Parameter Estimation for Bivariate Exponential Sums

Benedikt Diederichs and Armin Iske
Department of Mathematics, University of Hamburg, Germany
{benedikt.diederichs,armin.iske}@uni-hamburg.de

***Abstract*—Parameter estimation for exponential sums is a classical problem in signal processing. Recently, a new concept for estimating parameters of bivariate exponential sums has been proposed. The resulting method relies on parameter estimations for univariate exponential sums along several lines in the plane. These (univariate) parameter estimations are being used to first compute the projections of the unknown bivariate frequency vectors onto these lines, before they are combined to obtain estimations for the sought frequency vectors of the bivariate exponential sum. In this paper, we address theoretical questions concerning this new concept, namely (a) how many lines are needed for exact reconstruction, and (b) how to recover linear combinations of shifted positive definite functions.**

I. Introduction

Univariate exponential sums are commonly used model functions in many relevant applications. Often, one wishes to recover their parameters exactly from only a few given samples. This reconstruction problem is well-understood and, in fact, it has a long history. The first reconstruction method was devoloped by Gaspard Riche de Prony as early as 1795 [1]. A variety of numerical algorithms have been developed since then, including ESPRIT [2], APM [3] and matrix pencil methods [4].

A significantly more difficult reconstruction problem is that of parameter estimation for *bivariate* exponential sums, i.e., functions of the form

$$f(\mathbf{x}) = \sum_{j=1}^{M} c_j e^{i\mathbf{y}_j \cdot \mathbf{x}} \qquad \text{for } \mathbf{x} \in \mathbb{R}^2 \tag{1}$$

with pairwise distinct *frequency vectors* $\mathbf{y}_j \in \mathbb{R}^2$ and their corresponding coefficients $c_j \in \mathbb{C}^* := \mathbb{C} \setminus \{0\}$.

Several methods for parameter estimation of bivariate exponential sums rely on gridded data samples [5], [6]. Just very recently, a new concept has been proposed, where only a few sample points are needed [7], [8]. The basic idea in the new approach of [7], [8] is to first apply a univariate parameter estimation along several lines in the plane, before the resulting information are being combined to obtain estimations for the bivariate frequency vectors $\mathbf{y}_j$ in (1).

The outline of this paper is as follows. In Section II, we briefly explain parameter estimation for univariate exponential sums, before we turn to the bivariate case in Section III. In Section III, we analyse theoretical properties for the line-based method of [7], [8]. To this end, we first prove an upper bound for the number of lines being needed for exact reconstruction of exponential sums with fixed order (i.e., length). Moreover, we show that there is no set of finitely many lines for which the bivariate reconstruction problem has a unique solution. We then characterize a bivariate exponential sum f in (1) as the unique solution of a non-convex optimization problem, although the complexity for computing its solution is for practical purposes too large. In Section IV, we show how linear combinations of shifted positive definite functions $\Phi : \mathbb{R}^2 \to \mathbb{R}$ can be recovered from their Fourier data. Supporting numerical examples are finally provided in Section V.

II. Univariate Parameter Estimation

In this section, we give a short introduction to the problem of parameter estimation for univariate exponential sums. To this end, let

$$f(x) = \sum_{j=1}^{M} c_j e^{iy_j x} \qquad \text{for } x \in \mathbb{R} \tag{2}$$

denote a univariate exponential sum of *order* M with pairwise distinct frequencies $y_j \in (-\pi, \pi]$ and coefficients $c_j \in \mathbb{C}^*$. We collect all univariate exponential sums of finite order in the linear space

$$\mathcal{E}_1 := \left\{ \sum_{j=1}^{N} c_j e^{iy_j x} \;\middle|\; c_j \in \mathbb{C}^*, y_j \in (-\pi, \pi], N \in \mathbb{N} \right\}.$$

Parameter estimation on $\mathcal{E}_1$ requires computing, from given $2M$ samples $f(k)$, $k = 0, \ldots, 2M-1$, all frequencies y_j and coefficients c_j of f in (2). Note that there are infinitely many functions $g \in \mathcal{E}_1$ satisfying $g(k) = f(k)$, for all $k = 0, \ldots, 2M-1$, but f is the one with *minimal* order, M. In this sense, we are concerned with a problem of sparse approximation on $\mathcal{E}_1$.

