\mathbf{c}}\|_{\ell^2} \leq \sigma_{N+1}(\mathbf{\Gamma_c})$. Our method includes a numerical procedure to compute all nonzero singular values of $\mathbf{\Gamma_c}$. Knowing these singular values we will be able to give an a priori estimate of the error $\|\mathbf{c} - \tilde{\mathbf{c}}\|_{\ell^2}$ that can be achieved using the new sequence $\tilde{\mathbf{c}}^{(N)}$ built by the shorter exponential sum in (4.2) and thus of the L^2-error $\|f^+ - f_N^+\|_{L^2}$ achieved by the N-th adaptive Fourier sum f_N^+.

Step 1: Computation of the representation (4.1) by Prony's method. Let us shortly summarize Prony's method for the recovery of exponential sums. Assume, we are given the coefficients $c_k = c_k(f^+)$, $k = 0, 1, \ldots, L$, with $L \geq 2M + 3$, and we want to compute $\breve{a}_j$ and $\breve{z}_j$ in (4.1). First, we introduce the characteristic polynomial

$$P(z) := \prod_{j=0}^M (z - \breve{z}_j) = \sum_{k=0}^M p_k z^k + p_{M+1} z^{M+1}$$

with the help of the (unknown) zeros $\breve{z}_j$, where it is clear that a monomial representation with coefficients p_k of $P(z)$ exists and $p_{M+1} = 1$. Using (4.1), we observe now that

$$
\begin{aligned}
\sum_{k=0}^{M+1} p_k c_{k+m} &= \sum_{k=0}^{M+1} p_k \Big(\sum_{j=0}^M \breve{a}_j \breve{z}_j^{k+m} \Big) = \sum_{j=0}^M \breve{a}_j \breve{z}_j^m \Big(\sum_{k=0}^{M+1} p_k \breve{z}_j^k \Big) \\
&= \sum_{j=0}^M \breve{a}_j \breve{z}_j^m P(\breve{z}_j) = 0
\end{aligned}
$$

is true for all $m \geq 0$ and with the given coefficients c_k we obtain the homogeneous linear system

$$
\mathbf{H_c p} = \begin{pmatrix} c_0 & c_1 & \ldots & c_{M+1} \\ c_1 & c_2 & \ldots & c_{M+2} \\ \vdots & \vdots & & \vdots \\ c_{L-M-1} & c_{L-M} & \ldots & c_L \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \\ \vdots \\ p_{M+1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},
$$

where the coefficient matrix of size $(L - M) \times (M + 2)$ has Hankel structure, and $\mathbf{p} = (p_0, \ldots, p_{M+1})^T$ contains the coefficients of the characteristic polynomial $P(z)$. Since the zeros $\check{z}_j$ are assumed to be pairwise different, the Hankel matrix $\mathbf{H_c}$ has rank $M+1$ such that $\mathbf{p}$ is uniquely defined as the singular vector of $\mathbf{H_c}$ corresponding to the singular value 0 with the normalization $p_{M+1} = 1$. The rank condition for $\mathbf{H_c}$ follows from a factorization with Vandermonde matrices that can be directly obtained from (4.1),

$$\mathbf{H_c} = \mathbf{V}_{L-M}(\check{\mathbf{z}}) \operatorname{diag}(\check{a}_1, \ldots, \check{a}_M) \mathbf{V}_{M+2}(\check{\mathbf{z}})^T$$

where

$$\mathbf{V}_K(\check{\mathbf{z}}) := \begin{pmatrix} 1 & 1 & \ldots & 1 \\ \check{z}_0 & \check{z}_1 & \ldots & \check{z}_M \\ \vdots & \vdots & & \vdots \\ \check{z}_0^{K-1} & \check{z}_1^{K-1} & \ldots & \check{z}_M^{K-1} \end{pmatrix} \in \mathbb{C}^{K \times K}.$$

These observations lead to the following algorithm to recover the parameters $\check{z}_j$, $\check{a}_j$ in (4.1) from the coefficients c_k, $k = 0, \ldots, L$.

Algorithm 4.2. (Prony method)
Input: M, c_k, $k = 0, \ldots, L$ with $L \geq 2M + 3$.

1. Compute the right singular vector $\mathbf{p} = (p_0, \ldots, p_M, 1)^T$ of $\mathbf{H_c}$ corresponding to its smallest singular value.

2. Determine the polynomial $P(z)$ with coefficient vector $\mathbf{p}$ and compute its $M+1$ zeros $\check{z}_j$, $j = 0, \ldots, M$.

3. Solve the overdetermined linear system

$$\sum_{j=0}^{M} \check{a}_j \check{z}_j^k = c_k, \qquad k = 0, \ldots, L$$

using a least squares approach and determine $\check{a}_j$, $j = 0, \ldots, M$.

Output: Parameters $\check{a}_j$, $\check{z}_j$, $j = 0, \ldots, M$.

Remarks 4.3. While theoretically, we are searching for the singular vector of $\mathbf{H_c}$ to the singular value zero, in practice, we do not obtain the zero singular value exactly. Usually, the considered sequence of Fourier coefficients will not exactly be of the form (4.1), but we are interested in an approximation of the form (4.1). Therefore we employ the singular vector corresponding to the smallest singular value of $\mathbf{H_c}$ in the first step. The zeros of the Prony polynomial are numerically found by solving a corresponding eigenvalue problem for the companion matrix of $P(z)$. Stable algorithms for Prony's method usually employ more than $2M + 3$ values and use e.g. singular value decompositions of (rectangular) Hankel matrices as well as matrix pencil methods for evaluating the zeros $\check{z}_j$, $j = 0, \ldots, M$. In our numerical tests, we particularly use the APM method for stable evaluation of the parameters, see [22].

Step 2: New algorithm to compute the approximating sequence $\tilde{\mathbf{c}}$ in (4.2).
In the second step we will derive a method to compute an optimal approximation $\tilde{\mathbf{c}}$ of $\mathbf{c}$ using a shorter exponential sum. For that purpose, we want to apply the following special theorem for infinite Hankel matrices with finite rank, see [1]. To formulate

this result, we first recall the notation of con-eigenvalues and con-eigenvectors and its correspondence to singular pairs of matrices.

Observe that $\mathbf{\Gamma_c}$ 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. [13] for the finite-dimensional case.

For an infinite Hankel matrix $\mathbf{\Gamma_c}$ we call $\lambda \in \mathbb{C}$ a *con-eigenvalue* with the corresponding *con-eigenvector* $\mathbf{v} \in \ell^2$ if it satisfies

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

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

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

for all $\alpha \in \mathbb{R}$. Thus, for each con-eigenvalue λ of $\mathbf{\Gamma_c}$ 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.

The symmetric infinite Hankel matrix $\mathbf{\Gamma_c}$ with finite rank is compact and unitarily con-diagonalizable, see [13]. Since $\mathbf{\Gamma_c}\overline{\mathbf{v}} = \lambda\,\mathbf{v}$ implies

$$(\mathbf{\Gamma_c^*\Gamma_c})\mathbf{v} = \mathbf{\Gamma_c^*}\lambda\overline{\mathbf{v}} = \lambda\overline{\mathbf{\Gamma_c}}\,\overline{\mathbf{v}} = \lambda\overline{\mathbf{\Gamma_c}\overline{\mathbf{v}}} = |\lambda|^2\mathbf{v},$$

we directly observe that the nonnegative con-eigenvalues and con-eigenvectors of $\mathbf{\Gamma_c}$ are also singular values and corresponding singular vectors of $\mathbf{\Gamma_c}$. 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.

