PARAMETER ESTIMATION FOR MULTIVARIATE EXPONENTIAL SUMS

DANIEL POTTS* AND MANFRED TASCHE‡

Abstract. The recovery of signal parameters from noisy sampled data is an essential problem in digital signal processing. In this paper, we discuss the numerical solution of the following parameter estimation problem. Let h_0 be a multivariate exponential sum, i.e., h_0 is a finite linear combination of complex exponentials with distinct frequency vectors. Determine all parameters of h_0, i.e., all frequency vectors, all coefficients, and the number of exponentials, if finitely many sampled data of h_0 are given. Using Ingham–type inequalities, the Riesz stability of finitely many multivariate exponentials with well–separated frequency vectors is discussed in continuous as well as discrete norms. Further we show that a rectangular Fourier–type matrix has a bounded condition number, if the frequency vectors are well–separated and if the number of samples is sufficiently large. Then we reconstruct the parameters of an exponential sum h_0 by a novel algorithm, the so–called sparse approximate Prony method (SAPM), where we use only some data sampled along few straight lines. The first part of SAPM estimates the frequency vectors by using the approximate Prony method in the univariate case. The second part of SAPM computes all coefficients by solving an overdetermined linear Vandermonde–type system. Numerical experiments show the performance of our method.

Key words and phrases: Parameter estimation, multivariate exponential sum, multivariate exponential fitting problem, harmonic retrieval, sparse approximate Prony method, sparse approximate representation of signals.

AMS *Subject Classifications*: 65D10, 65T40, 41A45, 41A63, 65F20, 94A12.

1. Introduction. Let the dimension $d \in \mathbb{N}$ and a positive integer $M \in \mathbb{N}\backslash\{1\}$ be given. We consider a *d-variate exponential sum of order M* that is a linear combination

$$h_0(\boldsymbol{x}) := \sum_{j=1}^{M} c_j \, \mathrm{e}^{\mathrm{i}\boldsymbol{f}_j \cdot \boldsymbol{x}} \quad (\boldsymbol{x} = (x_l)_{l=1}^d \in \mathbb{R}^d) \tag{1.1}$$

of M complex exponentials with complex coefficients $c_j \neq 0$ and distinct frequency vectors $\boldsymbol{f}_j = (f_{j,l})_{l=1}^d \in \mathbb{T}^d \cong [-\pi,\, \pi)^d$. Assume that $|c_j| > \varepsilon_0$ $(j = 1, \ldots, M)$ for a convenient bound $0 < \varepsilon_0 \ll 1$. Here the torus $\mathbb{T}$ is identified with the interval $[-\pi,\, \pi)$. Further the dots in the exponents of (1.1) denote the usual scalar product in $\mathbb{R}^d$.
If h_0 is real–valued, then (1.1) can be represented as a linear combination of ridge functions

$$h_0(\boldsymbol{x}) = \sum_{j=1}^{M} |c_j| \, \cos\left(\boldsymbol{f}_j \cdot \boldsymbol{x} + \varphi_j\right)$$

with $c_j = |c_j| \, \mathrm{e}^{\mathrm{i}\varphi_j}$. Assume that the frequency vectors $\boldsymbol{f}_j \in \mathbb{T}^d$ $(j = 1, \ldots, M)$ fulfill the *gap condition on* $\mathbb{T}^d$

$$\mathrm{dist}(\boldsymbol{f}_j, \boldsymbol{f}_l) := \min\{\|(\boldsymbol{f}_j + 2\pi\boldsymbol{k}) - \boldsymbol{f}_l\|_\infty : \boldsymbol{k} \in \mathbb{Z}^d\} \geq q > 0 \tag{1.2}$$

for all $j, l = 1, \ldots, M$ with $j \neq l$. Let $N \in \mathbb{N}$ with $N \geq 2M + 1$ be given. In the following $\mathbb{G}$ is either the full grid $\mathbb{Z}_N^d := [-N, N]^d \cap \mathbb{Z}^d$ or a union of $2N + 1$ grid points $\boldsymbol{n} \in \mathbb{Z}^d$ lying on few straight lines. If $\mathbb{G}$ is chosen such that $|\mathbb{G}| \ll (2N + 1)^d$

*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 of Mathematics, D–18051 Rostock, Germany

for $d \geq 2$, then $\mathbb{G}$ is called a *sparse sampling grid.*
Suppose that perturbed sampled data

$$h(\boldsymbol{n}) := h_0(\boldsymbol{n}) + e(\boldsymbol{n}), \quad |e(\boldsymbol{n})| \leq \varepsilon$$

of (1.1) for all $\boldsymbol{n} \in \mathbb{G}$ are given, where the error terms $e(\boldsymbol{n}) \in \mathbb{C}$ are bounded by certain accuracy $\varepsilon > 0$. Then we consider the following *parameter estimation problem* for the d–variate exponential sum (1.1): Recover the distinct frequency vectors $\boldsymbol{f}_j \in [-\pi, \pi)^d$ and the complex coefficients c_j so that

$$|h(\boldsymbol{n}) - \sum_{j=1}^{M} c_j \, \mathrm{e}^{\mathrm{i}\boldsymbol{f}_j \cdot \boldsymbol{n}}| \leq \varepsilon \quad (\boldsymbol{n} \in \mathbb{G}) \tag{1.3}$$

for very small accuracy $\varepsilon > 0$ and for minimal order M. In other words, we are interested in sparse approximate representations of the given noisy data $h(\boldsymbol{n}) \in \mathbb{C}$ ($\boldsymbol{n} \in \mathbb{G}$) by sampled data of the exponential sum (1.1), where the condition (1.3) is fulfilled.
The approximation of data by finite linear combinations of complex exponentials has a long history, see [19, 20]. There exists a variety of applications, such as fitting nuclear magnetic resonance spectroscopic data [18] or the annihilating filter method [31, 6, 30]. Recently, the reconstruction method of [3] was generalized to bivariate exponential sums in [1]. In contrast to [1], we introduce a sparse approximate Prony method, where we use only some data on a sparse sampling grid $\mathbb{G}$. Further we remark the relation to a reconstruction method for sparse multivariate trigonometric polynomials, see Remark 6.3 and [14, 12, 32].

In this paper, we extend the approximate Prony method (see [23]) to multivariate exponential sums. Our approach can be described as follows:
(i) Solving a few reconstruction problems of univariate exponential sums, we determine a finite set of feasible frequency vectors $\boldsymbol{f}'_k$ ($k = 1, \ldots, M'$). For each reconstruction we use only data sampled along a straight line. As parameter estimation we use the univariate approximate Prony method which can be replaced by another Prony–like method [24], such as ESPRIT (Estimation of Signal Parameters via Rotational Invariance Techniques) [26, 27] or matrix pencil methods [10, 29].
(ii) Then we test, if a feasible frequency vector $\boldsymbol{f}'_k$ ($k = 1, \ldots, M'$) is an actual frequency vector of the exponential sum (1.1) too. Replacing the condition (1.3) by the overdetermined linear system

$$\sum_{k=1}^{M'} c'_k \, \mathrm{e}^{\mathrm{i}\boldsymbol{f}'_k \cdot \boldsymbol{n}} = h(\boldsymbol{n}) \quad (\boldsymbol{n} \in \mathbb{G}), \tag{1.4}$$

we compute the least squares solution $(c'_k)_{k=1}^{M'}$. Then we say that $\boldsymbol{f}'_k$ is an actual frequency vector of (1.1), if $|c'_k| > \varepsilon_0$. Otherwise, $\boldsymbol{f}'_k$ is interpreted as frequency vector of noise and is canceled. Let $\tilde{\boldsymbol{f}}_j$ ($j = 1, \ldots, M$) be all the actual frequency vectors.
(iii) In a final correction step, we solve the linear system

$$\sum_{j=1}^{M} \tilde{c}_j \, \mathrm{e}^{\mathrm{i}\tilde{\boldsymbol{f}}_j \cdot \boldsymbol{n}} = h(\boldsymbol{n}) \quad (\boldsymbol{n} \in \mathbb{G}).$$

As explained above, our reconstruction method uses the least squares solution of the linear system (1.4) with the rectangular coefficient matrix

$$\left(\mathrm{e}^{\mathrm{i}\tilde{\boldsymbol{f}}_j\cdot\boldsymbol{n}}\right)_{\boldsymbol{n}\in\mathbb{G},\, j=1,\ldots,M} \quad (|\mathbb{G}| > M)\,.$$

If this matrix has full rank M and if its condition number is moderately sized, then one can efficiently compute the least squares solution of (1.4), which is sensitive to permutations of the coefficient matrix and the sampled data (see [7, pp. 239 – 244]). In the special case $\mathbb{G} = \mathbb{Z}_N^d$, we can show that this matrix is uniformly bounded, if $N > \frac{\sqrt{d}\pi}{q}$. Then we use $(2N+1)^d$ sampled data for the reconstruction of M frequency vectors $\boldsymbol{f}_j$ and M complex coefficients c_j of (1.1).
But our aim is an efficient parameter estimation of (1.1) by a relatively low number of given sampled data $h(\boldsymbol{n})$ ($\boldsymbol{n} \in \mathbb{G}$) on a sparse sampling grid $\mathbb{G}$. The corresponding approach is called *sparse approximate Prony method* (SAPM). Numerical experiments for d–variate exponential sums with $d \in \{2,\, 3,\, 4\}$ show the performance of our parameter reconstruction.

This paper is divided into two parts. The first part consists of Sections 2 and 3, where we discuss the Riesz stability of finitely many multivariate exponentials. It is a known fact that an exponential sum (1.1) with well–separated frequency vectors can be well reconstructed. Otherwise, one also knows that the parameter estimation of an exponential sum with clustered frequency vectors is very difficult. What is the basic cause of these effects? In Section 2, we investigate the Riesz stability of multivariate exponentials with respect to the continuous norms of $L^2([-N,\, N]^d)$ and $C([-N,\, N]^d)$, respectively, where we assume that the frequency vectors fulfill the gap condition (2.1) (see Lemma 2.1 and Corollary 2.3). These results are mainly based on Ingham–type inequalities (see [15, pp. 59 – 66 and pp. 153 – 156]). Furthermore we present a result for the converse assertion, i.e., if finitely many multivariate exponentials are Riesz stable, then the corresponding frequency vectors are well–separated (see Lemma 2.2). In Section 3, we extend these stability results to draw conclusions for the discrete norm of $\ell^2(\mathbb{Z}_N^d)$. Further we prove that the condition number of the coefficient matrix of (1.4) is uniformly bounded, if we choose the full sampling grid $\mathbb{G} = \mathbb{Z}_N^d$ and if N is sufficiently large. By the results of Section 3, one can see that well–separated frequency vectors are essential for a successful parameter estimation of (1.1). Up to now, a corresponding result for a sparse sampling grid $\mathbb{G}$ is unknown.
The second part of this paper consists of Sections 4 – 7, where we present a novel efficient parameter recovery algorithm of (1.1) for a *sparse* sampling grid. In Section 4 we sketch the approximate Prony method in the univariate setting. Then we extend this method to bivariate exponential sums in Section 5. Here we suggest the new SAPM. The main idea is to project the bivariate reconstruction problem to several univariate problems and combine finally the results of the univariate reconstructions. We use only few data sampled along some straight lines in order to reconstruct a bivariate exponential sum. In Section 6, we extend this reconstruction method to d–variate exponential sums for moderately sized dimensions $d \geq 3$. Finally, various numerical examples are presented in Section 7.