The first method to solve this problem is Prony's method [1], which we briefly explain in this section. To this end, we let $z_k = e^{iy_k}$ and define the *Prony polynomial*

$$P(z) = \prod_{k=1}^{M} (z - z_k) = \sum_{j=0}^{M} p_j z^j \qquad \text{for } z \in \mathbb{C}, \tag{3}$$

where $p_M = 1$. Then, the M coefficients p_j, $j = 0, \ldots, M-1$, satisfy the set of M linear equations

$$\sum_{j=0}^{M} p_j f(j+m) = \sum_{k=1}^{M} c_k e^{imy_k} \sum_{j=0}^{M} p_j z_k^j = 0$$

for $m = 0, ..., M-1$, which can be written in matrix form as

$$\mathbf{H}_M \mathbf{p} = -\mathbf{f} \tag{4}$$

where

$$\begin{aligned} \mathbf{H}_M &= (f(j+k))_{j,k=0,...,M-1} \in \mathbb{C}^{M\times M}, \\ \mathbf{p} &= (p_0, ..., p_{M-1})^T \in \mathbb{C}^M, \\ \mathbf{f} &= (f(M), ..., f(2M-1))^T \in \mathbb{C}^M. \end{aligned}$$

Now, $\mathbf{H}_M$ can be diagonalized by the (regular) Vandermonde matrix $\mathbf{V}(\mathbf{z}) = (z_j^{k-1})_{k,j=1,...,M} \in \mathbb{C}^{M\times M}$,

$$\mathbf{H}_M = \mathbf{V}(\mathbf{z})\,\mathrm{diag}(c_1, ..., c_M)\,\mathbf{V}(\mathbf{z})^T, \tag{5}$$

so that $\mathbf{H}_M$ is regular. Hence, we can solve the system (4) uniquely for the (unknown) Prony coefficients $\mathbf{p}$, and so we get the Prony polynomial P in (3). Finally, from the roots z_k of P we obtain the frequencies y_k, and the coefficients c_j can then be obtained by solving the overdetermined linear system

$$\sum_{j=1}^{M} c_j z_j^k = f(k) \qquad \text{for } k = 0, ..., 2M-1. \tag{6}$$

On the down side, Prony's method is numerically unstable. Moreover, for the performance of the Prony method, the order M of f is required to be known beforehand. Many methods have been developed to improve on the shortcomings of the Prony method, one of which is *ESPRIT*. For the performance of the ESPRIT algorithm, only an upper bound $K \geq M$ of M is needed. Then, $2N \geq 2K$ equispaced samples of f are stored in the rectangular Hankel matrix

$$\mathbf{H}_{2N-K,K+1} = (f(j+k-2))_{\substack{1\leq j\leq 2N-K \\ 1\leq k\leq K+1}} \in \mathbb{R}^{(2N-K)\times(K+1)}.$$

We can factorize $\mathbf{H}_{2N-K,K+1}$ similarly as $\mathbf{H}_M$ in (5), so that the rank of $\mathbf{H}_{2N-K,K+1}$ is equal to the order M of f. In practice, we determine the order numerically by the singular value decomposition of $\mathbf{H}_{2N-K,K+1}$. More precisely, the numerical rank of $\mathbf{H}_{2N-K,K+1}$ is taken as the largest number M satisfying $\varepsilon\sigma_1 < \sigma_M$, where σ_j is the j-th singular value of $\mathbf{H}_{2N-K,K+1}$ and $\varepsilon > 0$ is chosen with respect to the expected noise level.

The frequencies y_j of f in (2) can then be calculated by solving an eigenvalue problem for a matrix computed from the right singular vectors of $\mathbf{H}_{2N-K,K+1}$, before the coefficients c_j in (2) are finally determined by the solution of an overdetermined system similar to that in (6). For more details on the method ESPRIT, we refer to [2].

III. Bivariate Parameter Estimation

Now we turn to parameter estimation for bivariate exponential sums. We collect all bivariate exponential sums (1) of finite order in the linear space

$$\mathcal{E}_2 := \left\{ \sum_{j=1}^{N} c_j e^{i\mathbf{y}_j\cdot\mathbf{x}} \;\middle|\; c_j \in \mathbb{C}^*, \mathbf{y}_j \in \mathbb{R}^2, N \in \mathbb{N} \right\}.$$

By the restriction of any $f \in \mathcal{E}_2$ to a straight line

$$\ell = \{\lambda\mathbf{v} + \beta\boldsymbol{\eta} \mid \lambda \in \mathbb{R}\}, \tag{7}$$

for fixed perpendicular unit vectors $\mathbf{v} \perp \boldsymbol{\eta}$ and $\beta \in \mathbb{R}$, we obtain a univariate exponential sum

$$f|_\ell(\lambda) = \sum_{j=1}^{M} c_j e^{i\beta\mathbf{y}_j\cdot\boldsymbol{\eta}} e^{i\lambda\mathbf{y}_j\cdot\mathbf{v}} = \sum_{j=1}^{M_\ell} d_j e^{iy_j\lambda} \in \mathcal{E}_1. \tag{8}$$

for frequencies $y_j = \mathbf{y}_j \cdot \mathbf{v} \in \mathbb{R}$ and coefficients $d_j \in \mathbb{C}^*$.

