Efficient spectral estimation by MUSIC and ESPRIT with application to sparse FFT

D. Potts, M. Tasche, T. Volkmer

Preprint 2015-14

Preprintreihe der Fakultät für Mathematik
ISSN 1614-8835

EFFICIENT SPECTRAL ESTIMATION BY MUSIC AND ESPRIT WITH APPLICATION TO SPARSE FFT

DANIEL POTTS*, MANFRED TASCHE†, AND TONI VOLKMER‡

Abstract. In the spectral estimation, one has to determine all parameters of an exponential sum, if only finitely many (noisy) sampled data of this exponential sum are given. Frequently used methods for spectral estimation are MUSIC (= MUltiple SIgnal Classification) and ESPRIT (= Estimation of Signal Parameters via Rotational Invariance Technique). For a trigonometric polynomial of large sparsity, we present a new sparse fast Fourier transform by shifted sampling and using MUSIC resp. ESPRIT, where the ESPRIT based methods will be faster. Later this technique is extended to the reconstruction of multivariate trigonometric polynomials of large sparsity, if (noisy) sampled values on a reconstructing rank-1 lattice are given. Numerical experiments illustrate the high performance of this procedure.

Key words and phrases: Spectral estimation, ESPRIT, MUSIC, exponential sum, sparsity, frequency analysis, parameter identification, rectangular Hankel matrix, sparse fast Fourier transform, sparse trigonometric polynomial.

AMS *Subject Classifications*: 65T50, 42A16, 94A12.

1. Introduction. The problem of spectral estimation resp. frequency analysis arises quite often in signal processing, electrical engineering, and mathematical physics (see e.g. the survey [19]) and reads as follows:

(P1) Recover the positive integer M, the distinct frequencies $\omega_j \in (-\frac{1}{2}, \frac{1}{2}]$, and the complex coefficients $c_j \neq 0$ $(j = 1, \ldots, M)$ in the exponential sum of sparsity M

$$h(x) := \sum_{j=1}^{M} c_j \, \mathrm{e}^{2\pi \mathrm{i} \omega_j x} \quad (x \in \mathbb{R}) \,, \tag{1.1}$$

if noisy sampled data $\tilde{h}_k := h(k) + e_k$ $(k = 0, \ldots, N-1)$ with $N \geq 2\,M$ are given, where $e_k \in \mathbb{C}$ are small error terms with $|e_k| \leq \varepsilon_1$ and $0 \leq \varepsilon_1 \ll \min_j |c_j|$.

Introducing so-called left/right signal spaces and noise spaces in Section 2, we explain the numerical solution of the problem (P1) by the MUSIC method (created by [26]) and the ESPRIT method (created by [25]). We show that both methods are based on singular value decomposition (SVD) of a rectangular Hankel matrix of given sampled data. The main disadvantages of MUSIC and ESPRIT are the high computational cost in the case of large sparsity M, caused mainly by the SVD. The computational cost of an algorithm is measured in the number of arithmetical operations, where all operations are counted equally. Often the computational cost of an algorithm is reduced to the leading term, i.e., all lower order terms are omitted.
Our aim is an efficient spectral estimation for a moderate number of samples with low computational cost, if one has to recover an exponential sum (1.1) of large sparsity M. Therefore we specialize the problem (P1). Let $S > 0$ be a large even integer. Replacing the variable x by Sx in (1.1), we consider the 1-periodic trigonometric

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

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

‡toni.volkmer@mathematik.tu-chemnitz.de, Technische Universität Chemnitz, Faculty of Mathematics, D–09107 Chemnitz, Germany

polynomial of sparsity M

$$g(x) := h(Sx) := \sum_{j=1}^{M} c_j \, \mathrm{e}^{2\pi \mathrm{i}\omega_j x} \quad (x \in \mathbb{R}) \,, \tag{1.2}$$

but now with integer frequencies $\omega_j \in (-\frac{S}{2}, \frac{S}{2}] \cap \mathbb{Z}$. Consequently, in Section 3 we investigate the following spectral estimation problem:

(P2) Recover the sparsity $M \in \mathbb{N}$, all integer frequencies $\omega_j \in (-\frac{S}{2}, \frac{S}{2}] \cap \mathbb{Z}$ as well as all non-zero coefficients $c_j \in \mathbb{C}$ of the trigonometric polynomial (1.2) from noisy sampled values $\tilde{g}_k := g(\frac{k}{S}) + e_k = h(k) + e_k$ $(k = 0, \ldots, N-1)$ with $N \geq 2\,M$, where $e_k \in \mathbb{C}$ are small error terms with $|e_k| \leq \varepsilon_1$ and $0 \leq \varepsilon_1 \ll \min_j |c_j|$. Often one considers the modified problem (P2*), where the sparsity M is known.

A numerical solution of problem (P2) or (P2*) with low computational cost is called sparse fast Fourier transform (sparse FFT). Using divide-and-conquer technique, the trigonometric polynomial (1.2) of large sparsity M is split into some trigonometric polynomials of lower sparsity and corresponding samples are determined by fast Fourier transform (FFT). Here we borrow an idea from sparse FFT in [15, 2] and use shifted sampling of (1.2), i.e., equidistant sampling with few equidistant shifts. Then the trigonometric polynomials of lower sparsity can be recovered by MUSIC resp. ESPRIT. The computational cost of the new sparse FFT is analyzed too.
A similar splitting technique is suggested in [15, 2], but with a different method to detect frequencies, when aliasing between two or more frequencies occurs. The method in [15, 2] follows an idea of [8], which is based on the Chinese Remainder Theorem, see also [1]. A different method for the sparse FFT, based on efficient filters is suggested in [7, 5]. We remark that there are two types of methods, deterministic (see [8]) and randomized (see [7, 15, 5]). Further related randomized methods based on compressed sensing can be found in the papers [24, 14, 6] and in the monograph [4]. Please note that the sparse FFT methods mentioned before solve the problem (P2*), i.e., one assumes that the sparsity is known, whereas our new deterministic sparse FFT also detects the sparsity M.
In Section 4, several numerical experiments with noiseless resp. noisy sampled data illustrate the high performance of the sparse FFT proposed. Note that in the case of successful recovery of the sparse trigonometric polynomial (1.2) all frequencies are correctly detected.
In Section 5, we extend our method to the reconstruction of multivariate trigonometric polynomials of large sparsity, where sampled data on a convenient rank-1 lattice are given. Numerical examples for 6-variate trigonometric polynomials of sparsity 256 are given. Here in the case of successful recovery of a sparse multivariate trigonometric polynomial, all frequency vectors are detected without errors.

2. Reconstruction of exponential sums. The main difficulty is the recovery of the *frequency set* $\Omega := \{\omega_1, \ldots, \omega_M\}$ in (1.1). Note that the distance between two frequencies $\omega_j,\, \omega_\ell \in (-\frac{1}{2},\, \frac{1}{2}]$ is measured by

$$d(\omega_j, \omega_\ell) := \min_{n \in \mathbb{Z}} |\omega_j + n - \omega_\ell| \,.$$

We introduce the *rectangular Fourier–type matrix* $\boldsymbol{F}_{N,M} := \big(\mathrm{e}^{2\pi\mathrm{i}\omega_j(k-1)}\big)_{k,j=1}^{N,M}$. Note that $\boldsymbol{F}_{N,M}$ coincides with the *rectangular Vandermonde matrix* $\boldsymbol{V}_{N,M}(\boldsymbol{z}) := (z_j^{k-1})_{k,j=1}^{N,M}$ with $\boldsymbol{z} := (z_j)_{j=1}^{M}$, where $z_j := \mathrm{e}^{2\pi\mathrm{i}\omega_j}$ $(j = 1, \ldots, M)$ are distinct nodes on the unit circle. Then the spectral estimation problem can be formulated in following matrix-vector form

$$\boldsymbol{V}_{N,M}(\boldsymbol{z})\, \boldsymbol{c} = \big(\tilde{h}_k\big)_{k=0}^{N-1}\,, \tag{2.1}$$

where $\tilde{\boldsymbol{h}} := (\tilde{h}_k)_{k=0}^{N-1}$ is the vector of noisy sampled data and $\boldsymbol{c} := (c_j)_{j=1}^{M}$ the vector of complex coefficients.
Under the natural assumption that the nodes z_j $(j = 1, \ldots, M)$ are well-separated on the unit circle, it can be shown that $\boldsymbol{F}_{P,M}$ is well conditioned for sufficiently large $P > M$.

Theorem 2.1 *(see [16, Theorem 2]) Assume that the frequencies $\omega_j \in (-\frac{1}{2},\, \frac{1}{2}]$ $(j = 1, \ldots, M)$ are well-separated by the separation distance*

$$q := \min_{j \neq \ell}\, d(\omega_j, \omega_\ell) > 0$$

and that $P > \max\{M,\, 2\pi + \frac{1}{q}\}$.
Then the discrete Ingham inequalities related to $\boldsymbol{F}_{P,M}$ indicate that for all $\boldsymbol{x} \in \mathbb{C}^M$

$$\alpha_1(P)\, \|\boldsymbol{x}\|_2^2 \le \|\boldsymbol{F}_{P,M}\, \boldsymbol{x}\|_2^2 \le \alpha_2(P)\, \|\boldsymbol{x}\|_2^2 \tag{2.2}$$

with

$$\alpha_1(P) := P\, \left(\frac{2}{\pi} - \frac{2}{\pi P^2 q^2} - \frac{4}{P}\right)\,, \quad \alpha_2(P) := P\, \left(\frac{4\sqrt{2}}{\pi} + \frac{\sqrt{2}}{\pi P^2 q^2} + \frac{3\sqrt{2}}{P}\right)\,.$$

Furthermore, the rectangular Fourier–type matrix $\boldsymbol{F}_{P,M}$ has a uniformly bounded spectral norm condition number

$$\mathrm{cond}_2\, \boldsymbol{F}_{P,M} \le \sqrt{\frac{4\, P^2 q^2 + 1 + 3\, \pi P q^2}{\sqrt{2}\, P^2 q^2 - \sqrt{2} - 2\sqrt{2}\pi P q^2}}\,.$$

Proof. The assumption $P > 2\pi + \frac{1}{q}$ is sufficient that the gap condition

$$q > \frac{1}{P}\, \left(1 - \frac{2\pi}{P}\right)^{-1/2} \tag{2.3}$$

is fulfilled. The gap condition (2.3) ensures that $\alpha_1(P) > 0$. For a proof of the discrete Ingham inequalities (2.2) under the gap condition (2.3) see [16, Theorem 2]. Let $\lambda_1 \ge \ldots \ge \lambda_M > 0$ be the ordered eigenvalues of $\boldsymbol{F}_{P,M}^{*} \boldsymbol{F}_{P,M} \in \boldsymbol{C}^{M \times M}$. Using the Raleigh–Ritz Theorem and (2.2), we obtain that for all $\boldsymbol{x} \in \mathbb{C}^M$

$$\alpha_1(P)\, \|\boldsymbol{x}\|_2^2 \le \lambda_M\, \|\boldsymbol{x}\|_2^2 \le \|\boldsymbol{F}_{P,M}\, \boldsymbol{x}\|_2^2 \le \lambda_1\, \|\boldsymbol{x}\|_2^2 \le \alpha_2(P)\, \|\boldsymbol{x}\|_2^2$$

and hence

$$0 < \alpha_1(P) \le \lambda_M \le \lambda_1 \le \alpha_2(P) < \infty\,. \tag{2.4}$$

