Robust and tractable multidimensional exponential analysis

H. N. Mhaskar,[*] S. Kitimoon[†] and Raghu G. Raj,[‡] *Senior Member IEEE*

April 18, 2024

Abstract

Motivated by a number of applications in signal processing, we study the following question. Given samples of a multidimensional signal of the form

$$f(\boldsymbol{\ell}) = \sum_{k=1}^{K} a_k \exp(-i\langle \boldsymbol{\ell}, \mathbf{w}_k \rangle), \qquad \mathbf{w}_1, \cdots, \mathbf{w}_k \in \mathbb{R}^q, \boldsymbol{\ell} \in \mathbb{Z}^q, \ |\boldsymbol{\ell}| < n,$$

determine the values of the number K of components, and the parameters a_k and $\mathbf{w}_k$'s. We develop an algorithm to recuperate these quantities accurately using only a subsample of size $\mathcal{O}(qn)$ of this data. For this purpose, we use a novel localized kernel method to identify the parameters, including the number K of signals. Our method is easy to implement, and is shown to be stable under a very low SNR range. We demonstrate the effectiveness of our resulting algorithm using 2 and 3 dimensional examples from the literature, and show substantial improvements over state-of-the-art techniques including Prony based, MUSIC and ESPRIT approaches.

Keywords: Exponential sums, localized kernels, digital signal separation.

1 Introduction

Multidimensional exponential analysis is a core problem in signal processing that appears in various applications such as tomographic imaging (including Computerized Tomography (CT), magnetic resonance imaging (MRI), radar and sonar imaging), wireless communication, antenna array processing, sensor networks, and automotive radar, among others. Mathematically, the problem can be formulated as follows. Given a multidimensional signal of the form

$$f(\mathbf{x}) = \sum_{k=1}^{K} a_k \exp(-i\langle \mathbf{x}, \mathbf{w}_k \rangle), \qquad \mathbf{x}, \mathbf{w}_1, \cdots, \mathbf{w}_k \in \mathbb{R}^q, \tag{1.1}$$

find the number K of components, and the parameters a_k and $\mathbf{w}_k$'s. Of course, this is a problem of inverse Fourier transform if we could observe the function f at **all** values of $\mathbf{x}$. In practice, however, one can observe (after some sampling and renaming of the variables) the values of f at only **finitely many** multi-integer values of $\mathbf{x}$. In this case, it is not possible to distinguish values of $\mathbf{w}_k$ which are equal modulo 2π in all variables. So, this is a special case of the ancient trigonometric moment problem [18], except that we do not have **all** the trigonometric moments (i.e., the samples $f(\boldsymbol{\ell})$) for all values of $\boldsymbol{\ell} \in \mathbb{Z}^q$. Thus, the problem is the ill-posed problem known often as the super-resolution problem: knowing the information in a finite domain of the frequency space, we need to extend it to the entire frequency space. The important problem in this connection is to determine the relationship between the number of samples $f(\boldsymbol{\ell})$ needed to recuperate the desired quantities up to a given accuracy.

Over the past several decades, several approaches to exponential analysis have been investigated. These can be categorized in three main groups–Fourier based, Prony based, and subspace based methods.

- **Fourier-based methods:** In order to utilize Fourier-based methods, it is necessary to obtain a substantial and densely sampled dataset in either a two-dimensional (2D) or three-dimensional (3D) format. However, the collection of such a dataset can be time-consuming. Additionally, these techniques face a trade-off between time and frequency resolution, making it challenging to distinguish closely located scatterers, as pointed out in reference [5].

- **Prony's methods:** Prony's spectral estimation or exponential analysis algorithms have attracted the interest of numerous researchers. In reference [14], the authors assert that these methods exhibit significantly higher accuracy compared to Fourier-based approaches. However, it is important to note that the effectiveness of exponential analysis techniques can be substantially compromised when dealing with a low signal-to-noise ratio (SNR), resulting in the misclassification of noise as actual signals. There are many recent efforts to stabilize the Prony method by taking more than the minimal number of samples. The number of samples required to recuperate the $\mathbf{w}_k$'s up to an accuracy of $\mathcal{O}(1/n)$ is typically on the order of $O(n^q)$ [4, 8, 11, 13], $O(2^q n)$ [16], $O(n \log^{q-1} n)$ [5], or at most $(q+1)n^2 \log^{2q-2} n$ [17].

- **Subspace methods:** Under the rubric of target estimation or localization, which is one of the fundamental problems in radar signal processing with many civilian and military applications including landmine detection and geolocations (cf. [20]), a broad set of techniques has emerged for solving exponential analysis problems. Most of these methods are categorized broadly as subspace methods [7], and are based on statistical considerations, rather than the nature of the signal itself. In fact, quite a few papers, e.g., [19, 15], are interested in testing a statistical hypothesis on whether or not there exists a signal at all. There is a theoretical limit on how much SNR can be tolerated, depending upon the number of antenna elements and number of observations [19], in spite of a huge computational cost. On the other hand, beamforming methods [7] take into account the nature of the signal but focus again on noise, and try to maximize the SNR.

In [3, 2], the authors proposed a method to solve the problem using a combination of Prony and subspece based methods, so that number of sample required is $\mathcal{O}(n)$. Our paper develops these ideas further to develop an algorithm that also utilizes $\mathcal{O}(qn)$ samples, but is far more robust under noise. Our method is based on localized trigonometric polynomial kernels developed in [9]. In contrast to the subspace based methods, our method takes into account the nature of the signal, resulting in a significant noise reduction with only a small number of observations per signal, and yields accurate results with theoretical guarantees. This is a completely different viewpoint from what we have seen in the literature, and therefore, not entirely comparable with other approaches.

The rest of this paper is organized as follows. Section 2 introduces the system model for tomographic imaging which illustrates how the problem of multidimensional exponential analysis arises in signal processing. In Section 3, we introduce the concept of localized trigonometric kernels, and provide the neccessary probabilistic background to formulate the main results of our paper which we state and prove our main results in Section 4. The algorithms to implement these theorems are given in Section 5, and demonstrated in the case of the three examples explored in [3].

2   System model

In this section, we explain how the problem of multidimensional exponential analysis arises in tomographic imaging.

In tomographic imaging, an object of interest being imaged is probed by a sequence of monochromatic tones is swept through a frequency range $[\Omega_{init}, \Omega_{fin}]$. The sensor transmits a signal onto a scene with respect to various angles $\{\boldsymbol{\theta}_m = (\theta_m, \phi_m)\}_m$, where $\theta_m \in [\Theta_{init}, \Theta_{fin}]$ and $\phi_m \in [\Phi_{init}, \Phi_{fin}]$. The scene itself is modeled as a distribution, μ, which is a linear combination of weighted and spatially shifted Dirac delta functons of dimensionality q. For the case of 3D imaging, $q = 3$; i.e., μ is supported on a cube–consisting of points $r = [r_1, r_2, r_3]^T$, for $1 \le r_1 \le D_1$, $1 \le r_2 \le D_2$, and $1 \le r_3 \le D_3$–whose side-lengths ($D1$, $D2$, and $D3$) are referred to as the dimensions of the scene reflectivity function.