2. Stability of exponentials. As known, the main difficulty is the reconstruction of frequency vectors with small separation distance $q > 0$ (see (1.2)). Therefore first we discuss the stability properties of the finitely many d–variate exponentials in dependence of q. We start with a generalization of the known Ingham inequalities (see [11]):

Lemma 2.1. *(see [15, pp. 153 – 156]). Let $d \in \mathbb{N}$, $M \in \mathbb{N} \setminus \{1\}$ and $N > 0$ be given. If the frequency vectors $\boldsymbol{f}_j \in \mathbb{R}^d$ $(j = 1, \ldots, M)$ fulfill the gap condition on $\mathbb{R}^d$*

$$\|\boldsymbol{f}_j - \boldsymbol{f}_l\|_\infty \ge q > \frac{\sqrt{d}\,\pi}{N} \quad (j,\, l = 1, \ldots, M;\, j \ne l), \tag{2.1}$$

then the exponentials $\mathrm{e}^{\mathrm{i}\boldsymbol{f}_j \cdot (\cdot)}$ $(j = 1, \ldots, M)$ are Riesz stable in $L^2([-N, N]^d)$, i.e., for all complex vectors $\boldsymbol{c} = (c_j)_{j=1}^M$

$$\gamma_1 \,\|\boldsymbol{c}\|_2^2 \le \|\sum_{j=1}^M c_j\, \mathrm{e}^{\mathrm{i}\boldsymbol{f}_j \cdot (\cdot)}\|_2^2 \le \gamma_2 \,\|\boldsymbol{c}\|_2^2 \tag{2.2}$$

with some positive constants γ_1, γ_2, independent of the particular choice of the coefficients c_j. Here $\|\boldsymbol{c}\|_2$ denotes the Euclidean norm of $\boldsymbol{c} \in \mathbb{C}^M$ and

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

For a proof see [15, pp. 153 – 156]. Note that for $d = 1$, we obtain exactly the classical Ingham inequalities (see [11]) with the positive constants

$$\gamma_1 = \frac{2}{\pi}\,\big(1 - \frac{\pi^2}{N^2 q^2}\big)\,, \quad \gamma_2 = \frac{4\sqrt{2}}{\pi}\,\big(1 + \frac{\pi^2}{4N^2 q^2}\big)\,.$$

In the case $d \ge 2$, the Lemma 2.1 provides only the existence of positive constants γ_1, γ_2 without corresponding explicit expressions.
Obviously, the exponentials

$$\mathrm{e}^{\mathrm{i}\boldsymbol{f}_j \cdot (\cdot)} \quad (j = 1, \ldots, M) \tag{2.3}$$

with distinct frequency vectors $\boldsymbol{f}_j \in \mathbb{R}^d$ $(j = 1, \ldots, M)$ are linearly independent and Riesz stable. Now we show that from the first inequality (2.2) it follows that the frequency vectors $\boldsymbol{f}_j$ are well–separated. The following lemma generalizes a former result [17] for univariate exponentials.

Lemma 2.2. *Let $d \in \mathbb{N}$, $M \in \mathbb{N} \setminus \{1\}$ and $N > 0$. Further let $\boldsymbol{f}_j \in \mathbb{R}^d$ $(j = 1, \ldots, M)$ be given. If there exists a constant $\gamma_1 > 0$ such that*

$$\gamma_1 \,\|\boldsymbol{c}\|_2^2 \le \|\sum_{j=1}^M c_j\, \mathrm{e}^{\mathrm{i}\boldsymbol{f}_j \cdot (\cdot)}\|_2^2$$

for all complex vectors $\boldsymbol{c} = (c_j)_{j=1}^M$, then the frequency vectors $\boldsymbol{f}_j$ are well–separated by

$$\|\boldsymbol{f}_j - \boldsymbol{f}_l\|_\infty \ge \frac{\sqrt{2\gamma_1}}{d\,N}$$

for all $j,\, l = 1, \ldots, M$ $(j \ne l)$. Moreover the exponentials (2.3) are Riesz stable in $L^2([-N, N]^d)$.

Proof. 1. In the following proof we use similar arguments as in [5, Theorem 7.6.5]. We choose $c_j = -c_l = 1$ for $j \ne l$. All the other coefficients are equal to 0. Then by the assumption, we obtain

$$\begin{aligned}
2\,\gamma_1 &\le \|\mathrm{e}^{\mathrm{i}\boldsymbol{f}_j\cdot(\cdot)} - \mathrm{e}^{\mathrm{i}\boldsymbol{f}_l\cdot(\cdot)}\|_2^2 \\
&= \frac{1}{(2\,N)^d}\int_{[-N,N]^d} |1-\mathrm{e}^{\mathrm{i}(\boldsymbol{f}_l-\boldsymbol{f}_j)\cdot\boldsymbol{x}}|^2\,\mathrm{d}\boldsymbol{x} \\
&= \frac{1}{(2\,N)^d}\int_{[-N,N]^d} 4\,\sin^2\big((\boldsymbol{f}_l-\boldsymbol{f}_j)\cdot\boldsymbol{x}/2\big)\,\mathrm{d}\boldsymbol{x} \\
&\le \frac{1}{(2\,N)^d}\int_{[-N,N]^d} \big|(\boldsymbol{f}_l-\boldsymbol{f}_j)\cdot\boldsymbol{x}\big|^2\,\mathrm{d}\boldsymbol{x} \\
&\le \frac{1}{(2\,N)^d}\int_{[-N,N]^d} \|\boldsymbol{f}_l-\boldsymbol{f}_j\|_1^2\,N^2\,\mathrm{d}\boldsymbol{x}\,, \qquad (2.4)
\end{aligned}$$

where we have used the Hölder estimate

$$|(\boldsymbol{f}_l-\boldsymbol{f}_j)\cdot\boldsymbol{x}| \le \|\boldsymbol{f}_l-\boldsymbol{f}_j\|_1\,\|\boldsymbol{x}\|_\infty \le \|\boldsymbol{f}_l-\boldsymbol{f}_j\|_1\,N$$

for all $\boldsymbol{x}\in[-N,\,N]^d$. Therefore (2.4) shows that

$$d\,\|\boldsymbol{f}_l-\boldsymbol{f}_j\|_\infty \ge \|\boldsymbol{f}_l-\boldsymbol{f}_j\|_1 \ge \frac{\sqrt{2\gamma_1}}{N}$$

for all $j,l=1,\ldots,M$ $(j\ne l)$.
2. We see immediately that M is an upper Riesz bound for the exponentials (2.3) in $L^2([-N,N]^d)$. By the Cauchy–Schwarz inequality we obtain

$$|\sum_{j=1}^{M} c_j\,\mathrm{e}^{\mathrm{i}\boldsymbol{f}_j\cdot\boldsymbol{x}}|^2 \le M\,\|\boldsymbol{c}\|_2^2$$

for all $\boldsymbol{c}=(c_j)_{j=1}^M\in\mathbb{C}^M$ and all $\boldsymbol{x}\in[-N,\,N]^d$ such that

$$\|\sum_{j=1}^{M} c_j\,\mathrm{e}^{\mathrm{i}\boldsymbol{f}_j\cdot(\cdot)}\|_2^2 \le M\,\|\boldsymbol{c}\|_2^2\,.$$

This completes the proof. ■

By the Lemmas 2.1 and 2.2, the Riesz stability of the exponentials (2.3) in $L^2([-N,\,N]^d)$ is equivalent to the fact that the frequency vectors $\boldsymbol{f}_j$ are well–separated. Now we show that in Lemma 2.1 the square norm can be replaced by the uniform norm of $C([-N,\,N]^d)$.

Corollary 2.3. *If the assumptions of Lemma* 2.1 *are fulfilled, then the exponentials* (2.3) *are Riesz stable in* $C([-N,N]^d)$, *i.e., for all complex vectors* $\boldsymbol{c}=(c_j)_{j=1}^M$

$$\sqrt{\frac{\gamma_1}{M}}\,\|\boldsymbol{c}\|_1 \le \|\sum_{j=1}^{M} c_j\,\mathrm{e}^{\mathrm{i}\boldsymbol{f}_j\cdot(\cdot)}\|_\infty \le \|\boldsymbol{c}\|_1$$

with the uniform norm

$$\|f\|_\infty := \max_{\boldsymbol{x}\in[-N,N]^d} |f(\boldsymbol{x})| \quad (f\in C([-N,N]^d)).$$

Proof. Let $h_0 \in C([-N,N]^d)$ be defined by (1.1). Then $\|h_0\|_2 \le \|h_0\|_\infty < \infty$. Using the triangle inequality, we obtain that

$$\|h_0\|_\infty \le \sum_{j=1}^{M} |c_j| \cdot 1 = \|\boldsymbol{c}\|_1 \,.$$

From Lemma 2.1 and $\|\boldsymbol{c}\|_1 \le \sqrt{M}\, \|\boldsymbol{c}\|_2$, it follows that

$$\sqrt{\frac{\gamma_1}{M}}\, \|\boldsymbol{c}\|_1 \le \sqrt{\gamma_1}\, \|\boldsymbol{c}\|_2 \le \|h_0\|_2 \,.$$

This completes the proof. ■

Now we use the uniform norm of $C([-N,\, N]^d)$ and estimate the error $\|h_0 - \tilde{h}\|_\infty$ between the original exponential sum (1.1) and its reconstruction

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

We obtain a small error $\|h_0 - \tilde{h}\|_\infty$ in the case $\sum_{j=1}^{M} |c_j - \tilde{c}_j| \ll 1$ and $\|\boldsymbol{f}_j - \tilde{\boldsymbol{f}}_j\|_\infty \le \delta \ll 1$ $(j = 1, \ldots, M)$.

THEOREM 2.4. *Let $M \in \mathbb{N}\setminus\{1\}$ and $N > 0$ be given. Let $\boldsymbol{c} = (c_j)_{j=1}^M$ and $\tilde{\boldsymbol{c}} = (\tilde{c}_j)_{j=1}^M$ be arbitrary complex vectors. If $\boldsymbol{f}_j,\, \tilde{\boldsymbol{f}}_j \in \mathbb{R}^d$ $(j = 1, \ldots, M)$ fulfill the conditions*

$$\|\boldsymbol{f}_j - \boldsymbol{f}_l\|_\infty \ge q > \frac{3\sqrt{d}\pi}{2N} \quad (j,\, l = 1,\, \ldots,\, M;\, j \ne l),$$
$$\|\tilde{\boldsymbol{f}}_j - \boldsymbol{f}_j\|_\infty \le \delta < \frac{\sqrt{d}\pi}{4N} \quad (j = 1,\, \ldots,\, M),$$

then both (2.3) *and*

$$\mathrm{e}^{\mathrm{i}\tilde{\boldsymbol{f}}_j \cdot (\cdot)} \quad (j = 1, \ldots, M)$$

are Riesz stable in $C([-N,\, N]^d)$. Further

$$\|h_0 - \tilde{h}\|_\infty \le \|\boldsymbol{c} - \tilde{\boldsymbol{c}}\|_1 + d\delta N \, \|\boldsymbol{c}\|_1 \,.$$

Proof. 1. By the gap condition on $\mathbb{R}^d$ we know that

$$\|\boldsymbol{f}_j - \boldsymbol{f}_l\|_\infty \ge q > \frac{3\sqrt{d}\pi}{2N} > \frac{\sqrt{d}\pi}{N} \quad (j,\, l = 1,\, \ldots,\, M;\, j \ne l).$$

Hence the original exponentials (2.3) are Riesz stable in $C([-N,\, N]^d)$ by Corollary 2.3. Using the assumptions, we conclude that

$$\begin{aligned} \|\tilde{\boldsymbol{f}}_j - \tilde{\boldsymbol{f}}_l\|_\infty &\ge \|\boldsymbol{f}_j - \boldsymbol{f}_l\|_\infty - \|\tilde{\boldsymbol{f}}_j - \boldsymbol{f}_j\|_\infty - \|\boldsymbol{f}_l - \tilde{\boldsymbol{f}}_l\|_\infty \\ &\ge q - 2\,\frac{\sqrt{d}\pi}{4N} > \frac{\sqrt{d}\pi}{N} \,. \end{aligned}$$