Thus $\boldsymbol{F}_{P,M}^{*} \boldsymbol{F}_{P,M}$ is positive definite and

$$\mathrm{cond}_2\, \boldsymbol{F}_{P,M} = \sqrt{\frac{\lambda_1}{\lambda_M}} \le \sqrt{\frac{\alpha_2(P)}{\alpha_1(P)}}\,.$$

This completes the proof. ■

Corollary 2.1 *Assume that the frequencies* $\omega_j \in (-\frac{1}{2}, \frac{1}{2}]$ $(j = 1, \ldots, M)$ *are well-separated by the separation distance* $q > 0$ *and that* $P > \max\{M, 2\pi + \frac{1}{q}\}$*.*
Then the discrete Ingham inequalities related to $\boldsymbol{F}_{P,M}^{\mathrm{T}}$ *indicate that for all* $\boldsymbol{y} \in \mathbb{C}^P$

$$\alpha_1(P)\, \|\boldsymbol{y}\|_2^2 \le \|\boldsymbol{F}_{P,M}^{\mathrm{T}}\, \boldsymbol{y}\|_2^2 \le \alpha_2(P)\, \|\boldsymbol{y}\|_2^2\,. \tag{2.5}$$

Proof. The matrices $\boldsymbol{F}_{P,M}$ and $\boldsymbol{F}_{P,M}^{\mathrm{T}}$ possess the same singular values λ_j $(j = 1, \ldots, M)$. By the Rayleigh–Ritz Theorem we obtain that

$$\lambda_M\, \|\boldsymbol{y}\|_2^2 \le \|\boldsymbol{F}_{P,M}^{\mathrm{T}}\, \boldsymbol{y}\|_2^2 \le \lambda_1\, \|\boldsymbol{y}\|_2^2$$

for all $\boldsymbol{y} \in \mathbb{C}^P$. Applying (2.4), it follows the discrete Ingham inequalities (2.5). ■

Remark 2.2 The Riesz stability of the exponentials $\exp(2\pi \mathrm{i}\omega_j x)$ $(j = 1, \ldots, M)$ in the Hilbert space $\ell^2(\mathbb{Z}_N)$ follows immediately from the discrete Ingham inequalities (2.2), where $\mathbb{Z}_N := \{0, \ldots, N-1\}$ denotes the sampling grid. If the assumptions of Theorem 2.1 are fulfilled for $P = N$, then the exponentials $\exp(2\pi \mathrm{i}\omega_j x)$ $(j = 1, \ldots, M)$ are Riesz stable with respect to the discrete norm of $\ell^2(\mathbb{Z}_N)$, i.e.

$$\alpha_1(N)\, \|\boldsymbol{c}\|_2^2 \le \sum_{k=0}^{N-1} |h(k)|^2 \le \alpha_2(N)\, \|\boldsymbol{c}\|_2^2$$

for all exponential sums (1.1) with arbitrary coefficient vectors $\boldsymbol{c} = (c_j)_{j=1}^M \in \mathbb{C}^M$. Note that the Riesz stability of these exponentials related to continuous norms was formerly discussed and applied in spectral estimation in [18, 21]. □

In practice, the sparsity M of the exponential sum (1.1) is often unknown. Assume that $L \in \mathbb{N}$ is a convenient upper bound of M with $M \le L \le N - M + 1$. In applications, such an upper bound L is mostly known *a priori*. If this is not the case, then one can choose $L \approx \frac{N}{2}$. Often the sequence $\{\tilde{h}_0, \tilde{h}_1, \ldots, \tilde{h}_{N-1}\}$ of sampled data is called a *time series of length* N. Then we form the *L-trajectory matrix* of this time series

$$\tilde{\boldsymbol{H}}_{L,N-L+1} := \big(\tilde{h}_{\ell+m}\big)_{\ell,m=0}^{L-1,N-L} \tag{2.6}$$

with the *window length* $L \in \{M, \ldots, N - M + 1\}$. Analogously, we define

$$\boldsymbol{H}_{L,N-L+1} := \big(h(\ell+m)\big)_{\ell,m=0}^{L-1,N-L}\,. \tag{2.7}$$

Obviously, (2.6) and (2.7) are $L \times (N - L + 1)$ Hankel matrices. For simplicity, we consider mainly the noiseless case, i.e. $\tilde{h}_k = h(k)$ $(k = 0, \ldots, N-1)$.
The main step in the solution of the frequency analysis problem (P1) is the determination of the sparsity M and the computation of the frequencies ω_j or alternatively of the nodes $z_j = \mathrm{e}^{2\pi \mathrm{i}\omega_j}$ $(j = 1, \ldots, M)$. Afterwards one can calculate the coefficient vector $\boldsymbol{c} \in \mathbb{C}^M$ as least squares solution of the overdetermined linear system (2.1), i.e., the coefficient vector $\boldsymbol{c}$ is the solution of the least squares problem

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

We denote square matrices with only one index and refer to the well known fact that the square Vandermonde matrix $\boldsymbol{V}_M(\boldsymbol{z})$ is invertible and the matrix $\boldsymbol{V}_{N,L}(\boldsymbol{z})$

with $L \in \{M, \ldots, N-M+1\}$ has full column rank. Additionally we introduce the rectangular Hankel matrices

$$\tilde{\boldsymbol{H}}_{L,N-L}(s) := \tilde{\boldsymbol{H}}_{L,N-L+1}(1:L,\, 1+s:N-L+s) \quad (s=0,\,1)\,. \tag{2.8}$$

In the case of noiseless data $\tilde{h}_k = h(k)$ $(k=0,\ldots,N-1)$, the related Hankel matrices (2.8) are denoted by $\boldsymbol{H}_{L,N-L}(s)$ $(s=0,\,1)$.

Remark 2.3 The Hankel matrix $\boldsymbol{H}_{L,N-L+1}$ has the rank M for each window length $L \in \{M, \ldots, N-M+1\}$ and the related Hankel matrices $\boldsymbol{H}_{L,N-L}(s)$ $(s=0,\,1)$ possess the same rank M for each window length $L \in \{M, \ldots, N-M\}$ (see [20, Lemma 2.1]). Consequently, the order M of the exponential sum (1.1) coincides with the rank of these Hankel matrices. □

By the Vandermonde decomposition of the Hankel matrix $\boldsymbol{H}_{L,N-L+1}$ we obtain that

$$\boldsymbol{H}_{L,N-L+1} = \boldsymbol{V}_{L,M}(\boldsymbol{z})\,(\operatorname{diag}\boldsymbol{c})\,(\boldsymbol{V}_{N-L+1,M}(\boldsymbol{z}))^{\mathrm{T}}\,. \tag{2.9}$$

Under mild conditions, the Hankel matrix $\boldsymbol{H}_{L,N-L+1}$ is well-conditioned too.

Theorem 2.2 *Let $L,\, N \in \mathbb{N}$ with $M \le L \le N-M+1$ and $\min\{L,\, N-L+1\} > 2\pi + \frac{1}{q}$ be given. Assume that the frequencies $\omega_j \in (-\frac{1}{2},\, \frac{1}{2}]$ $(j=1,\ldots,M)$ are well-separated by the separation distance $q>0$ and that the non-zero coefficients c_j $(j=1,\ldots,M)$ of the exponential sum (1.1) fulfil the condition*

$$0 < \gamma_1 \le |c_j| \le \gamma_2 < \infty \quad (j=1,\ldots,M)\,. \tag{2.10}$$

Then for all $\boldsymbol{y} \in \mathbb{C}^{N-L+1}$

$$\gamma_1^2\,\alpha_1(L)\,\alpha_1(N-L+1)\,\|\boldsymbol{y}\|_2^2 \le \|\boldsymbol{H}_{L,N-L+1}\,\boldsymbol{y}\|_2^2 \le \gamma_2^2\,\alpha_2(L)\,\alpha_2(N-L+1)\,\|\boldsymbol{y}\|_2^2\,. \tag{2.11}$$

Further, the lowest (non-zero) resp. largest singular value of $\boldsymbol{H}_{L,N-L+1}$ can be estimated by

$$0 < \gamma_1\,\sqrt{\alpha_1(L)\,\alpha_1(N-L+1)} \le \sigma_M \le \sigma_1 \le \gamma_2\,\sqrt{\alpha_2(L)\,\alpha_2(N-L+1)}\,.$$

The spectral norm condition number of $\boldsymbol{H}_{L,N-L+1}$ is bounded by

$$\operatorname{cond}_2 \boldsymbol{H}_{L,N-L+1} \le \frac{\gamma_2}{\gamma_1}\,\sqrt{\frac{\alpha_2(L)\,\alpha_2(N-L+1)}{\alpha_1(L)\,\alpha_1(N-L+1)}}\,.$$

Proof. By the Vandermonde decomposition (2.9) of the Hankel matrix $\boldsymbol{H}_{L,N-L+1}$, we obtain that for all $\boldsymbol{y} \in \mathbb{C}^{N-L+1}$

$$\|\boldsymbol{H}_{L,N-L+1}\,\boldsymbol{y}\|_2^2 = \|\boldsymbol{F}_{L,M}\,(\operatorname{diag}\boldsymbol{c})\,\boldsymbol{F}_{N-L+1,M}^{\mathrm{T}}\,\boldsymbol{y}\|_2^2\,.$$

By the discrete Ingham inequalities (2.2) and the assumption (2.10), it follows that

$$\gamma_1^2\,\alpha_1(L)\,\|\boldsymbol{F}_{N-L+1,M}^{\mathrm{T}}\,\boldsymbol{y}\|_2^2 \le \|\boldsymbol{H}_{L,N-L+1}\,\boldsymbol{y}\|_2^2 \le \gamma_2^2\,\alpha_2(N-L+1)\,\|\boldsymbol{F}_{N-L+1,M}^{\mathrm{T}}\,\boldsymbol{y}\|_2^2\,.$$

Using the discrete Ingham inequalities (2.5), we obtain the estimates (2.11). Finally, the estimates of the lowest resp. largest singular value and the spectral norm condition number of $\boldsymbol{H}_{L,N-L+1}$ arise from (2.11) and the Rayleigh–Ritz Theorem. ■

The ranges of $\boldsymbol{H}_{L,N-L+1}$ and $\boldsymbol{V}_{L,M}(\boldsymbol{z})$ coincide in the noiseless case with $M \le L \le N-M+1$ by (2.9). If $L > M$, then the range of $\boldsymbol{V}_{L,M}(\boldsymbol{z})$ is a proper subspace of $\mathbb{C}^L$. This subspace is called *left signal space* $\mathcal{S}_L$. The left signal space $\mathcal{S}_L$ is of dimension M and is generated by the M columns $\boldsymbol{e}_L(\omega_j)$ $(j = 1, \ldots, M)$, where

$$\boldsymbol{e}_L(\omega) := \big(\mathrm{e}^{2\pi \mathrm{i}\, \ell\omega}\big)_{\ell=0}^{L-1} \quad (\omega \in (-\frac{1}{2},\, \frac{1}{2}]) \,.$$

Note that $\|\boldsymbol{e}_L(\omega)\|_2 = \sqrt{L}$ for each $\omega \in (-\frac{1}{2},\, \frac{1}{2}]$. The *left noise space* $\mathcal{N}_L$ is defined as the orthogonal complement of $\mathcal{S}_L$ in $\mathbb{C}^L$. The dimension of $\mathcal{N}_L$ is equal to $L - M$.

