NONLINEAR APPROXIMATION BY SUMS OF EXPONENTIALS AND TRANSLATES

THOMAS PETER†, DANIEL POTTS‡, AND MANFRED TASCHE§

Dedicated to Professor Lothar Berg on the occasion of his 80th birthday

Abstract. In this paper, we discuss the numerical solution of two nonlinear approximation problems. Many applications in electrical engineering, signal processing, and mathematical physics lead to the following problem: Let h be a linear combination of exponentials with real frequencies. Determine all frequencies, all coefficients, and the number of summands, if finitely many perturbed, uniformly sampled data of h are given. We solve this problem by an approximate Prony method (APM) and prove the stability of the solution in the square and uniform norm. Further, an APM for nonuniformly sampled data is proposed too.
The second approximation problem is related to the first one and reads as follows: Let φ be a given 1–periodic window function as defined in Section 4. Further let f be a linear combination of translates of φ. Determine all shift parameters, all coefficients, and the number of translates, if finitely many perturbed, uniformly sampled data of f are given. Using Fourier technique, this problem is transferred into the above parameter estimation problem for an exponential sum which is solved by APM. The stability of the solution is discussed in the square and uniform norm too. Numerical experiments show the performance of our approximation methods.

Key words and phrases: Nonlinear approximation, exponential sum, exponential fitting, harmonic retrieval, sum of translates, approximate Prony method, nonuniform sampling, parameter estimation, least squares method, signal processing, signal recovery, singular value decomposition, matrix perturbation theory, perturbed rectangular Hankel matrix.

AMS *Subject Classifications*: 65D10, 65T40, 41A30, 65F15, 65F20, 94A12.

1. Introduction. The recovery of signal parameters from noisy sampled data is a fundamental problem in signal processing which can be considered as a nonlinear approximation problem. In this paper, we discuss the numerical solution of two nonlinear approximation problems. These problems arise for example in electrical engineering, signal processing, or mathematical physics and read as follows:

1. Recover the pairwise different frequencies $f_j \in (-\pi, \pi)$, the complex coefficients $c_j \neq 0$, and the number $M \in \mathbb{N}$ in the exponential sum

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

if perturbed sampled data $\tilde{h}_k := h(k) + e_k$ $(k = 0, \ldots, 2N)$ are given, where e_k are small error terms.

The second problem is related to the first one:

2. Let $\varphi \in C(\mathbb{R})$ be a given 1–periodic window function as defined in Section 4. Recover the pairwise different shift parameters $s_j \in (-\frac{1}{2}, \frac{1}{2})$, the complex coefficients $c_j \neq 0$, and the number $M \in \mathbb{N}$ in the sum of translates

$$f(x) := \sum_{j=1}^{M} c_j \, \varphi(x + s_j) \quad (x \in \mathbb{R}) \,, \tag{1.2}$$

†t.peter@math.uni-goettingen.de, University of Goettingen, Institute for Numerical and Applied Mathematics, D–37083 Goettingen, Germany

‡potts@mathematik.tu-chemnitz.de, Chemnitz University of Technology, Faculty of Mathematics, D–09107 Chemnitz, Germany

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

if perturbed sampled data $\tilde{f}_k := f(k/n) + e_k$ $(k = -n/2, \ldots, n/2-1)$ are given, where n is a power of 2 and e_k are small error terms.

The first problem can be solved by an *approximate Prony method* (APM). The APM is based on ideas of G. Beylkin and L. Monzón [3, 4]. Note that the emphasis in [3, 4] is placed on approximate compressed representation of functions by linear combinations with only few exponentials. See also the impressive results in [5, 6]. Recently, the two last named authors of this paper have investigated the properties and the numerical behavior of APM in [26], where only real–valued exponential sums (1.1) were considered. Further, the APM is generalized to the parameter estimation for a sum of nonincreasing exponentials in [27].

The first part of APM recovers the frequencies f_j of (1.1). Here we solve a singular value problem of the rectangular Hankel matrix $\tilde{\mathbf{H}} := (\tilde{h}_{k+l})_{k,l=0}^{2N-L,L}$ and find f_j via zeros of a convenient polynomial of degree L, where L denotes an a priori known upper bound of M. Note that there exists a variety of further algorithms to recover the exponents f_j like ESPRIT or least squares Prony method, see e.g. [15, 23, 24] and the references therein. The second part uses the obtained frequencies and computes the coefficients c_j of (1.1) by solving an overdetermined linear Vandermonde–type system in a weighted least squares sense. Therefore, the second part of APM is closely related to the theory of *nonequispaced fast Fourier transform* (NFFT) (see [12, 2, 28, 11, 25, 16, 21]).

In contrast to [3, 4], we prefer an approach to the APM by the perturbation theory for a singular value decomposition of $\tilde{\mathbf{H}}$ (see [26]). In this paper, we investigate the stability of the approximation of (1.1) in the square and uniform norm for the first time. It is a known fact that clustered frequencies f_j make some troubles for the nonlinear approximation. Therefore, the strong relation between the separation distance of f_j and the number $T = 2\,N$ is very interesting in Section 3. Furthermore we prove the simultaneous approximation property of the suggested method. More precisely, under suitable assumptions we show that the derivative of h in (1.1) can be also very well approximated, see the estimate (3.5) in Theorem 3.4.

The second approximation problem is transferred into the first one with the help of Fourier technique. We use oversampling and present a new APM–algorithm of a sum (1.2) of translates. Corresponding error estimates between the original function f and its reconstruction are given in the square and uniform norm. The critical case of clustered shift parameters s_j is discussed too. We show a relation between the separation distance of s_j and the number n of sampled data.

Further, an APM for nonuniformly sampled data is presented too. We overcome the uniform sampling in the first problem by using results from the theory of NFFT. Finally, numerical experiments show the performance of our approximation methods.

This paper is organized as follows. In Section 2, we sketch the classical Prony method and present the APM. In Section 3, we consider the stability of the exponential sum and estimate the error between the original exponential sum h and its reconstruction in the square norm (see Lemma 3.3) and more important in the uniform norm (see Theorem 3.4). The nonlinear approximation problem for a sum (1.2) of translates is discussed in Section 4. We present the Algorithm 4.7 in order to compute all shift parameters and all coefficients of a sum f of translates as given in (1.2). The stability for sums of translates is handled in Section 5, see Lemma 5.2 for an estimate in the square norm and Theorem 5.3 for an estimate in the uniform norm. In Section 6, we generalize the APM to a new parameter estimation for an exponential sum from nonuniform sampling. Various numerical examples are described in Section 7. Finally, conclusions are presented in Section 8.

In the following we use standard notations. By $\mathbb{R}$ and $\mathbb{C}$, we denote the set of all real and complex numbers, respectively. The complex unit circle is denoted by $\mathbb{T}$. Let $\mathbb{Z}$ be the set of all integers and let $\mathbb{N}$ be the set of all positive integers. The Kronecker symbol is δ_k. The linear space of all column vectors with N complex components is denoted by $\mathbb{C}^N$, where $(\cdot,\cdot)$ is the corresponding scalar product. The linear space of all complex $M \times N$ matrices is denoted by $\mathbb{C}^{M\times N}$. For a matrix $\mathbf{A} \in \mathbb{C}^{M\times N}$, its transpose is denoted by $\mathbf{A}^{\mathrm{T}}$ and its conjugate–transpose by $\mathbf{A}^{\mathrm{H}}$. For the maximum column sum norm, spectral norm and maximum row sum norm of $\mathbf{A} \in \mathbb{C}^{M\times N}$, we write $\|\mathbf{A}\|_1$, $\|\mathbf{A}\|_2$ and $\|\mathbf{A}\|_\infty$, respectively. Then $\ker \mathbf{A}$ is the null space of a matrix $\mathbf{A}$. For the sum norm, Euclidean norm and maximum norm of a vector $\mathbf{b} \in \mathbb{C}^N$, we use the notation $\|\mathbf{b}\|_1$, $\|\mathbf{b}\|_2$ and $\|\mathbf{b}\|_\infty$, respectively.

For $T > 0$, the Banach space of all continuous functions $f : [-T, T] \to \mathbb{C}$ with the uniform norm $\|f\|_\infty$ is denoted by $C[-T, T]$. The Hilbert space of all square integrable functions $f : [-T, T] \to \mathbb{C}$ with the corresponding square norm $\|f\|_2$ is denoted by $L^2[-T, T]$. For a 1–periodic, continuous function $\varphi : \mathbb{R} \to \mathbb{C}$, the jth complex Fourier coefficient is denoted by $c_j(\varphi)$.

Computed quantities and approximations wear a tilde. Thus $\tilde{f}$ denotes a computed approximation to a function f and $\tilde{\mathbf{A}}$ a computed approximation to a matrix $\mathbf{A}$. Definitions are indicated by the symbol $:=$. Other notations are introduced when needed.

2. Nonlinear approximation by exponential sums. We consider a linear combination (1.1) of complex exponentials with complex coefficients $c_j \neq 0$ and pairwise different, ordered frequencies $f_j \in (-\pi, \pi)$, i.e.

$$-\pi < f_1 < \ldots < f_M < \pi .$$

Then h is infinitely differentiable, bounded and almost periodic on $\mathbb{R}$ (see [10, pp. 9 – 23]). We introduce the *separation distance* q of these frequencies by

$$q := \min_{j=1,\ldots,M-1} (f_{j+1} - f_j) .$$

Hence $q\,(M-1) < 2\pi$. Let $N \in \mathbb{N}$ with $N \geq 2M+1$ be given. Assume that perturbed sampled data

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

are known, where the error terms $e_k \in \mathbb{C}$ are bounded by a certain accuracy $\varepsilon_1 > 0$. Furthermore we suppose that $|c_j| \gg \varepsilon_1$ $(j = 1, \ldots, M)$.

Then we consider the following *nonlinear approximation problem for an exponential sum* (1.1): Recover the pairwise different frequencies $f_j \in (-\pi, \pi)$ and the complex coefficients c_j in such a way that

$$\Big|\tilde{h}_k - \sum_{j=1}^{M} c_j\, \mathrm{e}^{\mathrm{i} f_j k}\Big| \leq \varepsilon \quad (k = 0, \ldots, 2N) \tag{2.1}$$

for very small accuracy $\varepsilon > 0$ and for minimal number M of nontrivial summands. With other words, we are interested in *approximate representations* of $\tilde{h}_k \in \mathbb{C}$ by uniformly sampled data $h(k)$ $(k = 0, \ldots, 2N)$ of an exponential sum (1.1). Since $|f_j| < \pi$ $(j = 1, \ldots, M)$, we infer that the Nyquist condition is fulfilled (see [7, p.

183]).
All reconstructed values of the frequencies f_j, the coefficients c_j, and the number M of exponentials depend on ε, ε_1 and N (see [4]). By the assumption $|c_j| \gg \varepsilon_1$ $(j = 1, \ldots, M)$, we will be able to recover the original integer M in the case of small error bounds ε and ε_1.
The classical Prony method solves this problem for *exact* sampled data $\tilde{h}_k = h(k)$, cf. [17, pp. 457 – 462]. This procedure is based on a *separate computation* of all frequencies f_j and then of all coefficients c_j. First we form the exact rectangular Hankel matrix

$$\mathbf{H} := \big(h(k+l)\big)_{k,l=0}^{2N-L,L} \in \mathbb{C}^{(2N-L+1)\times(L+1)}\,, \tag{2.2}$$

where $L \in \mathbb{N}$ with $M \le L \le N$ is an a priori known upper bound of M. If $\mathbb{T}$ denotes the complex unit circle, then we introduce the pairwise different numbers

$$w_j := \mathrm{e}^{\mathrm{i} f_j} \in \mathbb{T} \quad (j = 1, \ldots, M)\,.$$

Thus we obtain that

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

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

Lemma 2.1 *Let L, M, $N \in \mathbb{N}$ with $M \le L \le N$ be given. Furthermore let $h_k = h(k) \in \mathbb{C}$ $(k = 0, \ldots, 2N)$ be the exact sampled data of* (1.1) *with $c_j \in \mathbb{C} \setminus \{0\}$ and pairwise distinct frequencies $f_j \in (-\pi, \pi)$ $(j = 1, \ldots, M)$.*
Then the rectangular Hankel matrix (2.2) *has the singular value* 0*, where*

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

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

For a proof see [27]. The classical Prony method is based on the following result.

Lemma 2.2 *Under the assumptions of Lemma* 2.1 *the following assertions are equivalent:*
(i) *The polynomial*

$$\sum_{k=0}^{L} u_k\, z^k \quad (z \in \mathbb{C}) \tag{2.3}$$

with complex coefficients u_k $(k = 0, \ldots, L)$ has M different zeros $w_j = \mathrm{e}^{\mathrm{i} f_j} \in \mathbb{T}$ $(j = 1, \ldots, M)$.
(ii) 0 *is a singular value of the complex rectangular Hankel matrix* (2.2) *with a right singular vector $\mathbf{u} := (u_l)_{l=0}^{L} \in \mathbb{C}^{L+1}$.*

For a proof see [27].