Thus the reconstructed exponentials

$$\mathrm{e}^{\mathrm{i}\tilde{\boldsymbol{f}}_j \cdot (\cdot)} \quad (j = 1, \ldots, M)$$

are Riesz stable in $C([-N,\,N]^d)$ by Corollary 2.3 too.
2. Now we estimate the normwise error $\|h_0 - \tilde h\|_\infty$ by the triangle inequality. Then we obtain

$$\begin{aligned}\|h_0 - \tilde h\|_\infty &\le \|\sum_{j=1}^M (c_j - \tilde c_j)\,\mathrm{e}^{\mathrm{i}\tilde{\boldsymbol f}_j\cdot(\cdot)}\|_\infty + \|\sum_{j=1}^M c_j\,(\mathrm{e}^{\mathrm{i}\boldsymbol f_j\cdot(\cdot)} - \mathrm{e}^{\mathrm{i}\tilde{\boldsymbol f}_j\cdot(\cdot)})\|_\infty \\ &\le \sum_{j=1}^M |c_j - \tilde c_j| + \sum_{j=1}^M |c_j| \max_{\boldsymbol x\in[-N,N]^d} |\mathrm{e}^{\mathrm{i}\boldsymbol f_j\cdot\boldsymbol x} - \mathrm{e}^{\mathrm{i}\tilde{\boldsymbol f}_j\cdot\boldsymbol x}|\,.\end{aligned}$$

Since for $\boldsymbol d_j := \tilde{\boldsymbol f}_j - \boldsymbol f_j$ $(j = 1,\ldots,M)$ and arbitrary $\boldsymbol x \in [-N,\,N]^d$, we can estimate

$$\begin{aligned}|\mathrm{e}^{\mathrm{i}\boldsymbol f_j\cdot\boldsymbol x} - \mathrm{e}^{\mathrm{i}\tilde{\boldsymbol f}_j\cdot\boldsymbol x}| &= |1 - \mathrm{e}^{\mathrm{i}\boldsymbol d_j\cdot\boldsymbol x}| = \sqrt{2 - 2\,\cos(\boldsymbol d_j\cdot\boldsymbol x)} \\ &= 2\,|\sin\frac{\boldsymbol d_j\cdot\boldsymbol x}{2}| \le |\boldsymbol d_j\cdot\boldsymbol x| \le \|\boldsymbol d_j\|_\infty\,\|\boldsymbol x\|_1 \le d\delta\,N\end{aligned}$$

such that we obtain

$$\|h_0 - \tilde h\|_\infty \le \|\boldsymbol c - \tilde{\boldsymbol c}\|_1 + d\delta N\,\|\boldsymbol c\|_1\,.$$

This completes the proof. ■

3. Stability of exponentials on a grid. In the former section we have studied the Riesz stability of d–variate exponentials (2.3) with respect to continuous norms. Now we investigate the Riesz stability of d–variate exponentials restricted on the full grid $\mathbb Z_N^d$ with respect to the discrete norm of $\ell^2(\mathbb Z_N^d)$. First we will show that a discrete version of Lemma 2.1 is also true for d–variate exponential sums (1.1). If we sample an exponential sum (1.1) on the full grid $\mathbb Z_N^d$, then it is impossible to distinguish between the frequency vectors $\boldsymbol f_j$ and $\boldsymbol f_j + 2\pi\boldsymbol k$ with certain $\boldsymbol k \in \mathbb Z^d$, since by the periodicity of the complex exponential

$$\mathrm{e}^{\mathrm{i}\tilde{\boldsymbol f}_j\cdot\boldsymbol n} = \mathrm{e}^{\mathrm{i}\,(\tilde{\boldsymbol f}_j+2\pi\boldsymbol k)\cdot\boldsymbol n} \quad (\boldsymbol n \in \mathbb Z_N^d)\,.$$

Therefore we assume in the following that $\boldsymbol f_j \in [-\pi,\,\pi)^d$ $(j = 1,\ldots,M)$ and we measure the distance between two distinct frequency vectors $\boldsymbol f_j,\,\boldsymbol f_l \in [-\pi,\,\pi)^d$ $(j,l = 1,\ldots,M;\ j \ne l)$ by

$$\mathrm{dist}(\boldsymbol f_j, \boldsymbol f_l) := \min\{\|(\boldsymbol f_j + 2\pi\boldsymbol k) - \boldsymbol f_l\|_\infty : \boldsymbol k \in \mathbb Z^d\}\,.$$

Then the *separation distance* of the set $\{\boldsymbol f_j \in [-\pi,\,\pi)^d :\ j = 1,\ldots,M\}$ is defined by

$$\min\{\mathrm{dist}(\boldsymbol f_j, \boldsymbol f_l) : j,l = 1,\ldots,M;\ j \ne l\} \in (0,\pi].$$

The separation distance can be interpreted as the smallest gap between two distinct frequency vectors in the d–dimensional torus $\mathbb T^d$.
Since we restrict an exponential sum h_0 on the full sampling grid $\mathbb Z_N^d$, we use the norm

$$\frac{1}{(2N+1)^{d/2}}\Big(\sum_{\boldsymbol k\in\mathbb Z_N^d} |h_0(\boldsymbol k)|^2\Big)^{1/2}$$

in the Hilbert space $\ell^2(\mathbb Z_N^d)$.

LEMMA 3.1. *(see [16]). Let $q \in (0, \pi]$ and $M \in \mathbb{N} \setminus \{1\}$ be given. If the frequency vectors $\boldsymbol{f}_j \in (-\pi + \frac{q}{2},\, \pi - \frac{q}{2})^d$ $(j = 1, \ldots, M)$ satisfy*

$$\|\boldsymbol{f}_j - \boldsymbol{f}_l\|_\infty \geq q > \frac{\sqrt{d}\pi}{N} \quad (j,\, l = 1, \ldots, M;\, j \neq l)\,,$$

then the exponentials (2.3) are Riesz stable in $\ell^2(\mathbb{Z}_N^d)$, i.e., all complex vectors $\boldsymbol{c} = (c_j)_{j=1}^M$ satisfy the following Ingham–type inequalities

$$\gamma_3\, \|\boldsymbol{c}\|_2^2 \leq \frac{1}{(2N+1)^d} \sum_{\boldsymbol{k} \in \mathbb{Z}_N^d} |\sum_{j=1}^M c_j\, \mathrm{e}^{\mathrm{i}\, \boldsymbol{f}_j \cdot \boldsymbol{k}}\, |^2 \leq \gamma_4\, \|\boldsymbol{c}\|_2^2$$

with some positive constants γ_3 and γ_4, independent of the particular choice of $\boldsymbol{c}$.
For a proof see [16]. Note that the Lemma 3.1 delivers only the existence of positive constants γ_3, γ_4 without corresponding explicit expressions.

LEMMA 3.2. *Let $d \in \mathbb{N}$, $M \in \mathbb{N} \setminus \{1\}$ and $N \in \mathbb{N}$ with $N \geq 2M+1$ be given. Further let $\boldsymbol{f}_j \in [-\pi,\, \pi)^d$ $(j = 1, \ldots, M)$. If there exists a constant $\gamma_3 > 0$ such that*

$$\gamma_3\, \|\boldsymbol{c}\|_2^2 \leq \frac{1}{(2N+1)^d} \sum_{\boldsymbol{k} \in \mathbb{Z}_N^d} |\sum_{j=1}^M c_j\, \mathrm{e}^{\mathrm{i} \boldsymbol{f}_j \cdot \boldsymbol{k}}|^2$$

for all complex vectors $\boldsymbol{c} = (c_j)_{j=1}^M$, then the frequency vectors $\boldsymbol{f}_j$ are well–separated by

$$\mathrm{dist}(\boldsymbol{f}_j, \boldsymbol{f}_l) \geq \frac{\sqrt{2\gamma_3}}{d\, N}$$

for all $j, l = 1, \ldots, M$ with $j \neq l$. Moreover the exponentials (2.3) are Riesz stable in $\ell^2(\mathbb{Z}_N^d)$.
The proof follows similar lines as the proof of Lemma 2.2 and is omitted here. By Lemmas 3.1 and 3.2, the Riesz stability of the exponentials (2.3) in $\ell^2(\mathbb{Z}_N^d)$ is equivalent to the condition that the frequency vectors $\boldsymbol{f}_j$ are well–separated.

Introducing the *rectangular Fourier–type matrix*

$$\boldsymbol{F} := (2N+1)^{-d/2} \left(\mathrm{e}^{\mathrm{i}\, \boldsymbol{f}_j \cdot \boldsymbol{k}}\right)_{\boldsymbol{k} \in \mathbb{Z}_N^d,\, j=1,\ldots,M} \in \mathbb{C}^{(2N+1)^d \times M}\,,$$

we improve the result of [22, Theorem 4.3].
COROLLARY 3.3. *Under the assumptions of Lemma 3.1, the rectangular Fourier–type matrix $\boldsymbol{F}$ has a uniformly bounded condition number $\mathrm{cond}_2(\boldsymbol{F})$ for all integers $N > \frac{\sqrt{d}\,\pi}{q}$.*
Proof. By Lemma 3.1, we know that for all $\boldsymbol{c} \in \mathbb{C}^M$

$$\gamma_3\, \boldsymbol{c}^{\mathrm{H}} \boldsymbol{c} \leq \boldsymbol{c}^{\mathrm{H}} \boldsymbol{F}^{\mathrm{H}} \boldsymbol{F}\, \boldsymbol{c} \leq \gamma_4\, \boldsymbol{c}^{\mathrm{H}} \boldsymbol{c} \tag{3.1}$$

with positive constants γ_3, γ_4. Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_M \geq 0$ be the ordered eigenvalues of $\boldsymbol{F}^{\mathrm{H}} \boldsymbol{F} \in \mathbb{C}^{M \times M}$. Using the Rayleigh–Ritz Theorem and (3.1), we obtain that

$$\gamma_3\, \boldsymbol{c}^{\mathrm{H}} \boldsymbol{c} \leq \lambda_M\, \boldsymbol{c}^{\mathrm{H}} \boldsymbol{c} \leq \boldsymbol{c}^{\mathrm{H}} \boldsymbol{F}^{\mathrm{H}} \boldsymbol{F}\, \boldsymbol{c} \leq \lambda_1\, \boldsymbol{c}^{\mathrm{H}} \boldsymbol{c} \leq \gamma_4\, \boldsymbol{c}^{\mathrm{H}} \boldsymbol{c}$$

and hence

$$0 < \gamma_3 \le \lambda_M \le \lambda_1 \le \gamma_4 < \infty\,.$$

Thus $\boldsymbol{F}^{\rm H}\boldsymbol{F}$ is positive definite and

$$\mathrm{cond}_2(\boldsymbol{F}) = \sqrt{\frac{\lambda_1}{\lambda_M}} \le \sqrt{\frac{\gamma_4}{\gamma_3}}\,.$$

This completes the proof. ■

REMARK 3.4. Let us consider the parameter estimation problem (1.3) in the special case $\mathbb{G} = \mathbb{Z}_N^d$ with $(2N+1)^d$ given sampled data $h(\boldsymbol{n})$ ($\boldsymbol{n} \in \mathbb{Z}_N^d$). Assume that distinct frequency vectors $\boldsymbol{f}_j \in [-\pi,\, \pi)^d$ ($j = 1,\, \ldots,\, M$) with separation distance q are determined. If we replace (1.3) by the overdetermined linear system

$$\sum_{j=1}^{M} c_j\, \mathrm{e}^{\mathrm{i}\boldsymbol{f}'_j \cdot \boldsymbol{k}} = h(\boldsymbol{k}) \quad (\boldsymbol{k} \in \mathbb{Z}_N^d)\,,$$

then by Corollary 3.3 the coefficient matrix has a uniformly bounded condition number for all $N > \frac{\sqrt{d}\pi}{q}$. Further this matrix has full rank M. Hence the least squares solution $(c_j)_{j=1}^M$ can be computed and the sensitivity of the least squares solution to perturbations can be shown [7, pp. 239 – 244]. Unfortunately, this method requires too many sampled data. In Sections 5 and 6, we propose another parameter estimation method which uses only a relatively low number of sampled data. □