Theorem 4.4. *Let the Hankel matrix $\mathbf{\Gamma_c}$ of rank $M+1$ be generated by the sequence $\mathbf{c}$ of the form (4.1) with $1 > |\breve{z}_0| \geq \cdots \geq |\breve{z}_M| > 0$. Let the $M + 1$ nonzero singular values of $\mathbf{\Gamma_c}$ be ordered by size $\sigma_0 \geq \sigma_1 \ldots \geq \sigma_M > 0$. Then, for each $N \in \{0, \ldots, M-1\}$ satisfying $\sigma_{N+1} \neq \sigma_k$ for $N + 1 \neq k$ the Laurent polynomial of the con-eigenvector $\mathbf{v}^{(N+1)} = (v_l^{(N+1)})_{l=0}^\infty$ corresponding to σ_{N+1},*

$$P_{\mathbf{v}^{(N+1)}}(z) := \sum_{l=0}^\infty v_l^{(N+1)}\,z^l,$$

has exactly $N + 1$ zeros $z_0^{(N)}, \ldots, z_N^{(N)}$ in $\mathbb{D}$, repeated according to their multiplicity. Furthermore, if $z_0^{(N)}, \ldots, z_N^{(N)}$ are pairwise different, then there exist coefficients $a_0^{(N)}, \ldots, a_N^{(N)} \in \mathbb{C}$ such that for

$$\tilde{\mathbf{c}}^{(N)} = \left(\tilde{c}_l^{(N)}\right)_{l=0}^\infty = \left(\sum_{j=0}^N a_j^{(N)}(z_j^{(N)})^l\right)_{l=0}^\infty \tag{4.3}$$

we have

$$\|\mathbf{c} - \tilde{\mathbf{c}}^{(N)}\|_{\ell^2} \leq \sigma_{N+1}.$$

For a proof of this theorem we refer to [1] and to [3].

We will present now a computational method for the explicit computation of the parameters $a_j^{(N)}$ and $z_j^{(N)}$, $j = 0, \ldots, N$ of $\tilde{\mathbf{c}}^{(N)}$ in (4.3). In order to achieve this goal we need to perform the following steps.

Step i) We have to find a numerical procedure to compute the singular pairs $(\sigma_n, \mathbf{v}^{(n)})$ of the infinite Hankel matrix $\mathbf{\Gamma_c}$ for $n = 0, \ldots, M$.

Step ii) We have to find all zeros $z_j^{(N)}$ of the expansion $P_{\mathbf{v}^{(N+1)}}(z)$ lying inside $\mathbb{D}$.

Step iii) In a final step we have to compute the optimal coefficients $a_j^{(N)}$ in (4.3).

Step i) To derive the numerical procedure to compute the singular pairs of $\mathbf{\Gamma_c}$, we investigate the special structure of the con-eigenvectors of $\mathbf{\Gamma_c}$ corresponding to the non-zero con-eigenvalues (resp. singular values). We can show the following result that provides us with an algorithm to compute all non-zero singular values of $\mathbf{\Gamma_c}$ and the corresponding con-eigenvectors exactly by solving a con-eigenvalue problem of size $M + 1$.

Theorem 4.5. *Let* $\mathbf{c}$ *be of the form (4.1). Then the con-eigenvector* $\mathbf{v}^{(l)} = (v_k^{(l)})_{k=0}^\infty$ *of* $\mathbf{\Gamma_c}$ *corresponding to a single nonzero singular value* σ_l *of* $\mathbf{\Gamma_c}$, $l \in \{0, 1, \ldots, M\}$ *is given by*

$$\overline{v}_k^{(l)} = \frac{1}{\sigma_l} \sum_{j=0}^M \breve{a}_j b_j^{(l)} \breve{z}_j^k, \qquad k \in \mathbb{N}_0, \tag{4.4}$$

where the vector $\mathbf{b}^{(l)} = (b_0^{(l)}, \ldots, b_M^{(l)})^T$ *is the con-eigenvector of the finite con-eigenvalue problem*

$$\mathbf{A}_{M+1} \mathbf{Z}_{M+1} \mathbf{b} = \sigma_l \overline{\mathbf{b}} \tag{4.5}$$

with

$$\mathbf{A}_{M+1} := \begin{pmatrix} \breve{a}_0 & & & 0 \\ & \breve{a}_1 & & \\ & & \ddots & \\ 0 & & & \breve{a}_M \end{pmatrix}, \quad \mathbf{Z}_{M+1} := \begin{pmatrix} \frac{1}{1-|\breve{z}_0|^2} & \frac{1}{1-\breve{z}_0\overline{\breve{z}}_1} & \cdots & \frac{1}{1-\breve{z}_0\overline{\breve{z}}_M} \\ \frac{1}{1-\overline{\breve{z}}_0\breve{z}_1} & \frac{1}{1-|\breve{z}_1|^2} & \cdots & \frac{1}{1-\breve{z}_1\overline{\breve{z}}_M} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{1-\overline{\breve{z}}_0\breve{z}_M} & \frac{1}{1-\overline{\breve{z}}_1\breve{z}_M} & \cdots & \frac{1}{1-|\breve{z}_M|^2} \end{pmatrix}.$$

Proof. Let $(\sigma_l, \mathbf{v}^{(l)})$ with $\sigma_l \neq 0$ be a con-eigenpair of $\mathbf{\Gamma_c}$, i.e., $\mathbf{\Gamma_c}\overline{\mathbf{v}}^{(l)} = \sigma \mathbf{v}^{(l)}$. We determine $b_j^{(l)} := \sum_{r=0}^\infty v_r^{(l)} \breve{z}_j^r = P_{\mathbf{v}^{(l)}}(\breve{z}_j)$ for $j = 0, \ldots, M$. Then it follows by (4.1) that

$$\begin{aligned} \sigma_l \overline{v}_k^{(l)} &= (\mathbf{\Gamma_c}\mathbf{v}^{(l)})_k = \sum_{r=0}^\infty c_{k+r} v_r^{(l)} = \sum_{r=0}^\infty \sum_{j=0}^M \breve{a}_j \breve{z}_j^{k+r} v_r^{(l)} \\ &= \sum_{j=0}^M \breve{a}_j \Big(\sum_{r=0}^\infty v_r^{(l)} \breve{z}_j^r \Big) \breve{z}_j^k = \sum_{j=0}^M \breve{a}_j b_j^{(l)} \breve{z}_j^k \end{aligned}$$

for all $k \in \mathbb{N}_0$, and hence (4.4) is true. Relation (4.5) is now a consequence of (4.4) observing that

$$\sigma_l \overline{b}_k^{(l)} = \sigma_l \sum_{r=0}^\infty \overline{v}_r^{(l)} (\overline{\breve{z}}_k)^r = \sum_{r=0}^\infty \Big(\sum_{j=0}^M \breve{a}_j b_j^{(l)} \breve{z}_j^r \Big) (\overline{\breve{z}}_k)^r = \sum_{j=0}^M \frac{\breve{a}_j b_j^{(l)}}{1 - \breve{z}_j \overline{\breve{z}}_k} \tag{4.6}$$