We define the center reference point (CRP) to be the center of mass of the scene to be reconstructed, and the line of sight (los) as the unit vector that points from the transmitter to the CRP of the scene. The distance along the los from the transmitter to the range-bin, $\mathbf{r}_0$, of interest is called the 'downrange' $r_0 = \|\mathbf{r}_0\|$. Given this, it can be shown that the backscattered signal at downrange $\mathbf{r}_0$ from the sensor, when viewed at angle $\boldsymbol{\theta}$, is given by

$$\Upsilon(t)|_{\boldsymbol{\theta}, \mathbf{r}_0} = [R_{\boldsymbol{\theta}} \{\mu\} (r_0)]\chi \left(t - \frac{2r_0}{\nu_p} \right) + \aleph(t), \tag{2.1}$$

where $2r_0/\nu_p$ denotes the two-way time delay, ν_p represents the speed of wave propagation, $\aleph(t)$ is the measurement noise, and $R_{\boldsymbol{\theta}}$ is the Radon transform of the scene, μ, with respect to angle $\boldsymbol{\theta}$, and evaluated at the downrange

location r_0. This corresponds to the integral across sensor returns from all points along a hyperplane perpendicular to the downrange location r_0, commonly referred to as an 'iso-range contour'.

Therefore the complete response at time t from all ranges, along the line formed by intersecting the scene at $\boldsymbol{\theta}$, is rewritten as

$$\Upsilon_{\boldsymbol{\theta}}(t) = \int [R_{\boldsymbol{\theta}}\{\mu\}(r)]\chi\left(t - \frac{2r}{\nu_p}\right) dr + \aleph(t). \tag{2.2}$$

This can be reformulated in convolution form (after a choice of units, without loss of generality, so that $\nu_p = 2$) as:

$$\Upsilon_{\boldsymbol{\theta}}(t) = (R_{\boldsymbol{\theta}}(\mu) * \chi)(t) + \aleph(t), \tag{2.3}$$

where $*$ denotes the convolution operation. Equation (2.3) can be interpreted as the response to an Linear Time Invariant (LTI) system with an input signal $\chi(t)$ and an impulse response $\mu_{\boldsymbol{\theta}}(t)$ which characterizes the interaction between the transmit waveform and the scene with respect to sensing angle $\boldsymbol{\theta}$. The received signal $\Upsilon_{\boldsymbol{\theta}}(t)$, is the output of this LTI system and is subsequently sampled at the receiver. Given $\Upsilon_{\boldsymbol{\theta}}(t)$ for a finite grid $\boldsymbol{\theta} \in \{\boldsymbol{\theta}_m\}$ as described earlier, our problem is to estimate the underlying scene reflectivity function μ i.e. to form the image. In this paper we consider the basic case where $\chi(t) = \delta(t)$, where $\delta(t)$ is the Dirac delta function. In this case, the received signal at angle $\boldsymbol{\theta}$ is given by

$$\Upsilon_{\boldsymbol{\theta}}(t) = R_{\boldsymbol{\theta}}(\mu)(t) + \aleph(t). \tag{2.4}$$

Taking the Fourier transform, we obtain

$$\widehat{\Upsilon_{\boldsymbol{\theta}}}(x) = \widehat{R_{\boldsymbol{\theta}}(\mu)}(x) + \hat{\aleph}(x). \tag{2.5}$$

When $\boldsymbol{\theta} = (\theta, \phi)$, the Fourier projection slice theorem implies that the one dimensional Fourier transform $\widehat{R_{\boldsymbol{\theta}}(\mu)}(x)$ is given by $\mathfrak{F}(\mu)(x\cos(\theta)\cos(\phi), x\sin(\theta)cos(\phi), x\sin(\phi))$, where $\mathfrak{F}$ denotes three dimensional Fourier transform. Writing $\mathbf{x} = (x\cos(\theta)\cos(\phi), x\sin(\theta)cos(\phi), x\sin(\phi))$, equation (2.5) becomes

$$\widehat{\Upsilon_{\boldsymbol{\theta}}}(x) = \mathfrak{F}(\mu)(\mathbf{x}) + \hat{\aleph}(x). \tag{2.6}$$

Of course, as mentioned earlier, one deals with only a finite sample of x's and $\boldsymbol{\theta}$'s which are observed by the system. In tomographic imaging, this finite sample is typically chosen from a spherical grid. Since the goal is to find μ from (2.6), it is natural to use the inverse Fourier transform. For this reason, one has to approximate from these samples, samples at a Cartesian grid. This leads to the multidimensional exponential analysis problem described in Section 1. In this paper we present a novel approach for solving the multidimensional analysis problem by resampling the Fourier space $\mathfrak{F}$ in a computationally efficient manner that allows for accurate reconstruction in $O(\sqrt{(D_1 D_2 D_3)})$ samples, while showing substantial improvements over state-of-the-art techniques including Prony based, MUSIC and ESPRIT approaches.

It is important to note that this paper offers a general efficient tool for solving multidimensional exponential problems with applications beyond tomographic imaging such as signal source separation and direction-of-arrival estimation in multichannel radar systems.

3 Theoretical background

In this section, we consider the following univariate set up. Let $\mathbb{T} = \mathbb{R}/(2\pi\mathbb{Z})$, $|x| = |x \bmod 2\pi|$ for $x \in \mathbb{T}$. Let $K \geq 1$ be an integer, $\lambda_k \in \mathbb{T}$, $k = 1, \cdots, K$, $A_k \in \mathbb{C}$ for $k = 1, \cdots, K$. We define

$$\mu = \sum_{k=1}^{K} A_k \delta_{\lambda_k}, \quad \hat{\mu}(\ell) = \sum_{k=1}^{K} A_k \exp(-i\ell\lambda_k), \qquad \ell \in \mathbb{Z}, \tag{3.1}$$

where $i = \sqrt{-1}$, δ_λ denotes the Dirac delta supported at λ. In this section, we propose a solution to the following problem:

Point source separation problem.

Given finitely many noisy samples

$$\tilde{\mu}(\ell) = \hat{\mu}(\ell) + \epsilon_\ell, \qquad |\ell| < N, \tag{3.2}$$

where $N \geq 1$ is an integer, and ϵ_j are realizations of a sub-Gaussian random variable, determine K, A_k, and λ_k, $k = 1, \cdots, K$.

The harder part of the problem is to determine K and the λ_k's. The coefficients A_k can then be determined by solving a linear system of equations. Many algorithms to do this are known, e.g., [12]. Therefore, we will focus in this paper on the task for finding K and the λ_k's.

Our main idea is to use a low pass filter and the corresponding localized kernel Φ_n to be defined in Section 3.2 to observe that

$$\sigma_n(\hat{\mu})(x) = \hbar_n \sum_{|\ell| < n} H\left(\frac{|\ell|}{n}\right) \hat{\mu}(\ell) \exp(i\ell x) = \sum_{k=1}^{K} A_k \Phi_n(x - \lambda_k) \approx \mu(x). \tag{3.3}$$

Thus, the "prominent" peaks of $|\sigma_n(\hat{\mu})(x)|$ will occur at (or close to) the points λ_k and the value of σ_n at these points lead to the corresponding complex amplitudes A_k. It turns out that the values of A_k are actually determined fairly accurately simply by evaluating $\sigma_n(\mu)(x)$ at the estimated value of λ_k

The main difficulty is to make this more precise, and quantify the approximation error in terms of n and the properties of the noise.

In Section 3.1, we review certain properties of sub-Gaussian random varaibles. In Section 3.2, we introduce certain localized kernels and their properties which form the main ingredient in our construction. The main results are stated and proved in Section 4.