Note that $M_\ell \leq M$, since distinct frequencies $\mathbf{y}_j \neq \mathbf{y}_k$ may have the same projection onto ℓ, i.e., $\mathbf{y}_j\cdot\mathbf{v} = \mathbf{y}_k\cdot\mathbf{v}$. We choose M_ℓ, such that $d_j \neq 0$ for all $1 \leq j \leq M_\ell$. Then, we apply the univariate parameter estimation from the previous section to $f|_\ell$ in (8), where we take equispaced samples from $f|_\ell$. This way we obtain all parameters d_j and y_j of $f|_\ell$ in (8).

This line-based parameter estimation can be applied to f on any set $\mathcal{L} = \{\ell_1, ..., \ell_L\}$ of L pairwise non-parallel lines $\ell_1, ..., \ell_L$, yielding L pairs of projected frequencies $\mathbf{y} \equiv \mathbf{y}_{\ell_k} \in \mathbb{R}^{M_{\ell_k}}$ and corresponding coefficient vectors $\mathbf{d} \equiv \mathbf{d}_{\ell_k} \in \mathbb{R}^{M_{\ell_k}}$, i.e., one pair $(\mathbf{y}_{\ell_k}, \mathbf{d}_{\ell_k})$ for each restriction $f|_{\ell_k}$, for $k = 1, \ldots, L$.

Yet it remains to derive practical conditions under which f is uniquely determined. For the purpose of doing so, we can rely on the following useful result by Renyi [9].

Theorem 1: Any set of M points $\mathbf{y}_j \in \mathbb{R}^2$, associated with positive weights $c_j > 0$, is uniquely determined by the point projections onto $M+1$ distinct lines through the origin. □

We can adapt the proof in [9] to obtain a first result concerning the order of $f \in \mathcal{E}_2$. To this end, we define, for any set $\mathcal{L} = \{\ell_1, ..., \ell_L\}$ of L pairwise non-parallel lines ℓ_j the *restriction operator*

$$R_{\mathcal{L}} : \mathcal{E}_2 \to \mathscr{C}(\mathcal{L}, \mathbb{C}),$$

which maps any $f \in \mathcal{E}_2$ onto its restriction to the lines in $\mathcal{L}$.

Theorem 2: For a set $\mathcal{L} = \{\ell_1, ..., \ell_L\}$ of L pairwise non-parallel lines, let $f \in \ker R_{\mathcal{L}} \setminus \{0\}$ be a non-trivial element in the kernel of $R_{\mathcal{L}}$. Then, f is of order at least $2L$.

Proof: If all points are co-linear, they have distinct projections on all but at most one line. Otherwise we find for each line ℓ_j two perpendicular lines, passing through at least two frequency vectors, such that all other frequency vectors lie in the region between these two lines. The intersection of these regions forms a convex polygon, which contains all frequency vectors. This polygon has $2L$ edges, as at least two frequency vectors lie on each line and no line segment is in the interior of the polygon. In particular, at least $2L$ frequency vectors lie on the boundary of the polygon. ■

Corollary 3: Let $f_1, f_2 \in \mathcal{E}_2$ be of order at most M. If f_1 and f_2 coincide on a set $\mathcal{L}$ of $M+1$ pairwise non-parallel lines, $R_{\mathcal{L}}(f_1) \equiv R_{\mathcal{L}}(f_2)$, then they coincide on $\mathbb{R}^2$, $f_1 \equiv f_2$.

Proof: The function $f_1 - f_2 \in \mathcal{E}_2$ is of order at most $2M$ and hence cannot be a non-trivial kernel element of $R_{\mathcal{L}}$. ■

We remark that the estimate for the order of f in Theorem 2 is sharp. For instance, if we choose the vertices of a regular $2L$-gone $\mathcal{P}_{2L}$ as frequency vectors and associate them with alternating coefficients ± 1, then the corresponding exponential sum f vanishes along all lines through the origin which are perpendicular to the edges of $\mathcal{P}_{2L}$.

For a set $\mathcal{L}$ of arbitrarily chosen $2L$ lines, however, it is not clear whether such examples exist. But we can show that the restriction operator $R_{\mathcal{L}}$ has a non-trivial kernel $f \in \mathcal{E}_2$.

Theorem 4: For any finite set $\mathcal{L}$ of pairwise non-parallel lines there exists one non-trivial element $f \in \ker R_{\mathcal{L}}$ with arbitarily small frequency vectors and real coefficients.