for $k = 1, \ldots, M$. $\qquad\square$

Step ii) In particular, it follows from Theorem 4.5 for all $N = 0, \ldots, M$ that the Laurent polynomial

$$P_{\mathbf{v}^{(N+1)}}(z) := \sum_{k=0}^\infty v_k^{(N+1)} z^k = \frac{1}{\sigma_{N+1}} \sum_{k=0}^\infty \Big(\sum_{j=0}^M \overline{\breve{a}}_j \overline{b}_j^{(N+1)} \overline{\breve{z}}_j^k \Big) z^k = \frac{1}{\sigma_{N+1}} \sum_{j=0}^M \frac{\overline{\breve{a}}_j \overline{b}_j^{(N+1)}}{1 - \overline{\breve{z}}_j z} \tag{4.7}$$

is a rational function with numerator being a polynomial of degree at most M. Thus, in order to find the zeros of $P_{\mathbf{v}^{(N+1)}}(z)$ we only need to compute the M zeros of the numerator in this rational representation. By Theorem 4.4 we will find exactly $N+1$ zeros $z_0^{(N)}, \ldots, z_N^{(N)}$ in $\mathbb{D}$.

Step iii) Once we have computed the zeros $z_j^{(N)}$, $j = 0, \ldots, N$, we can find the optimal coefficients $a_j^{(N)}$ by solving a least squares problem.

We combine our observations with Theorem 4.4 and obtain the following algorithm to compute all parameters of the desired approximation $\widetilde{\mathbf{c}}^{(N)}$.

Algorithm 4.6. Algorithm to compute the representation (4.2).

Input: Representation (or approximation) (4.1) of $\mathbf{c}$ with $\breve{z}_j \in \mathbb{D}$ and $\breve{a}_j \in \mathbb{C} \setminus \{0\}$, $j = 0, \ldots, M$,

      target approximation error ϵ.

1. Solve the con-eigenproblem for the matrix $\mathbf{A}_{M+1}\mathbf{Z}_{M+1}$, determine N such that σ_{N+1} is the largest singular value with $\sigma_{N+1} < \epsilon$ and determine the corresponding con-eigenvector $\mathbf{v}^{(N+1)}$.

2. Compute the $N + 1$ zeros $z_j^{(N)} \in \mathbb{D}$, $j = 0, \ldots, N$, of the con-eigenpolynomial $P_{\mathbf{v}^{(N+1)}}(z)$ of $\mathbf{\Gamma_c}$ using its rational representation (4.7).

3. Compute the coefficients $a_j^{(N)}$ by solving the minimization problem

$$\min_{\tilde{a}_0, \ldots, \tilde{a}_N} \|\mathbf{c} - \tilde{\mathbf{c}}\|_{\ell^2}^2 = \min_{\tilde{a}_0, \ldots, \tilde{a}_N} \sum_{k=1}^{\infty} |c_k - \sum_{j=1}^{N} \tilde{a}_j (z_j^{(N)})^k|^2.$$

Output: parameters $z_j^{(N)}$, $a_j^{(N)}$, $j = 0, \ldots, N$, to determine $\tilde{\mathbf{c}}$ of the form (4.2) such that $\|\mathbf{c} - \tilde{\mathbf{c}}\|_{\ell^2} \leq \sigma_{N+1} < \epsilon$.

Remarks 4.7.
1. Note that a related idea of dimension reduction has been used by Beylkin and Monzón in [6]. But in contrast to the above approach, they considered the rank reduction of a finite Hankel matrix.
2. Since $\mathbf{A}_{M+1}\mathbf{Z}_{M+1}$ is con-diagonalizable by Theorem 4.5, it follows that

$$\mathbf{A}_{M+1}\mathbf{Z}_{M+1}\mathbf{B} = \overline{\mathbf{B}}\,\mathbf{\Sigma}, \quad \text{and} \quad \overline{\mathbf{A}_{M+1}\mathbf{Z}_{M+1}}\,\overline{\mathbf{B}} = \mathbf{B}\,\mathbf{\Sigma}$$

with $\mathbf{B} = (\mathbf{b}^{(0)} \ldots \mathbf{b}^{(M)})$ containing the con-eigenvectors of $\mathbf{A}_{M+1}\mathbf{Z}_{M+1}$ and with $\mathbf{\Sigma} = \operatorname{diag}(\sigma_0, \ldots, \sigma_M)$. Therefore

$$\overline{\mathbf{A}_{M+1}\mathbf{Z}_{M+1}}\mathbf{A}_{M+1}\mathbf{Z}_{M+1}\mathbf{B} = \overline{\mathbf{A}_{M+1}\mathbf{Z}_{M+1}}\,\overline{\mathbf{B}}\,\mathbf{B}^{-1}\mathbf{A}_{M+1}\mathbf{Z}_{M+1}\mathbf{B} = \mathbf{B}\mathbf{\Sigma}^2,$$

i.e., $\overline{\mathbf{A}_{M+1}\mathbf{Z}_{M+1}}\mathbf{A}_{M+1}\mathbf{Z}_{M+1}$ has only real nonnegative eigenvalues $\lambda_j = \sigma_j^2$.

Conversely, if $(\lambda_j, \mathbf{w}^{(j)})$ is an eigenpair of $\overline{\mathbf{A}_{M+1}\mathbf{Z}_{M+1}}\mathbf{A}_{M+1}\mathbf{Z}_{M+1}$, then $\mathbf{b}^{(j)} := \mathbf{A}_{M+1}\mathbf{Z}_{M+1}\mathbf{w}^{(j)} + \sigma_j\overline{\mathbf{w}}^{(j)}$ is a con-eigenvector of $\mathbf{A}_{M+1}\mathbf{Z}_{M+1}$ to the con-eigenvalue $\sigma_j = \sqrt{\lambda_j}$, since

$$\begin{aligned} \mathbf{A}_{M+1}\mathbf{Z}_{M+1}\overline{\mathbf{b}}^{(j)} &= \mathbf{A}_{M+1}\mathbf{Z}_{M+1}(\overline{\mathbf{A}}_{M+1}\overline{\mathbf{Z}}_{M+1}\overline{\mathbf{w}}^{(j)} + \sigma_j\mathbf{w}^{(j)}) \\ &= \overline{\overline{\mathbf{A}_{M+1}\mathbf{Z}_{M+1}}\mathbf{A}_{M+1}\mathbf{Z}_{M+1}\mathbf{w}^{(j)}} + \sigma_j\mathbf{A}_{M+1}\mathbf{Z}_{M+1}\mathbf{w}^{(j)} \end{aligned}$$

$$= \sigma_j^2 \overline{\mathbf{w}}^{(j)} + \sigma_j \mathbf{A}_{M+1}\mathbf{Z}_{M+1}\mathbf{w}^{(j)} = \sigma_j \mathbf{b}^{(j)}.$$

Thus, to solve the con-eigenvalue problem in step 2 of Algorithm 4.6, we can employ a usual eigenvalue-decomposition of $\overline{\mathbf{A}_{M+1}\mathbf{Z}_{M+1}}\mathbf{A}_{M+1}\mathbf{Z}_{M+1}$.