Algorithm 2.3 (*Classical Prony Method*)

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

1. *Compute a right singular vector* $\mathbf{u} = (u_l)_{l=0}^{L}$ *corresponding to the singular value* 0 *of* (2.2).
2. *For the polynomial* (2.3), *evaluate all zeros* $\tilde{z}_j \in \mathbb{C}$ *with* $||\tilde{z}_j| - 1| \le \varepsilon'$ ($j = 1, \ldots, \tilde{M}$). *Note that* $L \ge \tilde{M}$.
3. *For* $\tilde{w}_j := \tilde{z}_j/|\tilde{z}_j|$ ($j = 1, \ldots, \tilde{M}$), *compute* $\tilde{c}_j \in \mathbb{C}$ ($j = 1, \ldots, \tilde{M}$) *as least squares solution of the overdetermined linear Vandermonde–type system*

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

4. *Cancel all that pairs* $(\tilde{w}_l, \tilde{c}_l)$ ($l \in \{1, \ldots, \tilde{M}\}$) *with* $|\tilde{c}_l| \le \varepsilon$ *and denote the remaining set by* $\{(\tilde{w}_j, \tilde{c}_j) : j = 1, \ldots, M\}$ *with* $M \le \tilde{M}$. *Form* $\tilde{f}_j := \mathrm{Im}\,(\log \tilde{w}_j)$ ($j = 1, \ldots, M$), *where* log *is the principal value of the complex logarithm.*

Output: $M \in \mathbb{N}$, $\tilde{f}_j \in (-\pi, \pi)$, $\tilde{c}_j \in \mathbb{C}$ ($j = 1, \ldots, M$).

Note that we consider a rectangular Hankel matrix (2.2) with only $L + 1$ columns in order to determine the zeros of a polynomial (2.3) of relatively low degree L (see step 2 of Algorithm 2.3).
Unfortunately, the classical Prony method is notorious for its sensitivity to noise such that numerous modifications were attempted to improve its numerical behavior. The main drawback of this Prony method is the fact that 0 has to be a singular value of (2.2) (see Lemma 2.1 or step 1 of Algorithm 2.3). But in practice, only perturbed values $\tilde{h}_k = h(k) + e_k \in \mathbb{C}$ ($k = 0, \ldots, 2N$) of the exact sampled data $h(k)$ of an exponential sum (1.1) are known such that 0 is not a singular value of (2.2), in general. Here we assume that $|e_k| \le \varepsilon_1$ with certain accuracy $\varepsilon_1 > 0$. Then the error Hankel matrix

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

has a small spectral norm by

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

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

$$\tilde{\mathbf{H}} := \big(\tilde{h}_{k+l}\big)_{k,l=0}^{2N-L,L} = \mathbf{H} + \mathbf{E} \in \mathbb{C}^{(2N-L+1)\times(L+1)}\, . \tag{2.4}$$

By the singular value decomposition of the complex rectangular Hankel matrix $\tilde{\mathbf{H}}$ (see [18, pp. 414 – 415]), there exist two unitary matrices $\tilde{\mathbf{V}} \in \mathbb{C}^{(2N-L+1)\times(2N-L+1)}$, $\tilde{\mathbf{U}} \in \mathbb{C}^{(L+1)\times(L+1)}$ and a rectangular diagonal matrix $\tilde{\mathbf{D}} := \big(\tilde{\sigma}_k\, \delta_{j-k}\big)_{j,k=0}^{2N-L,L}$ with $\tilde{\sigma}_0 \ge \tilde{\sigma}_1 \ge \ldots \ge \tilde{\sigma}_L \ge 0$ such that

$$\tilde{\mathbf{H}} = \tilde{\mathbf{V}}\, \tilde{\mathbf{D}}\, \tilde{\mathbf{U}}^{\mathrm{H}}\, . \tag{2.5}$$

By (2.5), the orthonormal columns $\tilde{\mathbf{v}}_k \in \mathbb{C}^{2N-L+1}$ ($k = 0, \ldots, 2N - L$) of $\tilde{\mathbf{V}}$ and $\tilde{\mathbf{u}}_k \in \mathbb{C}^{L+1}$ ($k = 0, \ldots, L$) of $\tilde{\mathbf{U}}$ fulfill the conditions

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

i.e., $\tilde{\mathbf{u}}_k$ is a *right singular vector* and $\tilde{\mathbf{v}}_k$ is a *left singular vector* of $\tilde{\mathbf{H}}$ related to the *singular value* $\tilde{\sigma}_k \geq 0$ (see [18, p. 415]).
Note that $\sigma \geq 0$ is a singular value of the exact rectangular Hankel matrix $\mathbf{H}$ if and only if σ^2 is an eigenvalue of the Hermitian and positive semidefinite matrix $\mathbf{H}^{\mathrm{H}}\,\mathbf{H}$ (see [18, p. 414]). Thus all eigenvalues of $\mathbf{H}^{\mathrm{H}}\,\mathbf{H}$ are nonnegative. Let $\sigma_0 \geq \sigma_1 \geq \ldots \geq \sigma_L \geq 0$ be the ordered singular values of the exact Hankel matrix $\mathbf{H}$. Note that $\ker \mathbf{H} = \ker \mathbf{H}^{\mathrm{H}}\,\mathbf{H}$, since obviously $\ker \mathbf{H} \subseteq \ker \mathbf{H}^{\mathrm{H}}\,\mathbf{H}$ and since from $\mathbf{u} \in \ker \mathbf{H}^{\mathrm{H}}\,\mathbf{H}$ it follows that

$$0 = (\mathbf{H}^{\mathrm{H}}\,\mathbf{H}\,\mathbf{u},\, \mathbf{u}) = \|\mathbf{H}\,\mathbf{u}\|_2^2\,,$$

i.e., $\mathbf{u} \in \ker \mathbf{H}$. Then by Lemma 2.1, we know that $\dim(\ker \mathbf{H}^{\mathrm{H}}\,\mathbf{H}) = L - M + 1$, and hence $\sigma_{M-1} > 0$ and $\sigma_k = 0$ $(k = M, \ldots, L)$. Then the basic perturbation bound for the singular values σ_k of $\mathbf{H}$ reads as follows (see [18, p. 419])

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

Thus at least $L - M + 1$ singular values of $\tilde{\mathbf{H}}$ are contained in $[0,\, \|\mathbf{E}\|_2]$. We evaluate the smallest singular value $\tilde{\sigma} \in (0,\, \|\mathbf{E}\|_2]$ and a corresponding right singular vector of the matrix $\tilde{\mathbf{H}}$.

For noisy data we can not assume that our reconstruction yields roots $\tilde{z}_j \in \mathbb{T}$. Therefore we compute all zeros $\tilde{z}_j$ with $|\,|\tilde{z}_j| - 1| \leq \varepsilon_2$, where $0 < \varepsilon_2 \ll 1$. Now we can formulate the following APM–algorithm.

Algorithm 2.4 (APM)

Input: $L, N \in \mathbb{N}$ ($3 \leq L \leq N$, L *is upper bound of the number of exponentials*), $\tilde{h}_k = h(k) + e_k \in \mathbb{C}$ $(k = 0, \ldots, 2N)$ *with* $|e_k| \leq \varepsilon_1$, *accuracy bounds* $\varepsilon_1, \varepsilon_2 > 0$.

1. *Compute a right singular vector* $\tilde{\mathbf{u}} = (\tilde{u}_k)_{k=0}^{L}$ *corresponding to the smallest singular value* $\tilde{\sigma} > 0$ *of the perturbed rectangular Hankel matrix* (2.4).
2. *For the polynomial* $\sum_{k=0}^{L} \tilde{u}_k\, z^k$, *evaluate all zeros* $\tilde{z}_j$ $(j = 1, \ldots, \tilde{M})$ *with* $|\,|\tilde{z}_j| - 1| \leq \varepsilon_2$. *Note that* $L \geq \tilde{M}$.
3. *For* $\tilde{w}_j := \tilde{z}_j / |\tilde{z}_j|$ $(j = 1, \ldots, \tilde{M})$, *compute* $\tilde{c}_j \in \mathbb{C}$ $(j = 1, \ldots, \tilde{M})$ *as least squares solution of the overdetermined linear Vandermonde–type system*

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

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

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

with respect to the new set $\{\tilde{w}_j : j = 1, \ldots, M\}$ *again. Set* $\tilde{f}_j := \mathrm{Im}\,(\log \tilde{w}_j)$ $(j = 1, \ldots, M)$.

Output: $M \in \mathbb{N}$, $\tilde{f}_j \in (-\pi,\, \pi)$, $\tilde{c}_j \in \mathbb{C}$ $(j = 1, \ldots, M)$.

Remark 2.5 The convergence and stability properties of Algorithm 2.4 are discussed in [26]. In all numerical tests of Algorithm 2.4 (see Section 7 and [26]), we have obtained very good reconstruction results. All frequencies and coefficients can be computed such that

$$\max_{j=1,\ldots,M} |f_j - \tilde{f}_j| \ll 1, \quad \sum_{j=1}^{M} |c_j - \tilde{c}_j| \ll 1\,. \tag{2.6}$$

We have to assume that the frequencies f_j are separated, that $|c_j|$ are not too small, that the number $2N+1$ of samples is sufficiently large, that a convenient upper bound L of the number of exponentials is known, and that the error bound ε_1 of the sampled data is small. If none such upper bound L is known, one can always set $L = N$. Up to now, error estimates of $\max |f_j - \tilde{f}_j|$ and $\sum_{j=1}^{M} |c_j - \tilde{c}_j|$ are unknown.
If noiseless data are given, then the Algorithm 2.4 can be simplified by leaving step 5. But for perturbed data, the step 5 is essentially in general.
The steps 1 and 2 of Algorithm 2.4 can be replaced by the least squares ESPRIT method [23, p. 493], for corresponding numerical tests see [26]. Furthermore we can avoid the singular value decomposition by solving an overdetermined Hankel system, see [27, Algorithm 3.9]. Further we remark that in [3] the quadratic Toeplitz matrix

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

was considered instead of the rectangular Hankel matrix (2.2), where all coefficients c_j are positive such that $h(-k) = \overline{h(k)}$ for negative integers k. In this case one obtains an algorithm similar to [3, Algorithm 2]. In the step 3 (and analogously in step 5) of Algorithm 2.4, we use the diagonal preconditioner $\mathbf{D} = \mathrm{diag}\big(1 - |k|/(N+1)\big)_{k=-N}^{N}$. For very large $\tilde{M}$ and N, we can apply the CGNR method (conjugate gradient on the normal equations), where the multiplication of the rectangular Vandermonde–type matrix

$$\tilde{\mathbf{W}} := \big(\tilde{w}_j^k\big)_{k=0,j=1}^{2N,\,\tilde{M}} = \big(\mathrm{e}^{\mathrm{i}k\tilde{f}_j}\big)_{k=0,j=1}^{2N,\,\tilde{M}}$$

is realized in each iteration step by the NFFT (see [25, 21]). By [1, 26], the condition number of $\tilde{\mathbf{W}}$ is bounded for large N. Thus $\tilde{\mathbf{W}}$ is well conditioned, provided the frequencies $\tilde{f}_j$ $(j = 1, \ldots, \tilde{M})$ are not too close to each other or provided N is large enough, see also [22]. □

3. Stability of exponential sums. In this section, we discuss the stability of exponential sums. We start with the known Ingham inequality (see [19] or [29, pp. 162 – 164]).

Lemma 3.1 *Let $M \in \mathbb{N}$ and $T > 0$ be given. If the ordered frequencies f_j $(j = 1, \ldots, M)$ fulfill the gap condition*

$$f_{j+1} - f_j \ge q > \frac{\pi}{T} \quad (j = 1, \ldots, M-1),$$

then the exponentials $\mathrm{e}^{\mathrm{i}f_j x}$ $(j = 1, \ldots, M)$ are Riesz stable in $L^2[-T, T]$, i.e., for all complex vectors $\mathbf{c} = (c_j)_{j=1}^{M}$

$$\alpha(T)\,\|\mathbf{c}\|_2^2 \le \|\sum_{j=1}^{M} c_j\,\mathrm{e}^{\mathrm{i}f_j x}\|_2^2 \le \beta(T)\,\|\mathbf{c}\|_2^2$$

with positive constants

$$\alpha(T) := \frac{2}{\pi}\,\big(1-\frac{\pi^2}{T^2 q^2}\big), \quad \beta(T) := \frac{4\sqrt{2}}{\pi}\,\big(1+\frac{\pi^2}{4T^2 q^2}\big)$$

and with the square norm

$$\|f\|_2 := \Big(\frac{1}{2T}\int_{-T}^{T} |f(x)|^2\,\mathrm{d}x\Big)^{1/2} \quad (f \in L^2[-T,T])\,.$$

For a proof see [19] or [29, pp. 162 – 164]. The Ingham inequality for exponential sums can be considered as a far–reaching generalization of the Parseval equation for Fourier series. The constants $\alpha(T)$ and $\beta(T)$ are not optimal in general. Note that these constants are independently on M. The assumption $q > \frac{\pi}{T}$ is necessary for the existence of a positive constant $\alpha(T)$.