3.1   Probabilistic background

The material in this section is based on [1, Section 2.3]), with a slight change of notation. A mean zero real valued random variable X is called sub-Gaussian with parameter V ($X \in \mathcal{G}(V)$) if $\log \mathbb{E}(\exp(tX)) \leq (tV)^2/2$. Examples include Gaussian variables and all bounded random variables. For a sub-Gaussian variable, it is proved in [1, Section 2.3] that

$$\mathsf{Prob}(|X| > t) \leq 2\exp(-t^2/(2V^2)).$$

We will say that a complex valued random variable X is in $\mathcal{G}(V)$ if both the real and imaginary parts of X are in $\mathcal{G}(V)$. We observe that if $z \in \mathbb{C}$ and $|z| > t$ then $\max(|\Re ez|, |\Im mz|) > t/\sqrt{2}$. So, for such variables, we have

$$\mathsf{Prob}(|X| > t) \leq 4\exp(-t^2/(4V^2)). \tag{3.4}$$

It is not difficult to see that if $X_1, \cdots, X_n$ are i.i.d., complex valued variables all in $\mathcal{G}(V)$, $\mathbf{a} = (a_1, \cdots, a_n) \in \mathbb{R}^n$, $|\mathbf{a}|_n^2 = \sum_{\ell=1}^{n} a_\ell^2$, then $\sum_{\ell=1}^{n} a_\ell X_\ell \in \mathcal{G}(|\mathbf{a}|_n V)$. Therefore, (3.4) implies

$$\mathsf{Prob}\left(\left|\sum_{\ell=1}^{n} a_\ell X_\ell\right| > t\right) \leq 4\exp\left(-\frac{t^2}{4|\mathbf{a}|_n^2 V^2}\right), \qquad t > 0. \tag{3.5}$$

3.2   Localized kernels

A *low pass filter* is an infinitely differentiable function $H : \mathbb{R} \to [0, 1]$ such that $H(t) = H(-t)$ for all $t \in \mathbb{R}$, $H(t) = 1$ if $|t| \leq 1/2$, and $H(t) = 0$ if $|t| \geq 1$. In this paper, we fix a low pass filter. All constants may depend upon this filter, and the filter will be omitted from the notation.

We define

$$\hbar_n = \left\{\sum_{|\ell| < n} H\left(\frac{|\ell|}{n}\right)\right\}^{-1}, \qquad \Phi_n(x) = \hbar_n \sum_{k \in \mathbb{Z}} H\left(\frac{\ell}{n}\right) e^{i\ell x}, \qquad x \in \mathbb{T}, \ n > 0. \tag{3.6}$$

An important property of Φ_n is the following localization estimate: For every $S \geq 2$, there exists $L = L(H, S) > 0$ such that

$$|\Phi_n(x)| \leq \frac{L}{\max(1, (n|x|)^S)}, \qquad x \in \mathbb{T}, \ n > 0. \tag{3.7}$$

Explicit expressions for L in terms of H and S are given in [6]. The estimate (3.7) implies that $\Phi_n(x)$ is an approximation to the Dirac delta δ_0.

Constant convention

The letters $c, c_1, \cdots$ will denote generic positive constants depending on H and S alone. Their values might be different at different occurrences within single formula. The notation $A \lesssim B$ means $A \leq cB$, $A \gtrsim B$ means $B \lesssim A$, and $A \sim B$ means $A \lesssim B \lesssim A$. Constants denoted by capital letters, such as L, C, etc. will retain their values.

4 Main result

We assume the notation in (3.2), and make further notation as follows. Let

$$M = \sum_{k=1}^{K} |A_k|, \quad \mathfrak{m} = \min_{1 \leq k \leq K} |A_k|, \quad \eta = \min_{\ell \neq k} |\lambda_k - \lambda_\ell|, \tag{4.8}$$

We further assume that each ϵ_ℓ is a realization of a sub-Gaussian random variable in $\mathcal{G}(V)$.

We will write

$$\sigma_n(x) = \sigma_m(\tilde{\mu})(x) = \hbar_n \sum_{|\ell| < n} H\left(\frac{|\ell|}{n}\right) \tilde{\mu}(\ell) \exp(i\ell x), \qquad x \in \mathbb{T}. \tag{4.9}$$

Writing

$$E_n(x) = \sigma_n(\{\epsilon_\ell\})(x) = \hbar_n \sum_{|\ell| < n} H\left(\frac{|\ell|}{n}\right) \epsilon_\ell \exp(i\ell x), \qquad x \in \mathbb{T}, \tag{4.10}$$

we observe that

$$\sigma_n(\tilde{\mu})(x) = \sigma_n(\hat{\mu})(x) + E_n(x) = \sum_{k=1}^{K} A_k \Phi_n(x - \lambda_k) + E_n(x). \tag{4.11}$$

With this set up, our main theorem concerning the recuperation of point sources can be stated as follows.

Theorem 4.1 *Let*

$$\mathbb{G} = \{x \in [-\pi, \pi] : |\sigma_n(x)| \geq \mathfrak{m}/2\}. \tag{4.12}$$

and (cf. (3.7), (4.8))

$$C = \max\left(1, \left(\frac{16ML}{\mathfrak{m}}\right)^{1/S}\right). \tag{4.13}$$

For sufficiently large n (cf. (4.26)), each of the following statements holds with probability exceeding $1 - \delta$.

- *(**Disjoint union condition**) the set $\mathbb{G}$ is a disjoint union of exactly K subsets $\mathbb{G}_\ell$,*

- *(**Diameter condition**) for each $\ell = 1, \cdots, K$, $\mathsf{diam}(\mathbb{G}_\ell) \leq 2C/n$,*

- *(**Separtion**) $\mathsf{dist}(\mathbb{G}_\ell, \mathbb{G}_k) \geq \eta/2$ for $\ell \neq k$,*

- *(**Interval inclusion**) For each $\ell = 1, \cdots, K$, $I_\ell = \{x \in \mathbb{T} : |x - \lambda_\ell| \leq 1/(4n)\} \subseteq \mathbb{G}_\ell$.*

Moreover, if

$$\hat{\lambda}_\ell = \arg\max_{x \in \mathbb{G}_\ell} |\sigma_n(x)|, \tag{4.14}$$

then

$$|\hat{\lambda}_\ell - \lambda_\ell| \leq 2C/n. \tag{4.15}$$

The following inequality, known as the *Bernstein inequality*, plays an important role in our proofs. We recall that a trigonometric polynomial of order $< n$ is a function of the form $x \mapsto \sum_{|\ell| < n} b_\ell \exp(i\ell x)$.

Proposition 4.1 *Let $n \geq 1$ be an integer, T be any trigonometric polynomial of order $< n$. Then*

$$\max_{x \in \mathbb{T}} |T'(x)| \leq n \max_{x \in \mathbb{T}} |T(x)|. \tag{4.16}$$

In particular, if $N \geq 4\pi n$, then

$$\frac{1}{2} \max_{x \in \mathbb{T}} |T(x)| \leq \max_{0 \leq k \leq N} \left| T\left(\frac{2\pi k}{N}\right) \right| \leq \max_{x \in \mathbb{T}} |T(x)|. \tag{4.17}$$

For the proof of Theorem 4.1, we first estimate $E_n(x)$.

Lemma 4.1 *Let $\delta \in (0,1)$. There exist positive constants C_1, C_2, C_3, depending only on H such that for $n \geq C_1 (\geq 1)$, we have*

$$\mathsf{Prob}\left(\max_{x \in \mathbb{T}} |E_n(x)| \geq C_2 V \sqrt{\frac{\log(C_3 n/\delta)}{n}}\right) \leq \delta. \tag{4.18}$$

PROOF.