Remark 2.4 Let $M \le L < N - M + 1$ be given. If we use $\boldsymbol{H}^*_{L,N-L+1}$ instead of $\boldsymbol{H}_{L,N-L+1}$, then we can define the *right signal space* as the range of $\boldsymbol{V}_{N-L,M}(\bar{\boldsymbol{z}})$, where $\bar{\boldsymbol{z}}$ denotes the complex conjugate of $\boldsymbol{z}$. The right signal space is an M-dimensional subspace of $\mathbb{C}^{N-L+1}$ and is generated by the M linearly independent vectors $\overline{\boldsymbol{e}_{N-L+1}(\omega_j)}$ $(j = 1, \ldots, M)$. Then the corresponding *right noise space* is the orthogonal complement of the right signal space in $\mathbb{C}^{N-L+1}$. $\square$

By $\boldsymbol{Q}_L$ we denote the orthogonal projection onto the left noise space $\mathcal{N}_L$. Since $\boldsymbol{e}_L(\omega_j) \in \mathcal{S}_L$ $(j = 1, \ldots, M)$ and $\mathcal{N}_L \perp \mathcal{S}_L$, we obtain that

$$\boldsymbol{Q}_L\, \boldsymbol{e}_L(\omega_j) = \boldsymbol{0} \quad (j = 1, \ldots, M)\,.$$

If $\omega \in (-\frac{1}{2},\, \frac{1}{2}] \setminus \Omega$, then the vectors $\boldsymbol{e}_L(\omega_1), \ldots, \boldsymbol{e}_L(\omega_M), \boldsymbol{e}_L(\omega) \in \mathbb{C}^L$ are linearly independent, since the square Vandermonde matrix

$$\big(\boldsymbol{e}_L(\omega_1)\,|\, \ldots \,|\, \boldsymbol{e}_L(\omega_M)\,|\, \boldsymbol{e}_L(\omega)\big)(1:M+1,\, 1:M+1)$$

is invertible for each $L \ge M + 1$. Hence $\boldsymbol{e}_L(\omega) \notin \mathcal{S}_L = \mathrm{span}\,\{\boldsymbol{e}_L(\omega_1), \ldots, \boldsymbol{e}_L(\omega_M)\}$, i.e. $\boldsymbol{Q}_L \boldsymbol{e}_L(\omega) \ne \boldsymbol{0}$. Thus the frequency set Ω can be determined via the zeros of the *left noise-space correlation function*

$$N_L(\omega) := \frac{1}{\sqrt{L}}\, \|\boldsymbol{Q}_L\, \boldsymbol{e}_L(\omega)\|_2 \quad (\omega \in (-\frac{1}{2},\, \frac{1}{2}])\,,$$

since $N_L(\omega_j) = 0$ for each $j = 1, \ldots, M$ and $0 < N_L(\omega) \le 1$ for all $\omega \in (-\frac{1}{2},\, \frac{1}{2}] \setminus \Omega$, where $\boldsymbol{Q}_L \boldsymbol{e}_L(\omega)$ can be computed on an equispaced fine grid. Alternatively, one can seek the peaks of the *left imaging function*

$$J_L(\omega) := \sqrt{L}\, \|\boldsymbol{Q}_L\, \boldsymbol{e}_L(\omega)\|_2^{-1} \quad (\omega \in (-\frac{1}{2},\, \frac{1}{2}])\,.$$

In this approach, we prefer the zeros resp. the lowest local minima of the left noise-space correlation function $N_L(\omega)$.

In the next step we determine the orthogonal projection $\boldsymbol{Q}_L$ onto the left noise space $\mathcal{N}_L$. Here we can use the singular value decomposition (SVD) or the QR decomposition of the L-trajectory matrix $\boldsymbol{H}_{L,N-L+1}$. For an application of QR decomposition see [20]. Applying SVD, we obtain that

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

where $\boldsymbol{U}_L \in \mathbb{C}^{L\times L}$ and $\boldsymbol{W}_{N-L+1} \in \mathbb{C}^{(N-L+1)\times(N-L+1)}$ are unitary matrices and where $\boldsymbol{D}_{L,N-L+1} \in \mathbb{R}^{L\times(N-L+1)}$ is a rectangular diagonal matrix. The diagonal

entries of $\boldsymbol{D}_{L,N-L+1}$ are the singular values σ_j of the L-trajectory matrix arranged in nonincreasing order $\sigma_1 \ge \ldots \ge \sigma_M > \sigma_{M+1} = \ldots = \sigma_{\min\{L,N-L+1\}} = 0$. Thus we can determine the order M of the exponential sum (1.1) by the number of positive singular values σ_j.
Introducing the matrices

$$\boldsymbol{U}_{L,M}^{(1)} := \boldsymbol{U}_L(1:L,\, 1:M)\,, \quad \boldsymbol{U}_{L,L-M}^{(2)} := \boldsymbol{U}_L(1:L,\, M+1:L)$$

with orthonormal columns, we see that the columns of $\boldsymbol{U}_{L,M}^{(1)}$ form an orthonormal basis of $\mathcal{S}_L$ and that the columns of $\boldsymbol{U}_{L,L-M}^{(2)}$ are an orthonormal basis of $\mathcal{N}_L$. Hence the orthogonal projection onto the left noise space $\mathcal{N}_L$ has the form

$$\boldsymbol{Q}_L = \boldsymbol{U}_{L,L-M}^{(2)}\,(\boldsymbol{U}_{L,L-M}^{(2)})^*\,.$$

Consequently, we obtain that

$$\begin{aligned}\|\boldsymbol{Q}_L\,\boldsymbol{e}_L(\omega)\|_2^2 &= \langle \boldsymbol{Q}_L\,\boldsymbol{e}_L(\omega),\, \boldsymbol{Q}_L\,\boldsymbol{e}_L(\omega)\rangle = \langle (\boldsymbol{Q}_L)^2\,\boldsymbol{e}_L(\omega),\, \boldsymbol{e}_L(\omega)\rangle \\ &= \langle \boldsymbol{Q}_L\,\boldsymbol{e}_L(\omega),\, \boldsymbol{e}_L(\omega)\rangle = \langle \boldsymbol{U}_{L,L-M}^{(2)}\,(\boldsymbol{U}_{L,L-M}^{(2)})^*\,\boldsymbol{e}_L(\omega),\, \boldsymbol{e}_L(\omega)\rangle \\ &= \langle (\boldsymbol{U}_{L,L-M}^{(2)})^*\,\boldsymbol{e}_L(\omega),\, (\boldsymbol{U}_{L,L-M}^{(2)})^*\,\boldsymbol{e}_L(\omega)\rangle = \|(\boldsymbol{U}_{L,L-M}^{(2)})^*\,\boldsymbol{e}_L(\omega)\|_2^2\,.\end{aligned}$$

Hence the left noise-space correlation function can be represented by

$$N_L(\omega) = \frac{1}{\sqrt{L}}\,\|(\boldsymbol{U}_{L,L-M}^{(2)})^*\,\boldsymbol{e}_L(\omega)\|_2 \quad (\omega \in (-\frac{1}{2},\, \frac{1}{2}])\,.$$

In MUSIC, one determines the lowest local minima of the left noise-space correlation function, see e.g. [26, 17, 3, 13].

Algorithm 2.5 (MUSIC via SVD)
Input: $L,\, N \in \mathbb{N}$ ($N \gg 1$, $M < L \le N-M+1$, M is the unknown sparsity of (1.1)), $\tilde{h}_k = h(k) + e_k \in \mathbb{C}$ ($k = 0, \ldots, N-1$) noisy sampled values of (1.1), $0 < \varepsilon \ll 1$ tolerance.

1. Compute the SVD of the rectangular Hankel matrix (2.6), where the singular values $\tilde{\sigma}_\ell$ ($\ell = 1, \ldots, \min\{L,\, N-L+1\}$) are arranged in nonincreasing order. Determine the numerical rank M of (2.6) such that $\tilde{\sigma}_M \ge \varepsilon\,\tilde{\sigma}_1$ and $\tilde{\sigma}_{M+1} < \varepsilon\tilde{\sigma}_1$. Form the matrix $\tilde{\boldsymbol{U}}_{L,L-M}^{(2)}$.
2. Calculate the left noise-space correlation function $\tilde{N}_L(\omega) := \frac{1}{\sqrt{L}}\,\|(\tilde{\boldsymbol{U}}_{L,L-M}^{(2)})^*\boldsymbol{e}_L(\omega)\|_2$ on an equispaced grid $\{-\frac{1}{2} + \frac{1}{S},\, \ldots,\, \frac{1}{2} - \frac{1}{S},\, \frac{1}{2}\}$ for sufficiently large S.
3. The M lowest local minima of $\tilde{N}_L(\frac{2k-S}{2S})$ ($k = 1, \ldots, S$) form the frequency set $\tilde{\Omega} := \{\tilde{\omega}_1, \ldots, \tilde{\omega}_M\}$. Set $\tilde{z}_j := \mathrm{e}^{2\pi\mathrm{i}\tilde{\omega}_j}$ ($j = 1, \ldots, M$).
4. Compute the coefficient vector $\tilde{\boldsymbol{c}} := (\tilde{c}_j)_{j=1}^M \in \mathbb{C}^M$ as solution of the least squares problem

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

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

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

Let $L, N \in \mathbb{N}$ with $M < L \le N - M + 1$ be given. For noisy sampled data $\tilde{h}_k = h(k) + e_k$ $(k = 0, \ldots, N-1)$, the MUSIC Algorithm 2.5 is relatively insensitive to small perturbations on the data (see [16, Theorem 3]).
In opposite to the MUSIC Algorithm 2.5, the following ESPRIT Algorithm is based on orthogonal projection onto a right signal space. For details see [25, 20, 23].

Algorithm 2.6 (ESPRIT via SVD)
Input: $L, N \in \mathbb{N}$ $(N \gg 1,\ M < L \le N - M + 1$, M is the unknown sparsity of (1.1)), $\tilde{h}_k = h(k) + e_k \in \mathbb{C}$ $(k = 0, \ldots, N-1)$ noisy sampled values of (1.1), $0 < \varepsilon \ll 1$ tolerance.

1. Compute the SVD of the rectangular Hankel matrix (2.6), where the singular values $\tilde{\sigma}_\ell$ $(\ell = 1, \ldots, \min\{L,\, N-L+1\})$ are arranged in nonincreasing order. Determine the numerical rank M of (2.6) such that $\tilde{\sigma}_M \ge \varepsilon\, \tilde{\sigma}_1$ and $\tilde{\sigma}_{M+1} < \varepsilon \tilde{\sigma}_1$. Form the submatrices

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

2. Calculate the matrix

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

where $\left(\tilde{\boldsymbol{W}}_{N-L,M}(0)^*\right)^{\dagger}$ denotes the Moore–Penrose pseudoinverse.
3. Determine all eigenvalues z'_j $(j = 1, \ldots, M)$ of $\tilde{\boldsymbol{F}}_M$. Set

$$\tilde{\omega}_j := \frac{1}{2\pi}\, \mathrm{Arg}\, \frac{z'_j}{|z'_j|} \in (-\frac{1}{2},\, \frac{1}{2}] \quad (j = 1, \ldots, M)\,,$$

where $\mathrm{Arg}\, z \in (-\pi,\, \pi]$ means the principal value of the argument of $z \in \mathbb{C} \setminus \{0\}$.
4. Compute the coefficient vector $\tilde{\boldsymbol{c}} := (\tilde{c}_j)_{j=1}^M \in \mathbb{C}^M$ as solution of the least squares problem

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

where $\tilde{\boldsymbol{z}} := \big(\tilde{z}_j\big)_{j=1}^M$ denotes the vector of computed nodes $\tilde{z}_j := \mathrm{e}^{2\pi\mathrm{i}\,\tilde{\omega}_j}$.

