Resolution Limit for Line Spectral Estimation: Theory and Algorithm

Ping Liu* and Hai Zhang†

March 9, 2020

Abstract

Line spectral estimation is a classical signal processing problem aimed to estimate the spectral lines of a signal from its noisy (deterministic or random) measurements. Despite a large body of research on this subject, the theoretical understanding of the spectral line estimation is still elusive. In this paper, we quantitatively characterize the two resolution limits in the line spectral estimation problem: one is the minimum separation distance between the spectral lines that is required for an exact recovery of the number of spectral lines, and the other is the minimum separation distance between the spectral lines that is required for a stable recovery of the supports of the spectral lines. The quantitative characterization implies a phase transition phenomenon in each of the two recovery problems, and also the subtle difference between the two. Moreover, they give a sharp characterization to the resolution limit for the deconvolution problem as a consequence. Finally, we proposed a recursive MUSIC-type algorithm for the number recovery and an augmented MUSIC-algorithm for the support recovery, and analyze their performance both theoretically and numerically. The numerical results also confirm our results on the resolution limit and the phase transition phenomenon.

Keywords: Line spectral estimation, resolution limit, super-resolution, phase transition, MUSIC.

1 Introduction

This paper is concerned with recovering the number and supports of a collection of spectral lines from the noisy signals, which is usually termed as line spectral estimation (LSE) in statistical inference. It is at the core of diverse research fields such as traditional wireless communications, radar, sonar, seismology, astronomy and NMR imaging, and has received significant attention over the years. While the LSE was usually cast as a statistical parameter estimation problem with random noises in the measurements, we are interested in the

*Department of Mathematics, HKUST, Clear Water Bay, Kowloon, Hong Kong (pliuah@connect.ust.hk).
†Department of Mathematics, HKUST, Clear Water Bay, Kowloon, Hong Kong (haizhang@ust.hk). Hai Zhang was supported by Hong Kong RGC grant GRF 16306318 and GRF 16305419.
case of deterministic noise. To be specific, we consider our mathematical model as follows. Suppose the collection of spectral lines is a discrete measure:

$$\mu = \sum_{j=1}^{n} a_j \delta_{y_j},$$

where $y_j \in \mathbb{R}$, $j = 1, \cdots, n$ are the spectral lines and $a_j \in \mathbb{C}$, $j = 1, \cdots, n$ the amplitudes. We assume that $|y_j| \leq d$ for all $j$. The distances between $y_j$ are defined as $|y_p - y_j|$ where $|\cdot|$ is the Euclidean distance. We denote the lower bound and upper bound of the amplitude as

$$m_{\text{min}} = \min_{j=1,\cdots,n}|a_j|, \quad m = ||\mu||_{TV} = \sum_{j=1}^{n} |a_j|.$$

The available measurements are the Fourier transforms of $\mu$ sampled at $M$ equispaced points $x_1 = -\Omega, x_2 = -\Omega + h, \cdots, x_M = \Omega$:

$$Y(x_t) = \mathcal{F}\mu(x_t) + W(x_t) = \sum_{j=1}^{n} a_j e^{iy_j x_t} + W(x_t), \quad 1 \leq t \leq M$$

where $h = \frac{2\Omega}{M-1}$ is the sampling spacing and $W(x_t)$ is the noise with $|W(x_t)| < \sigma$. Here $\sigma$ is the noise level.

Throughout, we assume that $M > 2n$ and that $h \leq \frac{n}{2}$ so that the sampling obeys the classic Nyquist criterion. Note that the latter assumption also exclude the non-uniqueness of the spectral lines due to shift by multiples of $\frac{2\pi}{h}$.

Denote

$$Y = (Y(x_1), \cdots, Y(x_M))^T, \quad [\mu] = (\mathcal{F}\mu(x_1), \cdots, \mathcal{F}\mu(x_M))^T$$

and $W = (W(x_1), \cdots, W(x_M))^T$.

Then the noisy measurements can be written in the following form

$$Y = [\mu] + W.$$

The LSE problem is to recover the discrete measure $\mu$ from the above noisy measurements $Y$. We note that this problem is closely related to the deconvolution problem in imaging. Indeed, let $f(x)$ be a band-limited point spread function. The convolution of $f$ and $\mu$, given by $\mu \ast f(t) = \sum_{j=1}^{n} a_j f(t - y_j)$, can be viewed as the image of the point sources $y_j$'s. With presence of additive noise $\epsilon(t)$, the measured image is

$$y(t) = \mu \ast f(t) + \epsilon(t) = \sum_{j=1}^{n} a_j f(t - y_j) + \epsilon(t).$$

By taking the Fourier transform on both sides, we obtain

$$\mathcal{F}y(x) = \mathcal{F}f(x) \cdot \mathcal{F}\mu(x) + \mathcal{F}\epsilon(x) = \mathcal{F}f(x) \left( \sum_{j=1}^{n} a_j e^{iy_j x} \right) + \mathcal{F}\epsilon(x), \quad (1.1)$$

which reduced to our LSE problem.
It is well-known, since Rayleigh's work, that two sources or spectral lines can be resolved if they are separated more than the Rayleigh limit (or Rayleigh length in some literature) $\frac{\pi}{\Omega}$. On the other hand, as their separation distance decreases and below the Rayleigh limit, it becomes increasingly difficult to resolve them from the noisy measurements. In the so-called "super-resolution" problem people are interested in resolving the sources or spectral lines that are separated below the Rayleigh limit.