Let $x \in \mathbb{T}$. We will use (3.5) with $a_\ell = \hbar_n H(|\ell|/n) \exp(i\ell x)$, $|\ell| < n$. In view of the Euler-Mclaurin summation formula [10, Formula (2.01), p. 285], it is not difficult to show that for $n \geq C_1$,

$$n \int_{-1}^1 H(t)^2 dt \sim |\mathbf{a}|_n^2 = \sum_{|\ell| < n} H\left(\frac{|\ell|}{n}\right)^2, \qquad \hbar_n^{-1} \sim n; \tag{4.19}$$

i.e.,

$$|\mathbf{a}|_n^2 \sim 1/n. \tag{4.20}$$

Hence, (3.5) shows that

$$\mathsf{Prob}\left(|E_n(x)| > t\right) \leq 4 \exp\left(-c\frac{nt^2}{V^2}\right). \tag{4.21}$$

Applying this inequality for each $x = k/2n$, $k = 0, \cdots, \lceil 4\pi n \rceil - 1$, we see that

$$\mathsf{Prob}\left(\max_{0 \leq k \leq 4n-1} |E_n(2\pi/(4\pi n))| > t\right) \leq c_1 n \exp\left(-c\frac{nt^2}{V^2}\right). \tag{4.22}$$

We observe that E_n is a trigonometric polynomial of order $< n$. Hence, (4.17) shows that

$$\mathsf{Prob}\left(\max_{x \in \mathbb{T}} |E_n(x)| > 2t\right) \leq c_1 n \exp\left(-c\frac{nt^2}{V^2}\right). \tag{4.23}$$

We set the right hand side of the estimate (4.23) equal to δ, and solve for t to obtain

$$2t = C_2 V \sqrt{\frac{\log(C_3 n/\delta)}{n}}.$$

This leads to (4.18). ∎

We are now in a position to prove Theorem 4.1.

PROOF OF THEOREM 4.1

In this proof, we will denote

$$\varepsilon_n = \max_{x \in \mathbb{T}} |E_n(x)|.$$

We choose n so that Lemma 4.1 is applicable and yields with probability exceeding $1 - \delta$:

$$\varepsilon_n \leq C_2 V \sqrt{\frac{\log(C_3 n/\delta)}{n}} \leq \frac{\mathfrak{m}}{16}. \tag{4.24}$$

All the statements in the rest of the proof assume a realization of the ϵ_j's so that (4.24) holds; i.e., they all hold with probability exceeding $1 - \delta$.

We observe next that if $J \subseteq \{1, \cdots, K\}$, $d \geq c_1 n$, $x \in \mathbb{T}$, and $|x - \lambda_\ell| \geq d$ for all $\ell \notin J$, then

$$\left| \sigma_n(x) - \sum_{\ell \in J} A_\ell \Phi_n(x - \lambda_\ell) \right| \leq \frac{ML}{(nd)^S} + \varepsilon_n \leq \frac{ML}{(nd)^S} + \frac{\mathfrak{m}}{16}. \tag{4.25}$$

We now choose C as in (4.13) and assume that

$$n \geq \max(4C/\eta, C_1), \text{ and } C_2 V \sqrt{\frac{\log(C_3 n/\delta)}{n}} \leq \frac{\mathfrak{m}}{16}. \tag{4.26}$$

6

Then (4.25) implies that for $x \in \mathbb{T}$, $|x - \lambda_\ell| \geq C/n$ for all $\ell \notin J$, we have

$$\left| \sigma_n(x) - \sum_{\ell \in J} A_\ell \Phi_n(x - \lambda_\ell) \right| \leq \frac{ML}{(nd)^S} + \varepsilon_n \leq \frac{\mathfrak{m}}{8}. \tag{4.27}$$

Hence, if $|x - \lambda_\ell| \geq C/n$ for all λ_ℓ, $\ell = 1, \cdots, K$, (4.25) applied with $J = \emptyset$ implies that

$$|\sigma_n(x)| \leq \frac{\mathfrak{m}}{8}. \tag{4.28}$$

Consequently, if $x \in \mathbb{G}$, then there is some λ_ℓ such that $|x - \lambda_\ell| < C/n \leq \eta/4$. Necessarily, there is only one λ_ℓ with this property. We now define for $\ell = 1, \cdots, K$,

$$\mathbb{G}_\ell = \{x \in \mathbb{G} : |x - \lambda_\ell| < C/n\}. \tag{4.29}$$

Obviously, $\mathbb{G} = \bigcup_{\ell=1}^{K} \mathbb{G}_\ell$, and $\mathbb{G}_\ell$'s are all mutually disjoint. This proves the disjoint union condition. The diameter condition as well as the separation condition are obviously satisfied.

We prove next the interval inclusion property, which implies in particular, that none of the sets $\mathbb{G}_\ell$ is empty. In order to prove this, we observe that $1 = \max_{x \in \mathbb{T}} |\Phi_n(x)|$. Hence, for $x \in I_\ell$, the estimate (4.27) applied with $J = \{\ell\}$ implies that

$$|\sigma_n(x) - A_\ell \Phi_n(x - \lambda_\ell)| \leq \frac{\mathfrak{m}}{8}. \tag{4.30}$$

Since Φ_n is a trigonometric polynomial of order $< n$, the Bernstein inequality implies that for $x \in I_\ell$ (i.e., $|x - \lambda_\ell| < 1(4n)$),

$$|\Phi_n(x - \lambda_\ell) - 1| \leq n|x - \lambda_\ell| \leq (1/4).$$

So, (4.30) leads to

$$|\sigma_n(x)| \geq (3/4)|A_\ell| - \frac{\mathfrak{m}}{8} \geq (5/8)\mathfrak{m}, \qquad x \in I_\ell.$$

This proves that $I_\ell \subseteq \mathbb{G}_\ell$ for all $\ell = 1, \cdots, K$. The estimate (4.15) is clear from the diameter condition and the interval inclusion property. ∎

5 Algorithms and illustrations

We describe our main algorithms in Section 5.1. In Section 5.2, we illustrate the various steps in the case of a two dimensional dataset, and discuss the results. Section 5.3 discusses the adaptation of our algorithm in the three dimensional case. The corresponding results are discussed in Section 5.4.

5.1 Algorithms

We organize our algorithm in three parts. The first part is to implement Theorem 4.1 in the univariate case. This is given in Algorithm 1.

The next algorithm, Algorithm 2 uses this algorithm to a two dimensional problem, resulting in an accurate estimation of one component of the $\mathbf{w}_k$'s and an approximate estimation of the other component.

The final step, Algorithm 3 uses Algorithm 2 successively with pairs of components to obtain the final accurate estimation of all the components of $\mathbf{w}_k$'s.

Performance assessment:

After finding the estimate $\hat{\mathbf{w}}_k$ for $\mathbf{w}_k$ for each k, the next challenge is to determine how many of these estimates represent the actual points. For this reason, we fix a radius r and declare $\hat{\mathbf{w}}_k$ to be an accurate estimator of $\mathbf{w}_k$ if $|\mathbf{w}_k - \hat{\mathbf{w}}_k| < r$. We then count how many points were estimated accurately within this error margin.

5.2 Illustration in a two dimensional case

We will illustrate our algorithm by using an example of 2-d image which we obtain from [2], as shown in Figure 1. Following this paper, we take $\Delta_1 = (1.38, 4.14)$, $\Delta_2 = (-7.56, 5.67)$.

Algorithm 1 Given a univariate signal $\hat{\mu}(\ell) = \sum_{k=1}^{K} A_k \exp(-i\ell\lambda_k)$, find K, A_k's and λ_k's.

a) **Input:** $\mathfrak{m}$, η and signal $\tilde{\mu}(\ell)$.

b) **Output:** Estimation of $\hat{A}_k$ and $\hat{\lambda}_k$ for $k = 1, \ldots, K$.