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

The numbers of required samples and the computational costs of the Algorithms 2.5 and 2.6 are listed in Table 2.1, where L is chosen such that $M < L \approx \frac{N}{2}$. Thus the main disadvantages of these algorithms are the high computational costs for large sparsity M, caused mainly by the SVD. Therefore in [22], we have suggested to use a partial SVD (based on partial Lanczos bidiagonalization) instead of a complete SVD. For both Algorithms 2.5 and 2.6, one needs too many operations in the case of large sparsity M, see Table 2.1.

3. Sparse fast Fourier transform. In this section, we apply Algorithm 2.5 (MUSIC) resp. Algorithm 2.5 (ESPRIT) to the reconstruction of sparse trigonometric polynomials. Clearly, one can approximate the unknown frequencies of the exponential sum (1.1) by fractions. Therefore we assume that the unknown frequencies of (1.1) are fractions $\frac{\omega_j}{S}$ with $\omega_j \in (-\frac{S}{2}, \frac{S}{2}] \cap \mathbb{Z}$, where S is a large even integer.

method	samples	computational cost
Algorithm 2.5 (MUSIC)	$2N+1$	$\mathcal{O}(N^3 + N^2 S + S \log S)$
Algorithm 2.6 (ESPRIT)	$2N+1$	$\mathcal{O}(N^3)$

TABLE 2.1

Numbers of required samples and computational costs for the Algorithms 2.5 and 2.6 in the case $M < L \approx \frac{N}{2}$.

Replacing the variable x by Sx in (1.1), we obtain the new exponential sum

$$g(x) := h(Sx) := \sum_{j=1}^{M} c_j \, \mathrm{e}^{2\pi \mathrm{i} \omega_j x} \quad (x \in \mathbb{R})\,. \tag{3.1}$$

Then g is a 1-*periodic trigonometric polynomial* with *sparsity* M. Consequently we consider the spectral estimation problem (P2) of Section 1.
In the following, we propose a new deterministic sparse FFT for solving the problem (P2) of a trigonometric polynomial (3.1) with large sparsity M. Using divide–and–conquer technique, we split the trigonometric polynomial (3.1) of large sparsity M into some trigonometric polynomials of lower sparsity and determine corresponding samples. Here we borrow an idea from sparse FFT in [2] and use shifted sampling of (3.1). For a positive integer $P \le S$, we construct a discrete array of samples of size $P \times (2K+1)$ via

$$g_P[s,k] := g\left(\frac{s}{P} + \frac{k}{S}\right), \quad (s = 0, \ldots, P-1;\ k = 0, \ldots, 2K).$$

For each $k = 0, \ldots, 2K$ we form the discrete Fourier transform (DFT) of length P and obtain

$$\hat{g}_P[\ell,k] := \sum_{s=0}^{P-1} g_P[s,k] \, \mathrm{e}^{-2\pi \mathrm{i} s \ell/P} \quad (\ell = 0, \ldots, P-1)\,.$$

The fast Fourier transform (FFT) allows the rapid computation of this DFT of length P in $\mathcal{O}(P \log P)$ operations. Then for each $\ell = 0, \ldots, P-1$, it follows that

$$\begin{aligned} \hat{g}_P[\ell,k] &= \sum_{s=0}^{P-1} \sum_{j=1}^{M} c_j \, \mathrm{e}^{2\pi \mathrm{i} \omega_j (s/P + k/S)} \, \mathrm{e}^{-2\pi \mathrm{i} s \ell/P} \\ &= \sum_{j=1}^{M} c_j \, \mathrm{e}^{2\pi \mathrm{i} \omega_j k/S} \underbrace{\sum_{s=0}^{P-1} \mathrm{e}^{2\pi \mathrm{i} (\omega_j - \ell) s/P}}_{=0 \text{ or } P}. \end{aligned}$$

Now we define the index sets

$$I_P(\ell) := \big\{ j \in \{1, \ldots, M\} : \omega_j \equiv \ell \ (\mathrm{mod}\, P) \big\}$$

such that

$$\hat{g}_P[\ell,k] = P \sum_{j \in I_P(\ell)} c_j \, \mathrm{e}^{2\pi \mathrm{i} \omega_j k/S}.$$

We choose $K \in \mathbb{N}$ as sparsity cut-off parameter. For each $\ell = 0, \ldots, P-1$ with $|I_P(\ell)| < K$, we apply Algorithm 2.5 resp. 2.6 to the "samples" $\hat{g}_P[\ell, k]$ ($k = 0, \ldots, 2K$), which yields the corresponding frequencies ω_j and coefficients c_j for $j \in I_P(\ell)$. If the condition $|I_P(\ell)| < K$ is fulfilled for all $\ell = 0, \ldots, P-1$, i.e., the frequencies ω_j are almost uniformly distributed on the sets $I_P(\ell)$ ($\ell = 0, \ldots, P-1$), this approach requires $(2K+1)\,P$ samples of g and $2K+1$ FFTs of length P. The computational costs for the corresponding algorithms are listed in Table 3.1.
If we cannot uniquely identify all frequencies, i.e., if $|I_P(\ell)| \geq K$ for some ℓ, then we form iteratively the new trigonometric polynomial

$$g_1(x) := g(x) - \sum_{j \in I} c_j \, \mathrm{e}^{2\pi \mathrm{i} \omega_j x}\,, \tag{3.2}$$

where I is the union of all $I_P(\ell)$ with the property $|I_P(\ell)| < K$. In the next iteration step, we choose a positive integer $P_1 \leq S$ different from P and repeat the method on the trigonometric polynomial g_1. In doing so, we can compute the values

$$\sum_{j \in I} c_j \, \mathrm{e}^{2\pi \mathrm{i} \omega_j (\frac{s}{P_1} + \frac{k}{S})} = \sum_{j \in I} \left(c_j \, \mathrm{e}^{2\pi \mathrm{i} \frac{\omega_j}{P_1} s} \right) \mathrm{e}^{2\pi \mathrm{i} \frac{\omega_j}{S} k} \quad (s = 0, \ldots, P_1 - 1;\, k = 0, \ldots, 2K)$$

by the nonequispaced fast Fourier transform (NFFT) [12] in $\mathcal{O}\left(P_1 (K \log K + |I|)\right)$ arithmetic operations.
We perform additional iterations until all frequencies can be identified, i.e., if $|I_{P_1}(\ell)| < K$ for all $\ell = 0, \ldots, P_1 - 1$. Note that our algorithm is related to the sparse FFT proposed in [2]. But here we use the methods of Section 2, if aliasing with respect to modulo P occurs.
All of the methods described in Section 2 apply an SVD and use the tolerance ε as a relative threshold parameter to determine the local sparsity M_ℓ of the signal. A good choice of this parameter may depend without limitation on noise in the sampling values of the trigonometric polynomial g and on the smallest distance between two frequencies, where this distance may change for each $\ell \in \{0, \ldots, P-1\}$ in each iteration. We propose to use a (small) list of possible relative threshold parameters ε, which are tested for each $\ell \in \{0, \ldots, P-1\}$ in each iteration.
Our sparse FFT for recovery of a trigonometric polynomial with large sparsity reads as follows:

Algorithm 3.1 (Sparse FFT via MUSIC resp. ESPRIT,
see Algorithm A.1 for detailed listing with extended parameter list)
Input: $S \in 2\mathbb{N}$, K, $P \in \mathbb{N}$ ($P \leq S$, M is the unknown sparsity of (3.1)), $\tilde{g}_P[s, k]$ noisy sampled value of (3.1) at $\frac{s}{P} + \frac{k}{S}$ for $s = 0, \ldots, P-1$ and $k = 0, \ldots, 2K$.

I. For each $k = 0, \ldots, 2K$ compute $\left(\hat{g}_P[\ell, k]\right)_{\ell=0}^{P-1}$ by FFT of $\left(\tilde{g}_P[s, k]\right)_{s=0}^{P-1}$ and form the Hankel matrix $\tilde{\boldsymbol{H}}_{K,2K+1} := \left(\hat{g}_P[\ell + m, k]\right)_{\ell,m=0}^{K-1,K+1}$.
II. For $\ell = 0, \ldots, P-1$:
II.1 Apply Algorithm 2.5 resp. 2.6 (with $L = K$ and $N = 2K+1$) on the Hankel matrix $\tilde{\boldsymbol{H}}_{K,2K+1}$. Determine the local sparsity M_ℓ and compute the local fractions $\tilde{\omega}_{\ell,j} \in (-\frac{1}{2}, \frac{1}{2}]$ for $j = 1, \ldots, M_\ell$. Compute the local frequencies $\omega_{\ell,j} := \mathrm{round}\,(\tilde{\omega}_{\ell,j}\, S)$ by rounding to nearest integer.
II.2. Compute the local coefficients $c_{\ell,j}$ as least squares solution of the overdetermined

Vandermonde system

$$\left\| P \left(\exp \frac{2\pi \mathrm{i} k \omega_{\ell,j}}{S} \right)_{k=0,j=1}^{2K,M_\ell} (c_{\ell,j})_{j=1}^{M_\ell} - (\hat{g}_P[\ell,k])_{k=0}^{2K} \right\|_2 = \min .$$

II.3. If the residual is small and $|c_{\ell,j}|$ is not too small, then append the frequencies $\omega_{\ell,j}$ $(j = 1, \ldots, M_\ell)$ to the frequency set Ω.
III. If $|I_P(\ell)| \geq K$ for some $\ell = 0, \ldots, P-1$, then form the new trigonometric polynomial (3.2). In the next iteration step choose a positive integer $P_1 \leq S$ different from P, sample (3.2) on $\frac{s}{P_1} + \frac{k}{S}$ for $s = 0, \ldots, P_1 - 1$ and $k = 0, \ldots, 2K$, and repeat the above method.

Output: $\Omega \subset (-\frac{S}{2}, \frac{S}{2}] \cap \mathbb{Z}$ set of recovered frequencies ω_j $(j = 1, \ldots, M)$, $M := |\Omega|$ detected sparsity, $c_j \in \mathbb{C}$ coefficient related to ω_j.

The numbers of required samples and the computational cost for one iteration of Algorithm 3.1 are given in Table 3.1.

method	samples	computational cost
Alg. 3.1 via MUSIC	$(2K+1)\,P$	$\mathcal{O}\left(KP^2 + K^3P + K^2S + S\log\frac{S}{P}\right)$
Alg. 3.1 via ESPRIT	$(2K+1)\,P$	$\mathcal{O}\left(KP^2 + K^3P\right)$

TABLE 3.1
Numbers of required samples and computational cost of one iteration step of Algorithm 3.1.

By choosing the parameters $K = \mathcal{O}(M^{1/3})$ and $P = \mathcal{O}(M^{2/3})$, we compare the numbers of required samples and computational costs for different methods of spectral estimation in Table 3.2 such as sparse FFT via MUSIC, sparse FFT via ESPRIT, MUSIC, ESPRIT, and classical FFT. As we can see, the sparse FFT via ESPRIT is very useful for the spectral estimation by a relatively low number of samples and low computational cost.

method	samples	computational cost
Alg. 3.1 via MUSIC Alg. 2.5	$\mathcal{O}(M)$	$\mathcal{O}\left(M^{2/3}S + S\log\frac{S}{M^{2/3}}\right)$
Alg. 3.1 via ESPRIT Alg. 2.6	$\mathcal{O}(M)$	$\mathcal{O}\left(M^{5/3}\right)$
Alg. 2.5 (MUSIC)	$\mathcal{O}\left(M\right)$	$\mathcal{O}\left(M^2S + S\log S\right)$
Alg. 2.6 (ESPRIT)	$\mathcal{O}\left(M\right)$	$\mathcal{O}\left(M^3\right)$
FFT of length S	S	$\mathcal{O}(S\log S)$

TABLE 3.2
Numbers of required samples and computational costs using the splitting approach for one iteration of Algorithm 3.1 as well as for Algorithms 2.5 and 2.6 in the case $K = \mathcal{O}(M^{1/3})$, $P = \mathcal{O}(M^{2/3})$ and $M \approx L/2 \approx N/4$.