Now we show that a Ingham–type inequality is also true in the uniform norm of $C[-T,T]$.

Corollary 3.2 *If the assumptions of Lemma* 3.1 *are fulfilled, then the exponentials* $\mathrm{e}^{\mathrm{i} f_j x}$ $(j = 1,\ldots,M)$ *are Riesz stable in* $C[-T,T]$*, i.e., for all complex vectors* $\mathbf{c} = (c_j)_{j=1}^M$

$$\sqrt{\frac{\alpha(T)}{M}}\,\|\mathbf{c}\|_1 \le \|\sum_{j=1}^{M} c_j\,\mathrm{e}^{\mathrm{i} f_j x}\|_\infty \le \|\mathbf{c}\|_1$$

with the uniform norm

$$\|f\|_\infty := \max_{-T\le x\le T} |f(x)| \quad (f \in C[-T,T])\,.$$

Proof. Let $h \in C[-T,T]$ be given by (1.1). Then $\|h\|_2 \le \|h\|_\infty < \infty$. Using the triangle inequality, we obtain that

$$\|h\|_\infty \le \sum_{j=1}^{M} |c_j| \cdot 1 = \|\mathbf{c}\|_1\,.$$

From Lemma 3.1, it follows that

$$\sqrt{\frac{\alpha(T)}{M}}\,\|\mathbf{c}\|_1 \le \sqrt{\alpha(T)}\,\|\mathbf{c}\|_2 \le \|h\|_2\,.$$

This completes the proof. ■

Now we estimate the error $\|h-\tilde{h}\|_2$ between the original exponential sum (1.1) and its reconstruction

$$\tilde{h}(x) := \sum_{j=1}^{M} \tilde{c}_j\,\mathrm{e}^{\mathrm{i}\tilde{f}_j x} \quad (x \in [-T,T])\,. \tag{3.1}$$

Lemma 3.3 *Let $M \in \mathbb{N}$ and $T > 0$ be given. Let $\mathbf{c} = (c_j)_{j=1}^M$ and $\tilde{\mathbf{c}} = (\tilde{c}_j)_{j=1}^M$ be arbitrary complex vectors. If $(f_j)_{j=1}^M$, $(\tilde{f}_j)_{j=1}^M \in \mathbb{R}^M$ fulfill the conditions*

$$f_{j+1} - f_j \geq q > \frac{\pi}{T} \quad (j = 1, \ldots, M-1),$$
$$|\tilde{f}_j - f_j| \leq \delta < \frac{\pi}{4T} \quad (j = 1, \ldots, M),$$

then

$$\|h - \tilde{h}\|_2 \leq \sqrt{\beta(T)}\,\Big[\|\mathbf{c} - \tilde{\mathbf{c}}\|_2 + \|\mathbf{c}\|_2\big(1 - \cos(T\delta) + \sin(T\delta)\big)\Big]$$

in the norm of $L^2[-T, T]$. Note that

$$1 - \cos(T\delta) + \sin(T\delta) = 1 - \sqrt{2}\,\sin(\frac{\pi}{4} - T\delta) = T\delta + \mathcal{O}(\delta^2) \in [0, 1)\,.$$

Proof. 1. If $\delta = 0$, then $f_j = \tilde{f}_j$ $(j = 1, \ldots, M)$ and the assertion

$$\|h - \tilde{h}\|_2 \leq \sqrt{\beta(T)}\,\|\mathbf{c} - \tilde{\mathbf{c}}\|_2$$

follows directly from Lemma 3.1. Therefore we suppose that $0 < \delta < \frac{\pi}{4T}$. For simplicity, we can assume that $T = \pi$. First we use the ideas of [29, pp. 42 – 44] and estimate

$$\sum_{j=1}^{M} c_j\left(\mathrm{e}^{\mathrm{i} f_j x} - \mathrm{e}^{\mathrm{i} \tilde{f}_j x}\right) \quad (x \in [-\pi, \pi]) \tag{3.2}$$

in the norm of $L^2[-\pi, \pi]$. Here $\mathbf{c} = (c_j)_{j=1}^M$ is an arbitrary complex vector. Further let $(f_j)_{j=-M}^M$ and $(\tilde{f}_j)_{j=1}^M$ be real vectors with following properties

$$f_{j+1} - f_j \geq q > 1 \quad (j = 1, \ldots, M-1),$$
$$|\tilde{f}_j - f_j| \leq \delta < \frac{1}{4} \quad (j = 1, \ldots, M)\,.$$

Write

$$\mathrm{e}^{\mathrm{i} f_j x} - \mathrm{e}^{\mathrm{i} \tilde{f}_j x} = \mathrm{e}^{\mathrm{i} f_j x}\left(1 - \mathrm{e}^{\mathrm{i}\delta_j x}\right)$$

with $\delta_j := \tilde{f}_j - f_j$ and $|\delta_j| \leq \delta < \frac{1}{4}$ $(j = 1, \ldots, M)$.
2. Now we expand the function $1 - \mathrm{e}^{\mathrm{i}\delta_j x}$ $(x \in [-\pi, \pi])$ into a Fourier series relative to the orthonormal basis $\{1, \cos(kx), \sin(k - \frac{1}{2})x : k = 1, 2, \ldots\}$ in $L^2[-\pi, \pi]$. Note that $\delta_j \in [-\delta, \delta] \subset [-\frac{1}{4}, \frac{1}{4}]$. Then we obtain for each $x \in (-\pi, \pi)$ that

$$1 - \mathrm{e}^{\mathrm{i}\delta_j x} = \big(1 - \mathrm{sinc}(\pi\delta_j)\big) + \sum_{k=1}^{\infty} \frac{2\,(-1)^k \delta_j \sin(\pi\delta_j)}{\pi(k^2 - \delta_j^2)}\,\cos(kx)$$
$$+\,\mathrm{i}\sum_{k=1}^{\infty} \frac{2\,(-1)^k \delta_j \cos(\pi\delta_j)}{\pi((k - \frac{1}{2})^2 - \delta_j^2)}\,\sin(k - \frac{1}{2})x\,.$$

Interchanging the order of summation and then using the triangle inequality, we see that

$$\|\sum_{j=1}^{M} c_j\left(\mathrm{e}^{\mathrm{i} f_j x} - \mathrm{e}^{\mathrm{i} \tilde{f}_j x}\right)\|_2 \leq S_1 + S_2 + S_3$$

with

$$S_1 := \| \sum_{j=1}^{M} \big(1 - \mathrm{sinc}(\pi \delta_j)\big)\, c_j\, \mathrm{e}^{\mathrm{i} f_j x} \|_2\,,$$

$$S_2 := \sum_{k=1}^{\infty} \| \cos(kx) \sum_{j=1}^{M} \frac{2\,(-1)^k \delta_j \sin(\pi \delta_j)}{\pi(k^2 - \delta_j^2)}\, c_j\, \mathrm{e}^{\mathrm{i} f_j x} \|_2\,,$$

$$S_3 := \sum_{k=1}^{\infty} \| \sin(k - \frac{1}{2})x \sum_{j=1}^{M} \frac{2\,(-1)^k \delta_j \cos(\pi \delta_j)}{\pi((k - \frac{1}{2})^2 - \delta_j^2)}\, c_j\, \mathrm{e}^{\mathrm{i} f_j x} \|_2\,.$$

From Lemma 3.1 and $\delta_j \in [-\delta, \delta]$, it follows that

$$S_1 \le \sqrt{\beta(\pi)} \Big(\sum_{j=1}^{M} |c_j|^2 \big(1 - \mathrm{sinc}(\pi \delta_j)\big)^2 \Big)^{1/2} \le \sqrt{\beta(\pi)}\, \|\mathbf{c}\|_2 \big(1 - \mathrm{sinc}(\pi\delta)\big)\,.$$

Now we estimate

$$S_2 \le \sum_{k=1}^{\infty} \| \sum_{j=1}^{M} \frac{2\,(-1)^k \delta_j \sin(\pi \delta_j)}{\pi(k^2 - \delta_j^2)}\, c_j\, \mathrm{e}^{\mathrm{i} f_j x} \|_2 \le \sqrt{\beta(\pi)}\, \|\mathbf{c}\|_2 \sum_{k=1}^{\infty} \frac{2\delta}{\pi(k^2 - \delta^2)}\, \sin(\pi\delta)\,.$$

Using the known expansion

$$\pi \cot(\pi\delta) = \frac{1}{\delta} + \sum_{k=1}^{\infty} \frac{2\delta}{\delta^2 - k^2}\,,$$

we receive

$$S_2 \le \sqrt{\beta(\pi)}\, \|\mathbf{c}\|_2 \big(\mathrm{sinc}(\pi\delta) - \cos(\pi\delta)\big)\,.$$

Analogously, we estimate

$$\begin{aligned} S_3 &\le \sum_{k=1}^{\infty} \| \sum_{j=1}^{M} \frac{2\,(-1)^k \delta_j \cos(\pi \delta_j)}{\pi((k - \frac{1}{2})^2 - \delta_j^2)}\, c_j\, \mathrm{e}^{\mathrm{i} f_j x} \|_2 \\ &\le \sqrt{\beta(\pi)}\, \|\mathbf{c}\|_2 \sum_{k=1}^{\infty} \frac{2\delta}{\pi((k - \frac{1}{2})^2 - \delta^2)}\, \cos(\pi\delta)\,. \end{aligned}$$

Applying the known expansion

$$\pi \tan(\pi\delta) = \sum_{k=1}^{\infty} \frac{2\delta}{(k - \frac{1}{2})^2 - \delta^2}\,,$$

we obtain

$$S_3 \le \sqrt{\beta(\pi)}\, \|\mathbf{c}\|_2\, \sin(\pi\delta)\,.$$

Hence we conclude that

$$\| \sum_{j=1}^{M} c_j \left(\mathrm{e}^{\mathrm{i} f_j x} - \mathrm{e}^{\mathrm{i} \tilde{f}_j x}\right) \|_2 \le \sqrt{\beta(\pi)}\, \|\mathbf{c}\|_2 \big(1 - \cos(\pi\delta) + \sin(\pi\delta)\big)\,. \tag{3.3}$$

3. Finally, we estimate the normwise error by the triangle inequality. Then we obtain by Lemma 3.1 and (3.3) that

$$\begin{aligned}\|h - \tilde{h}\|_2 &\le \|\sum_{j=1}^{M} (c_j - \tilde{c}_j)\, \mathrm{e}^{\mathrm{i}\tilde{f}_j x}\|_2 + \|\sum_{j=1}^{M} c_j \left(\mathrm{e}^{\mathrm{i}f_j x} - \mathrm{e}^{\mathrm{i}\tilde{f}_j x}\right)\|_2 \\ &\le \sqrt{\beta(\pi)}\, \Big[\|\mathbf{c} - \tilde{\mathbf{c}}\|_2 + \|\mathbf{c}\|_2 \big(1 - \cos(\pi\,\delta) + \sin(\pi\,\delta)\big)\Big]\,.\end{aligned}$$

This completes the proof in the case $T = \pi$. If $T \ne \pi$, then we use the substitution $t = \frac{\pi}{T}\, x \in [-\pi, \pi]$ for $x \in [-T, T]$. ■

A similar result is true in the uniform norm of $C[-T, T]$.

Theorem 3.4 *Let $M \in \mathbb{N}$ and $T > 0$ be given. Let $\mathbf{c} = (c_j)_{j=1}^M$ and $\tilde{\mathbf{c}} = (\tilde{c}_j)_{j=1}^M$ be arbitrary complex vectors. If $(f_j)_{j=1}^M$, $(\tilde{f}_j)_{j=1}^M \in (-\pi,\, \pi)^M$ fulfill the conditions*

$$\begin{aligned} f_{j+1} - f_j \ge q &> \frac{3\pi}{2T} \quad (j = 1, \ldots, M-1), \\ |\tilde{f}_j - f_j| \le \delta &< \frac{\pi}{4T} \quad (j = 1, \ldots, M), \end{aligned}$$

then both $\mathrm{e}^{\mathrm{i}f_j x}$ $(j = 1, \ldots, M)$ and $\mathrm{e}^{\mathrm{i}\tilde{f}_j x}$ $(j = 1, \ldots, M)$ are Riesz stable in $C[-T,\, T]$. Further

$$\|h - \tilde{h}\|_\infty \le \|\mathbf{c} - \tilde{\mathbf{c}}\|_1 + 2\, \|\mathbf{c}\|_1\, \sin \frac{\delta T}{2}\,, \tag{3.4}$$

$$\|h' - \tilde{h}'\|_\infty \le \pi\, \|\mathbf{c} - \tilde{\mathbf{c}}\|_1 + \|\mathbf{c}\|_1 \left(\delta + 2\pi\, \sin \frac{\delta T}{2}\right) \tag{3.5}$$

in the norm of $C[-T, T]$.