3. Usually, the singular values of $\mathbf{A}_{M+1}\mathbf{Z}_{M+1}$ decay very fast. There are several possibilities to improve the numerical stability of the necessary SVD for this matrix product. Indeed, $\mathbf{A}_{M+1}\mathbf{Z}_{M+1}$ can be understood as a Cauchy matrix since

$$(\mathbf{A}_{M+1}\mathbf{Z}_{M+1})_{jk} = \frac{\breve{a}_j/\breve{z}_j}{1/\breve{z}_j - \overline{\breve{z}}_k}, \qquad j, k = 0, \dots, M.$$

For accurate singular value decomposition we refer to the algorithm proposed by Demmel [9] that is based on a rank revealing decomposition with complexity $\mathcal{O}(M^3)$. Further ideas to overcome instability issues can be found in [11] where the con-eigenvalue problem is explicitly solved.

4. Observing that the components c_k have the form (4.1) and the $\tilde{c}_k$ have the form (4.2), the ℓ^2-minimization problem for infinite sequences in step 3 of Algorithm 4.6 can be reformulated as a finite-dimensional least squares problem,

$$\min_{\tilde{a}_0,\dots,\tilde{a}_N} \|\mathbf{c} - \tilde{\mathbf{c}}\|_{\ell^2}^2 = \min_{\tilde{a}_0,\dots,\tilde{a}_N} \left(-2\mathrm{Re}\sum_{k=0}^{\infty} \tilde{c}_k \overline{c}_k + \sum_{k=0}^{\infty} |\tilde{c}_k|^2 \right)$$

$$= \min_{\tilde{a}_0,\dots,\tilde{a}_N} \left(-2\mathrm{Re}\sum_{j=0}^{N}\sum_{i=0}^{M} \tilde{a}_j \overline{\tilde{a}}_i \sum_{k=0}^{\infty} (z_j^{(N)}\overline{\breve{z}}_i)^k + \sum_{j=0}^{N}\sum_{i=0}^{N} \tilde{a}_j \overline{\tilde{a}}_i \sum_{k=0}^{\infty} (z_j^{(N)}\overline{z}_i^{(N)})^k \right)$$

$$= \min_{\tilde{a}_0,\dots,\tilde{a}_N} \left(-2\mathrm{Re}\sum_{j=0}^{N}\sum_{i=0}^{M} \frac{\tilde{a}_j \overline{\tilde{a}}_i}{1 - z_j^{(N)}\overline{\breve{z}}_i} + \sum_{j=0}^{N}\sum_{i=0}^{N} \frac{\tilde{a}_j \overline{\tilde{a}}_i}{1 - z_j^{(N)}\overline{z}_i^{(N)}} \right),$$

where the values $\overline{\breve{z}}_i$, $z_j^{(N)}$ and $\overline{\breve{a}}_i$ are known. In practice, if only $L+1$ sequence values c_k, $k = 0, \dots, L$ are given, then the minimization problem may be replaced by

$$\min_{\tilde{a}_0,\dots,\tilde{a}_N} \sum_{k=0}^{L} |c_k - \tilde{c}_k|^2 = \min_{\tilde{a}_0,\dots,\tilde{a}_K} \sum_{k=0}^{L} |c_k - \sum_{j=0}^{N} \tilde{a}_j (z_j^{(N)})^k|^2.$$

5 Greedy approach for adaptive Fourier sums

We want to compare the algorithm proposed in the last section with a greedy approach. This method is very simple to implement but less efficient as we will show in the numerical experiments in Section 6. Again, we assume that we know a vector $(c_k(f))_{k=0}^{L}$ of Fourier coefficients of f with sufficiently large L.

The idea is now to approximate the vector $\mathbf{c}^{(0)} := (c_k(f))_{k=0}^{L}$ by a sequence of the form $(a_0 z_0^k)_{k=0}^{L}$ where we try to choose a_0 and z_0 in an optimal way. Then we consider the residual

$$\mathbf{c}^{(1)} := \mathbf{c}^{(0)} - (a_0 z_0^k)_{k=0}^{L}$$

and repeat this process. The iteration stops once we have achieved

$$\|\mathbf{c}^{(N)}\|_2 < \epsilon$$

for some $N \leq L$ and predefined tolerance ϵ.

At each iteration step we do the following. For given data $\mathbf{c}^{(j)}$ we apply Prony's method in Algorithm 4.2 for the special case $M = 1$. In this case, the characteristic polynomial is given by

$$P^{(j)}(z) = (z - z_j) = p_1^{(j)} z + p_0^{(j)},$$

i.e., $p_0^{(j)} = -z_j$ and $p_1^{(j)} = 1$. The Hankel matrix $\mathbf{H}_{\mathbf{c}^{(j)}}$ is now of the form

$$\mathbf{H}_{\mathbf{c}^{(j)}} = \begin{pmatrix} c_0^{(j)} & c_1^{(j)} \\ c_1^{(j)} & c_2^{(j)} \\ \vdots & \vdots \\ c_{L-1}^{(j)} & c_L^{(j)} \end{pmatrix} \tag{5.1}$$

and we have to compute $(p_0^{(j)}, p_1^{(j)})^T$ as the singular vector of $\mathbf{H}_{\mathbf{c}^{(j)}}$ to the smallest singular value with the normalization $p_1^{(j)} = 1$. Observe that $(\mathbf{H}_{\mathbf{c}^{(j)}})^* \mathbf{H}_{\mathbf{c}^{(j)}}$ is only of size 2×2. Having fixed z_j in this way, we compute the constant a_j by solving the overdetermined equation system

$$\mathbf{c}^{(j)} = a_j \, (z_j^k)_{k=0}^L$$

with the result

$$a_j = \frac{(\mathbf{c}^{(j)})^T \mathbf{c}^{(j)}}{(\mathbf{c}^{(j)})^T (z_j^k)_{k=0}^L}. \tag{5.2}$$

Finally, we compute the residuum $\mathbf{c}^{(j+1)} := \mathbf{c}^{(j)} - (a_j z_j^k)_{k=0}^L$. The complete algorithm is summarized as follows.

Algorithm 5.1. Greedy algorithm for adaptive Fourier sums.
Input: c_k, $k = 0, \ldots, L$ for sufficiently large L.

1. Initialize $\mathbf{c} = \mathbf{c}^{(0)} := (c_k)_{k=0}^L$ and $j := 0$.
2. While $\|\mathbf{c}\|_2 > \epsilon$
   a) Compute the right singular vector $\mathbf{p} = (-z_j, 1)^T$ of $\mathbf{H}_{\mathbf{c}}$ in (5.1) corresponding to its smallest singular value.
   b) Compute the coefficient a_j according to (5.2).
   c) Update
   $$\mathbf{c} := \mathbf{c} - (a_j z_j^k)_{k=0}^L$$
   and $j := j + 1$.
3. Set $N := j - 1$.

Output: Parameters N, a_j, z_j, $j = 0, \ldots, N$.