4. Numerical experiments with sparse FFT. In this section, we present some numerical results for Algorithm 3.1. All computations are performed in MATLAB with IEEE double–precision arithmetic. First we consider noiseless sampled data and later the case, where the sampled data are perturbed by additive (white) Gaussian noise.

4.1. Noiseless sampled data.

Example 4.1 From noiseless sampled values, we reconstruct 100 trigonometric polynomials (3.1) of order $M = 256$ with random frequencies $\omega_j \in (-\frac{S}{2}, \frac{S}{2}] \cap \mathbb{Z}$ and random coefficients c_j on the unit circle. We set the array of relative SVD threshold values `epsilon_svd_list` $:= [10^{-1}, 10^{-2}, \ldots, 10^{-8}]$, the parameter $\varepsilon_{\text{spatial}} := 10^{-8}$, the absolute value of minimal non-zero coefficients $\varepsilon_{\text{fc\_min}} := 10^{-1} = 10^{-1} \cdot \min_j |c_j|$ and the maximal number of iterations $R := 10$, see Algorithm A.1 for the extended parameter list. Applying the sparse FFT Algorithm 3.1 with MUSIC in the case $S = 2^{16}$ with parameters $K \in \{6, 12, 16\}$ and $P \in \{16, 32, 64, 128\}$, we can successfully detect all integer frequencies ω_j. In Table 4.1, the column "iterations" depicts the maximal number of iterations actually used by the Algorithm 3.1 (computed over 100 trigonometric polynomials). The column "samples" contains the maximal number of sampled values used by the Algorithm 3.1. The column "ℓ^2–errors" shows two values. The first value is the maximal relative ℓ^2–error of the coefficients, which are locally computed in step II.2 of Algorithm 3.1. The second value is the maximal relative ℓ^2–error of the coefficients, which are determined by additional solving one large Vandermonde system at the end of Algorithm 3.1 with all frequencies as well as all samples of (3.1).

K	P	iterations	samples	ℓ^2-errors
6	16	10	3939	3.1e-09/2.8e-15
6	32	5	2600	8.5e-10/2.5e-15
6	64	3	2626	3.0e-10/2.5e-15
6	128	2	3367	5.7e-11/1.9e-14
12	16	5	2600	3.4e-10/2.4e-15
12	32	2	1725	1.7e-10/9.2e-15
12	64	2	3275	1.3e-10/7.9e-14
12	128	1	3200	1.4e-11/1.6e-14
16	16	3	1716	5.2e-10/2.3e-15
16	32	2	2277	2.3e-10/4.4e-14
16	64	2	4323	9.9e-11/3.3e-14
16	128	1	4224	7.4e-12/6.3e-15
850	1	1	1701	2.3e-13/2.3e-13

Table 4.1

Results for Algorithm 3.1 via MUSIC for frequency grid parameter $S = 2^{16}$ and sparsity $M = 256$.

Note that the sparse FFT Algorithm 3.1 via ESPRIT yields identical iteration and sampling numbers as well as similar error values, but it requires a shorter computational time. For comparison, the classical FFT of length 2^{16} requires 2^{16} samples and the resulting ℓ^2–error is 2.6e-16. The minimal number of samples for the cases $K \in \{6, 12, 16\}$ and $P \in \{16, 32, 64, 128\}$ is reached for $K = P = 16$ with 1716 samples, the next smallest number of samples is 1725 for $K = 12$ and $P = 32$. If we do not use the splitting approach ($P = 1$ and $R = 1$), we observe that the detection of some frequencies fails for exactly 1 of the 100 signals for $K = 750$ and the detection of all frequencies of all 100 signals succeeds for $K = 850$ requiring 1701 samples. This number of samples is very close to the minimum of 1716 samples from above. However, a direct application of MUSIC method (entry $K = 850$) requires distinctly more computational cost than with the sparse FFT Algorithm 3.1. □

Example 4.2 Now we apply Algorithm 3.1 with ESPRIT. From noiseless sampled values, we reconstruct 100 trigonometric polynomials (3.1) of order $M = 1024$ with random frequencies $\omega_j \in (-\frac{S}{2}, \frac{S}{2}] \cap \mathbb{Z}$ and random coefficients c_j on the unit circle. The results for the frequency grid parameter $S := 2^{22}$ are shown in Table 4.2. The minimal number of samples is about 6 times higher compared to the results in Table 4.1.
In general, we observe that the maximal number of used iterations decreases for increasing initial FFT length $P \in \{64, 128, 256, 512\}$ as well as for increasing values $K \in \{8, 10, 12\}$. In the cases, where all frequencies of all the 100 trigonometric polynomials are correctly detected, the number of required samples first decreases and later increases again for increasing initial FFT length P and fixed values K. The reason for this is that the number of samples per iteration increases for growing FFT length, while the number of used iterations decreases until its minimum one is reached. □

K	P	iterations	samples	ℓ^2-errors
8	64	10	14059	1.3e-09/5.2e-15
8	128	8	19635	1.3e-09/6.0e-15
8	256	3	13192	1.0e-09/5.0e-15
8	512	2	17561	8.3e-10/9.0e-15
10	64	10	17367	2.7e-09/5.7e-15
10	128	6	17535	1.2e-09/5.5e-15
10	256	2	10773	1.2e-09/4.7e-15
10	512	2	21693	6.2e-10/1.6e-14
12	64	10	20675	1.5e-09/6.0e-15
12	128	4	13375	1.0e-09/5.0e-15
12	256	2	12825	1.1e-09/5.1e-15
12	512	2	25825	1.1e-09/1.9e-14

TABLE 4.2
Results for Algorithm 3.1 via ESPRIT for frequency grid parameter $S = 2^{22}$ and sparsity $M = 1024$.

4.2. Noisy case. In this subsection, we test the robustness to noise of Algorithm 3.1. For this, we perturb the samples of the trigonometric polynomials g from (3.1) by additive complex white Gaussian noise with zero mean and standard deviation σ, i.e., we have measurements $\tilde{g}\left(\frac{k}{S} + \frac{s}{P}\right) = g(\frac{k}{S} + \frac{s}{P}) + \eta_{k,s}$, where $\eta_{k,s} \in \mathbb{C}$ are independent and identically distributed complex Gaussian. Then, we may approximately compute the signal-to-noise ratio (SNR) in our case by

$$\text{SNR} \approx \frac{\frac{1}{S}\sum_{k=0}^{S-1} |g(\frac{k}{S})|^2}{\frac{1}{S}\sum_{k=0}^{S-1} |\eta_{k,0}|^2} \approx \frac{\sum_{j=1}^{M} |c_j|^2}{\sigma^2}.$$

Correspondingly, we choose $\sigma := \|(c_j)_{j=1}^{M}\|_2/\sqrt{\text{SNR}}$ for a targeted SNR value. For the numerical computations in MATLAB, we generate the noise by $\eta_{k,s} := \sigma/\sqrt{2}$ `* (randn + 1i*randn)`. Moreover, we choose the parameter $\varepsilon_{\text{spatial}} := 5\sigma$ and this means that $|\eta_{k,s}| \le \varepsilon_{\text{spatial}}$ for more than 99.9998% of the noise values $\eta_{k,s}$.

Example 4.3 As in Example 4.1, we generate 100 trigonometric polynomials (3.1) of order $M = 256$ with random frequencies $\omega_j \in (-\frac{S}{2}, \frac{S}{2}] \cap \mathbb{Z}$ and random coefficients

c_j from the unit circle. We set the frequency grid parameter $S = 2^{16}$, the signal sparsity $M = 256$, the array of relative SVD threshold values `epsilon_svd_list` $:= [10^{-2}, 10^{-3}, \ldots, 10^{-8}]$ and the maximal number of iterations $R := 10$. Here, we set the absolute value of minimal non-zero coefficients $\varepsilon_{\text{fc\_min}} := 10^{-1} = 10^{-1} \cdot \min_j |c_j|$. We use the parameters $P \in \{32, 64, 128\}$ and $K \in \{12, 24\}$ with SNR values 10^8 and 10^{10}. The results of Algorithm 3.1 via ESPRIT are presented in Table 4.3. Additionally, we test the sparsity cut-off parameter $K_2 \in \mathbb{N}$ differently from the Hankel matrix size parameter K, see Algorithm A.1. Here, we use the parameter combinations $(K, K_2) \in \{(12, 6), (12, 12), (24, 12)\}$. In general, we observe that we require more samples for SNR $= 10^8$ than for SNR $= 10^{10}$ and that the relative errors are about one order of magnitude larger for SNR $= 10^8$, since the maximal noise values $\eta_{k,s}$ are larger by about one order of magnitude with high probability. Moreover, the maximal number of samples in the noisy case is higher than in the noiseless case, cf. Table 4.1. For some parameter combinations, exactly one of the 100 signals is not correctly detected and this is indicated by the entry "–" in the column "ℓ^2-errors". All parameters of (3.1) are correctly detected in the case SNR $= 10^{10}$ for the parameter combinations $(K, K_2) \in \{(12, 6), (12, 12), (24, 12)\}$ and $P = 32$. For the considered test parameters, the choices $(K, K_2) \in \{(12, 6), (24, 12)\}$, which yield a higher oversampling within the ESPRIT algorithm, give slightly better results compared to the choice $K = K_2 = 12$. □

SNR	K	K_2	P	iterations	samples	ℓ^2-errors
10^8	12	6	32	7	7800	–
10^8	12	6	64	4	6875	3.9e-04/2.8e-05
10^8	12	6	128	3	9900	2.0e-04/2.2e-05
10^8	12	12	32	6	6325	–
10^8	12	12	64	4	6875	–
10^8	12	12	128	3	9900	2.0e-04/2.2e-05
10^8	24	12	32	4	7497	3.9e-04/2.8e-05
10^8	24	12	64	3	9898	1.4e-04/2.2e-05
10^8	24	12	128	2	12691	2.4e-05/2.3e-05
10^8	3000	3000	1	1	6001	4.2e-05/4.2e-05
10^8	3500	3500	1	1	7001	2.9e-05/2.9e-05
10^{10}	12	6	32	6	6325	1.4e-05/2.8e-06
10^{10}	12	12	32	5	5000	1.7e-05/3.4e-06
10^{10}	24	12	32	3	5390	8.9e-06/3.0e-06
10^{10}	3000	3000	1	1	6001	–
10^{10}	3500	3500	1	1	7001	3.3e-06/3.3e-06

TABLE 4.3

Results for Algorithm 3.1 via ESPRIT for frequency grid parameter $S = 2^{16}$ and sparsity $M = 256$ with noisy data.