Proof. 1. By the gap condition we know that

$$f_{j+1} - f_j \ge q > \frac{3\pi}{2T} > \frac{\pi}{T}\,.$$

Hence the original exponentials $\mathrm{e}^{\mathrm{i}f_j x}$ $(j = 1, \ldots, M)$ are Riesz stable in $C[-T,\, T]$ by Corollary 3.2. Using the assumptions, we conclude that

$$\begin{aligned}\tilde{f}_{j+1} - \tilde{f}_j &= (f_{j+1} - f_j) + (\tilde{f}_{j+1} - f_{j+1}) + (f_j - \tilde{f}_j) \\ &\ge q - 2\, \frac{\pi}{4T} > \frac{\pi}{T}\,.\end{aligned}$$

Thus the reconstructed exponentials $\mathrm{e}^{\mathrm{i}\tilde{f}_j x}$ $(j = 1, \ldots, M)$ are Riesz stable in $C[-T,\, T]$ by Corollary 3.2 too.

2. Using (3.2), we estimate the normwise error $\|h - \tilde{h}\|_\infty$ by the triangle inequality. Then we obtain

$$\begin{aligned}\|h - \tilde{h}\|_\infty &\le \|\sum_{j=1}^{M} (c_j - \tilde{c}_j)\, \mathrm{e}^{\mathrm{i}\tilde{f}_j x}\|_\infty + \|\sum_{j=1}^{M} c_j \left(\mathrm{e}^{\mathrm{i}f_j x} - \mathrm{e}^{\mathrm{i}\tilde{f}_j x}\right)\|_\infty \\ &\le \|\mathbf{c} - \tilde{\mathbf{c}}\|_1 + \sum_{j=1}^{M} |c_j| \max_{-T \le x \le T} |\mathrm{e}^{\mathrm{i}f_j x} - \mathrm{e}^{\mathrm{i}\tilde{f}_j x}|\,.\end{aligned}$$

Since

$$|\mathrm{e}^{\mathrm{i}f_j x} - \mathrm{e}^{\mathrm{i}\tilde{f}_j x}| = |1 - \mathrm{e}^{\mathrm{i}\delta_j x}| = \sqrt{2 - 2\,\cos(\delta_j x)}$$
$$= 2\,|\sin \frac{\delta_j x}{2}| \le 2\,\sin \frac{\delta T}{2}$$

for all $x \in [-T,\, T]$ and for $\delta_j = \tilde{f}_j - f_j \in [-\delta,\, \delta]$ with $\delta T < \frac{\pi}{4}$, we receive (3.4).
3. The derivatives h' and $\tilde{h}'$ can be explicitly represented by

$$h'(x) = \mathrm{i} \sum_{j=1}^{M} f_j\, c_j\, \mathrm{e}^{\mathrm{i}f_j x}\,, \quad \tilde{h}'(x) = \mathrm{i} \sum_{j=1}^{M} \tilde{f}_j\, \tilde{c}_j\, \mathrm{e}^{\mathrm{i}\tilde{f}_j x}$$

for all $x \in [-T, T]$. From the triangle inequality it follows that

$$\|(\mathrm{i}\, f_j\, c_j)_{j=1}^M - (\mathrm{i}\, \tilde{f}_j\, \tilde{c}_j)_{j=1}^M\|_1 \le \pi\, \|\mathbf{c} - \tilde{\mathbf{c}}\|_1 + \delta\, \|\mathbf{c}\|_1\,.$$

Further we see immediately that

$$\|(\mathrm{i}\, \tilde{f}_j\, \tilde{c}_j)_{j=1}^M\|_1 \le \pi\, \|\tilde{\mathbf{c}}\|_1\,.$$

Then by (3.4) we receive the assertion (3.5). Note that similar estimates are also true for derivatives of higher order. ■

Remark 3.5 Assume that perturbed sampled data

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

of a exponential sum (1.1) are given. Then from [26, Lemma 5.1] it follows that $\|\mathbf{c} - \tilde{\mathbf{c}}\|_2 \le \sqrt{3}\,\varepsilon_1$ for each $N \ge \pi^2/q$. By Lemma 3.3, $\tilde{h}$ is a good approximation of h in $L^2[-T, T]$. Fortunately, by Theorem 3.4, $\tilde{h}$ is also a good approximation of h in $C^1[-T, T]$, if N is large enough. Thus we obtain a uniform approximation of h from given perturbed values at $2N + 1$ equidistant nodes. Since the approximation of h is again an exponential sum $\tilde{h}$ with computed frequencies and coefficients, we can use $\tilde{h}$ for an efficient determination of derivatives and integrals. See Example 7.2 for numerical results. □

Remark 3.6 The conclusions of Section 3 show the stability of exponential sums with respect to the square and uniform norm. All results are valid without the additional assumption (2.6). But if $\tilde{f}_j$, $\tilde{c}_j$ $(j = 1, \ldots, M)$ reconstructed by Algorithm 2.4 fulfill the condition (2.6), then we obtain small errors $\|h - \tilde{h}\|_2$ and $\|h - \tilde{h}\|_\infty$ by Lemma 3.3 and Theorem 3.4, respectively. With other words, the reconstruction by the Algorithm 2.4 is stable. □

4. APM for sums of translates. Let $N \in 2\,\mathbb{N}$ be fixed. We introduce an *oversampling factor* $\alpha > 1$ such that $n := \alpha\, N$ is a power of 2. Let $\varphi \in C(\mathbb{R})$ be a 1–periodic even, nonnegative function with a uniformly convergent Fourier expansion. Further we assume that all Fourier coefficients

$$c_k(\varphi) := \int_{-1/2}^{1/2} \varphi(x)\, \mathrm{e}^{-2\pi \mathrm{i} k x}\, \mathrm{d}x = 2 \int_0^{1/2} \varphi(x)\, \cos(2\pi k x)\, \mathrm{d}x \quad (k \in \mathbb{Z})$$

are nonnegative and that $c_k(\varphi) > 0$ for $k = 0, \ldots, N/2$. Such a function φ is called a *window function*. We can consider one of the following window functions.

Example 4.1 A well known window function is the 1–*periodization of a Gaussian function* (see [12, 28, 11])

$$\varphi(x) = \sum_{k=-\infty}^{\infty} \varphi_0(x+k), \quad \varphi_0(x) := \frac{1}{\sqrt{\pi b}}\, \mathrm{e}^{-(nx)^2/b} \quad (x \in \mathbb{R},\, b \geq 1)$$

with the Fourier coefficients $c_k(\varphi) = \frac{1}{n}\, \mathrm{e}^{-b(\pi k/n)^2} > 0$ $(k \in \mathbb{Z})$. □

Example 4.2 Another window function is the 1–*periodization of a centered cardinal B–spline* (see [2, 28])

$$\varphi(x) = \sum_{k=-\infty}^{\infty} \varphi_0(x+k), \quad \varphi_0(x) := M_{2m}(nx) \quad (x \in \mathbb{R};\, m \in \mathbb{N})$$

with the Fourier coefficients $c_k(\varphi) = \frac{1}{n}\left(\mathrm{sinc}\frac{k\pi}{n}\right)^{2m}$ $(k \in \mathbb{Z})$. With M_{2m} $(m \in \mathbb{N})$ we denote the centered cardinal B–spline of order $2m$. □

Example 4.3 Let $m \in \mathbb{N}$ be fixed. A possible window function is the 1–*periodization of the* $2m$*-th power of a* sinc*–function* (see [21])

$$\varphi(x) = \sum_{k=-\infty}^{\infty} \varphi_0(x+k)\,, \quad \varphi_0(x) := \frac{N\,(2\alpha-1)}{2m}\, \mathrm{sinc}^{2m}\Big(\frac{\pi N x\,(2\alpha-1)}{2m}\Big)$$

with the Fourier coefficients $c_k(\varphi) = M_{2m}\big(\frac{2mk}{(2\alpha-1)N}\big)$ $(k \in \mathbb{Z})$. □

Example 4.4 Let $m \in \mathbb{N}$ be fixed. As next window function we mention the 1–*periodization of a Kaiser–Bessel function* (see [20])

$$\begin{aligned} \varphi(x) &= \sum_{k=-\infty}^{\infty} \varphi_0(x+k)\,, \\ \varphi_0(x) &:= \begin{cases} \dfrac{\sinh(b\sqrt{m^2-n^2x^2})}{\pi\,\sqrt{m^2-n^2x^2}} & \text{for } |x| \leq \frac{m}{n} \qquad \big(b := \pi\left(2-\frac{1}{\alpha}\right)\big)\,, \\[2ex] \dfrac{\sin(b\sqrt{n^2x^2-m^2})}{\pi\,\sqrt{n^2x^2-m^2}} & \text{otherwise} \end{cases} \end{aligned}$$

with the Fourier coefficients

$$c_k(\varphi) = \begin{cases} \frac{1}{n}\, I_0\big(m\sqrt{b^2-(2\pi k/n)^2}\big) & \text{for } |k| \leq n\left(1-\frac{1}{2\alpha}\right), \\ 0 & \text{otherwise,} \end{cases}$$

where I_0 denotes the modified zero–order Bessel function. □

Now we consider a linear combination (1.2) of translates with complex coefficients $c_j \neq 0$ and pairwise different shift parameters s_j, where

$$-\frac{1}{2} < s_1 < \ldots < s_M < \frac{1}{2} \tag{4.1}$$

is fulfilled. Then $f \in C(\mathbb{R})$ is a complex–valued 1–periodic function. Further let $N \geq 2M+1$. Assume that perturbed, uniformly sampled data

$$\tilde{f}_l = f(\frac{l}{n}) + e_l, \quad |e_l| \leq \varepsilon_1 \quad (l = -n/2, \ldots, n/2-1)$$

are given, where the error terms $e_l \in \mathbb{C}$ are bounded by a certain accuracy ε_1 ($0 < \varepsilon_1 \ll 1$). Again we suppose that $|c_j| \gg \varepsilon_1$ ($j = 1, \ldots, M$).
Then we consider the following *nonlinear approximation problem for a sum* (1.2) *of translates*: Determine the pairwise different shift parameters $s_j \in (-\frac{1}{2}, \frac{1}{2})$ and the complex coefficients c_j in such a way that

$$\big|\tilde{f}_l - \sum_{j=1}^{M} c_j \, \varphi\big(\frac{l}{n} + s_j\big)\big| \leq \varepsilon \quad (l = -n/2, \ldots, n/2-1) \tag{4.2}$$

for very small accuracy $\varepsilon > 0$ and for minimal number M of translates. Note that all reconstructed values of the parameters s_j, the coefficients c_j, and the number M of translates depend on ε, ε_1, and n. By the assumption $|c_j| \gg \varepsilon_1$ ($j = 1, \ldots, M$), we will be able to recover the original integer M in the case of small error bounds ε and ε_1.
This nonlinear inverse problem (4.2) can be numerically solved in two steps. First we convert the given problem (4.2) into a parameter estimation problem (2.1) for an exponential sum by using Fourier technique. Then the parameters of the transformed exponential sum are recovered by APM. Thus this procedure is based on a *separate computation* of all shift parameters s_j and then of all coefficients c_j.
For the 1–periodic function (1.2), we compute the corresponding Fourier coefficients. By (1.2) we obtain for $k \in \mathbb{Z}$

$$c_k(f) = \int_{-1/2}^{1/2} f(x)\, \mathrm{e}^{-2\pi\mathrm{i}kx}\, \mathrm{d}x = \Big(\sum_{j=1}^{M} c_j \, \mathrm{e}^{2\pi\mathrm{i}k s_j}\Big)\, c_k(\varphi) = h(k)\, c_k(\varphi) \tag{4.3}$$

with the exponential sum

$$h(x) := \sum_{j=1}^{M} c_j \, \mathrm{e}^{2\pi\mathrm{i}x s_j} \quad (x \in \mathbb{R})\,. \tag{4.4}$$

In applications, the Fourier coefficients $c_k(\varphi)$ of the window function φ are often explicitly known, where $c_k(\varphi) > 0$ ($k = 0, \ldots, N/2$) by assumption. Further the function f is sampled on a fine grid, i.e., we know noisy sampled data $\tilde{f}_l = f(l/n) + e_l$ ($l = -n/2, \ldots, n/2-1$) on the fine grid $\{l/n : l = -n/2, \ldots, n/2-1\}$ of $[-1/2,\, 1/2]$, where e_l are small error terms. Then we can compute $c_k(f)$ ($k = -N/2, \ldots, N/2$) by discrete Fourier transform

$$\begin{aligned} c_k(f) &\approx \frac{1}{n} \sum_{l=-n/2}^{n/2-1} f\big(\frac{l}{n}\big)\, \mathrm{e}^{-2\pi\mathrm{i}kl/n} \\ &\approx \hat{f}_k := \frac{1}{n} \sum_{l=-n/2}^{n/2-1} \tilde{f}_l \, \mathrm{e}^{-2\pi\mathrm{i}kl/n}\,. \end{aligned}$$

For shortness we set

$$\tilde{h}_k := \hat{f}_k / c_k(\varphi) \quad (k = -N/2, \ldots, N/2)\,. \tag{4.5}$$

Lemma 4.5 *Let φ be a window function. Further let $\mathbf{c} = (c_j)_{j=1}^M \in \mathbb{C}^M$ and let $\tilde{f}_l = f(l/n) + e_l$ $(l = -n/2, \ldots, n/2-1)$ with $|e_l| \le \varepsilon_1$ be given.*
Then $\tilde{h}_k$ is an approximate value of $h(k)$ for each $k \in \{-N/2, \ldots, N/2\}$, where the following error estimate

$$|\tilde{h}_k - h(k)| \le \frac{\varepsilon_1}{c_k(\varphi)} + \|\mathbf{c}\|_1 \max_{j=0,\ldots,N/2} \sum_{\substack{l=-\infty \\ l\ne 0}}^{\infty} \frac{c_{j+ln}(\varphi)}{c_j(\varphi)}$$

is fulfilled.