Proof: Any line $\ell \in \mathcal{L}$ has the form (7) for some perpendicular unit vectors $\mathbf{v} \perp \boldsymbol{\eta}$ and $\beta \in \mathbb{R}$. Therefore,

$$f_\ell(\mathbf{x}) = e^{i\alpha\mathbf{v}\cdot\mathbf{x}} - e^{i(\alpha\mathbf{v}+\gamma\boldsymbol{\eta})\cdot\mathbf{x}} \in \mathcal{E}_2$$

is zero along ℓ for any $\alpha \in \mathbb{R}$, where $\gamma \in \mathbb{R} \setminus \{0\}$ satisfies $\gamma\beta \in 2\pi\mathbb{Z}$. If $\beta = 0$, we can choose γ arbitrarily small. Otherwise we use the ansatz

$$f_\ell(\mathbf{x}) = c_1 e^{i\alpha\mathbf{v}\cdot\mathbf{x}} + c_2 e^{i(\alpha\mathbf{v}+\gamma_1\boldsymbol{\eta})\cdot\mathbf{x}} + c_3 e^{i(\alpha\mathbf{v}+\gamma_2\boldsymbol{\eta})\cdot\mathbf{x}}$$

with $\gamma_1 \neq \gamma_2 \in \mathbb{R} \setminus \{0\}$. But then

$$f_\ell(\lambda\mathbf{v} + \beta\boldsymbol{\eta}) = e^{i\alpha\lambda}(c_1 + c_2 e^{i\gamma_1\beta} + c_3 e^{i\gamma_2\beta})$$

and so in this case it is always possible to choose non-vanishing real coefficients c_j such that $f_\ell(\lambda\mathbf{v}_1 + \beta\boldsymbol{\eta}_1) = 0$ for all $\lambda \in \mathbb{R}$. Therefore, in either case, we find $f\big|_\ell \equiv 0$ on ℓ. By construction, the product

$$f = \prod_{\ell \in \mathcal{L}} f_\ell \in \mathcal{E}_2$$

is a non-trivial element in $\ker R_{\mathcal{L}}$. ∎

It was conjectured in [8] that – under certain additional assumptions – it is possible to choose only *four* lines passing through the origin to guarantee a unique reconstruction of any $f \in \mathcal{E}_2$. The additional assumptions are in [8] stated as follows.

- All coefficients of f are assumed to be positive.
- The first line ℓ_1 be the x-axis, and ℓ_2 be the y-axis. The other two lines, ℓ_3 and ℓ_4, be perpendicular, $\ell_3 \perp \ell_4$. Moreover, ℓ_3 is assumed to be spanned by the unit vector $(\cos(\alpha), \sin(\alpha))^T$ for $\alpha \in (0, \pi/2)$, where α is required to satisfy $\tan(\alpha) \neq 1/n$ for all $n \in \mathbb{N}$.
- Arbitrarily many samples may be taken along the 4 lines.

We now use the construction in our proof of Theorem 4 to falsify that conjecture in [8]. To this end, let $f \not\equiv 0$ be an exponential sum which is zero along the four preselected lines with real coefficients and sufficently small frequency vectors. We can represent f as a difference

$$f(\mathbf{x}) = f_1(\mathbf{x}) - f_2(\mathbf{x})$$

between two functions f_1 and f_2 with positive coefficients. By construction, f is zero along all of the four lines, i.e., f_1 and f_2 are equal along all four lines. But f_1 and f_2 cannot be equal on $\mathbb{R}^2$, since $f \not\equiv 0$.

While Corollary 3 states that an exponential sum of order M is uniquely determined by its restriction to $M+1$ pairwise non-parallel lines, this does not give a construction to compute f from its samples. But we can characterize f as a solution of a non-convex optimization problem. Our characterization relies on the following Lemma.

Lemma 5: Let f be an exponential sum of order M and let $\mathcal{L} = \{\ell_1, \ldots, \ell_{M+1}\}$ be $M+1$ pairwise non-parallel lines with direction vectors $\mathbf{v}_1, \ldots, \mathbf{v}_{M+1}$. Then, for every frequency vector $\mathbf{y}$ of f, there are at least two distinct lines $\ell_{j_1}, \ell_{j_2} \in \mathcal{L}$, such that the frequency $\mathbf{y}\cdot\mathbf{v}_{\ell_{j_k}}$ appears in the representation (8) of $f|_{\ell_{j_k}} \in \mathcal{E}_1$, with $d_{j_k} \neq 0$, for $k = 1, 2$.