Example 4.4 Additionally, we generate 100 random trigonometric polynomials (3.1), where the coefficients c_j are drawn uniformly at random from $[-1, 1] + [-1, 1]\,\mathrm{i}$ with $|c_j| \geq 10^{-2}$. We set the absolute value of minimal non-zero coefficients $\varepsilon_{\text{fc\_min}} := 10^{-3} = 10^{-1} \cdot \min_j |c_j|$. In the case SNR $= 10^8$, we observe in each considered parameter combination that the correct detection of one or two frequencies fails for several of the 100 trigonometric polynomials. The most likely reason is the fact that

the smallest coefficient can be very close to the noise level. If we decrease the noise by one order of magnitude, i.e. SNR $= 10^{10}$, the frequency detection succeeds for all considered parameter combinations.
Furthermore, we generate 100 random trigonometric polynomials (3.1), where the coefficients c_j are drawn uniformly at random from $[-1,1]+[-1,1]\,\mathrm{i}$ with $|c_j| \ge 10^{-1}$. Then we set the absolute value of minimal non-zero coefficients $\varepsilon_{\text{fc\_min}} := 10^{-2} = 10^{-1} \cdot \min_j |c_j|$. This means that the smallest possible coefficient as well as the parameter $\varepsilon_{\text{fc\_min}}$ are by one order of magnitude as before. Now in both of the cases SNR $= 10^8$ and SNR $= 10^{10}$, we observe for each parameter combination $(K, K_2) \in \{(12,6),(12,12),(24,12)\}$ and $P \in \{32, 64, 128\}$ that all frequencies of all trigonometric polynomials are correctly detected. □

5. Reconstruction of multivariate trigonometric polynomials. Let $d,\, M \in \mathbb{N}$ with $d > 1$ be given. We consider the *d-variate exponential sum* of *sparsity* M

$$g(\boldsymbol{x}) := \sum_{j=1}^{M} c_j\, \mathrm{e}^{2\pi\mathrm{i}\boldsymbol{\omega}_j\cdot\boldsymbol{x}} \tag{5.1}$$

for $\boldsymbol{x} := (x_1, \ldots, x_d)^{\mathrm{T}} \in \mathbb{R}^d$ with non-zero coefficients $c_j \in \mathbb{C}$ and distinct frequency vectors $\boldsymbol{\omega}_j := (\omega_{j,1}, \ldots, \omega_{j,d})^{\mathrm{T}} \in (-\frac{S}{2}, \frac{S}{2}]^d \cap \mathbb{Z}^d$, where $S > 0$ is an even integer power. Here the dot in the exponent denotes the usual scalar product in $\mathbb{R}^d$. Note that the function (5.1) is a d-variate trigonometric polynomial of sparsity M which is 1-periodical with respect to each variable. Let $\boldsymbol{\Omega} := \{\boldsymbol{\omega}_1,\, \ldots,\, \boldsymbol{\omega}_M\}$ be the set of the frequency vectors.
We assume that we know *a priori* that $\boldsymbol{\omega}_j$ are contained in a frequency set $\Gamma \subset \mathbb{Z}^d$. Then the cardinality of Γ satisfies $|\Gamma| \ge M$. Examples of possible frequency sets Γ are the cube $\{\boldsymbol{k} \in \mathbb{Z}^d : \|\boldsymbol{k}\|_\infty \le N\}$ and the symmetric hyperbolic cross

$$\left\{ \boldsymbol{k} = (k_s)_{s=1}^d \in \mathbb{Z}^d : \prod_{s=1}^{d} \max\{1,\, |k_s|\} \le N \right\}.$$

For given $\boldsymbol{z} \in \mathbb{Z}^d$ and $T \in \mathbb{N}$, the set

$$\Lambda(\boldsymbol{z}, T) := \left\{ \boldsymbol{x}_k = \frac{k}{T}\, \boldsymbol{z} \bmod \boldsymbol{1};\; k = 0, \ldots, T-1 \right\} \subset \mathbb{T}^d \simeq [0,\,1)^d$$

is called *rank*–1 *lattice*, where $\boldsymbol{1} := (1, \ldots, 1)^{\mathrm{T}}$. Note that $\boldsymbol{x}_k = \boldsymbol{x}_{k+nT}$ for $k = 0, \ldots, T-1$ and $n \in \mathbb{Z}$. For given $\Gamma \subset \mathbb{Z}^d$, there exists a *reconstructing rank*-1 *lattice* $\Lambda(\boldsymbol{z}, T)$ such that the matrix

$$\boldsymbol{A}_{T,|\Gamma|} := \left(\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}\right)_{\boldsymbol{x}\in\Lambda(\boldsymbol{z},T),\, \boldsymbol{k}\in\Gamma}$$

fulfils the condition (see [9] and [10, Section 3.2])

$$\boldsymbol{A}^*_{T,|\Gamma|}\, \boldsymbol{A}_{T,|\Gamma|} = T\, \boldsymbol{I}_{|\Gamma|}\,. \tag{5.2}$$

Then we consider the following spectral estimation problem:
(P3) Assume that $\boldsymbol{\omega}_j \in \Gamma$ $(j = 1, \ldots, M)$ and that $\Lambda(\boldsymbol{z}, T)$ is a reconstructing rank–1 lattice with respect to Γ. Recover the sparsity $M \in \mathbb{N}$, all frequencies $\boldsymbol{\omega}_j \in \Gamma$ as

well as all non-zero coefficients $c_j \in \mathbb{C}$ of the d-variate exponential sum (5.1), if noisy sampled data

$$\tilde{g}_k := g(\boldsymbol{x}_k) + e_k \quad (|e_k| \le \varepsilon_1 \ll \min_j |c_j|)$$

for all $k = 0, \ldots, 2L-2$ are given, where $\boldsymbol{x}_k \in \Lambda(\boldsymbol{z}, T)$, $T \ge L > M$ and $e_k \in \mathbb{C}$ are small error terms.

For simplicity we discuss only noiseless data. Let $\boldsymbol{H}_L := \big(g(\boldsymbol{x}_{k+n})\big)_{k,n=0}^{L-1}$ be the *response matrix* of the given data. Then $\boldsymbol{H}_L$ is a Hankel matrix. Further we introduce the *rectangular Fourier-type matrix*

$$\boldsymbol{F}_{L,M} := \big(\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{\omega}_j \cdot \boldsymbol{x}_k}\big)_{k=0,j=1}^{L-1,M}.$$

From (5.2) it follows in the case $L = T$ that $\boldsymbol{F}_{T,M}^* \, \boldsymbol{F}_{T,M} = T \, \boldsymbol{I}_M$ and hence for all $\boldsymbol{x} \in \mathbb{C}^M$

$$\|\boldsymbol{F}_{T,M}\boldsymbol{x}\|_2^2 = \boldsymbol{x}^* \boldsymbol{F}_{T,M}^* \, \boldsymbol{F}_{T,M} \boldsymbol{x} = T \, \|\boldsymbol{x}\|_2^2 \, .$$

Consequently, all singular values of $\boldsymbol{F}_{T,M}$ are equal to $\sqrt{T}$ and $\mathrm{cond}_2 \boldsymbol{F}_{T,M} = 1$.
The matrix $\boldsymbol{H}_L$ can be represented in the form

$$\boldsymbol{H}_L = \boldsymbol{F}_{L,M} \left(\mathrm{diag}\,(c_j)_{j=1}^M\right) \boldsymbol{F}_{L,M}^{\mathrm{T}} \, . \tag{5.3}$$

The ranges of $\boldsymbol{H}_L$ and $\boldsymbol{F}_{L,M}$ coincide in the noiseless case by (5.3). The range of $\boldsymbol{F}_{L,M}$ is a proper subspace of $\mathbb{C}^L$. This subspace is called *left signal space* $\mathcal{S}_L$. The left signal space $\mathcal{S}_L$ is of dimension M and is generated by the M columns $\boldsymbol{e}_L(\boldsymbol{\omega}_j)$ $(j = 1, \ldots, M)$, where

$$\boldsymbol{e}_L(\boldsymbol{\omega}) := \big(\mathrm{e}^{2\pi\mathrm{i}\,\boldsymbol{\omega} \cdot \boldsymbol{x}_k}\big)_{k=0}^{L-1} \quad (\boldsymbol{\omega} \in \Gamma)\, .$$

Note that $\|\boldsymbol{e}_L(\boldsymbol{\omega})\|_2 = \sqrt{L}$ for each $\boldsymbol{\omega} \in \Gamma$. The *left noise space* $\mathcal{N}_L$ is defined as the orthogonal complement of $\mathcal{S}_L$ in $\mathbb{C}^L$. The dimension of $\mathcal{N}_L$ is equal to $L - M > 0$.
By $\boldsymbol{Q}_L$ we denote the orthogonal projection onto the left noise space $\mathcal{N}_L$. Since $\boldsymbol{e}_L(\boldsymbol{\omega}_j) \in \mathcal{S}_L$ $(j = 1, \ldots, M)$ and $\mathcal{N}_L \perp \mathcal{S}_L$, we obtain that

$$\boldsymbol{Q}_L \, \boldsymbol{e}_L(\boldsymbol{\omega}_j) = \boldsymbol{0} \quad (j = 1, \ldots, M)\, .$$

If $\boldsymbol{\omega} \in \Gamma \setminus \boldsymbol{\Omega}$, then the vectors $\boldsymbol{e}_L(\boldsymbol{\omega}_1), \ldots, \boldsymbol{e}_L(\boldsymbol{\omega}_M), \boldsymbol{e}_L(\boldsymbol{\omega}) \in \mathbb{C}^L$ are linearly independent for $T \ge L > M$. This can be seen as follows: For distinct $\boldsymbol{\omega}, \boldsymbol{\omega}' \in \Gamma$, it follows by [10, Lemma 3.1] that

$$\boldsymbol{\omega} \cdot \boldsymbol{z} \not\equiv \boldsymbol{\omega}' \cdot \boldsymbol{z} \; (\mathrm{mod}\, T).$$

Consequently the vectors

$$\boldsymbol{e}_L(\boldsymbol{\omega}_j) := \big(\mathrm{e}^{2\pi\mathrm{i}\,(\boldsymbol{\omega}_j \cdot \boldsymbol{z})\,\frac{k}{T}}\big)_{k=0}^{L-1} \quad (j = 1, \ldots, M)$$

and $\boldsymbol{e}_L(\boldsymbol{\omega})$ with $\boldsymbol{\omega} \in \Gamma \setminus \Omega$ are linearly independent for $T \ge L > M$, since the square Vandermonde matrix

$$\big(\boldsymbol{e}_L(\boldsymbol{\omega}_1) | \ldots | \boldsymbol{e}_L(\boldsymbol{\omega}_M) | \boldsymbol{e}_L(\boldsymbol{\omega})\big)(1 : M+1,\, 1 : M+1)$$

is invertible for each $L \ge M+1$. Hence

$$\boldsymbol{e}_L(\boldsymbol{\omega}) \notin \mathcal{S}_L = \operatorname{span}\{\boldsymbol{e}_L(\boldsymbol{\omega}_1),\, \ldots,\, \boldsymbol{e}_L(\boldsymbol{\omega}_M)\}\,,$$

i.e. $\boldsymbol{Q}_L\, \boldsymbol{e}_L(\boldsymbol{\omega}) \neq \mathbf{0}$.
Thus the frequency vectors can be determined via the M zeros resp. lowest local minima of the *left noise-space correlation function*

$$N_L(\boldsymbol{\omega}) := \frac{1}{\sqrt{L}}\, \|\boldsymbol{Q}_L\, \boldsymbol{e}_L(\boldsymbol{\omega})\|_2 \quad (\boldsymbol{\omega} \in \Gamma)$$

or via the M peaks of the *left imaging function*

$$J_L(\boldsymbol{\omega}) := \sqrt{L}\, \|\boldsymbol{Q}_L\, \boldsymbol{e}_L(\boldsymbol{\omega})\|_2^{-1} \quad (\boldsymbol{\omega} \in \Gamma)\,.$$