Proof. The function $f \in C(\mathbb{R})$ defined by (1.2) is 1–periodic and has a uniformly convergent Fourier expansion. Let $k \in \{-N/2, \ldots, N/2\}$ be an arbitrary fixed index. By the discrete Poisson summation formula (see [7, pp. 181 – 182])

$$\frac{1}{n} \sum_{j=-n/2}^{n/2-1} f\big(\frac{j}{n}\big)\, \mathrm{e}^{-2\pi \mathrm{i} kj/n} - c_k(f) = \sum_{\substack{l=-\infty \\ l\ne 0}}^{\infty} c_{k+ln}(f)$$

and by the simple estimate

$$\frac{1}{n} \big| \sum_{j=-n/2}^{n/2-1} e_j\, \mathrm{e}^{-2\pi \mathrm{i} kj/n} \big| \le \frac{1}{n} \sum_{j=-n/2}^{n/2-1} |e_j| \le \varepsilon_1\,,$$

we conclude that

$$|\hat{f}_k - c_k(f)| \le \varepsilon_1 + \sum_{\substack{l=-\infty \\ l\ne 0}}^{\infty} |c_{k+ln}(f)|\,.$$

From (4.3) and (4.5) it follows that

$$\tilde{h}_k - h(k) = \frac{1}{c_k(\varphi)} \big(\hat{f}_k - c_k(f)\big)$$

and hence

$$|\tilde{h}_k - h(k)| \le \frac{1}{c_k(\varphi)} \big(\varepsilon_1 + \sum_{\substack{l=-\infty \\ l\ne 0}}^{\infty} |c_{k+ln}(f)|\big)\,.$$

Using (4.3) and

$$|h(k+ln)| \le \sum_{j=1}^{M} |c_j| = \|\mathbf{c}\|_1 \quad (l \in \mathbb{Z})\,,$$

we obtain for all $l \in \mathbb{Z}$

$$|c_{k+ln}(f)| = |h(k+ln)|\, c_{k+ln}(\varphi) \le \|\mathbf{c}\|_1\, c_{k+ln}(\varphi)\,.$$

Thus we receive the estimate

$$
\begin{aligned}
|\tilde{h}_k - h(k)| &\le \frac{\varepsilon_1}{c_k(\varphi)} + \|\mathbf{c}\|_1 \sum_{\substack{l=-\infty\\ l\neq 0}}^{\infty} \frac{c_{k+ln}(\varphi)}{c_k(\varphi)} \\
&\le \frac{\varepsilon_1}{c_k(\varphi)} + \|\mathbf{c}\|_1 \max_{j=-N/2,\dots,N/2} \sum_{\substack{l=-\infty\\ l\neq 0}}^{\infty} \frac{c_{j+ln}(\varphi)}{c_j(\varphi)}\,.
\end{aligned}
$$

Since the Fourier coefficients of φ are even, we obtain the error estimate of Lemma 4.5. ■

Remark 4.6 For a concrete window function φ from the Examples 4.1 – 4.4, we can more precisely estimate the expression

$$
\max_{j=0,\dots,N/2} \sum_{\substack{l=-\infty\\ l\neq 0}}^{\infty} \frac{c_{j+ln}(\varphi)}{c_j(\varphi)}\,. \tag{4.6}
$$

Let $n = \alpha\, N$ be a power of 2, where $\alpha > 1$ is the oversampling factor. For the window function φ of Example 4.1,

$$
\mathrm{e}^{-b\pi^2(1-\frac{1}{\alpha})}\Big[1 + \frac{\alpha}{(2\alpha-1)b\pi^2} + \mathrm{e}^{-2b\pi^2/\alpha}\big(1 + \frac{\alpha}{(2\alpha+1)b\pi^2}\big)\Big]
$$

is an upper bound of (4.6) (see [28]). For φ of Example 4.2,

$$
\frac{4m}{2m-1}\,\big(\frac{1}{2\alpha-1}\big)^{2m}
$$

is an upper bound of (4.6) (see [28]). For φ of Examples 4.3 – 4.4, the expression (4.6) vanishes, since $c_k(\varphi) = 0$ $(|k| > n/2)$. □

Thus $\tilde{h}_k$ is an approximate value of $h(k)$ for $k \in \{-N/2, \ldots, N/2\}$. For the computed data $\tilde{h}_k$ $(k = -N/2, \ldots, N/2)$, we determine a minimal number M of exponential terms with frequencies $2\pi s_j \in (-\pi, \pi)$ and complex coefficients c_j $(j = 1, \ldots, M)$ in such a way that

$$
\big|\tilde{h}_k - \sum_{j=1}^{M} c_j\, \mathrm{e}^{2\pi \mathrm{i} k s_j}\big| \le \varepsilon \quad (k = -N/2, \ldots, N/2) \tag{4.7}
$$

for very small accuracy $\varepsilon > 0$. Our nonlinear approximation problem (4.2) is transferred into a parameter estimation problem (4.7) of an exponential sum. Starting from the given perturbed sampled data $\tilde{f}_l$ $(l = -n/2, \ldots, n/2-1)$, we obtain approximate values $\tilde{h}_k$ $(k = -N/2, \ldots, N/2)$ of the exponential sum (4.4). In the next step we use the APM–Algorithm 2.4 in order to determine the frequencies $2\pi\, s_j$ of h (= shift parameters s_j of f) and the coefficients c_j. Note that in the case $|c_j| \le \varepsilon_1$ for certain $j \in \{1, \ldots, M\}$, we often cannot recover the corresponding shift parameter s_j in (4.2).

Algorithm 4.7 (APM for sums of translates)
Input: $N \in 2\mathbb{N}$, $L \in \mathbb{L}$ $(3 \le L \le N/2$, *L is an upper bound of the number of translated functions*), $n = \alpha\, N$ *power of* 2 *with* $\alpha > 1$, $\tilde{f}_l = f(l/n) + e_l$ $(l = -n/2, \ldots, n/2-1)$

with $|e_l| \le \varepsilon_1$, $c_k(\varphi) > 0$ $(k = 0, \ldots, N/2)$, *accuracies* ε_1, $\varepsilon_2 > 0$.

1. *By fast Fourier transform compute*

$$\hat{f}_k := \frac{1}{n} \sum_{l=-n/2}^{n/2-1} \tilde{f}_l \, \mathrm{e}^{-2\pi \mathrm{i} k l/n} \quad (k = -N/2, \ldots, N/2)\,,$$

$$\tilde{h}_k := \hat{f}_k / c_k(\varphi) \quad (k = -N/2, \ldots, N/2)\,.$$

2. *Compute a right singular vector* $\tilde{\mathbf{u}} = (\tilde{u}_l)_{l=0}^{L}$ *corresponding to the smallest singular value* $\tilde{\sigma} > 0$ *of the perturbed rectangular Hankel matrix* $\tilde{\mathbf{H}} := (\tilde{h}_{k+l-N/2})_{k,l=0}^{N-L,L}$.

3. *For the corresponding polynomial* $\sum_{k=0}^{L} \tilde{u}_k \, z^k$, *evaluate all zeros* $\tilde{z}_j$ $(j = 1, \ldots, \tilde{M})$ *with* $|\,|\tilde{z}_j| - 1| \le \varepsilon_2$. *Note that* $L \ge \tilde{M}$.

4. *For* $\tilde{w}_j := \tilde{z}_j / |\tilde{z}_j|$ $(j = 1, \ldots, \tilde{M})$, *compute* $\tilde{c}_j \in \mathbb{C}$ $(j = 1, \ldots, \tilde{M})$ *as least squares solution of the overdetermined linear Vandermonde–type system*

$$\sum_{j=1}^{\tilde{M}} \tilde{c}_j \, \tilde{w}_j^k = \tilde{h}_k \quad (k = -N/2, \ldots, N/2)$$

with the diagonal preconditioner $\mathbf{D} = \mathrm{diag}\big(1 - |k|/(N/2+1)\big)_{k=-N/2}^{N/2}$. *For very large* $\tilde{M}$ *and* N *use the* CGNR *method, where the multiplication of the Vandermonde–type matrix* $\tilde{\mathbf{W}} := (\tilde{w}_j^k)_{k=-N/2, j=1}^{N/2, \tilde{M}}$ *is realized in each iteration step by* NFFT [*21*].

5. *Delete all the* $\tilde{w}_l$ $(l \in \{1, \ldots, \tilde{M}\})$ *with* $|\tilde{c}_l| \le \varepsilon_1$ *and denote the remaining set by* $\{\tilde{w}_j : j = 1, \ldots, M\}$ *with* $M \le \tilde{M}$. *Form* $\tilde{s}_j := \frac{1}{2\pi} \, \mathrm{Im}(\log \tilde{w}_j)$ $(j = 1, \ldots, M)$.

6. *Compute* $\tilde{c}_j \in \mathbb{C}$ $(j = 1, \ldots, M)$ *as least squares solution of the overdetermined linear system*

$$\sum_{j=1}^{M} \tilde{c}_j \, \varphi\big(\frac{l}{n} + \tilde{s}_j\big) = \tilde{f}_l \quad (l = -n/2, \ldots, n/2-1)\,.$$

Output: $M \in \mathbb{N}$, $\tilde{s}_j \in (-\frac{1}{2}, \frac{1}{2})$, $\tilde{c}_j \in \mathbb{C}$ $(j = 1, \ldots, M)$.

For corresponding numerical tests of Algorithm 4.7 see the Examples 7.4 and 7.5.

Remark 4.8 If further we assume that the window function φ is well–localized, i.e., there exists $m \in \mathbb{N}$ with $2m \ll n$ such that the values $\varphi(x)$ are very small for all $x \in \mathbb{R} \setminus (I_m + \mathbb{Z})$ with $I_m := [-m/n, m/n]$, then φ can be approximated by a 1–periodic function ψ supported in $I_m + \mathbb{Z}$. For the window function φ of Example 4.1 – 4.4, we construct its truncated version

$$\psi(x) := \sum_{k=-\infty}^{\infty} \varphi_0(x+k)\, \chi_m(x+k) \quad (x \in \mathbb{R})\,, \tag{4.8}$$

where χ_m is the characteristic function of $[-m/n, m/n]$. For the window function φ of Example 4.2, we see that $\psi = \varphi$. Now we can replace φ by its truncated version ψ in (4.2). For each $l \in \{-\frac{n}{2}, \ldots, \frac{n}{2} - 1\}$, we define the index set $J_{m,n}(l) := \{j \in \{1, \ldots, M\} : l - m \le n\, s_j \le l + m\}$. In this case, we can replace the window function

φ in step 6 of Algorithm 4.7 by the function ψ. Then the related linear system of equations

$$\sum_{j\in J_{m,n}(l)} \tilde{c}_j\, \psi\big(\frac{l}{n}+\tilde{s}_j\big) = \tilde{f}_l \quad (l=-n/2,\ldots,n/2-1)$$

is sparse. □

Remark 4.9 In some applications, one is interested in the reconstruction of a nonnegative function (1.2) with positive coefficients c_j. Then we can use a nonnegative least squares method in the steps 4 and 6 of Algorithm 4.7. □

5. Stability of sums of translates. In this section, we discuss the stability of linear combinations of translated window functions.