Proof: Assume that $\mathbf{y}\cdot\mathbf{v}_j$ is not a frequency of $f|_{\ell_j}$. Then there must be another frequency vector $\tilde{\mathbf{y}}$ such that $\mathbf{y} - \tilde{\mathbf{y}}$ is orthogonal to $\mathbf{v}_j$. But there are only $M-1$ possible choices for $\tilde{\mathbf{y}}$, since $\mathcal{L}$ contains only $M+1$ lines. ∎

Note that the result of Lemma 5 allows us to find a large set of possible frequency vectors. In fact, we can characterize the frequency vectors of $f \in \mathcal{E}_2$ by a non-convex optimization problem, whose formulation is given in the following theorem.

Theorem 6: Let $f \in \mathcal{E}_2$ be a bivariate exponential sum of order M and let $\mathcal{L} = \{\ell_1, ..., \ell_{M+1}\}$ be a set of be pairwise non-parallel lines with direction vectors $\mathbf{v}_1, ..., \mathbf{v}_{M+1}$. Moreover, let Y_j be the set containing the frequencies of $f|_{\ell_j}$, and let

$$\tilde{Y} := \{\mathbf{y} \in \mathbb{R}^2 \mid \exists j \neq k : \mathbf{y}\cdot\mathbf{v}_j \in Y_j \text{ and } \mathbf{y}\cdot\mathbf{v}_k \in Y_k\}.$$

Let G be a set containing $2M$ equispaced sample points along each of the $M+1$ lines in $\mathcal{L}$. Assume that $\|\mathbf{y}\|_2 < \pi/h_{\max}$ for every frequency vector $\mathbf{y}$ of f, where $h_{\max}$ is the largest stepsize of the equispaced samples taken from the lines in $\mathcal{L}$. Then, the solution of the constrained optimization problem

$$\min_{\mathbf{c}\in\mathbb{R}^{|\tilde{Y}|}} \|\mathbf{c}\|_0 \text{ subject to } \sum_{\mathbf{y}\in\tilde{Y}} \mathbf{c}_{\mathbf{y}} e^{i\mathbf{w}\cdot\mathbf{y}} = f(\mathbf{w}) \text{ for all } \mathbf{w} \in G$$

determines f, where the frequency vectors of f are all those $\mathbf{y} \in \tilde{Y}$ with $\mathbf{c}_{\mathbf{y}} \neq 0$, $\mathbf{c}_{\mathbf{y}}$ are the corresponding coefficients, and $\|\mathbf{c}\|_0$ gives the number of nonzero entries of $\mathbf{c}$.

Proof: By Lemma 5, f corresponds to a coefficient vector $\mathbf{c}(f) \in \mathbb{R}^{|\tilde{Y}|}$ with $\mathbf{c}(f)_{\mathbf{y}} = c_j$ if $\mathbf{y} = \mathbf{y}_j$ and zero otherwise. Hence, every solution corresponds to an exponential sum with at most M summands. The equality constraint ensures that the solution is equal to f along the chosen lines. Applying Corollary 3 yields the claim. ∎

Unfortunately, the minimization problem in Theorem 6 is NP-hard (cf. [10]). Moreover, we cannot expect that the system matrix satisfies the restricted isometric property, which would allow to relax the zero semi-norm to the convex 1-norm. The reason is that frequency vectors may have projections which are close, leading to close points in $\tilde{Y}$. Unfortunately, $|\tilde{Y}|$ is of order M^4, which is too large for practical purposes.

A. Reconstruction Algorithms

But is there a more efficient method to calculate the frequency vectors from the projections onto the lines? It turns out that under a weak assumption there is one, namely the *sparse approximate Prony method* (SAPM) of [7]. Reconstruction by SAPM relies on the assumption, that the projections of all frequency vectors on all lines do not vanish, i.e., for $k = 1, ..., L$ we have

$$\{\mathbf{y}_j \cdot \mathbf{v}_k \mid j = 1, ..., M\} = \{y_j^{(\ell_k)} \mid j = 1, ..., M_{\ell_k}\}.$$

Note that frequency vectors with the same projection on a chosen line are allowed, as long as their coeffcients do not sum up to zero.