1: $\hbar_n \leftarrow \left\{ \sum_{|\ell| < n} H\left(\frac{|\ell|}{n}\right) \right\}^{-1}$

2: $\sigma_n(x) \leftarrow \hbar_n \sum_{|\ell| < n} H\left(\frac{|\ell|}{n}\right) \tilde{\mu}(\ell) e^{i\ell x}$

3: $\mathcal{G} \leftarrow \{x \in [-\pi, \pi] : |\sigma_n(x)| \geq \mathfrak{m}/2\}$

4: **for** $k = 1$ to K **do**

5:     $\mathcal{G}_k \leftarrow Partition(\mathcal{G})$ with minimal separation $\eta/4$

6:     $\hat{\lambda}_k \leftarrow \arg\max_{x \in \mathcal{G}_k}(|\sigma_n(x)|)$

7:     $\hat{A}_k \leftarrow |\sigma_n(\hat{\lambda}_k)|$

8: **end for**

    Note: step 3 - 8 can be computed by using `findpeaks` in MATLAB with parameters `MinPeakDistance` $\eta/4$ and `MinPeakHeight` $\mathfrak{m}/2$.

9: **Return:** $\hat{A}_k, \hat{\lambda}_k$

Algorithm 2 Extension of univariate algorithm to for a two dimensional signal. Given $\hat{\mu}(\Delta_2 + \ell\Delta_1) = \sum_{k=1}^{K} A_k \exp(-i\langle\Delta_2, \mathbf{w}_k\rangle) \exp(-i\ell\langle\Delta_1, \mathbf{w}_k\rangle)$ and $\tilde{\mu}(\Delta_2 + \ell\Delta_1) = \hat{\mu}(\Delta_2 + \ell\Delta_1) + \epsilon_\ell$. Find K, $\langle\Delta_1, \mathbf{w}_k\rangle$, $\langle\Delta_2, \mathbf{w}_k\rangle$.

a) **Input:** $\Delta_1, \Delta_2, \mathfrak{m}, \eta$ and signal $\tilde{\mu}(\Delta_2 + \ell\Delta_1)$.

b) **Output:** Estimation of $\hat{A}_k, \langle\Delta_1, \hat{\mathbf{w}}_k\rangle$, and $\langle\Delta_2, \hat{\mathbf{w}}_k\rangle$ for $k = 1, \ldots, K$.

1: $\hbar_n \leftarrow \left\{ \sum_{|\ell| < n} H\left(\frac{|\ell|}{n}\right) \right\}^{-1}$

2: $\sigma_n(x) \leftarrow \hbar_n \sum_{|\ell| < n} H\left(\frac{|\ell|}{n}\right) \tilde{\mu}(\Delta_2 + \ell\Delta_1) e^{i\ell x}$

3: $\mathcal{G} \leftarrow \{x \in [-\pi, \pi] : |\sigma_n(x)| \geq \mathfrak{m}/2\}$

4: **for** $k = 1$ to K **do**

5:     $\mathcal{G}_k \leftarrow Partition(\mathcal{G})$ with minimal separation $\eta/4$

6:     $\langle\Delta_1, \hat{\mathbf{w}}_k\rangle \leftarrow \arg\max_{x \in \mathcal{G}_k}(|\sigma_n(\langle\Delta_1, \hat{\mathbf{w}}_k\rangle)|)$

7:     $\langle\Delta_2, \hat{\mathbf{w}}_k\rangle \leftarrow Phase(\sigma_n(\langle\Delta_1, \hat{\mathbf{w}}_k\rangle))$

8:     $\hat{A}_k \leftarrow |\sigma_n(\langle\Delta_1, \hat{\mathbf{w}}_k\rangle)|$

9: **end for**

    Note: step 3 - 9 can be computed by using `findpeaks` in MATLAB with parameters `MinPeakDistance` $\eta/4$ and `MinPeakHeight` $\mathfrak{m}/2$.

10: **Return:** $\hat{A}_k, \langle\Delta_1, \hat{\mathbf{w}}_k\rangle, \langle\Delta_2, \hat{\mathbf{w}}_k\rangle$

Let $\mathbf{w}_1, \ldots, \mathbf{w}_{12} \in \mathbb{R}^2$ and $\{\Delta_1, \Delta_2\}$ be a basis for $\mathbb{R}^2$. Here, we have

$$\hat{\mu}(\Delta_2 + \ell\Delta_1) = \sum_{k=1}^{12} A_k \exp(-i\langle\Delta_2, \mathbf{w}_k\rangle) \exp(-i\ell\langle\Delta_1, \mathbf{w}_k\rangle) \tag{5.1}$$

$$\hat{\mu}(\Delta_1 + \ell\Delta_2) = \sum_{k=1}^{12} A_k \exp(-i\langle\Delta_1, \mathbf{w}_k\rangle) \exp(-i\ell\langle\Delta_2, \mathbf{w}_k\rangle) \tag{5.2}$$

The number of samples required is $\mathcal{O}(qn)$, where q is the dimension of the exponential sample, and n is the degree of the localized kernel. We then apply our low pass filter and obtain

$$\sigma_{n,1}(x) = \hbar_n \sum_{k=1}^{12} A_k \exp(-i\langle\Delta_2, \mathbf{w}_k\rangle) \Phi_n(x - \langle\Delta_1, \mathbf{w}_k\rangle) \tag{5.3}$$

$$\sigma_{n,2}(x) = \hbar_n \sum_{k=1}^{12} A_k \exp(-i\langle\Delta_1, \mathbf{w}_k\rangle) \Phi_n(x - \langle\Delta_2, \mathbf{w}_k\rangle) \tag{5.4}$$

From the Theorem 4.1, $\langle\Delta_1, \mathbf{w}_k\rangle$ will be x where the peaks occurs in $|\sigma_n(x)|$, $\langle\Delta_2, \mathbf{w}_k\rangle \approx Phase(\sigma_n(x))$, and $A_k \approx |\sigma_n(x)|$. Now, we can obtain the accurate estimation of $\langle\Delta_1, \mathbf{w}_k\rangle$ corresponding to less accurate estimation

Algorithm 3 Parameter estimation in a multidimensional signal.
___
a) **Input:** Δ_d for $d = 1, \ldots, q$, $\mathfrak{m}$, η, and signal $\tilde{\mu}(\Delta_{d_2} + \ell\Delta_{d_1})$, $1 \leq d_1 < d_2 \leq q$.
b) **Output:** Estimation of $\hat{A}_k$ and $\hat{\mathbf{w}}_k$ for $k = 1, \ldots, K$.
1: **for** $d_1 = 1$ to $q - 1$ **do**
2:     **for** $d_2 = d_1 + 1$ to q **do**
3:         Run Algorithm 2 with parameters $\Delta_{d_1}, \Delta_{d_2}$, $\mathfrak{m}$, η and signal $\tilde{\mu}(\Delta_{d_2} + \ell\Delta_{d_1})$
        Note: From the above step, we will obtain A_k and highly accurate result of $\langle \Delta_{d_1}, \hat{\mathbf{w}}_k \rangle$ together with corresponding less accurate result of $\langle \Delta_{d_2}, \hat{\mathbf{w}}_k \rangle$.
4:         Run Algorithm 2 with parameters $\Delta_{d_2}, \Delta_{d_1}$ $\mathfrak{m}$, η and signal $\tilde{\mu}(\Delta_{d_1} + \ell\Delta_{d_2})$
        Note: From the above step, we will obtain A_k and highly accurate result of $\langle \Delta_{d_2}, \hat{\mathbf{w}}_k \rangle$ together with corresponding less accurate result of $\langle \Delta_{d_1}, \hat{\mathbf{w}}_k \rangle$.
5:         Use nearest neighbor algorithm to obtain the highly accurate pair for both $\langle \Delta_{d_1}, \hat{\mathbf{w}}_k \rangle$ and $\langle \Delta_{d_2}, \hat{\mathbf{w}}_k \rangle$
6:     **end for**
7: **end for**
8: We can write the result as $\Delta \hat{\mathbf{w}}$, where $\Delta = [\Delta_1, \ldots, \Delta_q]^T$ and $\hat{\mathbf{w}} = [\hat{\mathbf{w}}_1, \ldots, \hat{\mathbf{w}}_k]$.
9: **Return:** We then obtain $\hat{\mathbf{w}}_k$ by computing $\hat{\mathbf{w}} = \Delta^{-1}\Delta\hat{\mathbf{w}}$.
10: **Return:** $\hat{A}_k$ and $\hat{\mathbf{w}}_k$ for $k = 1, \ldots, K$.
___