4. Approximate Prony method for $d = 1$. Here we sketch the *approximate Prony method* (APM) in the case $d = 1$. For details see [3, 23, 21]. Let $M \in \mathbb{N} \setminus \{1\}$ and $N \in \mathbb{N}$ with $N \ge 2M + 1$ be given. By $\mathbb{Z}_N$ we denote the finite set $[-N,\, N] \cap \mathbb{Z}$. We consider a univariate exponential sum

$$h_0(x) := \sum_{j=1}^{M} c_j\, \mathrm{e}^{\mathrm{i} f_j x} \quad (x \in \mathbb{R})$$

with distinct, ordered frequencies

$$-\pi \le f_1 < f_2 < \ldots < f_M < \pi$$

and complex coefficients $c_j \neq 0$. Assume that these frequencies are well–separated in the sense that

$$\mathrm{dist}(f_j, f_l) := \min\{|(f_j + 2\pi k) - f_l| : k \in \mathbb{Z}\} > \frac{\pi}{N}$$

for all $j, l = 1, \ldots, M$ with $j \neq l$. Suppose that noisy sampled data $h(k) := h_0(k) + e(k) \in \mathbb{C}$ ($k \in \mathbb{Z}_N$) are given, where the magnitudes of the error terms $e(k)$ are uniformly bounded by a certain accuracy $\varepsilon_1 > 0$. Further we assume that $|c_j| > \varepsilon_0$ ($j = 1, \ldots, M$) for a convenient bound $0 < \varepsilon_0 \ll 1$.

Then we consider the following *nonlinear approximation problem*: Recover the distinct frequencies $f_j \in [-\pi, \pi)$ and the complex coefficients c_j so that

$$|h(k) - \sum_{j=1}^{M} c_j\, \mathrm{e}^{\mathrm{i} f_j k}| \le \varepsilon \quad (k \in \mathbb{Z}_N)$$

for very small accuracy $\varepsilon > 0$ and for minimal number M of nontrivial summands. This problem can be solved by the following

Algorithm 4.1. (APM)

Input: $L,\, N \in \mathbb{N}$ ($3 \le L \le N$, L is an upper bound of the number of exponentials), $h(k) = h_0(k) + e(k) \in \mathbb{C}$ ($k \in \mathbb{Z}_N$) with $|e(k)| \le \varepsilon_1$, and bounds $\varepsilon_l > 0$ ($l = 0, 1, 2$).

1. Determine the smallest singular value of the rectangular Hankel matrix

$$\boldsymbol{H} := (h(k+l))_{k=-N,\, l=0}^{N-L,\, L}$$

and related right singular vector $\boldsymbol{u} = (u_l)_{l=0}^{L}$ by singular value decomposition.
2. Compute all zeros of the polynomial $\sum_{l=0}^{L} u_l\, z^l$ and determine all that zeros $\tilde{z}_j$ ($j = 1, \ldots, \tilde{M}$) that fulfill the property $|\,|\tilde{z}_j| - 1| \le \varepsilon_2$. 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

$$\sum_{j=1}^{\tilde{M}} \tilde{c}_j\, \tilde{w}_j^k = h(k) \quad (k \in \mathbb{Z}_N)\,.$$

For large $\tilde{M}$ and N, we can apply the CGNR method (conjugate gradient on the normal equations), where the multiplication of the rectangular Fourier–type matrix $(\tilde{w}_j^k)_{k=-N, j=1}^{N, \tilde{M}}$ is realized in each iteration step by the nonequispaced fast Fourier transform (NFFT) (see [13]).
4. Delete all the $\tilde{w}_l$ ($l \in \{1, \ldots, \tilde{M}\}$) with $|\tilde{c}_l| \le \varepsilon_0$ and denote the remaining entries by $\tilde{w}_j$ ($j = 1, \ldots, M$) with $M \le \tilde{M}$.
5. Repeat step 3 and compute $\tilde{c}_j \in \mathbb{C}$ ($j = 1, \ldots, M$) as least squares solution of the overdetermined linear Vandermonde–type system

$$\sum_{j=1}^{M} \tilde{c}_j\, \tilde{w}_j^k = h(k) \quad (k \in \mathbb{Z}_N)$$

with respect to the new set $\{\tilde{w}_j \,:\, j = 1, \ldots, M\}$ again. Set $\tilde{f}_j := \operatorname{Im}(\log \tilde{w}_j)$ ($j = 1, \ldots, M$), where log is the principal value of the complex logarithm.

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

Remark 4.2. The convergence and stability properties of Algorithm 4.1 are discussed in [23]. In all numerical tests of Algorithm 4.1 (see Section 7 and [23, 21]), we have obtained very good reconstruction results. All frequencies and coefficients can be computed such that

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

We have to assume that the frequencies f_j are well–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. Up to now, useful error estimates of $\max_{j=1,\ldots,M} |f_j - \tilde{f}_j|$ and $\sum_{j=1}^{M} |c_j - \tilde{c}_j|$ are unknown. □

REMARK 4.3. The above algorithm has been tested for $M \le 100$ and $N \le 10^5$ in MATLAB with double precision. For fixed upper bound L and variable N, the computational cost of this algorithm is very moderate with about $\mathcal{O}(N\,\log N)$ flops. In step 1, the singular value decomposition needs $14\,(2N-L+1)(L+1)^2+8\,(L+1)^2$ flops. In step 2, the QR decomposition of the companion matrix requires $\frac{4}{3}\,(L+1)^3$ flops (see [9, p. 337]). For large values N and $\tilde{M}$, one can use the nonequispaced fast Fourier transform iteratively in steps 3 and 5. Since the condition number of the Fourier–type matrix $(\tilde{w}_j^k)_{k=-N,j=1}^{N,\tilde{M}}$ is uniformly bounded by Corollary 3.3, we need finitely many iterations of the CGNR method. In each iteration step, the product between this Fourier–type matrix and an arbitrary vector of length $\tilde{M}$ can be computed with the NFFT by $\mathcal{O}(N\,\log N + L\,|\log\varepsilon|)$ flops, where $\varepsilon > 0$ is the wanted accuracy (see [13]). □

REMARK 4.4. In this paper, we use the Algorithm 4.1 for parameter estimation of univariate exponential sums. But we can replace this procedure also by another Prony–like method [24], such as ESPRIT [26, 27] or matrix pencil method [10, 29]. □

REMARK 4.5. By similar ideas, we can reconstruct also all parameters of an *extended exponential sum*

$$h_0(x) = \sum_{j=1}^{M} p_j(x)\,\mathrm{e}^{\mathrm{i}\,f_j x} \quad (x \in \mathbb{R})\,,$$

where p_j $(j = 1,\ldots,M)$ is an algebraic polynomial of degree $m_j \ge 0$ (see [4, p. 169]). Then we can interpret the exactly sampled values

$$h_0(n) = \sum_{j=1}^{M} p_j(n)\,z_j^n \quad (n \in \mathbb{Z}_N)$$

with $z_j := \mathrm{e}^{\mathrm{i}\,f_j}$ as a solution of a homogeneous linear difference equation

$$\sum_{k=0}^{M_0} p_k\,h_0(j+k) = 0 \quad (j \in \mathbb{Z})\,, \tag{4.1}$$

where the coefficients p_k $(k = 0,\ldots,M_0)$ are defined by

$$\prod_{j=1}^{M}(z - z_j)^{m_j+1} = \sum_{k=0}^{M_0} p_k\,z^k\,, \quad M_0 := \sum_{j=1}^{M}(m_j+1)\,.$$

Note that in this case z_j is a zero of order m_j of the above polynomial and we can cover multiple zeros with this approach. Consequently, (4.1) has the general solution

$$h_0(k) = \sum_{j=1}^{M}\Big(\sum_{l=0}^{m_j} c_{j,l}\,k^l\Big)\,z_j^k \quad (k \in \mathbb{Z})\,.$$

Then we determine the coefficients $c_{j,l}$ $(j = 1,\ldots,M;\ l = 0,\ldots,m_j)$ in such a way that

$$\sum_{j=1}^{M}\Big(\sum_{l=0}^{m_j} c_{j,l}\,k^l\Big)\,z_j^k = h(k) \quad (k \in \mathbb{Z}_N)\,,$$

where we assume that $N \ge 2M_0 + 1$. To this end, we compute the least squares solution of the above overdetermined linear system. □

5. Sparse approximate Prony method for $d = 2$. Let $M \in \mathbb{N} \setminus \{1\}$ and $N \in \mathbb{N}$ with $N \ge 2M+1$ be given. The aim of this section is to present a new efficient parameter estimation method for a bivariate exponential sum of order M using only $\mathcal{O}(N)$ sampling points. The main idea is to project the bivariate reconstruction problem to several univariate problems and to solve these problems by methods from the previous Section 4. Finally we combine the results from the univariate problems. Note that it is not necessary to sample the bivariate exponential sum

$$h_0(x_1,\, x_2) = \sum_{j=1}^{M} c_j\, \mathrm{e}^{\mathrm{i}\,(f_{j,1}x_1 + f_{j,2}x_2)}\,.$$

on the full sampling grid $\mathbb{Z}_N^d$. Assume that the distinct frequency vectors

$$\boldsymbol{f}_j = (f_{j,1},\, f_{j,2})^\top \in [-\pi,\, \pi)^2 \quad (j = 1, \ldots, M)$$

are *well–separated* by

$$\mathrm{dist}(f_{j,l}, f_{k,l}) > \pi/N \tag{5.1}$$

for all $j, k = 1, \ldots, M$ and $l = 1,\, 2$, if $f_{j,l} \ne f_{k,l}$. We solve the corresponding parameter estimation problem stepwise and call this new procedure *sparse approximate Prony method* (SAPM). Here we use only noisy values $h(n, 0)$, $h(0, n)$, $h(n, \alpha n + \beta)$ $(n \in \mathbb{Z}_N)$ sampled along straight lines, where $\alpha \in \mathbb{Z} \setminus \{0\}$ and $\beta \in \mathbb{Z}$ are conveniently chosen.

First we consider the given noisy data $h(n, 0)$ $(n \in \mathbb{Z}_N)$ of

$$h_0(n, 0) = \sum_{j=1}^{M} c_j\, \mathrm{e}^{\mathrm{i} f_{j,1} n} = \sum_{j_1=1}^{M_1} c_{j_1,1}\, \mathrm{e}^{\mathrm{i} f'_{j_1,1} n}\,, \tag{5.2}$$

where $1 \le M_1 \le M$, $f'_{j_1,1} \in [-\pi,\, \pi)$ $(j_1 = 1, \ldots, M_1)$ are the distinct values of $f_{j,1}$ $(j = 1, \ldots, M)$ and $c_{j_1,1} \in \mathbb{C}$ are certain linear combinations of the coefficients c_j. Assume that $c_{j_1,1} \ne 0$. Using the Algorithm 4.1, we compute the distinct frequencies $f'_{j_1,1} \in [-\pi,\, \pi)$ $(j_1 = 1, \ldots, M_1)$.