Lemma 5.1 (cf. [8, *pp.* 155 − 156]). *Let φ be a window function. Under the assumption (4.1), the translates $\varphi(x+s_j)$ $(j=1,\ldots,M)$ are linearly independent. Further for all $\mathbf{c}=(c_j)_{j=1}^M \in \mathbb{C}^M$*

$$\|\sum_{j=1}^M c_j\, \varphi(x+s_j)\big\|_2 \le \|\varphi\|_2\, \|\mathbf{c}\|_1 \le \sqrt{M}\, \|\varphi\|_2\, \|\mathbf{c}\|_2\,.$$

Proof. 1. Assume that for some complex coefficients a_j $(j=1,\ldots,M)$,

$$g(x) = \sum_{j=1}^M a_j\, \varphi(x+s_j) = 0 \quad (x\in\mathbb{R}).$$

Then the Fourier coefficients of g read as follows

$$c_k(g) = c_k(\varphi) \sum_{j=1}^M a_j\, \mathrm{e}^{2\pi\mathrm{i}s_j k} = 0 \quad (k\in\mathbb{Z}).$$

Since by assumption $c_k(\varphi)>0$ for all $k=0,\ldots,N/2$ and since $N\ge 2M+1$, we obtain the homogeneous system of linear equations

$$\sum_{j=1}^M a_j\, \mathrm{e}^{2\pi\mathrm{i}s_j k} = 0 \quad (k=0,\ldots,M-1).$$

By (4.1), we conclude that for $j\ne l$ $(j,\, l=1,\ldots M)$, $\mathrm{e}^{2\pi\mathrm{i}s_j} \ne \mathrm{e}^{2\pi\mathrm{i}s_l}$. Thus the Vandermonde matrix $\big(\mathrm{e}^{2\pi\mathrm{i}s_j k}\big)_{k=0,\,j=1}^{M-1,\,M}$ is nonsingular and hence $a_j=0$ $(j=1,\ldots,M)$.
2. Using the uniformly convergent Fourier expansion

$$\varphi(x) = \sum_{k=-\infty}^{\infty} c_k(\varphi)\, \mathrm{e}^{2\pi\mathrm{i}kx}\,,$$

we receive that

$$\sum_{j=1}^M c_j\, \varphi(x+s_j) = \sum_{k=-\infty}^{\infty} c_k(\varphi)\, h(k)\, \mathrm{e}^{2\pi\mathrm{i}kx}$$

with

$$h(k) = \sum_{j=1}^{M} c_j \, \mathrm{e}^{2\pi \mathrm{i} k s_j} \, .$$

We estimate

$$|h(k)| \le \|\mathbf{c}\|_1 \le \sqrt{M}\, \|\mathbf{c}\|_2 \, .$$

Applying the Parseval equation

$$\|\varphi\|_2^2 = \sum_{k=-\infty}^{\infty} c_k(\varphi)^2 \, ,$$

we obtain that

$$\big\| \sum_{j=1}^{M} c_j \, \varphi(x+s_j) \big\|_2^2 = \sum_{k=-\infty}^{\infty} c_k(\varphi)^2 \, |h(k)|^2 \le \|\varphi\|_2^2 \; \|\mathbf{c}\|_1^2 .$$

This completes the proof. ■

Now we estimate the error $\|f - \tilde f\|_2$ between the original function (1.2) and the reconstructed function

$$\tilde f(x) = \sum_{j=1}^{M} \tilde c_j \, \varphi(x+\tilde s_j) \quad (x \in \mathbb{R})$$

in the case $\sum_{j=1}^{M} |c_j - \tilde c_j| \le \varepsilon \ll 1$ and $|s_j - \tilde s_j| \le \delta \ll 1$ $(j=1,\ldots,M)$ with respect to the norm of $L^2[-\frac{1}{2},\, \frac{1}{2}]$.

Lemma 5.2 *Let φ be a window function. Further let $M \in \mathbb{N}$. Let $\mathbf{c} = (c_j)_{j=1}^M$ and $\tilde{\mathbf{c}} = (\tilde c_j)_{j=1}^M$ be two complex vectors with $\|\mathbf{c} - \tilde{\mathbf{c}}\|_1 \le \varepsilon \ll 1$. Assume that $N \in 2\,\mathbb{N}$ is sufficiently large that*

$$\sum_{|k|>N/2} c_k(\varphi)^2 < \varepsilon_1^2$$

for given accuracy $\varepsilon_1 > 0$. If $(s_j)_{j=1}^M$, $(\tilde s_j)_{j=1}^M \in [-\frac{1}{2},\, \frac{1}{2}]^M$ fulfill the conditions

$$s_{j+1} - s_j \ge \frac{q}{2\pi} > \frac{3}{2\,N} \quad (j=1,\ldots,M-1), \tag{5.1}$$

$$|s_j - \tilde s_j| \le \frac{\delta}{2\pi} < \frac{1}{4\,N} \quad (j=1,\ldots,M), \tag{5.2}$$

then

$$\|f - \tilde f\|_2 \le \|\varphi\|_2 \left(\varepsilon + 2\, \|\mathbf{c}\|_1 \, \sin \frac{\delta N}{4}\right) + \left(2\, \|\mathbf{c}\|_1 + \varepsilon\right) \varepsilon_1 .$$

in the square norm of $L^2[-\frac{1}{2},\, \frac{1}{2}]$.

Proof. 1. Firstly, we compute the Fourier coefficients of f and $\tilde{f}$. By (4.3) – (4.4) we obtain that

$$c_k(f) - c_k(\tilde{f}) = c_k(\varphi)\,(h(k) - \tilde{h}(k)) \quad (k \in \mathbb{Z})$$

with the exponential sum

$$\tilde{h}(x) := \sum_{j=1}^{M} \tilde{c}_j\, \mathrm{e}^{2\pi \mathrm{i} \tilde{s}_j x}\,.$$

Using the Parseval equation, we receive for sufficiently large N that

$$\begin{aligned}
\|f - \tilde{f}\|_2^2 &= \sum_{k=-\infty}^{\infty} |c_k(f) - c_k(\tilde{f})|^2 = \sum_{k=-\infty}^{\infty} c_k(\varphi)^2\, |h(k) - \tilde{h}(k)|^2 \\
&= \sum_{|k|\le N/2} c_k(\varphi)^2\, |h(k) - \tilde{h}(k)|^2 + \sum_{|k|> N/2} c_k(\varphi)^2\, |h(k) - \tilde{h}(k)|^2 \\
&\le \|\varphi\|_2^2 \big(\max_{|k|\le N/2} |h(k) - \tilde{h}(k)|\big)^2 + \big(\|\mathbf{c}\|_1 + \|\tilde{\mathbf{c}}\|_1\big)^2\, \varepsilon_1^2\,.
\end{aligned}$$

2. By Theorem 3.4 we know that for all $x \in [-N/2,\, N/2]$

$$|h(x) - \tilde{h}(x)| \le \|\mathbf{c} - \tilde{\mathbf{c}}\|_1 + 2\,\|\mathbf{c}\|_1\, \sin\frac{\delta N}{4}\,.$$

This completes the proof. ∎

Theorem 5.3 *Let φ be a window function. Further let $M \in \mathbb{N}$. Let $\mathbf{c} = (c_j)_{j=1}^M$ and $\tilde{\mathbf{c}} = (\tilde{c}_j)_{j=1}^M$ be two complex vectors with $\|\mathbf{c} - \tilde{\mathbf{c}}\|_1 \le \varepsilon \ll 1$. Assume that $N \in 2\,\mathbb{N}$ is sufficiently large that*

$$\sum_{|k|>N/2} c_k(\varphi) < \varepsilon_1$$

for given accuracy $\varepsilon_1 > 0$. If further the assumptions (5.1) *and* (5.2) *are fulfilled, then*

$$\|f - \tilde{f}\|_\infty \le \sqrt{N+1}\, \|\varphi\|_2\, \big(\varepsilon + 2\,\|\mathbf{c}\|_1\, \sin\frac{\delta N}{4}\big) + (2\,\|\mathbf{c}\|_1 + \varepsilon)\, \varepsilon_1$$

in the norm of $C[-\frac{1}{2}, \frac{1}{2}]$.

Proof. Using first the triangle inequality and then the Cauchy–Schwarz inequality, we obtain that

$$\begin{aligned}
\|f - \tilde{f}\|_\infty &\le \sum_{k=-\infty}^{\infty} |c_k(f) - c_k(\tilde{f})| = \sum_{k=-\infty}^{\infty} c_k(\varphi)\, |h(k) - \tilde{h}(k)| \\
&= \sum_{|k|\le N/2} c_k(\varphi)\, |h(k) - \tilde{h}(k)| + \sum_{|k|>N/2} c_k(\varphi)\, |h(k) - \tilde{h}(k)| \\
&\le \big(\sum_{|k|\le N/2} c_k(\varphi)^2\big)^{1/2} \big(\sum_{|k|\le N/2} |h(k) - \tilde{h}(k)|^2\big)^{1/2} + (\|\mathbf{c}\|_1 + \|\tilde{\mathbf{c}}\|_1)\, \varepsilon_1\,.
\end{aligned}$$

From the Bessel inequality and Theorem 3.4 it follows that

$$\begin{aligned}\|f-\tilde f\|_\infty &\le \|\varphi\|_2 \Big(\sum_{|k|\le N/2} |h(k)-\tilde h(k)|^2\Big)^{1/2} + (2\,\|\mathbf{c}\|_1+\varepsilon)\,\varepsilon_1\\ &\le \sqrt{N+1}\,\|\varphi\|_2 \max_{|k|\le N/2} |h(k)-\tilde h(k)| + (2\,\|\mathbf{c}\|_1+\varepsilon)\,\varepsilon_1\\ &\le \sqrt{N+1}\,\|\varphi\|_2 \big(\varepsilon + 2\,\|\mathbf{c}\|_1 \sin\frac{\delta N}{4}\big) + (2\,\|\mathbf{c}\|_1+\varepsilon)\,\varepsilon_1\,.\end{aligned}$$

This completes the proof. ■

6. APM for nonuniform sampling. In this section we generalize the APM to nonuniformly sampled data. More precisely, as in Section 2 we recover all parameters of a linear combination h of complex exponentials. But now we assume that the sampled data $h(x_k)$ at the nonequispaced, pairwise different nodes $x_k \in (-\frac{1}{2},\, \frac{1}{2})$ $(k=1,\ldots,K)$ are given, i.e., $Nx_k \in (-\frac{N}{2},\, \frac{N}{2})$. We consider the exponential sum

$$h(x) := \sum_{j=1}^{M} c_j\, \mathrm{e}^{2\pi \mathrm{i} x N s_j}\,, \tag{6.1}$$

with complex coefficients $c_j \ne 0$ and pairwise different parameters

$$-\frac{1}{2} < s_1 < \ldots < s_M < \frac{1}{2}\,.$$

Note that $2\pi N s_j \in (-\pi N,\, \pi N)$ are the frequencies of h.
We regard the following *nonlinear approximation problem for an exponential sum* (6.1): Recover the pairwise different parameters $s_j \in (-\frac{1}{2},\, \frac{1}{2})$ and the complex coefficients c_j in such a way that

$$\big|h(x_k) - \sum_{j=1}^{M} c_j\, \mathrm{e}^{2\pi \mathrm{i} x_k N s_j}\big| \le \varepsilon \quad (k=1,\ldots,K)$$

for very small accuracy $\varepsilon > 0$ and for minimal number M of nontrivial summands. Note that all reconstructed values of the shift parameters s_j, the coefficients c_j, and the number M of exponentials depend on ε and K. By the additional assumption $|c_j| \gg \varepsilon$ $(j=1,\ldots,M)$, we will be able to recover the original integer M for a small target accuracy ε.
The fast evaluation of the exponential sum (6.1) at the nodes x_k $(k=1,\ldots,K)$ is known as NFFT of type 3 [16]. A corresponding fast algorithm presented first by B. Elbel and G. Steidl in [13] (see also [21, Section 4.3]) requires only $\mathcal{O}(N\log N + K + M)$ arithmetic operations. Here N is called the nonharmonic bandwith.
Note that a Prony–like method for nonuniform sampling was already proposed in [9]. There the unknown parameters were estimated by a linear regression equation which uses filtered signals. We use the approximation schema of the NFFT of type 3 in order to develop a new algorithm. As proven in [13], the exponential sum (6.1) can be approximated with the help of a truncated window function ψ (see (4.8)) in the form

$$\tilde h(x) = \sum_{l=1}^{L} h_l\, \psi(x - \frac{l}{L})\,. \tag{6.2}$$

with $L > N$. From this observation, we immediately obtain the following algorithm:

Algorithm 6.1 (APM for nonuniform sampling)
Input: K, N, $L \in \mathbb{N}$ *with* $K \geq L > N$, $h(x_k) \in \mathbb{C}$ *with nonequispaced, pairwise different nodes* $x_k \in (-\frac{1}{2}, \frac{1}{2})$ $(k = 1, \ldots, K)$, *truncated window function* ψ.

1. *Solve the least squares problem*

$$\sum_{l=1}^{L} h_l \, \psi(x_k - \frac{l}{L}) = h(x_k) \quad (k = 1, \ldots, K)$$

to obtain the coefficients h_l $(l = 1, \ldots, L)$.
2. *Compute the values* $\tilde{h}(n/N)$ $(n = -N/2, \ldots, N/2 - 1)$ *of* (6.2) *and use Algorithm* 2.4 *in order to compute all parameters* s_j *and all coefficients* c_j $(j = 1, \ldots, M)$.

Output: $M \in \mathbb{N}$, $\tilde{s}_j \in (-\frac{1}{2}, \frac{1}{2})$, $\tilde{c}_j \in \mathbb{C}$ $(j = 1, \ldots, M)$.

See the Example 7.6 for a numerical test.

7. Numerical experiments. Finally, we apply the suggested algorithms to various examples. We have implemented our algorithms in MATLAB with IEEE double precision arithmetic.