Similar to Section 2, one can determine the left noise-space correlation function resp. the left imaging function on Γ by using SVD of the response matrix $\boldsymbol{H}_L$.
Now, we proceed analogously to Section 3 replacing the parameter S by the rank-1 lattice size T. For a positive integer $P \le T$, we construct the sampling array of (5.1) of size $P \times (2K+1)$ via

$$g_P[s,k] := g\left(\left(\frac{s}{P} + \frac{k}{T}\right)\boldsymbol{z}\right) \quad (s = 0, \ldots, P-1;\, k = 0, \ldots, 2K).$$

As in the univariate case, for each $k = 0, \ldots, 2K$ we form the DFT of length P

$$\hat{g}_P[\ell,k] := \sum_{s=0}^{P-1} g_P[s,k]\, \mathrm{e}^{-2\pi \mathrm{i} s\ell/P} \quad (\ell = 0, \ldots, P-1)\,.$$

For each $\ell = 0, \ldots, P-1$, we obtain that

$$\begin{aligned}\hat{g}_P[\ell,k] &= \sum_{s=0}^{P-1} \sum_{j=1}^{M} c_j\, \mathrm{e}^{2\pi \mathrm{i}(s/P + k/T)\,\boldsymbol{\omega}_j\cdot\boldsymbol{z}}\, \mathrm{e}^{-2\pi \mathrm{i} s\ell/P} \\ &= \sum_{j=1}^{M} c_j\, \mathrm{e}^{2\pi \mathrm{i} k\,\boldsymbol{\omega}_j\cdot\boldsymbol{z}/T} \sum_{s=0}^{P-1} \mathrm{e}^{2\pi \mathrm{i}((\boldsymbol{\omega}_j\cdot\boldsymbol{z})-\ell)s/P}\,.\end{aligned}$$

Introducing the index sets

$$I_P(\ell) := \Big\{ j \in \{1, \ldots, M\} : \boldsymbol{\omega}_j \cdot \boldsymbol{z} \equiv \ell \pmod{P} \Big\} \quad (\ell = 0, \ldots, P-1)\,,$$

it follows that

$$\hat{g}_P[\ell,k] = P \sum_{j \in I_P(\ell)} c_j\, \mathrm{e}^{2\pi \mathrm{i} k\,\boldsymbol{\omega}_j\cdot\boldsymbol{z}/T}\,.$$

Now we apply Algorithm 2.5 resp. 2.6 for each $\ell = 0, \ldots, P-1$ and we compute the one-dimensional frequencies $\omega_{\ell,j} \in (-\frac{T}{2},\, \frac{T}{2}] \cap \mathbb{Z}$. We transform these one-dimensional frequencies $\omega_{\ell,j}$ into their d-dimensional counterparts $\boldsymbol{\omega}_{\ell,j} \in \Gamma$ using the relation $\boldsymbol{\omega}_{\ell,j} \cdot \boldsymbol{z} \equiv \omega_{\ell,j} \pmod{T}$ given by the reconstructing rank-1 lattice $\Lambda(\boldsymbol{z}, T)$. Then we compute coefficients c_j from the samples $\hat{g}_P[\ell,k]$ $(k = 0, \ldots, 2K)$ by solving the corresponding overdetermined Vandermonde system. If we cannot identify all the frequencies, i.e., if $|I_P(\ell)| \geq K$ for some indices ℓ, we consider the new trigonometric polynomial

$$g_1(\boldsymbol{x}) := g(\boldsymbol{x}) - \sum_{j\in I} c_j \,\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{\omega}_j\cdot\boldsymbol{x}} = \sum_{j=1}^{M} c_j \,\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{\omega}_j\cdot\boldsymbol{x}} - \sum_{j\in I} c_j \,\mathrm{e}^{2\pi\mathrm{i}\boldsymbol{\omega}_j\cdot\boldsymbol{x}} \quad (\boldsymbol{x}\in\mathbb{T}^d) \tag{5.4}$$

in an additional iteration, where the index set I contains all index sets $I_P(\ell)$ with $|I_P(\ell)| < K$. In the next iteration, we choose a positive integer $P_1 \leq S$ different from P and repeat the method for the trigonometric polynomial g_1. In doing so, we compute the values

$$\sum_{j\in I} c_j \,\mathrm{e}^{2\pi\mathrm{i}(\frac{s}{P_1}+\frac{k}{T})\,\boldsymbol{\omega}_j\cdot\boldsymbol{z}} = \sum_{j\in I}\left(c_j\,\mathrm{e}^{2\pi\mathrm{i}\frac{\omega_j}{P_1}s}\right)\mathrm{e}^{2\pi\mathrm{i}\frac{\boldsymbol{\omega}_j\cdot\boldsymbol{z}}{T}k}\ (s=0,\ldots,P_1-1;\ k=0,\ldots,2K)$$

of the second sum in (5.4) evaluated at the nodes $\boldsymbol{x} = (\frac{s}{P_1}+\frac{k}{T})\boldsymbol{z}$ with the univariate NFFT [12] in $\mathcal{O}\left(P_1(K\log K + |I|)\right)$ arithmetic operations. We perform additional iterations until all frequencies can be identified, i.e. $|I_{P_1}(\ell)| < K$ for all $\ell = 0,\ldots,P_1-1$.

We modify Algorithm 3.1 from Section 3 as described above and additionally in the following way. Here, we describe the changes in the detailed listing (see Algorithm A.1) of Algorithm 3.1. In step 1, we sample the multivariate trigonometric polynomial at the nodes $\left(\frac{s}{P}+\frac{k}{T}\right)\cdot\boldsymbol{z} \quad (s=0,\ldots,P-1, \quad k=0,\ldots,2K)$. In step 3.3.3, we compute the discrete frequencies $\omega_{\ell,j} := \mathrm{round}\left(\tilde{\omega}_{\ell,j}\,T\right)$ for $j=1,\ldots,M_\ell$. Next, we compute the d-dimensional counterparts $\boldsymbol{\omega}_{\ell,j}$ of the one-dimensional frequencies $\omega_{\ell,j}$ using the relation $\boldsymbol{\omega}_{\ell,j}\cdot\boldsymbol{z} \equiv \omega_{\ell,j} \pmod{T}$. In step 3.3.4, we filter the frequencies $\boldsymbol{\omega}_{\ell,j}$ by removing non-unique ones and by keeping only those with $\boldsymbol{\omega}_{\ell,j}\cdot\boldsymbol{z} \equiv \ell \pmod{P}$. We remark that we have to modify step 3.3.4 and that we have to perform the conversion of one-dimensional frequencies $\omega_{\ell,j}$ to their d-dimensional counterparts $\boldsymbol{\omega}_{\ell,j}$ before the filtering, since the conditions $\boldsymbol{\omega}_{\ell,j}\cdot\boldsymbol{z} \equiv \ell \pmod{P}$ and $\omega_{\ell,j} \equiv \ell \pmod{P}$ are not equivalent in general if P is not a divisor of T.

In the following example, we present some numerical results for the modified Algorithm 3.1 for dimension $d = 6$.

Example 5.1 We choose the index set Γ of possible frequency vectors as the 6-dimensional hyperbolic cross $\Gamma := \{\boldsymbol{k}\in\mathbb{Z}^6 : \prod_{s=1}^{6}\max\{1,\,|k_s|\} \leq 16\}$ of cardinality 169209. Further we use the reconstructing rank-1 lattice $\Lambda(\boldsymbol{z},T)$ with generating vector $\boldsymbol{z} = (1, 33, 579, 3628, 21944, 169230)^\top$ and rank-1 lattice size $T = 1105193$, see [11, Table 6.2]. We generate 100 random trigonometric polynomials (5.1) with sparsity $M = 256$, where the frequency vectors $\boldsymbol{\omega}_j$ $(j = 1,\ldots,M)$ are chosen uniformly at random from Γ (without repetition) and the corresponding coefficients c_j are randomly chosen on the unit circle. We set the array of relative SVD threshold values `epsilon_svd_list` := $[10^{-2}, 10^{-3}, \ldots, 10^{-8}]$, the absolute value of minimal non-zero coefficients $\varepsilon_{\mathrm{fc\_min}} := 10^{-1}$, and the maximal number of iterations $R := 10$. In the noiseless case, we set the parameter $\varepsilon_{\mathrm{spatial}} := 10^{-8}$, and in the noisy case as described in Subsection 4.2. The results of the modified Algorithm 3.1 via ESPRIT are presented in Table 5.1.

The columns of Table 5.1 have the same meaning as in Section 4. For the noiseless case, i.e. "SNR $=\infty$", we observe the same behavior as in the one-dimensional case in Subsection 4.1. The detection of all frequency vectors of all 100 trigonometric polynomials (5.1) succeeds for $K = K_2 \in \{8, 10, 12\}$ and $P \in \{8, 16, 32, 64, 128\}$. For the noisy case, the results are worse than in Table 4.3 of the one-dimensional

SNR	K	K_2	P	iterations	samples	ℓ^2-errors
∞	8	8	8	10	3893	5.1e-08/6.2e-11
∞	8	8	16	10	5151	5.1e-08/6.4e-11
∞	8	8	32	10	8687	6.6e-09/9.3e-11
∞	8	8	64	3	3434	2.6e-09/3.2e-10
∞	8	8	128	2	4403	2.8e-09/1.8e-09
∞	10	10	8	9	3948	5.0e-09/6.5e-11
∞	10	10	16	10	6363	1.0e-08/6.8e-11
∞	10	10	32	10	10731	3.9e-09/1.1e-10
∞	10	10	64	10	17367	6.2e-09/5.5e-10
∞	10	10	128	2	5439	2.3e-09/1.7e-09
∞	12	12	8	10	5725	8.6e-09/6.8e-11
∞	12	12	16	10	7575	1.1e-08/7.1e-11
∞	12	12	32	3	2750	2.9e-09/1.3e-10
∞	12	12	64	4	6875	3.0e-09/8.5e-10
∞	12	12	128	2	6475	2.2e-09/1.9e-09
10^8	24	12	32	10	25039	8.6e-05/1.4e-05
10^8	24	12	64	7	25774	5.1e-05/1.5e-05
10^{10}	12	12	64	4	6875	2.2e-05/2.5e-06
10^{10}	24	12	32	5	9800	6.2e-05/2.4e-06
10^{10}	24	12	64	3	9898	8.4e-06/2.2e-06

TABLE 5.1

Results of the modified Algorithm 3.1 via ESPRIT with sparsity $M = 256$*, frequency vectors within* 6*-dimensional hyperbolic cross index set* $\Gamma = \{\boldsymbol{k} \in \mathbb{Z}^6 : \prod_{s=1}^{6} \max\{1, |k_s|\} \leq 16\}$*, and resulting rank-1 lattice* $\Lambda(\boldsymbol{z}, T)$ *with* $\boldsymbol{z} = (1, 33, 579, 3628, 21944, 169230)^\top$ *and* $T = 1105193$.

case. The reason for this is that we have a bijective mapping between 6-dimensional $\boldsymbol{\omega}_j$ and one-dimensional frequencies ω_j by means of the reconstructing rank-1 lattice, $\boldsymbol{\omega}_j \cdot \boldsymbol{z} \equiv \omega_j (\bmod T)$, and the rank-1 lattice size T influences how close two distinct one-dimensional fractional frequencies ω_j'/T and ω_j''/T may get in the ESPRIT algorithm. For the lower SNR $= 10^8$ value, the detection of all frequencies of all 100 signals succeeds for the parameters $K = 24$, $K_2 = 12$ and $P \in \{32, 64\}$. For the higher SNR $= 10^{10}$ value, the detection succeeds for all depicted parameter combinations. □

Acknowledgment. The first and last named author gratefully acknowledge the support by the German Research Foundation within the project PO 711/10–2. These results were first presented during the Dagstuhl Seminar 15251 on "Sparse modeling and multi-exponential analysis" (June 14 – 19, 2015).