Alternatively, the update of the coefficients a_j in iteration step 2b) can be done in a global way to obtain the best approximation that can be achieved with an adaptive basis with the zeros $z_0, \ldots, z_j$ found so far. In this case, we replace the step 2b) by solving the least squares problem

$$\min_{a_0, \ldots, a_j} \sum_{k=0}^L |c_k - \sum_{\ell=0}^j a_\ell \, z_\ell^k|^2.$$

In our numerical experiments, we will compare this simple approach with the results of Algorithms 4.2 and 4.6 in Section 4.
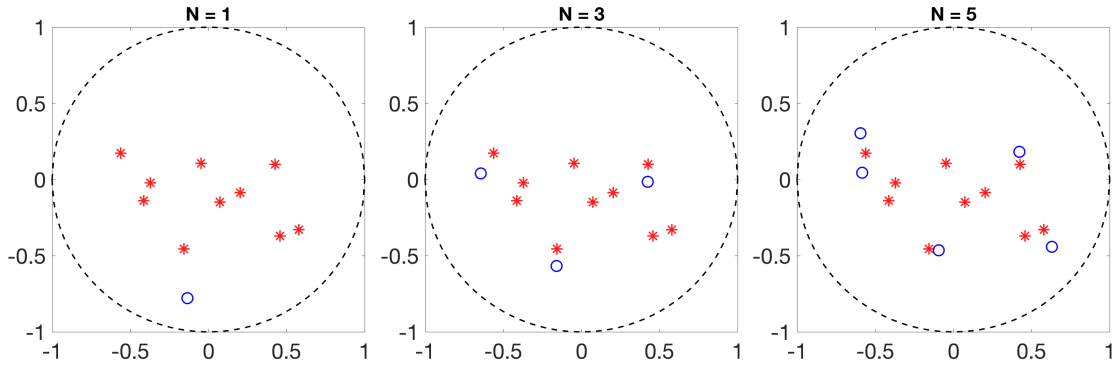
Figure 1 *The original nodes $z_1, \ldots, z_{10}$ (red stars) and the new nodes $z_1^{(N)}, \ldots, z_N^{(N)}$ obtained by Algorithm 4.6 (blue circles)*

6 Numerical Examples

In this section we present two numerical examples demonstrating the performance of our algorithms. In both examples we set a priori $z_0^{(N)} = 0$ and take the Fourier coefficient $c_0(f)$ independently. Then we approximate the Fourier coefficients $c_k(f)$ as described in the previous subsections, but starting with $k = 1$.

Example 1. We approximate the rational function $f(t) = \sum\limits_{\ell=1}^{10} \frac{\breve{a}_\ell}{(1 - \breve{z}_\ell e^{it})}$ with the classical Fourier expansion $f(t) = \sum_{k=0}^{\infty} c_k(f) e^{ikt}$ where

$$c_k(f) = \sum_{\ell=1}^{10} \breve{a}_\ell \, \breve{z}_\ell^k, \quad k \in \mathbb{N}_0. \tag{6.1}$$

In this example, we have $\breve{a}_0 = 0$, $\breve{z}_0 = 0$, and the further parameters $\breve{z}_j$, $\breve{a}_j$, $j = 1, \ldots, 10$, in the representation (6.1) have been obtained by applying a generator of random numbers in $\mathbb{D}$ and are given as follows:

$$
\begin{pmatrix} \breve{z}_1 \\ \breve{z}_2 \\ \breve{z}_3 \\ \breve{z}_4 \\ \breve{z}_5 \\ \breve{z}_6 \\ \breve{z}_7 \\ \breve{z}_8 \\ \breve{z}_9 \\ \breve{z}_{10} \end{pmatrix}
=
\begin{pmatrix}
-0.5609 + 0.1737i \\
0.0734 - 0.1485i \\
0.4582 - 0.3709i \\
0.2030 - 0.0861i \\
-0.3715 - 0.0216i \\
-0.1573 - 0.4553i \\
-0.0471 + 0.1074i \\
0.5780 - 0.3286i \\
-0.4123 - 0.1385i \\
0.4266 + 0.0996i
\end{pmatrix},
\quad
\begin{pmatrix} \breve{a}_1 \\ \breve{a}_2 \\ \breve{a}_3 \\ \breve{a}_4 \\ \breve{a}_5 \\ \breve{a}_6 \\ \breve{a}_7 \\ \breve{a}_8 \\ \breve{a}_9 \\ \breve{a}_{10} \end{pmatrix}
=
\begin{pmatrix}
-0.2978 + 0.4876i \\
-0.2515 + 0.4556i \\
0.1405 + 0.3813i \\
-0.2817 - 0.2871i \\
0.3893 + 0.1132i \\
0.5792 - 0.5730i \\
0.2763 - 0.0897i \\
-0.1873 - 0.2247i \\
0.1009 - 0.4062i \\
-0.4707 - 0.3855i
\end{pmatrix}.
$$

We denote by $S_N(f)$ the N-th classical Fourier sum

$$S_N(f)(t) = \sum_{k=0}^{N} c_k(f) e^{ikt},$$

and by

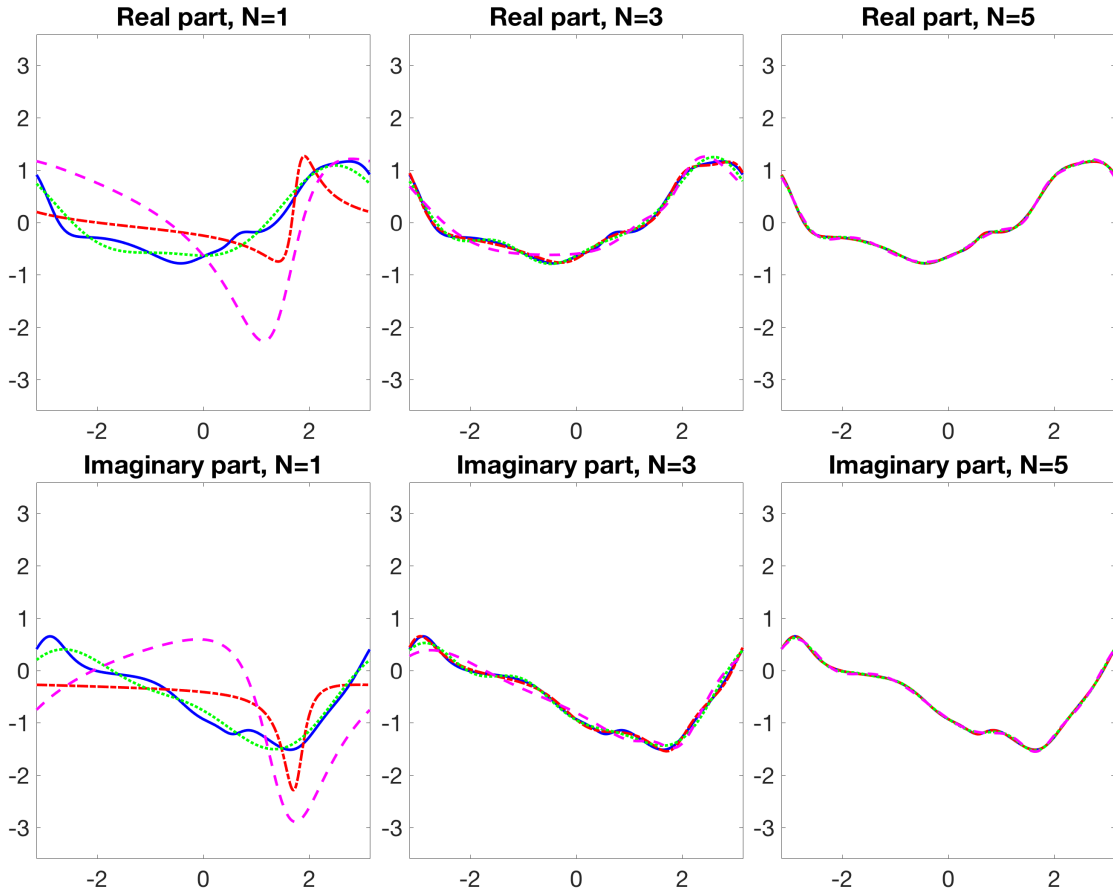$$f_N := c_0(f) + \sum_{j=1}^{N} a_j^{(N)} P_j$$