Analogously, we consider the given noisy data $h(0, n)$ $(n \in \mathbb{Z}_N)$ of

$$h_0(0, n) = \sum_{j=1}^{M} c_j\, \mathrm{e}^{\mathrm{i} f_{j,2} n} = \sum_{j_2=1}^{M_2} c_{j_2,2}\, \mathrm{e}^{\mathrm{i} f'_{j_2,2} n}, \tag{5.3}$$

where $1 \le M_2 \le M$, $f'_{j_2,2} \in [-\pi,\, \pi)$ $(j_2 = 1, \ldots, M_2)$ are the distinct values of $f_{j,2}$ $(j = 1, \ldots, M)$ and $c_{j_2,2} \in \mathbb{C}$ are certain linear combinations of the coefficients c_j. Assume that $c_{j_2,2} \ne 0$. Using the Algorithm 4.1, we compute the distinct frequencies $f'_{j_2,2} \in [-\pi,\, \pi)$ $(j_2 = 1, \ldots, M_2)$.

Then we form the Cartesian product

$$F = \{(f'_{j_1,1}, f'_{j_2,2})^\top \in [-\pi,\, \pi)^2 : j_1 = 1, \ldots, M_1,\, j_2 = 1, \ldots, M_2\} \tag{5.4}$$

of the sets $\{f'_{j_1,1} : j_1 = 1, \ldots, M_1\}$ and $\{f'_{j_2,2} : j_2 = 1, \ldots, M_2\}$. Now we test, if $(f'_{j_1,1}, f'_{j_2,2})^\top \in F$ is an approximation of an actual frequency vector $\boldsymbol{f}_j = (f_{j,1}, f_{j,2})^\top$

($j=1,\ldots,M$). Choosing further parameters $\alpha \in \mathbb{Z}\setminus\{0\}$, $\beta \in \mathbb{Z}$, we consider the given noisy data $h(n, \alpha n+\beta)$ ($n \in \mathbb{Z}_N$) of

$$h_0(n, \alpha n+\beta) = \sum_{j=1}^{M} c_j\, \mathrm{e}^{\mathrm{i}\beta f_{j,2}}\, \mathrm{e}^{\mathrm{i}(f_{j,1}+\alpha f_{j,2})n} = \sum_{k=1}^{M_2'} c_{k,3}\, \mathrm{e}^{\mathrm{i} f_k(\alpha) n}\,, \tag{5.5}$$

where $1 \le M_2' \le M$, $f_k(\alpha) \in [-\pi,\,\pi)$ ($k=1,\ldots,M_2'$) are the distinct values of $(f_{j,1}+\alpha f_{j,2})_{2\pi}$ ($j=1,\ldots,M$). Here $(f_{j,1}+\alpha f_{j,2})_{2\pi}$ is the symmetric residuum of $f_{j,1}+\alpha f_{j,2}$ modulo 2π, i.e. $f_{j,1}+\alpha f_{j,2} \in (f_{j,1}+\alpha f_{j,2})_{2\pi} + 2\pi\,\mathbb{Z}$ and $(f_{j,1}+\alpha f_{j,2})_{2\pi} \in [-\pi,\,\pi)$. Note that $f_k(\alpha) \in [-\pi,\,\pi)$ and that $f_{j,1}+\alpha f_{j,2}$ can be located outside of $[-\pi,\,\pi)$. The coefficients $c_{k,3} \in \mathbb{C}$ are certain linear combinations of the coefficients $c_j\, \mathrm{e}^{\mathrm{i}\beta f_{j,2}}$. Assume that $c_{k,3} \neq 0$. Using the Algorithm 4.1, we compute the distinct frequencies $f_k(\alpha) \in [-\pi,\,\pi)$ ($k=1,\ldots,M_2'$).

Then we form the set $\tilde F$ of all those $(f'_{j_1,1}, f'_{j_2,2})^\top \in F$ so that there exists a frequency $f_k(\alpha)$ ($k=1,\ldots,M_2'$) with

$$|f_k(\alpha) - (f'_{j_1,1} + \alpha f'_{j_2,2})_{2\pi}| < \varepsilon_1\,,$$

where $\varepsilon_1 > 0$ is an accuracy bound. Clearly, one can repeat the last step with other parameters $\alpha \in \mathbb{Z}\setminus\{0\}$ and $\beta \in \mathbb{Z}$ to obtain a smaller set $\tilde F := \{\tilde{\boldsymbol f}_j = (\tilde f_{j,1}, \tilde f_{j,2})^\top : j=1,\ldots,|\tilde F|\}$.

Finally we compute the coefficients $\tilde c_j$ ($j=1,\ldots,|\tilde F|$) as least squares solution of the overdetermined linear system

$$\sum_{j=1}^{|\tilde F|} \tilde c_j\, \mathrm{e}^{\mathrm{i}\tilde{\boldsymbol f}_j\cdot \boldsymbol n} = h(\boldsymbol n) \quad (\boldsymbol n \in \mathbb{G})\,, \tag{5.6}$$

where $\mathbb{G} := \{(n,0),\,(0,n),\,(n,\alpha\,n+\beta);\, n \in \mathbb{Z}_N\}$ is the sparse sampling grid. In other words, this linear system (5.6) reads as follows

$$\sum_{j=1}^{|\tilde F|} \tilde c_j\, \mathrm{e}^{\mathrm{i}\tilde f_{j,1} n} = h(n,0) \quad (n \in \mathbb{Z}_N)\,,$$

$$\sum_{j=1}^{|\tilde F|} \tilde c_j\, \mathrm{e}^{\mathrm{i}\tilde f_{j,2} n} = h(0,n) \quad (n \in \mathbb{Z}_N)\,,$$

$$\sum_{j=1}^{|\tilde F|} \tilde c_j\, \mathrm{e}^{\mathrm{i}\beta\tilde f_{j,2}}\, \mathrm{e}^{\mathrm{i}(\tilde f_{j,1}+\alpha\tilde f_{j,2})n} = h(n,\alpha n+\beta) \quad (n \in \mathbb{Z}_N)\,.$$

Unfortunately, these three system matrices can possess equal columns. Therefore we represent these matrices as products $\boldsymbol F_l\, \boldsymbol M_l$ ($l=1,\,2,\,3$), where $\boldsymbol F_l$ is a nonequispaced Fourier matrix with distinct columns and where all entries of $\boldsymbol M_l$ are equal to 0 or 1 and only one entry of each column is equal to 1. By [22, Theorem 4.3] the nonequispaced Fourier matrices

$$\boldsymbol F_l := \left(\mathrm{e}^{\mathrm{i}\tilde f_{j,1} n}\right)_{n\in\mathbb{Z}_N,\, j=1,\ldots,|\tilde F|} \quad (l=1,\,2)\,,$$

$$\boldsymbol F_3 := \left(\mathrm{e}^{\mathrm{i}(\tilde f_{j,1}+\alpha\tilde f_{j,2})n}\right)_{n\in\mathbb{Z}_N,\, j=1,\ldots,|\tilde F|}$$

possess left inverses $\boldsymbol{L}_l$. If we introduce the vectors $\boldsymbol{h}_1 := (h(n,0))_{n=-N}^N$, $\boldsymbol{h}_2 := (h(0,n))_{n=-N}^N$, $\boldsymbol{h}_3 := (h(n,\alpha n+\beta))_{n=-N}^N$, $\tilde{\boldsymbol{c}} := (\tilde{c}_j)_{j=1}^{|\tilde{F}|}$, and the diagonal matrix $\boldsymbol{D} := \mathrm{diag}\,(\exp(\mathrm{i}\beta \tilde{f}_{j,2}))_{j=1}^{|\tilde{F}|}$, we obtain the linear system

$$\begin{pmatrix} \boldsymbol{M}_1 \\ \boldsymbol{M}_2 \\ \boldsymbol{M}_3\,\boldsymbol{D} \end{pmatrix} \tilde{\boldsymbol{c}} = \begin{pmatrix} \boldsymbol{L}_1\,\boldsymbol{h}_1 \\ \boldsymbol{L}_2\,\boldsymbol{h}_2 \\ \boldsymbol{L}_3\,\boldsymbol{h}_3 \end{pmatrix}. \tag{5.7}$$

By a convenient choice of the parameters $\alpha \in \mathbb{Z}\setminus\{0\}$ and $\beta \in \mathbb{Z}$, the rank of the above system matrix is equal to $|\tilde{F}|$. If this is not the case, we can use sampled values of h_0 along another straight line. We summarize:

Algorithm 5.1. (SAPM for $d=2$)
Input: $h(n,0), h(0,n) \in \mathbb{C}$ $(n \in \mathbb{Z}_N)$, bounds ε_0, $\varepsilon_1 > 0$,
m number of additional straight lines, parameters $\alpha_l \in \mathbb{Z}\setminus\{0\}$, $\beta_l \in \mathbb{Z}$ $(l=1,\ldots,m)$, $h(n,\alpha_l n+\beta_l) \in \mathbb{C}$ $(n \in \mathbb{Z}_N;\, l=1,\ldots,m)$.

1. From the noisy data $h(n,0)$ $(n \in \mathbb{Z}_N)$ and $h(0,n)$ $(n \in \mathbb{Z}_N)$ compute by Algorithm 4.1 the distinct frequencies $f'_{j_1,1} \in [-\pi,\,\pi)$ $(j_1 = 1,\ldots,M_1)$ in (5.2) and $f'_{j_2,2} \in [-\pi,\,\pi)$ $(j_2 = 1,\ldots,M_2)$ in (5.3), respectively. Set $\mathbb{G} := \{(n,0),\,(0,n) : n \in \mathbb{Z}_N\}$.
2. Form the Cartesian product (5.4).
3. For $l = 1,\ldots,m$ do:
   From the noisy data $h(n,\alpha_l n+\beta_l)$ $(n \in \mathbb{Z}_N)$, compute the distinct frequencies $f_k(\alpha_l) \in [-\pi,\,\pi)$ $(k=1,\ldots,M'_2)$ in (5.5) by Algorithm 4.1. Form the set $F' := \{\boldsymbol{f}'_j : j = 1,\ldots,|F'|\}$ of all those $(f'_{j_1,1}, f'_{j_2,2})^\top \in F$ so that there exists a frequency $f_k(\alpha_l)$ $(k=1,\ldots,M'_2)$ with

$$|f_k(\alpha_l) - (f'_{j_1,1} + \alpha_l f'_{j_2,2})_{2\pi}| < \varepsilon_1\,.$$

   Set $\mathbb{G} := \mathbb{G} \cup \{(n,\alpha_l n+\beta_l) : n \in \mathbb{Z}_N\}$.
4. Compute the least squares solution of the overdetermined linear system

$$\sum_{j=1}^{|F'|} c'_j\,\mathrm{e}^{\mathrm{i}\boldsymbol{f}'_j\cdot\boldsymbol{n}} = h(\boldsymbol{n}) \quad (\boldsymbol{n} \in \mathbb{G})$$

for the frequency set F'.
5. Form the subset $\tilde{F} = \{\tilde{\boldsymbol{f}}_j : j = 1,\ldots,|\tilde{F}|\}$ of F' of all those $\boldsymbol{f}'_k \in F'$ $(k = 1,\ldots,|F'|)$ with $|c'_k| > \varepsilon_0$.
6. Compute the least squares solution of the overdetermined linear system (5.6) corresponding to the new frequency set $\tilde{F}$.

Output: $M := |\tilde{F}| \in \mathbb{N}$, $\tilde{\boldsymbol{f}}_j \in [-\pi,\pi)^2$, $\tilde{c}_j \in \mathbb{C}$ $(j=1,\ldots,M)$.
Note that it can be useful in some applications to choose grid points $(n,\alpha_l n+\beta_l)$ $(n \in \mathbb{Z}_N)$ at random straight lines.