Example 7.1 We start with a short comparison between the Algorithm 2.4 and an algorithm proposed in [6, Appendix A.1]. If noiseless data are given and if $L = N$, then the Algorithm 2.4 is very similar to the algorithm in [6]. The advantage of the Algorithm 2.4 is the fact that it works for perturbed data and that its arithmetic cost is very low for conveniently chosen L with $M \leq L \ll N$.
We sample the trigonometric sum

$$\begin{aligned} h(x) &:= 14 - 8\cos(0.453\,x) + 9\sin(0.453\,x) + 4\cos(0.979\,x) + 8\sin(0.979\,x) \\ &-2\cos(0.981\,x) + 2\cos(1.847\,x) - 3\sin(1.847\,x) + 0.1\cos(2.154\,x) - 0.3\sin(2.154\,x) \end{aligned} \tag{7.1}$$

with $M = 11$ at the equidistant nodes $x = k$ $(k = 0, \ldots, 2N)$. We set $\mathbf{f} := (f_j)_{j=1}^{M}$, $\mathbf{c} := (c_j)_{j=1}^{M}$, and $\mathbf{h} := h(2Nj\,10^{-4})_{j=0}^{10^4}$. Further we denote the computed values by $\tilde{\mathbf{f}} := (\tilde{f}_j)_{j=1}^{M}$, $\tilde{\mathbf{c}} := (\tilde{c}_j)_{j=1}^{M}$, and $\tilde{\mathbf{h}}$. Let $\tilde{\sigma}$ be the smallest singular value of the (quadratic or rectangular) Hankel matrix (2.2). We emphasize that step 2 of Algorithm 2.4 leads to a polynomial of low degree $L \ll N$ instead of degree N, if one uses the full $(N+1) \times (N+1)$ Hankel matrix as in the algorithm [6]. In both cases we apply the singular value decomposition from MATLAB but remark that one can avoid the rather expensive singular value decomposition for very large N and M by using an iterative algorithm for the solution of the rectangular Hankel matrix, see [27, Algorithm 3.9]. The linear Vandermonde–type system can be solved by the inverse NFFT, see [21]. We get the following results, see Table 7.1. □

Example 7.2 Next we confirm the uniform approximation property, see Theorem 3.4. We sample the trigonometric sum (7.1) at the equidistant nodes $x = k/2$ $(k = 0, \ldots, 120)$, where we add uniformly distributed pseudo–random numbers $e_k \in [-2.5, 2.5]$ to $h(k/2)$. The points $(k/2, h(k/2) + e_k)$ $(k = 0, \ldots, 120)$ are the centers of the red circles in Figure 7.1. In Figure 7.1 we plot the functions $h + 2.5$ and $h - 2.5$ by blue dashed lines. Finally the function $\tilde{h}$ reconstructed by Algorithm 2.4 is represented as a green line. We observe that $\|h - \tilde{h}\|_\infty \leq 2.5$. Furthermore, we can

	N	L	$\tilde{\sigma}$	$\|\tilde{\mathbf{f}} - \mathbf{f}\|_2$	$\|\tilde{\mathbf{c}} - \mathbf{c}\|_2$	$\|\tilde{\mathbf{h}} - \mathbf{h}\|_\infty$
Alg. 2.4	50	20	1.2e–13	2.3e–11	2.5e–7	5.3e–7
Alg. [6]	50	50	1.1e–14	4.1e–11	1.3e–7	2.3e–8
Alg. 2.4	500	20	5.7e–12	2.3e–11	6.8e–7	2.1e–7
Alg. 2.4	500	100	1.6e–12	2.2e–12	4.8e–8	3.4e–8
Alg. 2.4	500	200	1.1e–12	3.9e–13	6.1e–9	2.2e–8
Alg. [6]	500	500	3.4e–15	5.8e–13	7.0e–9	9.1e–9
Alg. 2.4	1000	20	1.0e–11	1.5e–11	1.2e–7	3.1e–7
Alg. 2.4	1000	100	5.2e–12	1.4e–12	5.3e–8	4.5e–8
Alg. 2.4	1000	500	2.1e–12	6.7e–14	4.8e–9	7.6e–9
Alg. [6]	1000	1000	1.5e–15	4.1e–14	2.2e–9	1.8e–9

Table 7.1
Errors of Example 7.1.

improve the approximation results, if we choose only uniformly distributed pseudo–random numbers $e_k \in [-0.5,\, 0.5]$ $(k = 0, \ldots, 120)$. Then we obtain $\|h - \tilde{h}\|_\infty \le 0.68$, see Figure 7.2 (left). In Figure 7.2 (right), the derivative h' is shown as a blue dashed line. The derivative $\tilde{h}'$ of the reconstructed function is drawn as a green line, cf. Theorem 3.4. We remark that further examples for the recovery of signal parameters in (1.1) from noisy sampled data are given in [26], which support also the new stability results in Section 3. In the Example 7.2 we have recovered the frequency 2.154 with the corresponding coefficient $0.05 + 0.15\,\mathrm{i}$ of small absolute value 0.158114. By adding much more noise, we cannot distinguish between frequencies of the noise and true frequencies with small coefficients. In that case one can repeat the experiment and average the results. These methods are behind the scope of this paper, but see [14] for corresponding results. □

Example 7.3 In the following, we test our Algorithm 2.4 for a relatively large number M of exponentials. For this we consider the exponential sum (1.1) with $M = 150$, $f_j = \pi \cos(j\pi/151)$, $c_j = \pi \sin(j\pi/151) + \mathrm{i}\,\pi \cos(j\pi/151)$ $(j = 1, \ldots, 150)$. We sample this function at the equidistant nodes $x = k$ $(k = 0, \ldots, 2N)$ and apply the Algorithm 2.4. The relative error of the frequencies is computed by

$$e(\mathbf{f}) := \Big(\sum_{j=1}^{M} |f_j - \tilde{f}_j|^2\Big)^{1/2} \Big(\sum_{j=1}^{M} |f_j|^2\Big)^{-1/2},$$

where $\tilde{f}_j$ are the frequencies computed by our Algorithm 2.4. Analogously, the relative error of the coefficients is defined by

$$e(\mathbf{c}) := \Big(\sum_{j=1}^{M} |c_j - \tilde{c}_j|^2\Big)^{1/2} \Big(\sum_{j=1}^{M} |c_j|^2\Big)^{-1/2},$$

where $\tilde{c}_j$ are the coefficients computed by the Algorithm 2.4. Let h be the original exponential sum (1.1) and let $\tilde{h}$ be the exponential sum (3.1) recovered by Algorithm 2.4. Then we determine the error $\|h - \tilde{h}\|_\infty = \max |h(x) - \tilde{h}(x)|$, where the maximum is built from 10000 equidistant points of $[0,\, 2N]$.

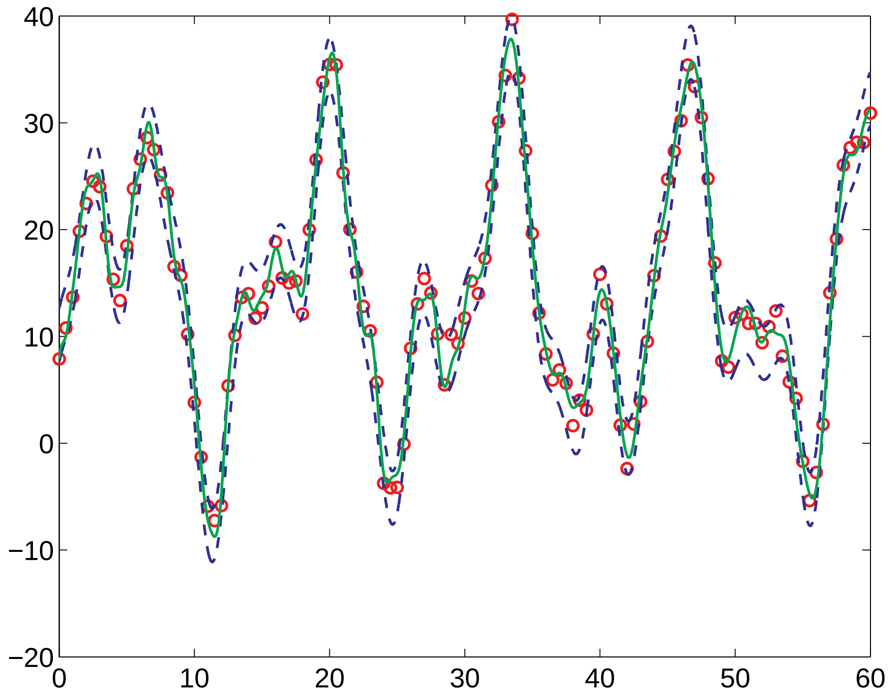

Fig. 7.1. *The functions $h + 2.5$ and $h - 2.5$ from Example 7.2 are shown as blue dashed lines. The perturbed sampling points $(k/2, h(k/2) + e_k)$ with $e_k \in [-2.5,\, 2.5]$ $(k = 0, \ldots, 120)$ are the centers of the red circles. The reconstructed function $\tilde{h}$ is shown as a green line.*

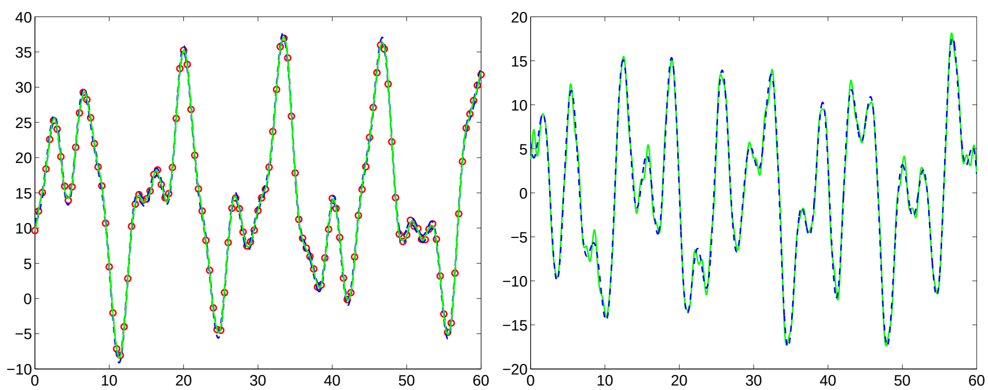

Fig. 7.2. *Left: The functions $h + 0.5$ and $h - 0.5$ from Example 7.2 are shown as blue dashed lines. The perturbed sampling points $(k/2, h(k/2) + e_k)$ with $e_k \in [-0.5,\, 0.5]$ $(k = 0, \ldots, 120)$ are the centers of the red circles. The reconstructed function $\tilde{h}$ is shown as a green line. Right: The function h' from Example 7.2 is shown as a blue dashed line. The derivative $\tilde{h}'$ of the reconstructed function is shown as a green line.*

We present also the results of the modified Algorithm 2.4, if we replace the steps 1 and 2 by the ESPRIT method, see Remark 2.5. Note that $q \approx 0.00204$. As pointed out in [26, Lemma 4.1], then the Vandermonde–type matrix $\tilde{\mathbf{W}}$ is left invertible for all $N > 2793$. However for smaller $N \in \{500,\, 1000,\, 1500\}$ we still get good results of the parameter reconstruction, see Table 7.2. $\square$

N	L	$e(\mathbf{f})$	$e(\mathbf{c})$	$\|h-\tilde{h}\|_\infty$
Algorithm 2.4				
500	150	9.5e–03	5.5e–01	1.2e–01
1000	150	2.5e–08	1.2e–04	2.4e–08
1500	150	6.4e–13	3.3e–09	2.2e–09
Algorithm 2.4 based on ESPRIT				
500	150	2.5e–02	3.2e–01	2.2e–00
1000	150	6.8e–10	2.1e–06	8.2e–06
1500	150	1.3e–13	8.6e–10	6.8e–09

TABLE 7.2
Errors of Example 7.3.