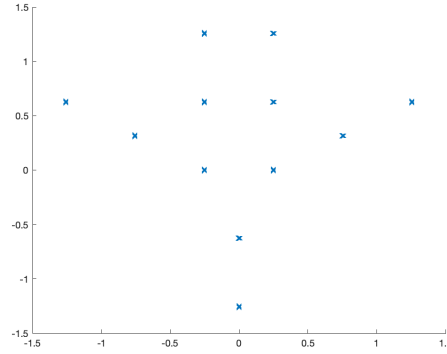| k	ω_k	a_k
1	$(-1.2566, 0.6283)$	50
2	$(-0.7540, 0.3142)$	50
3	$(-0.2513, 1.2566)$	50
4	$(-0.2513, 0.6283)$	50
5	$(-0.2513, 0)$	50
6	$(0, -0.6283)$	50
7	$(0, -1.2566)$	50
8	$(0.2513, 1.2566)$	50
9	$(0.2513, 0.6283)$	50
10	$(0.2513, 0)$	50
11	$(0.7540, 0.3142)$	50
12	$(1.2566, 0.6283)$	50

Figure 1: The two dimensional data comprising 12 points [2]. Left: The acutal points and amplitudes, Right: A graphic representation.

of $\langle \Delta_2, \mathbf{w}_k \rangle$ (Figure 2). Then, we apply the same method in Δ_2 direction to obtain the accurate estimation of $\langle \Delta_2, \mathbf{w}_k \rangle$ corresponding to less accurate estimation of $\langle \Delta_1, \mathbf{w}_k \rangle$ (Figure 3 left and middle).

Finally, we can use nearest neighbor to obtain accurate estimation for both $\langle \Delta_1, \mathbf{w}_k \rangle$ corresponding to $\langle \Delta_2, \mathbf{w}_k \rangle$ and compute A_k. The final result is showed as the rightmost part of Figure 3.

In comparison, we show the results obtained by MUSIC and ESPRIT algorithms in a graphic manner in Figure 4, and as a table in Table 1. It is obvious that both of these algorithms fail drastically even at high SNR levels.

5.3 Algorithm adaption in the three dimensional case

To extend our algorithm to 3-d problem, we can simply add another 3-d projection to our algorithm to find the points in the third coordinate that corresponding to the first and the second coordinates.

Let $\mathbf{w}_1, \ldots, \mathbf{w}_K \in \mathbb{R}^3$ and $\{\Delta_1, \Delta_2, \Delta_3\}$ be a basis for $\mathbb{R}^3$. Here, we have

$$\hat{\mu}(\Delta_2 + \ell\Delta_1) = \sum_{k=1}^{K} A_k \exp(-i\langle \Delta_2, \mathbf{w}_k \rangle) \exp(-i\ell\langle \Delta_1, \mathbf{w}_k \rangle) \tag{5.5}$$

$$\hat{\mu}(\Delta_1 + \ell\Delta_2) = \sum_{k=1}^{K} A_k \exp(-i\langle \Delta_1, \mathbf{w}_k \rangle) \exp(-i\ell\langle \Delta_2, \mathbf{w}_k \rangle) \tag{5.6}$$

$$\hat{\mu}(\Delta_1 + \ell\Delta_3) = \sum_{k=1}^{K} A_k \exp(-i\langle \Delta_1, \mathbf{w}_k \rangle) \exp(-i\ell\langle \Delta_3, \mathbf{w}_k \rangle) \tag{5.7}$$
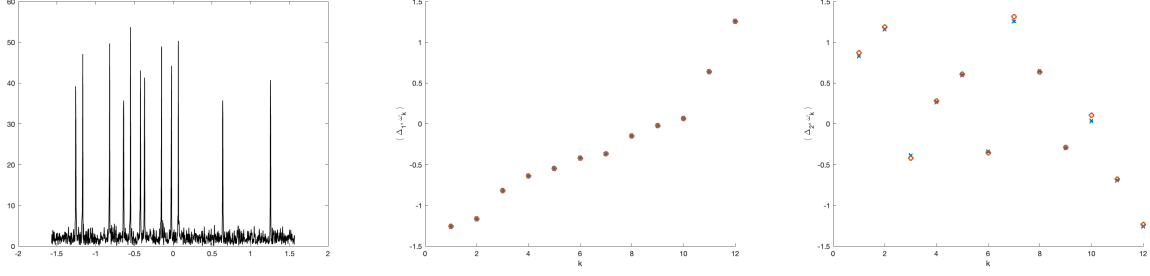
Figure 2: Left: $|\sigma_{n,1}(x)|$, middle: Accurate determination of $\langle \Delta_1, \mathbf{w}_k \rangle$, right: Approximate estimation of $\langle \Delta_2, \mathbf{w}_k \rangle$.
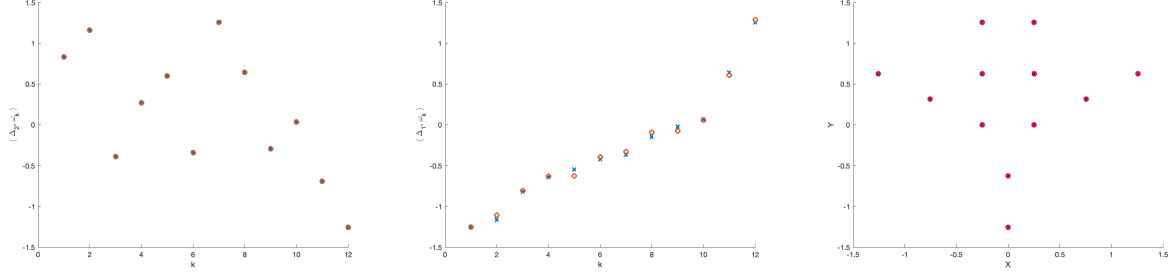

Figure 3: Left: Accurate estimation of $\langle \Delta_2, \mathbf{w}_k \rangle$., Middle: approximate estimation of $\langle \Delta_2, \mathbf{w}_k \rangle$., Right: Final reconstruction

By applying the low pass filter to the signals, we will get

$$\sigma_{n,1}(x) = \hbar_n \sum_{k=1}^{K} A_k \exp(-i\langle \Delta_2, \mathbf{w}_k \rangle) \Phi_n(x - \langle \Delta_1, \mathbf{w}_k \rangle) \tag{5.8}$$

$$\sigma_{n,2}(x) = \hbar_n \sum_{k=1}^{K} A_k \exp(-i\langle \Delta_1, \mathbf{w}_k \rangle) \Phi_n(x - \langle \Delta_2, \mathbf{w}_k \rangle) \tag{5.9}$$

$$\sigma_{n,3}(x) = \hbar_n \sum_{k=1}^{K} A_k \exp(-i\langle \Delta_1, \mathbf{w}_k \rangle) \Phi_n(x - \langle \Delta_3, \mathbf{w}_k \rangle) \tag{5.10}$$

From the Theorem 4.1, $\langle \Delta_1, \mathbf{w}_k \rangle$ will be x where the peaks occurs in $|\sigma_n(x)|$, $\langle \Delta_2, \mathbf{w}_k \rangle \approx Phase\,(\sigma_n(x))$, and $A_k \approx |\sigma_n(x)|$. Now, we can obtain the accurate estimation of $\langle \Delta_1, \mathbf{w}_k \rangle$ corresponding to less accurate estimation of $\langle \Delta_2, \mathbf{w}_k \rangle$.