Remark 5.2. For the above parameter reconstruction, we have used sampled values of a bivariate exponentialsum h_0 on $m+2$ straight lines. We have determined in the step 3 of Algorithm 5.1 only a set F' which contains the set $\tilde{F}$ of all exact frequency vectors as a subset. This method is related to a result of A. Rényi [25] which is known in discrete tomography: M distinct points in $\mathbb{R}^2$ are completely determined, if their orthogonal projections onto $M+1$ arbitrary distinct straight lines through the origin are known. Let us additionally assume that $\|\boldsymbol{f}_j\|_2 < \pi$ $(j = 1, \ldots, M)$. Further let $\varphi_\ell \in [0, \pi)$ $(\ell = 0, \ldots, M)$ be distinct angles. From sampled data $h_0(n \cos\varphi_\ell, n \sin\varphi_\ell)$ $(n \in \mathbb{Z}_N)$ we reconstruct the parameters $f_{j,1} \cos\varphi_\ell + f_{j,2} \sin\varphi_\ell$ for $j = 1, \ldots, M$ and $\ell = 0, \ldots, M$. Since $|f_{j,1} \cos\varphi_\ell + f_{j,2} \sin\varphi_\ell| < \pi$, we have

$$(f_{j,1} \cos\varphi_\ell + f_{j,2} \sin\varphi_\ell)_{2\pi} = f_{j,1} \cos\varphi_\ell + f_{j,2} \sin\varphi_\ell \,.$$

Thus $f_{j,1} \cos\varphi_\ell + f_{j,2} \sin\varphi_\ell$ is equal to the distance between $\boldsymbol{f}_j$ and the line $x_1 \cos\varphi_\ell + x_2 \sin\varphi_\ell = 0$, i.e., we know the orthogonal projection of $\boldsymbol{f}_j$ onto the straight line $x_1 \cos\varphi_\ell - x_2 \sin\varphi_\ell = 0$. Hence we know that $m \le M-1$. □

6. Sparse approximate Prony method for $d \ge 3$. Now we extend the Algorithm 5.1 to the parameter estimation of a d–variate exponential sum (1.1), where the dimension $d \ge 3$ is moderately sized. Let $M \in \mathbb{N}\setminus\{1\}$ and $N \in \mathbb{N}$ with $N \ge 2M+1$ be given. Assume that the distinct frequency vectors $\boldsymbol{f}_j = (f_{j,l})_{l=0}^d$ are well–separated by the condition

$$\mathrm{dist}(f_{j,l}, f_{k,l}) > \pi/N$$

for all $j, k = 1, \ldots, M$ and $l = 1, \ldots, d$ with $f_{j,l} \ne f_{k,l}$.
Our strategy for parameter recovery of (1.1) is based on a stepwise enhancement of the dimension from 2 to d.
For $r = 2, \ldots, d$, we introduce the matrices

$$\boldsymbol{\alpha}^{(r)} := \begin{pmatrix} \alpha^{(r)}_{1,1} & \cdots & \alpha^{(r)}_{1,r-1} \\ \vdots & \ddots & \vdots \\ \alpha^{(r)}_{m_r,1} & \cdots & \alpha^{(r)}_{m_r,r-1} \end{pmatrix} \in (\mathbb{Z}\setminus\{0\})^{m_r \times (r-1)} ,$$

$$\boldsymbol{\beta}^{(r)} := \begin{pmatrix} \beta^{(r)}_{1,1} & \cdots & \beta^{(r)}_{1,r-1} \\ \vdots & \ddots & \vdots \\ \beta^{(r)}_{m_r,1} & \cdots & \beta^{(r)}_{m_r,r-1} \end{pmatrix} \in \mathbb{Z}^{m_r \times (r-1)} ,$$

where $\alpha^{(r)}_{l,1}, \ldots, \alpha^{(r)}_{l,r-1}$ and $\beta^{(r)}_{l,1}, \ldots, \beta^{(r)}_{l,r-1}$ are the parameters of the grid points

$$(n, \alpha^{(r)}_{l,1} n + \beta^{(r)}_{l,1}, \ldots, \alpha^{(r)}_{l,r-1} n + \beta^{(r)}_{l,r-1}, 0, \ldots, 0) \in \mathbb{Z}^d \quad (n \in \mathbb{Z}^d)$$

lying at the l-th straight line $(l = 1, \ldots, m_r)$. By $\boldsymbol{\alpha}^{(r)}_l$ $(l = 1, \ldots, m_r)$, we denote the l–th row of the matrix $\boldsymbol{\alpha}^{(r)}$.

Using the given values $h(n,0,0,\ldots,0)$, $h(0,n,0,\ldots,0)$, $h(n, \alpha^{(2)}_{l,1} n + \beta^{(2)}_{l,1}, 0, \ldots, 0)$ $(l = 1, \ldots, m_2)$ for $n \in \mathbb{Z}_N$, we determine frequency vectors $(f'_{j,1}, f'_{j,2})^\top \in [-\pi, \pi)^2$ $(j = 1, \ldots, M')$ by Algorithm 5.1.

Then we consider the noisy data $h(0,0,n,0,\ldots,0)$ $(n \in \mathbb{Z}_N)$ of

$$h_0(0,0,n,0,\ldots,0) = \sum_{j=1}^{M} c_j \, \mathrm{e}^{\mathrm{i} f_{j,3} n} = \sum_{j_3=1}^{M_3} c_{j_3,3} \, \mathrm{e}^{\mathrm{i} f'_{j_3,3} n},$$

where $1 \leq M_3 \leq M$, where $f'_{j_3,3} \in [-\pi, \pi)$ $(j_3 = 1, \ldots, M_3)$ the distinct values of $f_{j,3}$ $(j = 1, \ldots, M)$, and where $c_{j_3,3} \in \mathbb{C}$ are certain linear combinations of the coefficients c_j. Assume that $c_{j_3,3} \neq 0$. Using the Algorithm 4.1, we compute the distinct frequencies $f'_{j_3,3} \in [-\pi, \pi)$ $(j_3 = 1, \ldots, M_3)$. Now we form the Cartesian product

$$F := \{(f'_{j,1}, f'_{j,2}, f'_{j_3,3})^\top \in [-\pi, \pi)^3 : j = 1, \ldots, M'; j_3 = 1, \ldots, M_3\}$$

of the sets $\{(f'_{j,1}, f'_{j,2})^\top : j = 1, \ldots, M'\}$ and $\{f'_{j_3,3} : j = 1, \ldots, M_3\}$. Now we form a subset of F by using the data

$$h(n, \alpha^{(3)}_{l,1} n + \beta^{(3)}_{l,1}, \alpha^{(3)}_{l,2} n + \beta^{(3)}_{l,2}, 0, \ldots, 0) \quad (l = 1, \ldots, m_3)\,.$$

Since

$$\begin{aligned} & h_0(n, \alpha^{(3)}_{l,1} n + \beta^{(3)}_{l,1}, \alpha^{(3)}_{l,2} n + \beta^{(3)}_{l,2}, 0, \ldots, 0) \\ &= \sum_{j=1}^{M} c_j\, \mathrm{e}^{\mathrm{i}(\beta^{(3)}_{l,1} f'_{j,2} + \beta^{(3)}_{l,2} f'_{j,3})}\, \mathrm{e}^{\mathrm{i}(f'_{j_1,1} + f'_{j_2,2}\alpha^{(3)}_{l,1} + f'_{j_3,3}\alpha^{(3)}_{l,2})n} \\ &= \sum_{k=1}^{M'_3} c_{k,3}\, \mathrm{e}^{\mathrm{i} f_k(\boldsymbol{\alpha}^{(3)}_l) n}\,, \end{aligned}$$

where $1 \leq M'_3 \leq M$ and where $f_k(\boldsymbol{\alpha}^{(3)}_l) \in [-\pi, \pi)$ $(k = 1, \ldots, M'_3)$ are the distinct values of $(f'_{j,1} + \alpha^{(3)}_{l,1} f'_{j,2} + \alpha^{(3)}_{l,2} f'_{j,3})_{2\pi}$. The coefficients $c_{k,3} \in \boldsymbol{C}$ are certain linear combinations of the coefficients $c_j\, \mathrm{e}^{\mathrm{i}(\beta^{(3)}_{l,1} f'_{j,2} + \beta^{(3)}_{l,2} f'_{j,3})}$. Then we form the set $F' := \{\boldsymbol{f}'_j : j = 1, \ldots, |F'|\}$ of all those $(f'_{j_1,1}, f'_{j_2,2}, f'_{j_3,3})^\top \in F$ so that there exists a frequency $f_k(\boldsymbol{\alpha}^{(3)}_l)$ $(k = 1, \ldots, M'_3)$ with

$$|f_k(\boldsymbol{\alpha}^{(3)}_l) - (f'_{j_1,1} + f'_{j_2,2}\alpha^{(3)}_{l,1} + f'_{j_3,3}\alpha^{(3)}_{l,2})_{2\pi}| < \varepsilon_1\,.$$

Continuing analogously this procedure, we obtain

Algorithm 6.1. (SAPM for $d \geq 3$)

Input: $h(n, 0, \ldots, 0), h(0, n, 0, \ldots, 0), \ldots, h(0, \ldots, 0, n)$ $(n \in \mathbb{Z}_N)$, bounds ε_0, $\varepsilon_1 > 0$, m_r number of straight lines for dimension $r = 2, \ldots, d$, parameters of straight lines $\boldsymbol{\alpha}^{(r)}$, $\boldsymbol{\beta}^{(r)} \in \mathbb{Z}^{m_r \times (r-1)}$.

1. From the noisy data $h(n, 0, \ldots, 0)$, $h(0, n, 0, \ldots, 0)$, $\ldots$, $h(0, \ldots, 0, n)$ $(n \in \mathbb{Z}_N)$ compute by Algorithm 4.1 the distinct frequencies $f'_{j_r,r} \in [-\pi, \pi)$ $(j_r = 1, \ldots, M_r)$ for $r = 1, \ldots, d$.
Set $\mathbb{G} := \{(n, 0, \ldots, 0), \ldots, (0, \ldots, 0, n) : n \in \mathbb{Z}_N\}$.
2. Set $F := \{f'_{j_1,1} : j_1 = 1, \ldots, M_1\}$.
3. For $r = 2, \ldots, d$ do:

Form the Cartesian product

$$F := F \times \{f'_{j_r,r} : j_r = 1, \ldots, M_r\} = \{(\boldsymbol{f}_l^\top, f'_{j,r})^\top : l = 1, \ldots |F|, j = 1, \ldots, M_r\}\,.$$

For $l = 1, \ldots, m_r$ do:

For the noisy data

$$h(n, \alpha^{(r)}_{l,1} n + \beta^{(r)}_{l,1}, \ldots, \alpha^{(r)}_{l,r-1} n + \beta^{(r)}_{l,r-1}, 0, \ldots, 0) \quad (n \in \mathbb{Z}_N)\,,$$

compute the distinct frequencies $f_k(\boldsymbol{\alpha}_l^{(r)}) \in [-\pi,\, \pi)$ $(k = 1, \ldots, M'_r)$ by Algorithm 4.1. Form the set $\tilde{F}$ of all those $(f'_{j_1,1}, f'_{j_2,2}, \ldots, f'_{j_r,r})^\top \in F$ so that there exists a frequency $f_k(\boldsymbol{\alpha}_l^{(r)})$ with

$$|f_k(\boldsymbol{\alpha}_l^{(r)}) - (f'_{j_1,1} + \alpha_{l,1}^{(r)} f_{j_2,2} + \cdots + \alpha_{l,r-1}^{(r)} f'_{j_r,r})_{2\pi}| < \varepsilon_1\,.$$

Set $F := \tilde{F}$ and

$$\mathbb{G} := \mathbb{G} \cup \{(n, \alpha_{l,1}^{(r)} n + \beta_{l,1}^{(r)}, \ldots, \alpha_{l,r-1}^{(r)} n + \beta_{l,r-1}^{(r)}, 0, \ldots, 0) : n \in \mathbb{Z}_N\}.$$

4. Compute the least squares solution of the overdetermined linear system

$$\sum_{j=1}^{|F|} c'_j\, \mathrm{e}^{\mathrm{i}\boldsymbol{f}_j \cdot \boldsymbol{n}} = h(\boldsymbol{n}) \quad (\boldsymbol{n} \in \mathbb{G}) \tag{6.1}$$

for the frequency set $F = \{\boldsymbol{f}_j : j = 1, \ldots, |F|\}$.
5. Form the set $\tilde{F} := \{\tilde{\boldsymbol{f}}_j : j = 1, \ldots, |\tilde{F}|\}$ of all those $\boldsymbol{f}_k \in F$ $(k = 1, \ldots, |F|)$ with $|c'_k| > \varepsilon_0$.
6. Compute the least squares solution of the overdetermined linear system

$$\sum_{j=1}^{|\tilde{F}|} \tilde{c}_j\, \mathrm{e}^{\mathrm{i}\tilde{\boldsymbol{f}}_j \cdot \boldsymbol{n}} = h(\boldsymbol{n}) \quad (\boldsymbol{n} \in \mathbb{G}) \tag{6.2}$$

corresponding to the new frequency set $\tilde{F} = \{\tilde{\boldsymbol{f}}_j : j = 1, \ldots, |\tilde{F}|\}$.