REFERENCES

[1] M. Ben-Or and P. Tiwari. A deterministic algorithm for sparse multivariate polynomial interpolation. In *Twentieth Annual ACM Symp. Theory Comput.*, pages 301 – 309. ACM Press New York, 1988.
[2] A. Christlieb, D. Lawlor, and Wang. A multiscale sub-linear time Fourier algorithm for noisy data. *Appl. Comput. Harmon. Anal.*, 2015. accepted.
[3] A. C. Fannjiang. The MUSIC algorithm for sparse objects: a compressed sensing analysis. *Inverse Problems*, 27(3):035013, 32, 2011.

[4] S. Foucart and H. Rauhut. *A mathematical introduction to compressive sensing*. Applied and Numerical Harmonic Analysis. Birkhäuser/Springer, New York, 2013.
[5] A. Gilbert, P. Indyk, M. Iwen, and L. Schmidt. Recent developments in the sparse Fourier transform: A compressed Fourier transform for big data. *IEEE Signal Proc. Mag.*, 31(5):91–100, Sept 2014.
[6] K. Gröchenig, B. M. Pötscher, and H. Rauhut. Learning trigonometric polynomials from random samples and exponential inequalities for eigenvalues of random matrices. *arXiv:math/0701781v2*, 2010.
[7] H. Hassanieh, P. Indyk, D. Katabi, and E. Price. Nearly optimal sparse Fourier transform. In *Proceedings of the Forty-fourth Annual ACM Symposium on Theory of Computing*, pages 563 – 578. ACM, 2012.
[8] M. A. Iwen. Combinatorial sublinear-time Fourier algorithms. *Found. Comput. Math.*, 10:303 – 338, 2010.
[9] L. Kämmerer. Reconstructing multivariate trigonometric polynomials from samples along rank-1 lattices. In G. E. Fasshauer and L. L. Schumaker, editors, *Approximation Theory XIV: San Antonio 2013*, pages 255 – 271. Springer International Publishing, 2014.
[10] L. Kämmerer. *High Dimensional Fast Fourier Transform Based on Rank-1 Lattice Sampling*. Dissertation. Universitätsverlag Chemnitz, 2014.
[11] L. Kämmerer, D. Potts, and T. Volkmer. Approximation of multivariate periodic functions by trigonometric polynomials based on rank-1 lattice sampling. *J. Complexity*, 31:543 – 576, 2015.
[12] 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.
[13] A. Kirsch. The MUSIC algorithm and the factorization method in inverse scattering theory for inhomogeneous media. *Inverse Problems*, 18(4):1025 – 1040, 2002.
[14] S. Kunis and H. Rauhut. Random sampling of sparse trigonometric polynomials II, Orthogonal matching pursuit versus basis pursuit. *Found. Comput. Math.*, 8:737 – 763, 2008.
[15] D. Lawlor, Y. Wang, and A. Christlieb. Adaptive sub-linear time Fourier algorithms. *Adv. Adapt. Data Anal.*, 5(1):1350003, 25, 2013.
[16] W. Liao and A. Fannjiang. MUSIC for single-snapshot spectral estimation: Stability and super-resolution. *Appl. Comput. Harmon. Anal.*, in press, 2014.
[17] D. G. Manolakis, V. K. Ingle, and S. M. Kogon. *Statistical and Adaptive Signal Processing*. McGraw-Hill, Boston, 2005.
[18] T. Peter, D. Potts, and M. Tasche. Nonlinear approximation by sums of exponentials and translates. *SIAM J. Sci. Comput.*, 33:1920 – 1947, 2011.
[19] G. Plonka and M. Tasche. Prony methods for recovery of structured functions. *GAMM–Mitt.* 37:239 – 258, 2014.
[20] D. Potts and M. Tasche. Parameter estimation for nonincreasing exponential sums by Prony-like methods. *Linear Algebra Appl.*, 439:1024 – 1039, 2013.
[21] D. Potts and M. Tasche. Parameter estimation for multivariate exponential sums. *Electron. Trans. Numer. Anal.*, 40:204 – 224, 2013.
[22] D. Potts and M. Tasche. Fast ESPRIT algorithms based on partial singular value decompositions. *Appl. Numer. Math.*, 88:31 – 45, 2015.
[23] D. Potts and M. Tasche. Error estimates for the ESPRIT algorithm. *Submitted*.
[24] H. Rauhut. Random sampling of sparse trigonometric polynomials. *Appl. Comput. Harmon. Anal.*, 22:16 – 42, 2007.
[25] R. Roy and T. Kailath. ESPRIT – estimation of signal parameters via rotational invariance techniques. *IEEE Trans. Acoust. Speech Signal Process.*, 37:984 – 994, 1989.
[26] R.O. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Trans. Antennas and Propagation*, 34:276 – 280, 1986.

Appendix A. Detailed sparse FFT algorithm.

Algorithm A.1 (Detailed listing of Algorithm 3.1 with extended parameter list)
Input: $S \in 2\mathbb{N}$ frequency grid parameter, $K \in \mathbb{N}$ Hankel matrix size parameter, $K_2 \in \mathbb{N}$ sparsity cut-off parameter (default value K), $P \in \mathbb{N}$ initial FFT length, g 1-periodic sparse trigonometric polynomial of unknown sparsity $M \in \mathbb{N}$ with frequencies in $(-\frac{S}{2}, \frac{S}{2}] \cap \mathbb{Z}$, `epsilon_svd_list` array of relative SVD threshold values $0 < \varepsilon_{\mathrm{SVD}} < 1$ in descending order, $\varepsilon_{\mathrm{spatial}} > 0$ estimate for maximal noise value, $\varepsilon_{\mathrm{fc\_min}} > 0$ lower bound of absolute values of non-zero coefficients, $R \in \mathbb{N}$ maximal number of iterations.

Create empty index set array Ω and coefficient array C.
for iteration $r := 1, \ldots, R$
1. Construct the discrete array of samples of g of length $P \times (2K+1)$ via
$g_P[s,k] := g\left(\frac{s}{P} + \frac{k}{S}\right) - \sum_{j'=1}^{|\Omega|} C[j']\, \mathrm{e}^{2\pi\mathrm{i}\Omega[j']\left(\frac{s}{P}+\frac{k}{S}\right)} \quad (s = 0, \ldots, P-1;\ k = 0, \ldots, 2K)$.
2. Compute for each $k = 0, \ldots, 2K$ an FFT of length P and obtain array $\hat{g}_P$ of length $P \times (2K+1)$, $\hat{g}_P[\ell, k] := \sum_{s=0}^{P-1} g_P[s,k]\, \mathrm{e}^{-2\pi\mathrm{i}s\ell/P}$ for $\ell = 0, \ldots, P-1$, if $P > 1$. Otherwise if $P = 1$, then set $\hat{g}_P[\ell, k] := g_P[\ell, k]$ for $\ell = 0, \ldots, P-1$.
3. **for** $\ell := 0, \ldots, P-1$
3.1. If $\|\hat{g}_P[\ell, 0:2K]\|_\infty / P < \varepsilon_{\mathrm{spatial}}$, then go to 3. and continue with next ℓ.
3.2. Set variable `found_svd` $:= 0$.
3.3. **for** $\varepsilon_{\mathrm{SVD}}$ in `epsilon_svd_list`
3.3.1. Apply Algorithm 2.5 resp. 2.6 with $L := K$, $N := 2K+1$ and $\varepsilon := \varepsilon_{\mathrm{SVD}}$ on the vector $\hat{g}_P[\ell, 0:2K]$ and obtain (local) sparsity M_ℓ, frequencies $\tilde{\omega}_{\ell,j} \in (-\frac{1}{2}, \frac{1}{2}]$ for $j = 1, \ldots, M_\ell$.
3.3.2. If $M_\ell \geq K_2$ then go to 3.3. and continue with next (smaller) $\varepsilon_{\mathrm{SVD}}$.
3.3.3. Compute discrete frequencies $\omega_{\ell,j} := \mathrm{round}\left(\tilde{\omega}_{\ell,j}\, S\right)$ for $j = 1, \ldots, M_\ell$.
3.3.4. Filter frequencies $\omega_{\ell,j}$ by removing non-unique ones and by keeping only those where $\omega_{\ell,j} \equiv \ell \pmod{P}$. Set M_ℓ to number of resulting frequencies $\omega_{\ell,j}$.
3.3.5. Compute (local) Fourier coefficients $c_{\ell,j}$ as least squares solution from the overdetermined Vandermonde system $(\hat{g}_P[\ell, 0:2K])^\top \approx (\mathrm{e}^{2\pi\mathrm{i}k\omega_{\ell,j}/S})_{k=0;\, j=1}^{2K;\ M_\ell}\ (P \cdot c_{\ell,j})_{j=1}^{M_\ell}$.
3.3.6. If residual $\|(\mathrm{e}^{2\pi\mathrm{i}k\omega_{\ell,j}/S})_{k=0,j=1}^{2K,M_\ell}\ (c_{\ell,j})_{j=1}^{M_\ell} - (\hat{g}_P[\ell, 0:2K])^\top / P\|_\infty > 10 \cdot \varepsilon_{\mathrm{spatial}}$, then go to 3.3. and continue with next (smaller) $\varepsilon_{\mathrm{SVD}}$.
Otherwise, set variable `found_svd` $:= 1$, leave **for** $\varepsilon_{\mathrm{SVD}}$ loop and go to 3.5.
3.3. **end for** $\varepsilon_{\mathrm{SVD}}$
3.4. If `found_svd` $\neq 1$, then go to 3. and continue with next ℓ.
3.5. If a frequency has already been found, i.e., $\omega_{\ell,j} = \Omega[j']$ for any $j = 1, \ldots, M_\ell$, then update the corresponding coefficient $C[j']$ by computing $C[j'] := C[j'] + c_{\ell,j}$.
3.6. Append new frequencies of $\omega_{\ell,j}$, $j = 1, \ldots, M_\ell$, to array Ω and append corresponding coefficients to array C.
3. **end for** ℓ
4. Remove small coefficients $\left|C[j']\right| < \varepsilon_{\mathrm{fc\_min}}$ from array C and remove corresponding frequencies from array Ω for any j'.
5. If the residual $\max\limits_{\substack{s=0,\ldots,P-1\\ k=0,\ldots,2K}} \left|\sum_{j'=1}^{|\Omega|} C[j']\, \mathrm{e}^{2\pi\mathrm{i}\Omega[j']\left(\frac{s}{P}+\frac{k}{S}\right)} - g\left(\frac{s}{P} + \frac{k}{S}\right)\right| < 10 \cdot \varepsilon_{\mathrm{spatial}}$,
then set $R_{\mathrm{used}} := r$ and exit r-loop.
Otherwise, determine next prime number larger than current FFT length P and use this larger prime as P in the next iteration.
end for iteration r

Output: Detected sparsity $M := |\Omega| \in \mathbb{N}$, array $\Omega \subset (-\frac{S}{2}, \frac{S}{2}] \cap \mathbb{Z}$ of detected frequencies, array $C \in \mathbb{C}^M$ of corresponding coefficients.