Then, we apply the same method in Δ_2 and Δ_3 directions to obtain the accurate estimation of $\langle \Delta_2, \mathbf{w}_k \rangle$ and $\langle \Delta_3, \mathbf{w}_k \rangle$ corresponding to less accurate estimation of $\langle \Delta_1, \mathbf{w}_k \rangle$ respectively.

SNR (dB)	Method	Total Points	number of points reconstructed	Run-time (seconds)	Memory (MB)	Accuracy (meters)	RMSE (meters)	Standard Deviation
90	Localized	12	12	0.0045	0.42	0.3	0.0038	0
90	MUSIC	12	1	11.99	0.42	0.3	0.0376	0.0007
90	ESPRIT	12	12	0.0443	0.42	0.3	0.0279	0.0137
60	Localized	12	12	0.0045	0.42	0.3	0.0038	0
60	MUSIC	12	1	12.09	0.42	0.3	0.0378	0.0005
60	ESPRIT	12	8	0.04	0.42	0.3	0.0981	0.0250
30	Localized	12	12	0.0045	0.42	0.3	0.0038	0
30	MUSIC	12	1	10.96	0.42	0.3	0.0367	0.0058
30	ESPRIT	12	5	0.04	0.42	0.3	0.1355	0.0540

Table 1: The tables above compare results between our algorithm, MUSIC, and ESPRIT on 1024 number of samples.
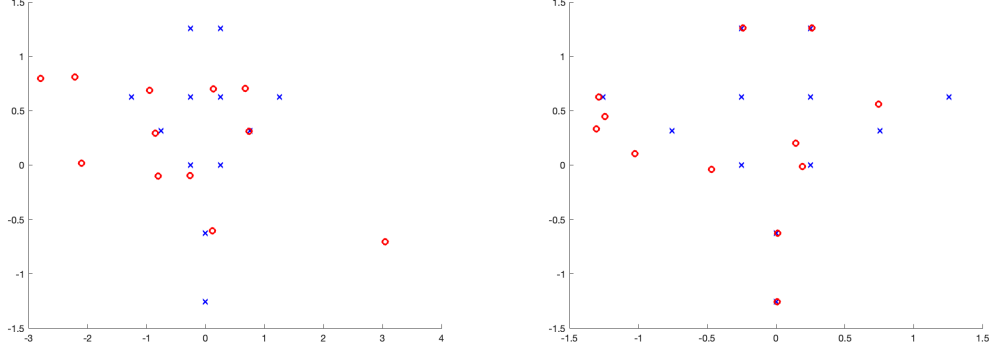
Figure 4: The performance of (Left:) MUSIC and (Right:) ESPRIT algorithms on the 12 point 2 dimensional data set at SNR level of 60dB with 1024 samples.

Figure 5: ESPRIT approximation at 60dB.

Finally, we can use nearest neighbor to obtain accurate estimation for all $\langle \Delta_1, \mathbf{w}_k \rangle$, $\langle \Delta_2, \mathbf{w}_k \rangle$, $\langle \Delta_3, \mathbf{w}_k \rangle$ and compute $A_k = |\sigma_n(x)|$.

5.4 Results in three dimensional experiments

We tested our method described in Section 5.3 on two three dimensional data sets as described in [2], one comprising 29 points and the other comprising 1000 points. Figure 6 shows the results in a special case in a graphical manner.
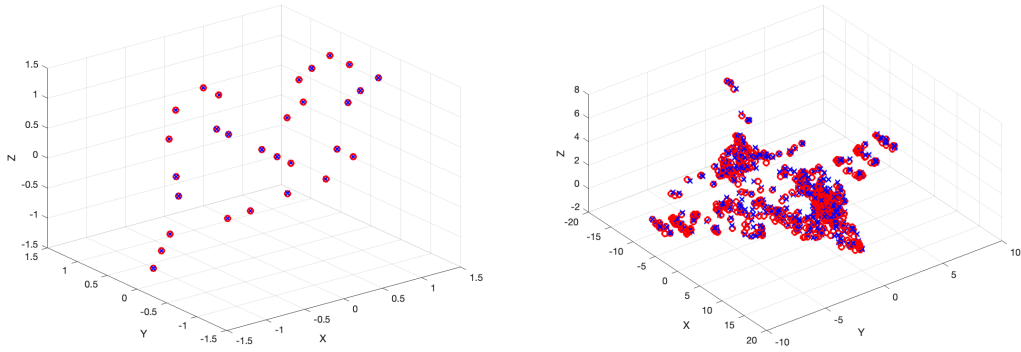

Figure 6: Left: Recuperation of 29 points with 65536 samples at -10dB noise level, Right: Recuperation of 1000 points with 65536 samples at 5dB noise level.

Our algorithm is able to reconstruct all 29 points of 3-d tomographic image accurately as shown in figure 6. The samples required by our method are 65,536 samples with -10 dB noise added to the data. As comparison to the original paper [2], the author required 42,875 samples to reconstruct all 29 points with noise levels varying from 40 dB SNR to 5 dB SNR.

In the 1000 points of 3-d fighter jet image, our algorithm cannot separate the signals that are really close together due to high density data points on a cluster. We can only reconstruct 903 data points of out 1000 points with 65536 samples. This method requires approximately 35MB of memory and the run-time is less than 1 second. Comparing to our baseline result from [2], the author has experimented with 72000, 90000, and 180000 samples with the result of 71% of the scatterers is reconstructed within an error of at most 10 cm and 93% within 30 cm, 81% within 10 cm and 95% within 30 cm, 94% within 10 cm and 98% within 30 cm respectively. One may see that the advantage

of our algorithm in terms of speed and accuracy.

Figure 7 shows the dependence on the SNR of the accuracy and the number of points recuperated out of the 1000 points.
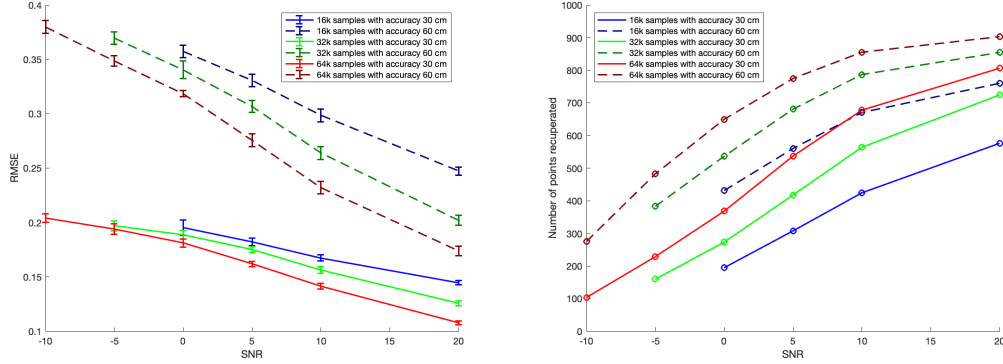

Figure 7: For the 1000 point data set, dependence on SNR for (Left:) accuracy and (Right:) number of points recuperated.