Figure 2 *The real and imaginary parts of the original function f (blue solid line), the Fourier approximation $S_N(f)$ (green dotted line), the Greedy adaptive Fourier sum $f_N^{(G)}$ (magenta dashed line) and adaptive Fourier sum f^N (red dash-dot line) for Example 1.*

the N-th adaptive Fourier sum. Here we have taken $z_0^{(N)} = 0$, $a_0^{(N)} = c_0(f)$. The parameters $a_j^{(N)}$ and the zeros $z_j^{(N)}$, $j = 1, \ldots, N$, determining the adaptive Fourier basis $\{P_j\}_{j=1}^N$ have been computed by Algorithm 4.6 for $N = 1, \ldots, 9$. The obtained nodes $z_j^{(N)}$ are displayed for $N = 1$, $N = 3$ and $N = 5$ in Figure 1. Further, we denote the sequence of adaptive Fourier sums obtained by the greedy Algorithm 5.1 by

$$f_N^{(G)} := c_0(f) + \sum_{k=1}^N a_k^{(G)} P_k,$$

where again $z_0 = 0$, $a_0^{(G)} = c_0(f)$, and the consecutively obtained zeros z_j, $j = 1, \ldots, N$ are taken to determine the adaptive Fourier basis $\{P_k\}_{k=0}^N$. In Figures 2 and 3 we compare the function f with the N-th adaptive Fourier sum f_N, the N-th adaptive Fourier sum $F_N^{(G)}$ and the classical Fourier sum $S_N(f)$. In Figure 3 the corresponding error functions are displayed. Furthermore, in Table 1 we compare the errors of the computed Fourier sums $\|f - f_N\|_{L^2}$ and $\|f - f_N^{(G)}\|_{L^2}$ for $N = 1, \ldots, 9$ and the singular values σ_N of the matrix $\mathbf{A}_{10}\mathbf{Z}_{10}$.

Example 2. We approximate the real function $f(t) = \sum_{k=1}^\infty c_k(f)\sin(kt)$ with $c_k(f) = 1/k^2$, $k \in \mathbb{N}$. First, using the samples $c_k(f)$ for $k = 1, \ldots, 101$, we apply
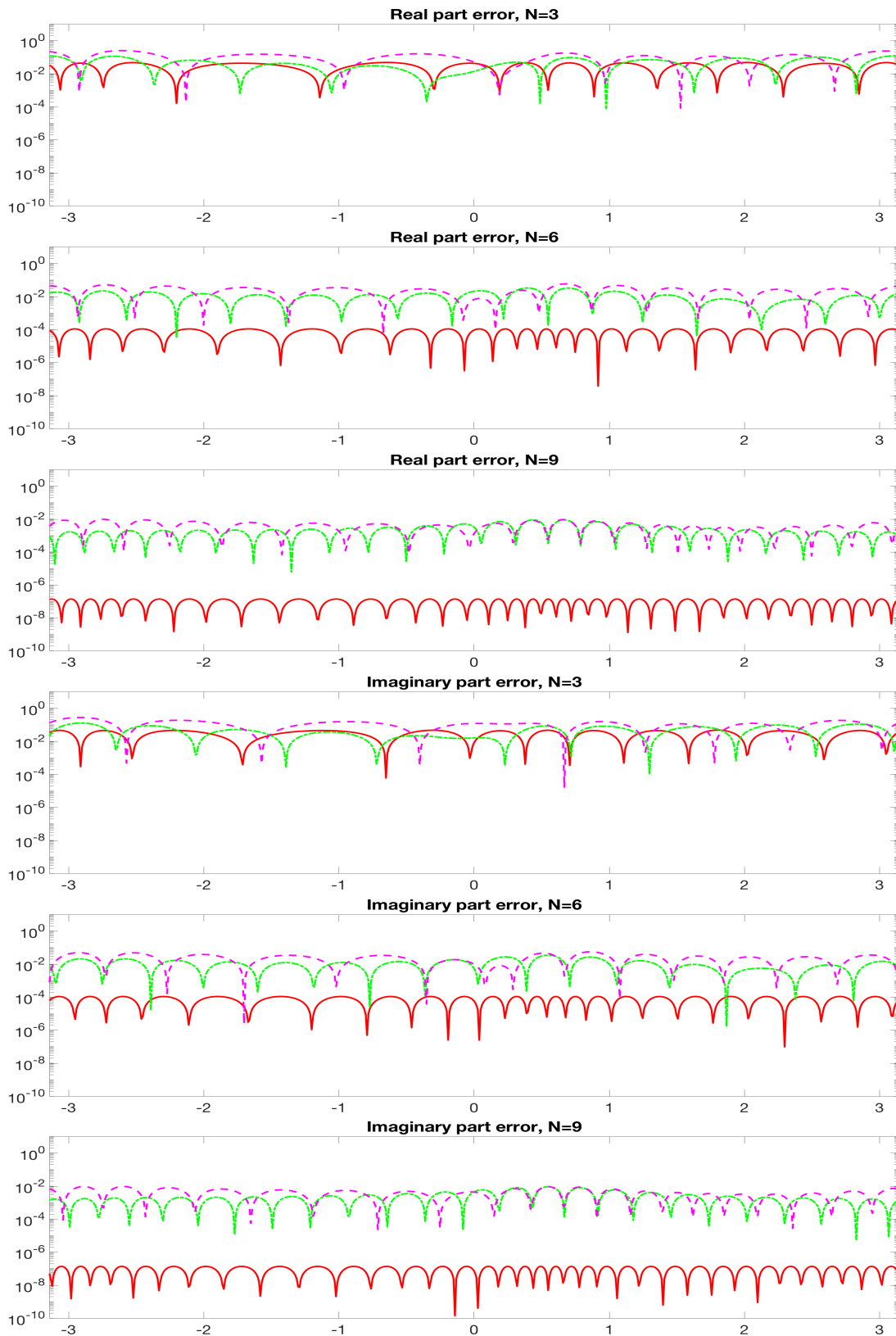
Figure 3 *The error $|f(t) - S_N(f)(t)|$ (green dotted line), $|f(t) - f_N^G(t)|$ (magenta dashed line) and $|f(t) - f_N(t)|$ (red solid line) for the real and imaginary part of the functions for $t \in [-\pi, \pi]$ on the x-axis for Example 1.*