Without loss of generality we assume that ℓ_1 is the x-axis and ℓ_2 is the y-axis. We start with the large set of possible frequency vectors

$$Y^{(1)} = \{(y_j^{(\ell_1)}, y_k^{(\ell_2)}) \in \mathbb{R}^2 \mid j = 1, ..., M_{\ell_1},\ k = 1, ..., M_{\ell_2}\}.$$

Our above assumption ensures that all frequency vectors are contained in this set. Then one compares the projection of the points in $Y^{(1)}$ with the projected frequencies $y_j^{(\ell_3)}$, then let

$$Y^{(2)} = \{\mathbf{y} \in Y^{(1)} \mid \exists j \in \{1, \dots, M_3\} : |\mathbf{y}\cdot\mathbf{v}_3 - y_j^{(\ell_3)}| < \varepsilon^{(2)}\},$$

where $M_3 = M_{\ell_3}$ and $\varepsilon^{(2)} > 0$ is an accuracy bound. We then repeat this reduction step for all available lines, obtaining the set $Y^{(L-1)}$. If L is larger than the order of f, then $Y^{(L-1)}$ will be equal to the set of all frequency vectors. Often, it will be sufficient to take significantly less lines to obtain all frequencies of f. Having determined all frequencies of f, we can compute the coefficients c_j. To this end, we compute the coefficients as the least squares solution of the overdetermined linear system

$$\sum_{\mathbf{y}\in Y^{(L-1)}} c_{\mathbf{y}} e^{i\mathbf{y}\cdot\mathbf{x}} = f(\mathbf{x}) \qquad \text{for all } \mathbf{x} \in G.$$

Finally, we remove all frequency vectors from $Y^{(L-1)}$ which are corresponding to small coefficients. This gives, for some $\tilde{\varepsilon} > 0$, the smaller set of frequencies

$$\tilde{Y} = \{\mathbf{y} \in Y^{(L-1)} \mid |c_{\mathbf{y}}| \geq \tilde{\varepsilon}\} \subset Y^{(L-1)}$$

on which we solve the linear system (rather than on $Y^{(L-1)}$) to obtain the corresponding coefficients $c_{\mathbf{y}}$ (cf. [7] for details).

A posteriori chosen lines: Often, SAPM only needs data taken along three lines, namely when the points in $Y^{(1)}$ have pairwise distinct projections on the third line. This observation is used in [8], where the following algorithm is proposed.

(1) Apply ESPRIT along two lines ℓ_1, ℓ_2 and calculate $Y^{(1)}$.
(2) Find a third line, ℓ_3, on which all frequencies in $Y^{(1)}$ have a distinct projection.
(3) Apply ESPRIT along ℓ_3. This gives the frequency vectors and the corresponding coefficients of $f \in \mathcal{E}_2$.

We remark that the above reconstruction scheme relies on the same assumption as SAPM, since it uses data taken on the first two lines to find a set which contains the frequency vectors.

IV. Reconstruction from Fourier Data

In this section, we show how the reconstruction method of the previous section can be used to reconstruct linear combinations of shifted basis functions from their Fourier data (see [8]). Here, we define the (continuous) Fourier transform $\hat{f}$ of a function $f \in L^1(\mathbb{R}^2)$ by

$$\hat{f}(\mathbf{w}) = \frac{1}{2\pi}\int_{\mathbb{R}^d} f(\mathbf{x}) e^{i\mathbf{x}\cdot\mathbf{w}} d\mathbf{x} \qquad \text{for } \mathbf{w} \in \mathbb{R}^2.$$

For an even function $\Phi : \mathbb{R}^2 \to \mathbb{R}$, we consider the model

$$f(\mathbf{x}) = \sum_{j=1}^{M} c_j \Phi(\mathbf{x} - \mathbf{x}_j). \tag{9}$$

For $\Phi \in L^1(\mathbb{R}^2)$, we get the Fourier transform of f by

$$\hat{f}(\mathbf{w}) = \hat{\Phi}(\mathbf{w}) \sum_{j=1}^{M} c_j e^{i\mathbf{w}\cdot\mathbf{x}_j}.$$

Sampling $\hat{f}$ at a finite point set G leads us to a reconstruction problem for bivariate exponential sums, provided that $\hat{\Phi} \not\equiv 0$ on G. As Φ is assumed to be even, $\hat{\Phi}$ is real-valued. Therefore, to allow all possible choices for G, we require $\hat{\Phi}$ to be positive (or negative) on $\mathbb{R}^2$. By Bochner's theorem, the condition

$$\hat{\Phi}(\mathbf{w}) > 0 \qquad \text{for all } \mathbf{w} \in \mathbb{R}^2$$

guarantees Φ to be positive definite. Positive definite functions are an important tool in approximation theory. Prototypical examples for positive definite functions are the *Gaussians* $\Phi(\mathbf{x}) = e^{-\alpha\|\mathbf{x}\|_2^2}$, for $\alpha > 0$, whose Fourier transform is

$$\hat{\Phi}(\mathbf{w}) = \frac{1}{2\alpha} e^{-\|\mathbf{w}\|_2^2/(4\alpha)} > 0.$$

Other examples are the *inverse multiquadrics*

$$\Phi(\mathbf{x}) = \left(1 + \|\mathbf{x}\|_2^2\right)^{\beta} \qquad \text{for } -2 < \beta < 0.$$

We finally summarize our proposed reconstruction method for model functions of the form (9) briefly as follows.