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

Remark 6.2. Note that we solve the overdetermined linear systems (6.1) and (6.2) only by using the values $h(\boldsymbol{n})$ $(\boldsymbol{n} \in \mathbb{G})$, which we have used to determine the frequencies $\tilde{\boldsymbol{f}}_j$. If more values $h(\boldsymbol{n})$ are available, clearly one can use further values as well in the final step to ensure a better least squares solvability of the linear systems, see (5.7) for the case $d = 2$ and Corollary 3.3. In addition we mention that there are various possibilities to combine the different dimensions, see e.g. Example 7.4. □

Remark 6.3. Our method can be interpreted as a reconstruction method for sparse multivariate trigonometric polynomials from few samples, see [14, 12, 32] and the references therein. More precisely, let Π_N^d denote the space of all d–variate trigonometric polynomials of maximal order N. An element $p \in \Pi_N^d$ can be represented in the form

$$p(\boldsymbol{y}) = \sum_{\boldsymbol{k} \in \mathbb{Z}_N^d} c_{\boldsymbol{k}}\, \mathrm{e}^{2\pi\mathrm{i}\,\boldsymbol{k}\cdot\boldsymbol{y}} \quad (\boldsymbol{y} \in [-\frac{1}{2}, \frac{1}{2}]^d)$$

with $c_{\boldsymbol{k}} \in \mathbb{C}$. There exist completely different methods for the reconstruction of "sparse trigonometric polynomials", i.e., one assumes that the number M of the nonzero coefficients $c_{\boldsymbol{k}}$ is much smaller than the dimension of Π_N^d. Therefore our method can be used with

$$h(\boldsymbol{x}) := p(\frac{\boldsymbol{x}}{2N}) = \sum_{j=1}^{M} c_j\, \mathrm{e}^{\mathrm{i}\,\boldsymbol{f}_j\cdot\boldsymbol{x}} \quad (\boldsymbol{x} \in [-N, N]^d),$$

and $\boldsymbol{x} = 2N\boldsymbol{y}$ and $\boldsymbol{f}_j = \pi\boldsymbol{k}/N$ if $c_{\boldsymbol{k}} \neq 0$. Using Algorithm 6.1, we find the frequency vectors $\boldsymbol{f}_j$ and the coefficients c_j and we finally set $\boldsymbol{k} := \mathrm{round}(N\boldsymbol{f}_j/\pi)$, $c_{\boldsymbol{k}} := c_j$. By [8] one knows sharp versions of L^2–norm equivalences for trigonometric polynomials under the assumption that the sampling set contains no holes larger than the inverse polynomial degree, see also [2]. $\square$

7. Numerical experiments. Finally, we apply the algorithms suggested in Section 5 to various examples. We have implemented our algorithms in MATLAB with IEEE double precision arithmetic. We compute the relative error of the frequencies given by

$$e(\boldsymbol{f}) := \max_{l=1,\ldots,d} \frac{\max\limits_{j=1,\ldots,M} |f_{j,l} - \tilde{f}_{j,l}|}{\max\limits_{j=1,\ldots,M} |f_{j,l}|} \,,$$

where $\tilde{f}_{j,l}$ are the frequency components computed by our algorithms. Analogously, the relative error of the coefficients is defined by

$$e(\boldsymbol{c}) := \frac{\max\limits_{j=1,\ldots,M} |c_j - \tilde{c}_j|}{\max\limits_{j=1,\ldots,M} |c_j|} \,,$$

where $\tilde{c}_j$ are the coefficients computed by our algorithms. Further we determine the relative error of the exponential sum by

$$e(h) := \frac{\max |h(\boldsymbol{x}) - \tilde{h}(\boldsymbol{x})|}{\max |h(\boldsymbol{x})|} \,,$$

where the maximum is built from approximately 10000 equispaced points from a grid of $[-N, N]^d$, and where

$$\tilde{h}(\boldsymbol{x}) := \sum_{j=1}^{M} \tilde{c}_j \, \mathrm{e}^{\tilde{\boldsymbol{f}}_j \cdot \boldsymbol{x}}$$

is the exponential sum recovered by our algorithms. We remark that the approximation property of h and $\tilde{h}$ in the uniform norm of the univariate method was shown in [21, Theorem 3.4]. We begin with an example previously considered in [28].

EXAMPLE 7.1. The bivariate exponential sum (1.1) taken from [28, Example 1] possesses the following parameters

$$(\boldsymbol{f}_j^\top)_{j=1}^3 = \begin{pmatrix} 0.48\pi & 0.48\pi \\ 0.48\pi & -0.48\pi \\ -0.48\pi & 0.48\pi \end{pmatrix} , \quad (c_j)_{j=1}^3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} .$$

We sample this exponential sum (1.1) at the nodes $h(k, 0)$, $h(0, k)$ and $h(k, \alpha k + \beta)$, $(k \in \mathbb{Z}_N)$, where α, $\beta \in \mathbb{Z}$ are given in Table 7.1. Therefore the number of total sampling points used in our method are only $3(2N + 1)$ or $4(2N + 1)$. Then we apply our Algorithm 5.1 for exact sampled data and for noisy sampled data $\tilde{h}(\boldsymbol{k}) = h(\boldsymbol{k}) + 10^{-\delta}\, e_{\boldsymbol{k}}$, where $e_{\boldsymbol{k}}$ is uniformly distributed in $[-1, 1]$. The notation $\delta = \infty$ means that exact data are given. We present the chosen parameters and the results in Table 7.1. We choose same bounds $\varepsilon_0 = \varepsilon_1$ in Algorithm 5.1 and obtain very precise results even in the case, where the unknown number $M = 3$ is estimated by L. $\square$

L	N	ε_0	α	β	δ	$e(\boldsymbol{f})$	$e(\boldsymbol{c})$	$e(h)$
5	6	10^{-4}	1	0	∞	1.7e−15	5.9e−14	3.2e−13
10	20	10^{-4}	1	0	∞	5.4e−15	4.5e−14	4.5e−14
5	25	10^{-3}	1	0	6	5.6e−09	1.6e−07	2.5e−07
5	25	10^{-3}	1, 2	0, 0	6	1.0e−08	5.9e−07	7.4e−07
5	25	10^{-3}	1	0	5	1.7e−08	1.2e−06	1.3e−06

Table 7.1
Results of Example 7.1.

Example 7.2. We consider the bivariate exponential sum (1.1) with following parameters

$$(\boldsymbol{f}_j^\top)_{j=1}^8 = \begin{pmatrix} 0.1 & 1.2 \\ 0.19 & 1.3 \\ 0.3 & 1.5 \\ 0.35 & 0.3 \\ -0.1 & 1.2 \\ -0.19 & 0.35 \\ -0.3 & -1.5 \\ -0.3 & 0.3 \end{pmatrix}, \quad (c_j)_{j=1}^8 = \begin{pmatrix} 1+\mathrm{i} \\ 2+3\,\mathrm{i} \\ 5-6\,\mathrm{i} \\ 0.2-\mathrm{i} \\ 1+\mathrm{i} \\ 2+3\,\mathrm{i} \\ 5-6\,\mathrm{i} \\ 0.2-\mathrm{i} \end{pmatrix}.$$

For given exact data, the results are presented in Table 7.2. Note that the condition (5.1) is not fulfilled, but the reconstruction is still possible in some cases. In order to fulfill (5.1), one has to choose $N > \frac{\pi}{0.05}$, i.e., $N \geq 63$.
The dash − in Table 7.2 means that we are not able to reconstruct the signal parameters. In the case $L = 15$, $N = 30$, $\alpha = 1$, $\beta = 0$, we are not able to find the 8 given frequency vectors and coefficients. There are other solutions of the reconstruction problem with 15 frequency vectors and coefficients. However, if we choose one more line with $\alpha = 2$, $\beta = 0$ or if we choose more sampling points with $N = 80$, then we obtain good parameter estimations.
Furthermore, we use noisy sampled data $h(\boldsymbol{k}) = h_0(\boldsymbol{k}) + 10^{-\delta}\, e_{\boldsymbol{k}}$, where $e_{\boldsymbol{k}}$ is uniformly distributed in $[-1, 1]$. Instead of predeterminated values α and β, we choose these values randomly. We use only one additional line for sampling and present the results in Table 7.3, where $e(\boldsymbol{f})$, $e(\boldsymbol{c})$ and $e(h)$ are the averages of 100 runs. Note that in this case we use only $3(2N + 1)$ sampling points for the parameter estimation. □

L	N	ε_0	α	β	$e(\boldsymbol{f})$	$e(\boldsymbol{c})$	$e(h)$
8	15	10^{-4}	1	0	2.7e−09	5.7e−09	3.4e−09
8	15	10^{-4}	1, 2, 3	0, 1, 2	2.7e−09	5.9e−09	3.3e−09
15	30	10^{-4}	1	0	1.4e−13	3.4e−13	6.5e−13
15	30	$2 \cdot 10^{-1}$	1	0	–	–	–
15	30	$2 \cdot 10^{-1}$	1, 2	0, 0	1.4e−13	4.0e−13	6.0e−13
15	80	$2 \cdot 10^{-1}$	1	0	3.5e−15	3.2e−14	7.5e−14

Table 7.2
Results of Example 7.2 with exact data.

L	N	ε_0	δ	$e(\boldsymbol{f})$	$e(\boldsymbol{c})$	$e(h)$
8	35	10^{-3}	6	1.4e−06	3.9e−06	5.5e−06
15	30	10^{-3}	6	1.2e−05	3.9e−05	5.3e−05
15	50	10^{-3}	5	4.0e−07	4.1e−06	3.8e−06
15	50	10^{-3}	6	3.8e−08	3.6e−07	3.3e−07

TABLE 7.3
Results of Example 7.2 with noisy data.