Example 7.4 Let φ be the 1–periodized Gaussian function (4.1) with $n = 128$ and $b = 5$. We consider the following sum of translates

$$f(x) = \sum_{j=1}^{12} \varphi(x + s_j) \tag{7.2}$$

with the shift parameters

$$(s_j)_{j=1}^{12} = (-0.44, -0.411, -0.41, -0.4, -0.2, -0.01, 0.01, 0.02, 0.05, 0.15, 0.2, 0.215)^{\mathrm{T}}.$$

Note that all coefficients c_j $(j = 1, \ldots, 12)$ are equal to 1. The separation distance of the shift parameters is very small with 0.001. We work with exact sampled data $\tilde{f}_k = f(\frac{k}{128})$ $(k = -64, \ldots, 63)$. By Algorithm 4.7, we can compute the shift parameters $\tilde{s}_j$ with high accuracy

$$\max_{j=1,\ldots,12} |s_j - \tilde{s}_j| = 4.8 \cdot 10^{-10}\,.$$

For the coefficients we observe an error of size

$$\max_{j=1,\ldots,12} |1 - \tilde{c}_j| = 8.8 \cdot 10^{-7}\,.$$

We determine the error $\|f - \tilde{f}\|_\infty = 6.19 \cdot 10^{-13}$ as the discretized uniform norm $\max |f(x) - \tilde{f}(x)|$ on 8192 equidistant points $x \in [-\frac{1}{2}, \frac{1}{2}]$ with

$$\tilde{f}(x) := \sum_{j=1}^{12} \tilde{c}_j\, \varphi(x + \tilde{s}_j)\,.$$

□

Example 7.5 Now we consider the function (7.2) with the shift parameters $s_7 = -s_6 = 0.09$, $s_8 = -s_5 = 0.11$, $s_9 = -s_4 = 0.21$, $s_{10} = -s_3 = 0.31$, $s_{11} = -s_2 = 0.38$, $s_{12} = -s_1 = 0.41$. The 1–periodic function f and the 64 sampling points are shown in Figure 7.3. The separation distance of the shift parameters is now 0.02. Using exact

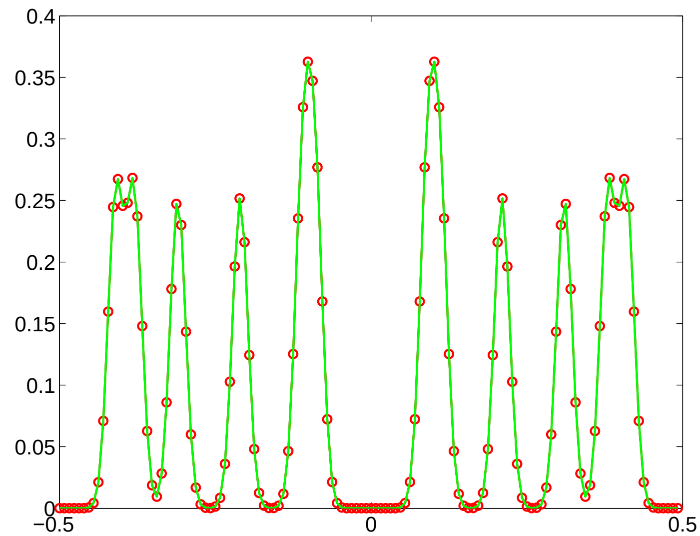

FIG. 7.3. *The function f from Example 7.5 with exact sampled data.*

sampled data $\tilde{f}_k = f(\frac{k}{128})$ $(k = -64, \ldots, 63)$, we expect a more accurate solution, see Section 5. By Algorithm 4.7, we can compute the shift parameters $\tilde{s}_j$ with high accuracy

$$\max_{j=1,\ldots,12} |s_j - \tilde{s}_j| = 2.81 \cdot 10^{-14} .$$

For the coefficients we observe an error of size

$$\max_{j=1,\ldots,12} |1 - \tilde{c}_j| = 1.71 \cdot 10^{-13} .$$

Now we consider the same function f with perturbed sampled data $\tilde{f}_k = f(\frac{k}{128}) + e_k$ $(k = -64, \ldots, 63)$, where $e_k \in [0,\, 0.01]$ are uniformly distributed random error terms. Then the computed shift parameters $\tilde{s}_j$ have an error of size

$$\max_{j=1,\ldots,12} |s_j - \tilde{s}_j| \approx 4.82 \cdot 10^{-4} .$$

For the coefficients we obtain an error

$$\max_{j=1,\ldots,12} |1 - \tilde{c}_j| \approx 5.52 \cdot 10^{-2} .$$

If the error $\|f - \tilde{f}\|_\infty$ is defined as in the Example 7.4, then we receive $\|f - \tilde{f}\|_\infty = 3.19 \cdot 10^{-2}$. $\square$

Example 7.6 Finally we estimate the parameters of an exponential sum (6.1) from nonuniform sampling points. We use the function (7.2) with the same shift parameters s_j $(j = 1, \ldots, 12)$ as in Example 7.5. The coefficients c_j $(j = 1, \ldots, 12)$ are uniformly distributed pseudo–random numbers in [0,1]. Then we choose 128 uniformly distributed pseudo–random numbers $x_k \in [-0.5,\, 0.5]$ as sampling nodes and set $N = 32$, see Figure 7.4. Using Algorithm 6.1, we compute the coefficients h_l $(l = 1, \ldots, 40)$ in

(6.2) and then the values $\tilde{h}(n/32)$ at the equidistant points $n/32$ ($n = -16, \dots, 15$). By Algorithm 2.4 we compute the shift parameters $\tilde{s}_j$ with an error of size

$$\max_{j=1,\dots,12} |s_j - \tilde{s}_j| \approx 8.24 \cdot 10^{-3}\,.$$

For the coefficients we obtain an error of size

$$\max_{j=1,\dots,12} |c_j - \tilde{c}_j| \approx 6.43 \cdot 10^{-2}\,.$$

If the error $\|f - \tilde{f}\|_\infty$ is defined as in the Example 7.4, then we receive $\|f - \tilde{f}\|_\infty = 1.29 \cdot 10^{-2}$. □

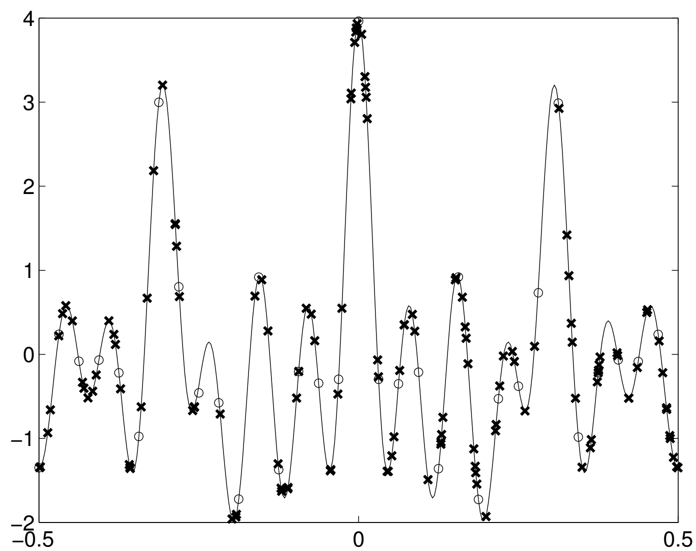

FIG. 7.4. *The function f from Example 7.6 with 128 nonequispaced sampling points $\times$ and with 32 equidistant sampling points $\circ$ computed by Algorithm 6.1.*

8. Conclusions. In this paper, we have presented nonlinear approximation techniques in order to recover all significant parameters of a signal from given noisy sampled data. Our first problem was devoted to the parameter reconstruction of a linear combination h of complex exponentials with real frequencies, where finitely many noisy uniformly sampled data of h are given. All parameters of h (i.e., all frequencies, all coefficients, and the number of exponentials) are computed by an approximate Prony method based on original ideas of G. Beylkin and L. Monzón [3, 4] and developed further by two of the authors [26, 27]. The emphasis of G. Beylkin and L. Monzón is placed on an approximate compressed representation of a function, i.e., finding a minimal number of exponentials and convenient coefficients to fit a function within a given accuracy. Our approach is mainly motivated by methods of signal recovery. Our question reads as follows: What is the significant part (in the form (1.1)) of a signal, where only finitely many perturbed sampled data of this signal are given? In our approach, the perturbation of the signal is bounded by a small constant and contains both the measurement error and the error of signal terms.

The nonlinear problem of finding all frequencies and coefficients of h was solved in

two steps. First we have computed the smallest singular value of a rectangular Hankel matrix formed by the sampled data. Then we have found the frequencies via zeros of a convenient polynomial. In the second step, we have used the obtained frequencies to solve an overdetermined linear Vandermonde–type system in a weighted least squares sense. It is interesting that this second step is closely related to the nonequispaced fast Fourier transform. In contrast to [3, 4], we have presented a new approach based on the perturbation theory for the singular value decomposition of a rectangular Hankel matrix. Using the Ingham inequalities, we have investigated the stability of the approximation by exponential sums with respect to the square and uniform norm. An extension to a parameter reconstruction of an exponential sum from given noisy nonuniformly sampled data is proposed too.
In a second problem, we have considered the parameter reconstruction of a linear combination f of shifted versions of a 1–periodic window function. For given noisy uniformly sampled data of f, we have recovered all parameters of f (i.e., all shift parameters, all coefficients, and the number of translates). The second problem is related to the first one by using Fourier technique such that this problem can be also solved by the approximate Prony method. Several numerical experiments have shown the performance of the algorithms proposed in this paper.

Acknowledgment. The second named author gratefully acknowledges the support by the German Research Foundation within the project KU 2557/1-1. The authors thank the reviewers for their helpful comments to improve this paper.

REFERENCES

[1] F. S. V. Bazán. Conditioning of rectangular Vandermonde matrices with nodes in the unit disk. *SIAM J. Matrix Anal. Appl.*, 21:679 – 693, 1999.
[2] G. Beylkin. On the fast Fourier transform of functions with singularities. *Appl. Comput. Harmon. Anal.*, 2:363 – 381, 1995.
[3] G. Beylkin and L. Monzón. On generalized Gaussian quadratures for exponentials and their applications. *Appl. Comput. Harmon. Anal.*, 12:332 – 373, 2002.
[4] G. Beylkin and L. Monzón. On approximations of functions by exponential sums. *Appl. Comput. Harmon. Anal.*, 19:17 – 48, 2005.
[5] G. Beylkin and L. Monzón. Nonlinear inversion of a band-limited Fourier transform. *Appl. Comput. Harmon. Anal.*, 27:351 – 366, 2009.
[6] G. Beylkin and L. Monzón. Approximation by exponential sums revisited. *Appl. Comput. Harmon. Anal.*, 28:131 – 149, 2010.
[7] W. L. Briggs and V. E. Henson. *The* DFT. *An Owner's Manual for the Discrete Fourier Transform.* SIAM, Philadelphia, 1995.
[8] O. Christensen. *An Introduction to Frames and Riesz Bases.* Applied and Numerical Harmonic Analysis. Birkhäuser, Boston, 2003.
[9] L. Coluccio, A. Eisinberg, and G. Fedele. A Prony–like method for non–uniform sampling. *Signal Process.*, 87:2484 – 2490, 2007.
[10] C. Corduneanu. *Almost Periodic Functions.* Chelsea, New York, 1989.
[11] A. J. W. Duijndam and M. A. Schonewille. Nonuniform fast Fourier transform. *Geophysics*, 64:539 – 551, 1999.
[12] A. Dutt and V. Rokhlin. Fast Fourier transforms for nonequispaced data. *SIAM J. Sci. Stat. Comput.*, 14:1368 – 1393, 1993.
[13] B. Elbel and G. Steidl. Fast Fourier transform for nonequispaced data. In C. K. Chui and L. L. Schumaker, editors, *Approximation Theory IX*, pages 39 – 46, Vanderbilt Univ. Press, Nashville, 1998.
[14] F. Filbir, H. N. Mhaskar, and J. Prestin. On the problem of parameter estimation in exponential sums. Preprint 2010 – 25, Institute of Biomathematics and Biometry, Helmholtz Center Munich, 2010.
[15] G. H. Golub, P. Milanfar, and J. Varah. A stable numerical method for inverting shape from moments. *SIAM J. Sci. Comput.*, 21:1222 – 1243, 1999.

[16] L. Greengard and J.-Y. Lee. Accelerating the nonuniform fast Fourier transform. *SIAM Rev.*, 46:443 – 454, 2004.
[17] F. B. Hildebrand. *Introduction to Numerical Analysis.* McGraw–Hill, New York, 1974.
[18] R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambridge Univ. Press, Cambridge, 1985.
[19] A. E. Ingham. Some trigonometrical inequalities with applications to the theory of series. *Math. Z.*, 41(1):367 – 379, 1936.
[20] J. I. Jackson, C. H. Meyer, D. G. Nishimura, and A. Macovski. Selection of a convolution function for Fourier inversion using gridding. IEEE *Trans. Med. Imag.*, 10:473 – 478, 1991.
[21] J. Keiner, S. Kunis, and D. Potts. Using NFFT3 – a software library for various nonequispaced fast Fourier transforms. ACM *Trans. Math. Software*, 36:Article 19, 1 – 30, 2009.
[22] S. Kunis and D. Potts. Stability results for scattered data interpolation by trigonometric polynomials. *SIAM J. Sci. Comput.*, 29:1403 – 1419, 2007.
[23] D. G. Manolakis, V. K. Ingle, and S. M. Kogon. *Statistical and Adaptive Signal Processing.* McGraw–Hill, Boston, 2005.
[24] J. M. Papy, L. De Lathauwer, and S. Van Huffel. Exponential data fitting using multilinear algebra: the decimative case. *J. Chemometrics*, 23:341 – 351, 2004.
[25] D. Potts, G. Steidl, and M. Tasche. Fast Fourier transforms for nonequispaced data: A tutorial. In J. J. Benedetto and P. J. S. G. Ferreira, editors, *Modern Sampling Theory: Mathematics and Applications*, pages 247 – 270, Birkhäuser, Boston, 2001.
[26] D. Potts and M. Tasche. Parameter estimation for exponential sums by approximate Prony method. *Signal Process.*, 90:1631 – 1642, 2010.
[27] D. Potts and M. Tasche. Nonlinear approximation by sums of nonincreasing exponentials. *Appl. Anal.*, in print.
[28] G. Steidl. A note on fast Fourier transforms for nonequispaced grids. *Adv. Comput. Math.*, 9:337 – 353, 1998.
[29] R. M. Young. *An Introduction to Nonharmonic Fourier Series.* Academic Press, New York, 1980.