N	σ_N	$\|f - f_N\|_{L_2}$	$\|f - f_N^{(G)}\|_{L_2}$
1	7.0672e-01	6.9094e-01	2.9125e+00
2	1.3593e-01	1.3484e-01	5.5597e-01
3	4.5874e-02	4.5843e-02	1.7501e-01
4	9.8703e-03	9.8537e-03	1.6970e-01
5	3.6531e-03	3.6526e-03	3.8253e-02
6	1.1020e-04	1.1020e-04	1.8900e-02
7	1.0701e-05	1.0701e-05	1.3875e-02
8	8.0077e-07	8.0077e-07	6.4196e-03
9	1.3831e-07	1.3831e-07	5.8672e-03

Table 1 *Comparison of the errors of adaptive Fourier sums for Example 1.*

the APM method in [22] and obtain beside the a priori fixed $\check{z}_0 = 0$ and $\breve{a}_0 = 0$ for $M = 11$ the zeros $\check{z}_j$ and coefficients $\breve{a}_j$, $j = 1, \dots, 11$,

$$
\begin{pmatrix} \check{z}_1 \\ \check{z}_2 \\ \check{z}_3 \\ \check{z}_4 \\ \check{z}_5 \\ \check{z}_6 \\ \check{z}_7 \\ \check{z}_8 \\ \check{z}_9 \\ \check{z}_{10} \\ \check{z}_{11} \end{pmatrix} = \begin{pmatrix} 0.9890 \\ 0.9621 \\ 0.9166 \\ 0.8493 \\ 0.7577 \\ 0.6415 \\ 0.5042 \\ 0.3543 \\ 0.2078 \\ 0.0870 \\ 0.0148 \end{pmatrix}, \quad \begin{pmatrix} \breve{a}_1 \\ \breve{a}_2 \\ \breve{a}_3 \\ \breve{a}_4 \\ \breve{a}_5 \\ \breve{a}_6 \\ \breve{a}_7 \\ \breve{a}_8 \\ \breve{a}_9 \\ \breve{a}_{10} \\ \breve{a}_{11} \end{pmatrix} = \begin{pmatrix} 0.0002 \\ 0.0014 \\ 0.0049 \\ 0.0129 \\ 0.0289 \\ 0.0567 \\ 0.0997 \\ 0.1571 \\ 0.2163 \\ 0.2438 \\ 0.1780 \end{pmatrix},
$$

such that

$$
\frac{1}{k^2} = c_k(f) \approx \tilde{c}_k(f) = \sum_{j=1}^{11} \breve{a}_j \, \check{z}_j^k. \tag{6.2}
$$

We compute the zeros $z_j^{(N)}$ and the coefficients $a_j^{(N)}$, $j = 1, \dots, N$ of the shorter sums of the form (4.2) by applying Algorithm 4.6, again with $z_0^{(N)} = 0$, $a_0^{(N)} = c_0(f) = 0$. The obtained parameters are given in Tables 2 and 3.

The corresponding Fourier sums f_N, $f_N^{(G)}$ and $S_N(f)$ are given in Figure 4 for $N = 1, 3$ and 5. The errors on the interval $[0, 2\pi]$ are displayed in Figure 5. The errors $\|f - f_N\|_{L_2}$ and $\|f - f_N^{(G)}\|_{L_2}$ for the Fourier coefficients as well as the singular values σ_{N+1} of the matrix $\mathbf{A}_{11}\mathbf{Z}_{11}$ are given in Table 4.

Acknowledgement

N	1	2	3	4	5	6	7	8	9	10
$z_1^{(N)}$	0.3535	0.2115	0.1501	0.1128	0.0855	0.0649	0.0494	0.0376	0.0286	0.0213
$z_2^{(N)}$		0.8110	0.6595	0.5395	0.4351	0.3458	0.2721	0.2124	0.1640	0.1239
$z_3^{(N)}$			0.9421	0.8601	0.7619	0.6557	0.5508	0.4532	0.3652	0.2856
$z_4^{(N)}$				0.9760	0.9272	0.8553	0.7666	0.6684	0.5667	0.4636
$z_5^{(N)}$					0.9853	0.9504	0.8943	0.8191	0.7284	0.6244
$z_6^{(N)}$						0.9881	0.9587	0.9102	0.8417	0.7528
$z_7^{(N)}$							0.9888	0.9614	0.9154	0.8484
$z_8^{(N)}$								0.9890	0.9620	0.9165
$z_9^{(N)}$									0.9890	0.9621
$z_{10}^{(N)}$										0.9890

Table 2 *Adaptive zeros $z_j^{(N)}$ obtained for $N = 1, \ldots, 10$ to determine the optimal basis of the adaptive Fourier sum f_N in Example 2.*

N	1	2	3	4	5	6	7	8	9	10
$a_1^{(N)}$	0.9677	0.9115	0.8043	0.7020	0.6029	0.5101	0.4272	0.3547	0.2914	0.2347
$a_2^{(N)}$		0.0865	0.1875	0.2662	0.3226	0.3538	0.3612	0.3498	0.3250	0.2904
$a_3^{(N)}$			0.0080	0.0304	0.0666	0.1109	0.1544	0.1895	0.2124	0.2216
$a_4^{(N)}$				0.0013	0.0074	0.0218	0.0451	0.0744	0.1053	0.1334
$a_5^{(N)}$					0.0004	0.0030	0.0101	0.0235	0.0434	0.0686
$a_6^{(N)}$						0.0003	0.0018	0.0064	0.0157	0.0316
$a_7^{(N)}$							0.0002	0.0015	0.0052	0.0133
$a_8^{(N)}$								0.0002	0.0014	0.0049
$a_9^{(N)}$									0.0002	0.0014
$a_{10}^{(N)}$										0.0002

Table 3 *Coefficients $a_j^{(N)}$ for adaptive Fourier sum f_N obtained for $N = 1, \ldots, 10$ in Example 2.*

N	σ_N	$\|f - f_N\|_{L_2}$	$\|f - f_N^{(G)}\|_{L_2}$
1	1.1078e-01	1.1057e-01	8.2932e-02
2	2.2710e-02	2.2693e-02	6.0175e-02
3	6.0753e-03	6.0538e-03	1.1511e-02
4	1.6706e-03	1.6635e-03	6.8640e-03
5	3.9237e-04	3.8919e-04	1.5419e-03
6	7.4187e-05	7.3599e-05	1.0114e-03
7	1.1081e-05	1.0949e-05	6.1344e-04
8	1.2559e-06	1.2382e-06	4.5046e-04
9	9.8620e-08	9.6716e-08	2.9683e-04
10	4.2914e-09	9.6368e-09	2.8196e-04