Finally, the details of all the results are summarized in Table 2.

6   Conclusions

The problem of multidimensional exponential analysis arises in many areas of applications including tomographic imaging, ISAR imaging, antenna array processing, etc. The problem is essentially to develop an efficient algorithm to obtain the inverse Fourier transform of a multidimensional signal, based on finitely many equidistant samples of the signal. We have given a very simple algorithm based on localized trigonometric kernel, which reduces the problem to a series of one dimensional problems. Our algorithm works with a tractable number of samples, gives high accuracy, and is very robust even in the presence of noise as high as -10dB. We have proved theoretical guarantees under the assumption that the noise is sub-Gaussian, a significantly weaker assumption that the assumption of white noise common in the literature.

References

[1] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities: A nonasymptotic theory of independence*. Oxford university press, 2013.

[2] Annie Cuyt, Yuan Hou, Ferre Knaepkens, and Wen-shin Lee. Sparse multidimensional exponential analysis with an application to radar imaging. *SIAM Journal on Scientific Computing*, 42(3):B675–B695, 2020.

[3] Annie Cuyt and Wen-shin Lee. Multivariate exponential analysis from the minimal number of samples. *Advances in Computational Mathematics*, 44:987–1002, 2018.

[4] Benedikt Diederichs. *Sparse frequency estimation: Stability and algorithms*. PhD thesis, Staats-und Universitätsbibliothek Hamburg Carl von Ossietzky, 2018.

[5] Benedikt Diederichs, Mihail N Kolountzakis, and Effie Papageorgiou. How many fourier coefficients are needed? *Monatshefte für Mathematik*, 200(1):23–42, 2023.

[6] F. Filbir, H. N. Mhaskar, and J Prestin. On the problem of parameter estimation in exponential sums. *Constructive Approximation*, 35(3):323–343, 2012.

[7] H. Krim and M. Viberg. Two decades of array signal processing research: the parametric approach. *Signal Processing Magazine, IEEE*, 13(4):67–94, 1996.

SNR (dB)	Number of samples	Total Points	number of points reconstructed	Run-time (seconds)	Memory (MB)	Accuracy (meters)	RMSE (meters)	Standard Deviation
-10	8192	12	12	0.04	2.58	0.3	0.0003	1.09×10^{-5}
-10	65536	29	29	0.36	20.53	0.3	0.0007	0.0003
20	16384	1000	575	0.63	14.88	0.3	0.1445	0.0020
10	16384	1000	424	0.75	14.68	0.3	0.1673	0.0029
5	16384	1000	307	0.67	14.52	0.3	0.1819	0.0035
0	16384	1000	194	0.95	14.44	0.3	0.1952	0.0068
20	16384	1000	760	0.62	14.85	0.6	0.2473	0.0037
10	16384	1000	670	0.87	14.81	0.6	0.2984	0.0060
5	16384	1000	559	0.69	14.58	0.6	0.3305	0.0054
0	16384	1000	431	0.87	14.45	0.6	0.3572	0.0055
20	32768	1000	724	0.52	21.97	0.3	0.1256	0.0023
10	32768	1000	563	0.81	21.82	0.3	0.1561	0.0030
5	32768	1000	417	1.06	21.78	0.3	0.1752	0.0026
0	32768	1000	273	1.31	21.57	0.3	0.1886	0.0039
-5	32768	1000	159	1.50	21.20	0.3	0.1969	0.0045
20	32768	1000	854	0.64	21.99	0.6	0.2020	0.0045
10	32768	1000	786	0.75	21.95	0.6	0.2638	0.0057
5	32768	1000	680	1.04	21.91	0.6	0.3067	0.0052
0	32768	1000	536	1.26	21.77	0.6	0.3404	0.0080
-5	32768	1000	382	1.75	21.46	0.6	0.3695	0.0057
20	65536	1000	806	0.58	35.64	0.3	0.1077	0.0018
10	65536	1000	677	0.68	35.59	0.3	0.1413	0.0025
5	65536	1000	536	0.90	35.49	0.3	0.1618	0.0025
0	65536	1000	367	1.24	35.34	0.3	0.1810	0.0037
-5	65536	1000	228	1.95	35.20	0.3	0.1939	0.0048
-10	65536	1000	103	7.33	35.71	0.3	0.2040	0.0040
20	65536	1000	903	0.56	35.62	0.6	0.1736	0.0043
10	65536	1000	855	0.74	35.61	0.6	0.2320	0.0057
5	65536	1000	774	0.83	35.56	0.6	0.2755	0.0060
0	65536	1000	649	1.34	35.50	0.6	0.3185	0.0030
-5	65536	1000	482	2.07	35.26	0.6	0.3485	0.0047
-10	65536	1000	275	7.07	35.25	0.6	0.3798	0.0058

Table 2: The tables above show full performances of our algorithm.

[8] Stefan Kunis, Thomas Peter, Tim Römer, and Ulrich von der Ohe. A multivariate generalization of prony's method. *Linear Algebra and its Applications*, 490:31–47, 2016.

[9] H. N. Mhaskar and J. Prestin. On local smoothness classes of periodic functions. *Journal of Fourier Analysis and Applications*, 11(3):353–373, 2005.

[10] Frank WJ Olver. *Asymptotics and special functions*. Academic press, 2014.

[11] Thomas Peter, Gerlind Plonka, and Robert Schaback. Prony's method for multivariate signals. *Pamm*, 15(1):665–666, 2015.

[12] D. Potts and M. Tasche. Parameter estimation for exponential sums by approximate Prony method. *Signal Processing*, 90(5):1631–1642, 2010.

[13] Daniel Potts and Manfred Tasche. Parameter estimation for multivariate exponential sums. *Electron. Trans. Numer. Anal*, 40(204-224):94, 2013.

[14] André Quinquis, Emanuel Radoi, and F-C Totir. Some radar imagery results using superresolution techniques. *IEEE Transactions on Antennas and Propagation*, 52(5):1230–1244, 2004.

[15] Ramachandran S Raghavan. A generalized version of ace and performance analysis. *IEEE Transactions on Signal Processing*, 68:2574–2585, 2020.

[16] Souleymen Sahnoun, Konstantin Usevich, and Pierre Comon. Multidimensional esprit for damped and undamped signals: Algorithm, computations, and perturbation analysis. *IEEE Transactions on Signal Processing*, 65(22):5897–5910, 2017.

[17] Tomas Sauer. Prony's method in several variables: symbolic solutions by universal interpolation. *Journal of Symbolic Computation*, 84:95–112, 2018.

[18] James Alexander Shohat and Jacob David Tamarkin. *The problem of moments*, volume 1. American Mathematical Society (RI), 1950.

[19] Shyam Venkatasubramanian, Sandeep Gogineni, Bosung Kang, Ali Pezeshki, Muralidhar Rangaswamy, and Vahid Tarokh. Toward data-driven radar stap. *arXiv preprint arXiv:2209.02890*, 2022.

[20] Zhihui Zhu, Gongguo Tang, Pawan Setlur, Sandeep Gogineni, Michael B Wakin, and Muralidhar Rangaswamy. Super-resolution in sar imaging: Analysis with the atomic norm. In *2016 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pages 1–5. IEEE, 2016.