(1) Take equispaced samples from $\hat{f}$ on enough lines.
(2) Calculate

$$g(\mathbf{w}) = \frac{\hat{f}(\mathbf{w})}{\hat{\Phi}(\mathbf{w})} = \sum_{j=1}^{M} c_j e^{i\mathbf{w}\cdot\mathbf{x}_j}$$

for all sample points.
(3) Use SAPM to reconstruct g, and so obtain the shift vectors $\mathbf{x}_j \in \mathbb{R}^2$ and coefficients $c_j \in \mathbb{R}$ of f in (9).

Let us finally make one remark concerning the stability of the proposed reconstruction scheme. As we divide by $\hat{\Phi}(\mathbf{w})$, we require $\hat{\Phi}$ to be uniformly bounded away from zero, i.e.,

$$\hat{\Phi}(\mathbf{w}) > C > 0 \tag{10}$$

for some sufficently large constant C. Otherwise noise gets overamplified. Due to the Riemann-Lebesgue lemma, we have

$$\hat{\Phi}(\mathbf{w}) \to 0 \quad \text{for } \mathbf{w} \to \infty,$$

for $\Phi \in L^1(\mathbb{R}^2)$, and so (10) can only hold on a bounded set.

V. Numerical Examples

For the purpose of illustration, we provide two numerical examples, one for the reconstruction by SAPM (as proposed in Section III) and another one for the reconstruction from Fourier data (as proposed in Section IV).

TABLE I
RESULTS OF THE FIRST EXAMPLE

N	K	δ	ε	$e(\mathbf{y})$	$e(\mathbf{c})$	$e(f)$
5	5	∞	1e-7	3.06e-12	2.25e-13	7.75e-13
20	10	∞	1e-7	3.28e-15	1.11e-15	3.35e-15
5	5	8	1e-7	2.31e-05	1.71e-06	6.31e-06
20	10	8	1e-7	1.31e-09	2.49e-10	4.00e-10
5	5	6	1e-5	-	-	0.30
10	10	6	1e-5	7.53e-06	5.56e-07	1.90e-06
20	10	6	1e-5	2.35e-07	1.39e-08	7.58e-08
40	20	6	1e-5	1.07e-08	1.36e-08	1.50e-08
10	10	4	1e-3	-	-	0.23
20	10	4	1e-3	1.10e-05	2.56e-06	5.01e-06
40	20	4	1e-3	1.10e-06	1.38e-06	1.55e-06

TABLE II
RESULTS OF THE SECOND EXAMPLE

N	K	δ	ε	$e(\mathbf{y})$	$e(\mathbf{c})$	$e(f)$
5	5	∞	1e-7	5.08e-12	3.17e-13	8.36e-13
20	10	∞	1e-7	1.44e-15	1.07e-15	5.15e-15
5	5	8	1e-7	1.20e-03	8.43e-05	2.17e-04
20	10	8	1e-7	5.63e-06	4.21e-07	8.36e-07
5	5	6	1e-5	-	-	> 1
10	10	6	1e-5	4.56e-04	3.32e-05	7.41e-05
20	10	6	1e-5	3.47e-05	4.04e-06	1.11e-05
40	20	6	1e-5	-	-	> 1
10	10	4	1e-3	-	-	> 1
20	10	4	1e-3	3.30e-03	3.72e-04	2.80e-03
30	10	4	1e-3	-	-	> 1

First Example: Regard frequency vectors and coefficients

$$\begin{pmatrix}\mathbf{y}_1 & \mathbf{y}_2 & \mathbf{y}_3 & \mathbf{y}_4 & \mathbf{y}_5\end{pmatrix} = \begin{pmatrix} 0 & 2 & 2 & 0.5 & 1 \\ 0 & 1 & 2 & 1 & 2.5 \end{pmatrix}$$
$$\mathbf{c} = \begin{pmatrix}-2 & 5 & 1.7 & -0.2 & 3.3\end{pmatrix}$$