Table 4 *Comparison of the errors of adaptive Fourier sums for Example 2.*
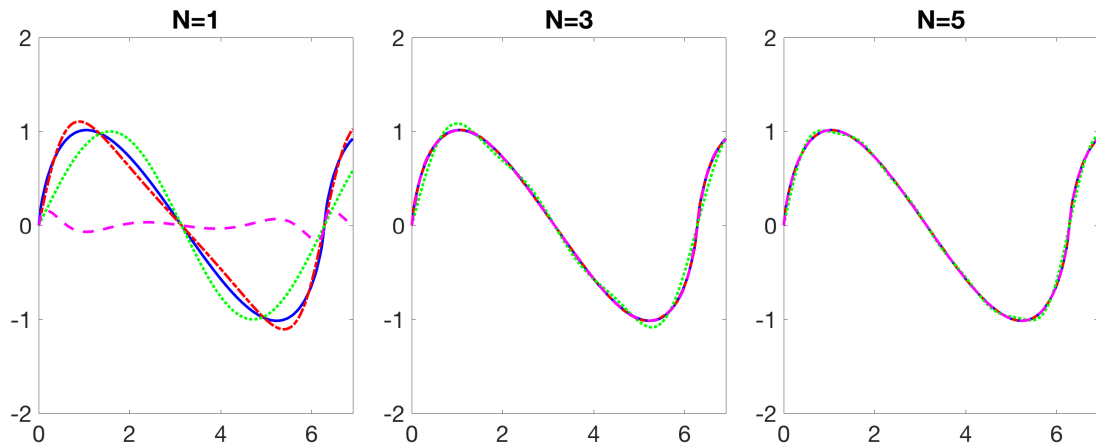
Figure 4 *The original function f (blue solid line), the Fourier approximation $S_N(f)$ (green dotted line), the greedy adaptive Fourier sum $f_N^{(G)}$ (magenta dashed line) and the adaptive Fourier sum f_N (red dash-dot line).*

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] N.I. Akhiezer, *Lectures on Approximation Theory*, Nauka, Moscow 1965 (Russian). English transl.: Dover Publications, Inc., New York, 1992.

[3] 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.

[4] L. Berg, Lineare Gleichungssysteme mit Bandstruktur, Carl Hanser, München/Wien, 1986.

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

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

[7] A. Bultheel and P. Carrette, *Algebraic and spectral properties of generalized Toeplitz matrices*, SIAM J. Control Optim. **41**(5) (2003), 1413–1439.

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

[9] J. Demmel, *Accurate singular value decomposition of structured matrices*, SIAM J. Matrix Ana. Appl. **21**(2) (2000), 562–580.

[10] P. Hartman, *On completely continuous Hankel matrices*, Proc. Amer. Math. Soc. **9** (1958), 862–866.

[11] T. S. Haut and G. Beylkin, *Fast and accurate con-eigenvalue algorithm for optimal rational approximations*, SIAM J. Matrix Anal. Appl. **33**(4) (2012), 1101–1125.

[12] P.S.C. Heuberger, P.M.J. Van den Hof, and B. Wahlberg, Modelling and Identification with rational orthogonal basis functions, Springer, London, 2005.

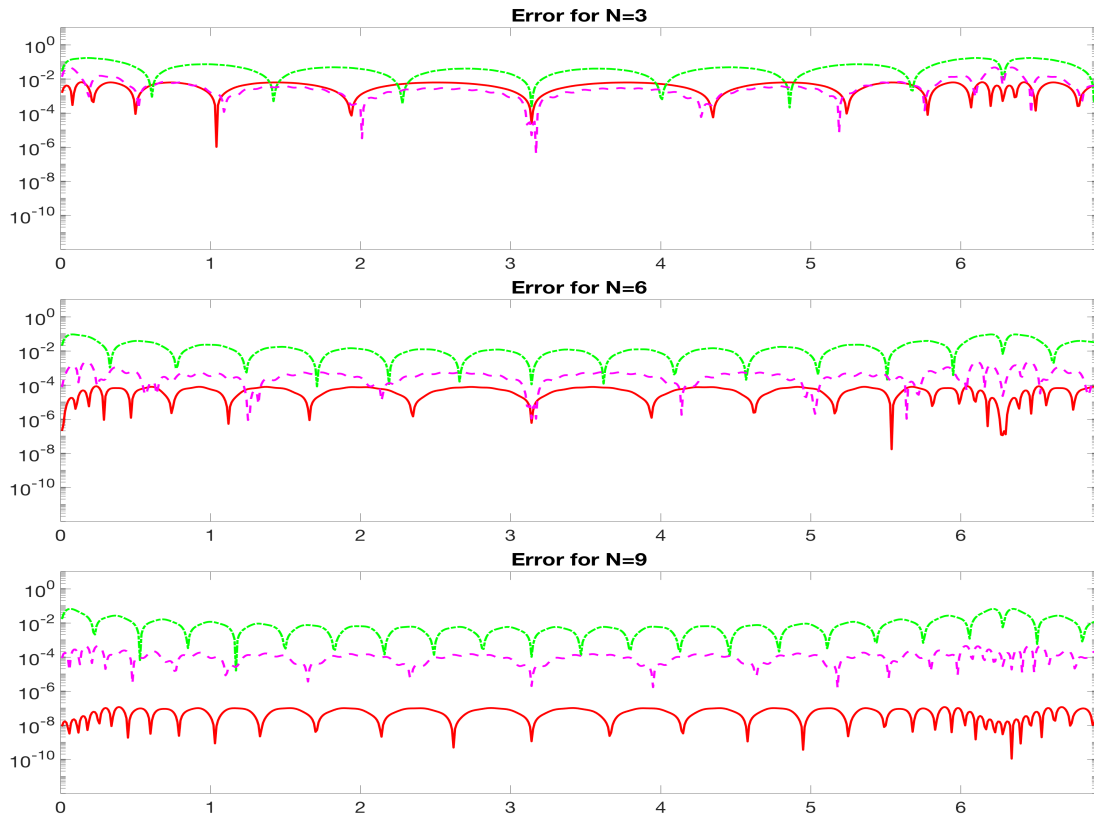[13] R.A. Horn and C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, 1990.

Figure 5 *The error* $|f(t) - S_N(f)(t)|$ *(green dotted line),* $|f(t) - f_N^{(G)}(t)|$ *(magenta dashed line) and* $|f(t) - f_N(t)|$ *(red solid line) for* $t \in [-\pi, \pi]$ *on the x-axis.*

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

[15] 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.

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

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

[18] B. Ninnes, H. Hjalmarsson, and F. Gustafsson, The fundamental role of general orthogonal bases in system identification, IEEE Trans. Automat. Control **44** (199), 1384–1407.

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

[20] 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

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

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

[23] A. Pushnitski and D. Yafaev, *Sharp estimates for singular values of Hankel operators*, Integr. Equ. Oper. Theory **83** (2015), 393–411.

[24] T. Qian and Y.-B. Wang, *Adaptive Fourier series - a variation of a greedy algorithm*, Adv. Comput. Math. **34** (2011), 279–293.

[25] T. Qian, L. Zhang, and Z. Li, *Algorithm of adaptive Fourier decomposition*, IEEE Trans. Signal Process. **59**(12) (2011), 5899–5906.

[26] 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, IMA Vol. Math. Appl. **23** (Springer, New York, 1990), 369 – 411.

[27] J.L. Walsh, *Interpolation and Approximation by Rational Functions in the Complex Domain*, Amer. Math. Soc. Colloq. Publ. 20, Amer. Math. Soc., Providence, RI, 1956.

[28] H. Widom, *Hankel matrices*, Trans. Amer. Math. Soc. **121**(1) (1966), 1–35.

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