EXAMPLE 7.3. We consider the trivariate exponential sum (1.1) with following parameters

$$(\boldsymbol{f}_j^\top)_{j=1}^8 = \begin{pmatrix} 0.1 & 1.2 & 0.1 \\ 0.19 & 1.3 & 0.2 \\ 0.4 & 1.5 & 1.5 \\ 0.45 & 0.3 & -0.3 \\ -0.1 & 1.2 & 0.1 \\ -0.19 & 0.35 & -0.5 \\ -0.4 & -1.5 & 0.25 \\ -0.4 & 0.3 & -0.3 \end{pmatrix}, \quad (c_j)_{j=1}^8 = \begin{pmatrix} 1+\mathrm{i} \\ 2+3\,\mathrm{i} \\ 5-6\,\mathrm{i} \\ 0.2-\mathrm{i} \\ 1+\mathrm{i} \\ 2+3\,\mathrm{i} \\ 5-6\,\mathrm{i} \\ 0.2-\mathrm{i} \end{pmatrix}.$$

and present the results in Table 7.4. We use only $5(2N+1)$ or $6(2N+1)$ sampling points for the parameter estimation. □

L	N	ε_0	$\boldsymbol{\alpha}^{(1)}$	$\boldsymbol{\alpha}^{(2)}$	$\boldsymbol{\beta}^{(1)}$	$\boldsymbol{\beta}^{(2)}$	δ	$e(\boldsymbol{f})$	$e(\boldsymbol{c})$	$e(h)$
8	15	10^{-4}	(1)	$\begin{pmatrix}1 & 1\end{pmatrix}$	(0)	$\begin{pmatrix}0 & 0\end{pmatrix}$	∞	1.5e–10	1.7e–10	8.2e–11
8	15	10^{-4}	(1)	$\begin{pmatrix}1 & 1\end{pmatrix}$	(1)	$\begin{pmatrix}1 & 1\end{pmatrix}$	∞	1.5e–10	1.7e–10	8.1e–11
10	30	10^{-3}	(1)	$\begin{pmatrix}1 & 1\end{pmatrix}$	(0)	$\begin{pmatrix}0 & 0\end{pmatrix}$	6	8.7e–07	1.5e–06	2.9e–06
10	30	10^{-3}	(1)	$\begin{pmatrix}1 & 1 \\ 1 & 2\end{pmatrix}$	(0)	$\begin{pmatrix}0 & 0 \\ 0 & 0\end{pmatrix}$	6	7.8e–08	1.1e–06	1.5e–06
10	30	10^{-3}	(1)	$\begin{pmatrix}1 & 1 \\ 1 & 2\end{pmatrix}$	(0)	$\begin{pmatrix}0 & 0 \\ 0 & 0\end{pmatrix}$	5	4.5e–06	1.0e–05	1.6e–05
10	30	10^{-3}	(1)	$\begin{pmatrix}1 & 1 \\ 1 & 2\end{pmatrix}$	(0)	$\begin{pmatrix}0 & 0 \\ 0 & 0\end{pmatrix}$	4	1.2e–05	2.5e–05	5.2e–05

TABLE 7.4
Results of Example 7.3.

EXAMPLE 7.4. Now we consider the 4–variate exponential sum (1.1) with following parameters

$$(\boldsymbol{f}_j^\top)_{j=1}^8 = \begin{pmatrix} 0.1 & 1.2 & 0.1 & 0.45 \\ 0.19 & 1.3 & 0.2 & 1.5 \\ 0.3 & 1.5 & 1.5 & -1.3 \\ 0.45 & 0.3 & -0.3 & 0.4 \\ -0.1 & 1.2 & 0.1 & -1.5 \\ -0.19 & 0.35 & -0.5 & -0.45 \\ -0.4 & -1.5 & 0.25 & 1.3 \\ -0.4 & 0.3 & -0.3 & 0.4 \end{pmatrix}, \quad (c_j)_{j=1}^8 = \begin{pmatrix} 1+\mathrm{i} \\ 2+3\,\mathrm{i} \\ 5-6\,\mathrm{i} \\ 0.2-\mathrm{i} \\ 1+\mathrm{i} \\ 2+3\,\mathrm{i} \\ 5-6\,\mathrm{i} \\ 0.2-\mathrm{i} \end{pmatrix}.$$

Instead of using Algorithm 6.1 directly, we apply the Algorithm 5.1 for the first two variables and then for the last variables with the parameters $\boldsymbol{\alpha}^{(2)}$ and $\boldsymbol{\beta}^{(2)}$. Then we take the tensor product of the obtained two parameter sets and use the additional parameters from $\boldsymbol{\alpha}^{(4)}$ and $\boldsymbol{\beta}^{(4)}$ in order to find a reduced set. Finally we solve the overdetermined linear system. The results are presented in Table 7.5. We use only $7(2N+1)$ or $10(2N+1)$ sampling points for the parameter estimation. □

L	N	ε_0	$\boldsymbol{\alpha}^{(2)}$	$\boldsymbol{\alpha}^{(4)}$	$\boldsymbol{\beta}^{(2)}$	$\boldsymbol{\beta}^{(4)}$	δ	$e(\boldsymbol{f})$	$e(\boldsymbol{c})$	$e(h)$
8	15	10^{-4}	1	$\begin{pmatrix}1 & 1 & 1\end{pmatrix}$	0	$\begin{pmatrix}0 & 0 & 0\end{pmatrix}$	∞	1.7e-10	2.5e-11	1.6e-10
8	15	10^{-4}	1	$\begin{pmatrix}1 & 1 & 1\end{pmatrix}$	1	$\begin{pmatrix}1 & 1 & 1\end{pmatrix}$	∞	1.7e-10	2.4e-11	1.6e-10
15	30	10^{-4}	1	$\begin{pmatrix}1 & 1 & 1\end{pmatrix}$	0	$\begin{pmatrix}0 & 0 & 0\end{pmatrix}$	∞	1.3e-14	6.4e-15	8.8e-14
15	30	10^{-3}	1	$\begin{pmatrix}1 & 1 & 1\end{pmatrix}$	0	$\begin{pmatrix}0 & 0 & 0\end{pmatrix}$	6	1.0e-06	3.2e-07	3.0e-06
15	30	10^{-3}	1	$\begin{pmatrix}1 & 1 & 1\end{pmatrix}$	0	$\begin{pmatrix}0 & 0 & 0\end{pmatrix}$	5	1.3e-05	3.4e-06	4.2e-05
15	30	10^{-3}	$\begin{pmatrix}1 \\ -1\end{pmatrix}$	$\begin{pmatrix}1 & 1 & 1 \\ -1 & 1 & -1\end{pmatrix}$	$\begin{pmatrix}0 \\ 0\end{pmatrix}$	$\begin{pmatrix}0 & 0 & 0 \\ 0 & 0 & 0\end{pmatrix}$	6	1.1e-06	2.7e-07	3.9e-06
15	30	10^{-3}	$\begin{pmatrix}1 \\ -1\end{pmatrix}$	$\begin{pmatrix}1 & 1 & 1 \\ -1 & 1 & -1\end{pmatrix}$	$\begin{pmatrix}0 \\ 0\end{pmatrix}$	$\begin{pmatrix}0 & 0 & 0 \\ 0 & 0 & 0\end{pmatrix}$	5	8.8e-06	1.9e-06	3.3e-05
15	50	10^{-3}	$\begin{pmatrix}1 \\ -1\end{pmatrix}$	$\begin{pmatrix}1 & 1 & 1 \\ -1 & 1 & -1\end{pmatrix}$	$\begin{pmatrix}0 \\ 0\end{pmatrix}$	$\begin{pmatrix}0 & 0 & 0 \\ 0 & 0 & 0\end{pmatrix}$	5	4.5e-07	1.2e-07	1.6e-06
15	50	10^{-3}	$\begin{pmatrix}1 \\ -1\end{pmatrix}$	$\begin{pmatrix}1 & 1 & 1 \\ -1 & 1 & -1\end{pmatrix}$	$\begin{pmatrix}0 \\ 0\end{pmatrix}$	$\begin{pmatrix}0 & 0 & 0 \\ 0 & 0 & 0\end{pmatrix}$	4	8.0e-07	2.4e-07	1.1e-05

TABLE 7.5
Results of Example 7.4.

Acknowledgment. The authors thank Franziska Nestler for the numerical experiments. The first named author gratefully acknowledges the support by the German Research Foundation within the project PO 711/10–1. Furthermore, the authors would like to thank the reviewers for their helpful comments to improve this paper.

REFERENCES

[1] F. Andersson, M. Carlsson, and M.V. de Hoop. Nonlinear approximation of functions in two dimensions by sums of exponential functions. *Appl. Comput. Harmon. Anal.* 29:156 – 181, 2010.

[2] R. F. Bass and K. Gröchenig. Random sampling of multivariate trigonometric polynomials. SIAM *J. Math. Anal.* 36:773 – 795, 2004.

[3] G. Beylkin and L. Monzón. On approximations of functions by exponential sums. *Appl. Comput. Harmon. Anal.* 19:17 – 48, 2005.

[4] D. Braess. Nonlinear Approximation Theory. Springer, Berlin, 1986.

[5] O. Christensen. An Introduction to Frames and Riesz Bases. Birkhäuser, Boston, 2003.

[6] P.L. Dragotti, M. Vetterli, and T. Blu. Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang–Fix. IEEE *Trans. Signal Process.* 55:1741 – 1757, 2007.

[7] G.H. Golub and C.F. Van Loan. Matrix Computations. 3rd ed. Johns Hopkins Univ. Press, Baltimore, 1996.

[8] K. Gröchenig. Reconstruction algorithms in irregular sampling. *Math. Comput.* 59:181 – 194, 1992.

[9] N.J. Higham. Functions of Matrices: Theory and Computation. SIAM, Philadelphia, 2008.

[10] Y. Hua and T.K. Sarkar. Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. IEEE *Trans. Acoust. Speech Signal Process.* 38:814 – 824, 1990.

[11] A.E. Ingham. Some trigonometrical inequalities with applications to the theory of series. *Math. Z.* 41:367 – 379, 1936.

[12] M.A. Iwen. Combinatorial sublinear–time Fourier algorithms. *Found. Comput. Math.*, 10:303–338, 2010.

[13] 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.

[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] V. Komornik and P. Loreti. Fourier Series in Control Theory. Springer, New York, 2005.

[16] V. Komornik and P. Loreti. Semi–discrete Ingham–type inequalities. *Appl. Math. Optim.* 55:203 – 218, 2007.

[17] A.M. Lindner. A universal constant for exponential Riesz sequences. *Z. Anal. Anw.* 19:553 – 559, 2000.

[18] 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.

[19] V. Pereyra and G. Scherer. Exponential data fitting. In V. Pereyra and G. Scherer, editors. *Exponential Data Fitting and its Applications*, pp. 1 – 26. Bentham Science Publ., Sharjah, 2010.

[20] V. Pereyra and G. Scherer. Exponential Data Fitting and its Applications. Bentham Science Publ., Sharjah, 2010.

[21] T. Peter, D. Potts, and M. Tasche. Nonlinear approximation by sums of exponentials and translates. SIAM *J. Sci. Comput.* 33:1920 – 1947, 2011.

[22] D. Potts and M. Tasche. Numerical stability of nonequispaced fast Fourier transforms. *J. Comput. Appl. Math.* 222:655 – 674, 2008.

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

[24] D. Potts and M. Tasche. Parameter estimation for nonincreasing exponential sums by Prony–like methods. Preprint, Chemnitz University of Technology, 2012.

[25] A. Rényi. On projections of probability distributions. *Acta Math. Acad. Sci. Hung.*, 3:131 – 142, 1952.

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

[27] R. Roy and T. Kailath. ESPRIT – estimation of signal parameters via rotational invariance techniques. In L. Auslander, F.A. Grünbaum, J.W. Helton, T. Kailath, P. Khargoneka, S. Mitter, editors. *Signal Processing, Part* II: *Control Theory and Applications*, pp. 369 – 411. Springer, New York, 1990.

[28] J.J. Sacchini, W.M. Steedy and R.L. Moses. Two–dimensional Prony modeling parameter estimation. IEEE *Trans. Signal Process.* 41:3127 – 3137, 1993.

[29] T.K. Sarkar and O. Pereira. Using the matrix pencil method to estimate the parameters of a sum of complex exponentials. IEEE *Antennas and Propagation Magazine* 37:48 – 55, 1995.

[30] P. Shukla and P.L. Dragotti. Sampling schemes for multidimensional signals with finite rate of innovation. IEEE *Trans. Signal Process.* 55:3670 – 3686, 2007.

[31] M. Vetterli, P. Marziliano, and T. Blu. Sampling signals with finite rate of innovation. IEEE *Trans. Signal Process.* 50:1417 – 1428, 2002.

[32] Z. Xu. Deterministic sampling of sparse trigonometric polynomials. *J. Complexity* 27:133 – 140, 2011.