We let ℓ_1 be the x-axis and ℓ_2 be the y-axis. Moreover, we let $\ell_3 = \{\lambda(1/2, \sqrt{3}/2)^T \mid \lambda \in \mathbb{R}\}$ and $h = 0.5$ for the sampling size. For univariate parameter estimation, we use ESPRIT. In our numerical experiments, we recorded the relative errors

$$e(\mathbf{y}) = \frac{\max_j |\mathbf{y}_j - \tilde{\mathbf{y}}_j|}{\max_j |\mathbf{y}_j|} \quad \text{and} \quad e(\mathbf{c}) = \frac{\max_j |\mathbf{c}_j - \tilde{\mathbf{c}}_j|}{\max_j |\mathbf{c}_j|}$$

for the frequency and coefficient vectors, and

$$e(f) = \frac{\max_{[0,4]^2} |f - \tilde{f}|}{\max_{[0,4]^2} |f|}$$

to measure the relative error of the reconstruction $\tilde{f}$. Our numerical results are in Table I, where N is the number of samples taken along each line, K is an upper bound for the order of f, and ε is the parameter used in ESPRIT to determine the rank of the Hankel matrix and so the order of f.

Moreover, we considered choosing $\varepsilon^{(2)} = \hat{\varepsilon} = 10^{-3}$ (cf. the definitions of $Y^{(2)}, \tilde{Y}$ in Section III). Furthermore, we have added noise, uniformly distributed in $[10^{-\delta}, 10^{\delta}]$. At the absense of noise we let $\delta = \infty$. All our numerical results (as shown in Table I) are averaged values over 50 runs.

Our numerical results of Table I support the good performance of the algorithm SAPM in [7]. Given enough samples, the algorithm is stable with respect to noise. Often, the error is even smaller than the added noise.

Second Example: We consider using the same frequency and coefficient vectors as in our first example. We take samples from a sum of shifted Gaussians

$$f(\mathbf{x}) = \sum_{j=1}^{5} c_j \Phi(\mathbf{x} - \mathbf{y}_j),$$

where $\Phi(\mathbf{x}) = e^{-25\|\mathbf{x}\|_2^2}$, and so $\hat{\Phi}(\mathbf{w}) = \frac{1}{50} e^{-\|\mathbf{x}\|_2^2/100}$.

Our numerical results are summarized in Table II. We see that for any sampling point of modulus greater than ten, the error is overamplified. The error is then magnified by factor $50e^{\|\mathbf{w}\|_2^2/100}$, i.e., any sample taken at $\|\mathbf{w}\|_2 > 10$ is very sensitive w.r.t. noise. In such cases, a larger set of samples may lead to even worser reconstructions. Nevertheless, we believe that our numerical results in Table II are quite promising.

VI. CONCLUSION

We have investigated the reconstruction problem for bivariate exponential sums f from samples taken along a few lines. Samples on at least $M + 1$ lines are needed to guarantee unique reconstruction for any f of order M. The reconstruction can be characterized by a non-convex optimization problem. Under rather mild assumptions on f, an efficient reconstruction method is discussed. A method to recover a sum of shifted positive definite functions from Fourier data is finally explained.

REFERENCES

[1] B. G. R. de Prony, "Essai éxperimental et analytique: sur les lois de la dilatabilité de fluides élastique et sur celles de la force expansive de la vapeur de lalkool,a différentes températures," *Journal de lécole polytechnique*, vol. 1, no. 22, pp. 24–76, 1795.

[2] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 37, no. 7, pp. 984–995, 1989.

[3] D. Potts and M. Tasche, "Parameter estimation for exponential sums by approximate prony method," *Signal Processing*, vol. 90, no. 5, pp. 1631–1642, 2010.

[4] T. K. Sarkar and O. Pereira, "Using the matrix pencil method to estimate the parameters of a sum of complex exponentials," *Antennas and Propagation Magazine, IEEE*, vol. 37, no. 1, pp. 48–55, 1995.

[5] Y. Hua, "Estimating two-dimensional frequencies by matrix enhancement and matrix pencil," *Signal Processing, IEEE Transactions on*, vol. 40, no. 9, pp. 2267–2280, 1992.

[6] S. Rouquette and M. Najim, "Estimation of frequencies and damping factors by two-dimensional ESPRIT type methods," *Signal Processing, IEEE Transactions on*, vol. 49, no. 1, pp. 237–245, 2001.

[7] D. Potts and M. Tasche, "Parameter estimation for multivariate exponential sums," *Electronic Transactions on Numerical Analysis*, vol. 40, pp. 204–224, 2013.

[8] G. Plonka and M. Wischerhoff, "How many Fourier samples are needed for real function reconstruction?" *Journal of Applied Mathematics and Computing*, vol. 42, no. 1-2, pp. 117–137, 2013.

[9] A. Renyi, "On projections of probability distributions," *Acta Mathematica Hungarica*, vol. 3, no. 3, pp. 131–142, 1952.

[10] S. Foucart and H. Rauhut, *A mathematical introduction to compressive sensing*. Springer, 2013.