Despite much progress over the years, the theoretical understanding of the super-resolution in LSE is still elusive. A particular puzzle is the gap between the physical (classical) resolution limit and the limit from a mathematical recovery problem view of point. Precisely, the empirical Rayleigh limit is not a rigorous limit for the LSE problem (see for instance [29] which discussed that the Rayleigh limit is not well applicable for data subjected to elaborate processing). To our knowledge, Donoho [15] first addressed this resolution limit question. He considered measures supported on the lattice $\{k\Delta\}_{k=-\infty}^{+\infty}$ and regularized by the so-called "Rayleigh index". He showed that the minimax error for the amplitude recovery with noise of size $\sigma$ scales like $SRF^\alpha \sigma$, where $SRF = \frac{1}{\Omega \Delta}$ is the super-resolution factor, and $\alpha$ depends on the Rayleigh index. This result highlights the importance of sparsity and signal-to-noise ratio (SNR) in the ill-posedness of this inverse problem. Further discussed in [13], Demanet and Nguyen considered the case of $n$-sparse signals supported on a grid and showed that the scaling of the noise level for the minimax error should be $SRF^{2n-1}$. See also a similar result for the multi-clumps case in [24]. In [28], using novel extremal functions, Moitra establishes a sharp phase transition for the amplitude recovery in the relation between cutoff frequency ($\Omega$) and separation distance ($\Delta$). However, these works mostly deal with the grid setting and do not address the recovery of source supports. Recently, a novel framework [3, 6] employing "Prony mapping" and "quantitative inverse function theorem" resolved the puzzle in the off-the-grid setting. It is shown in [3] that the minimum separation distance required for stable support recovery is of the order $O\left(\frac{1}{\Omega} \cdot \left(\frac{1}{SNR} \right)^{\frac{1}{2n-1}}\right)$. When the sources are separated beyond this distance, [6] demonstrated that the minimax error of support recovery scales as $SRF^{2n-1} \sigma \Omega$.

In addition to the above theoretical work, many algorithms are proposed to solve the LSE and the super-resolution problem. A large group of methods named subspace methods are shown to have favourable performance. Typical subspace methods, MUSIC and ESPRIT, estimate the spectral lines based on the singular value decomposition of the data matrix. These approaches usually assume a priori information of the model order (the number of spectral lines). In the statistical setting, the model order is usually determined by generic information theoretic criteria (e.g. AIC [1, 2], BIC [31]) or methods based on eigenvalues of the estimated signal covariance matrix (e.g. SORTE [21]). While the performance of these methods depends sensitively on the a priori estimate of the model order [20], their advantage is also evident from their appealing performance in the super-resolution region. See for instance the theoretical work in [23] for ESPRIT and the numerical work for the Matrix Pencil method in [6].

In recent years, inspired by the idea of sparse modeling and compressed sensing, many sparsity promoting algorithms are proposed for the sparse spectral recovery problem, e.g., [16, 27]. They restrict the sources or spectral lines to a grid and solve a sparse reconstruction problem. However, the grid setting will incur the basis mismatch issue [9, 12, 18] and
the granularity of the grid results in a non-trivial trade-off between accuracy and computational cost \cite{20}. To remedy these issues, one should consider the off-the-grid setting. In the work by Candès and Fernandez-Granda \cite{8}, it is proved that off-the-grid sources can be exactly recovered from their low-frequency measurements by TV minimization under a minimum separation condition. It invokes active researches in the off-the-grid algorithms, among them we would like to mention the BLASSO \cite{4,17,30} and the atomic norm minimization method \cite{33,34}. Both methods were proved to be able to stably recover the sources under a minimum separation condition or a non-degeneracy condition. BLASSO (Beurling LASSO) is an off-the-grid generalization of $l^1$ regularization (LASSO) and exhibits excellent performance in the off-the-grid source recovery \cite{17,30}. The atomic norm minimization method (originated from \cite{10}) is shown to form a nearly minimax optimal estimator when tackling LSE \cite{7,33} and is an appealing approach for the blind source separation \cite{22,35}. We refer to \cite{11} for a comprehensive overview of this method. Despite the abundant success, these convex optimization algorithms usually require a minimum separation distance of several Rayleigh limits (see \cite{32,23,30}) for the general source recovery, which may limit their applicability to the super-resolution regime. On the other hand, we note that for the positive sources, the minimum separation distance may be relaxed, for example, a signal-noise-ratio scaling like $(\frac{1}{\sqrt{\Omega}})^{2n-1}$ leads to stable support recovery in 1-D BLASSO case \cite{14}.

In this paper, we investigate a cluster of closely spaced spectral lines without the grid setting and aim to quantitatively characterize the resolution limits to both the number recovery problem and the support recovery problem in LSE. Our theoretical results address the issue when can one recover the spectral number exactly and when can one stably recover the supports with a given SNR. They imply a phase transition phenomenon in each of the two recovery problems, and also the subtle difference between the two. Moreover, they give a sharp characterization to the resolution limit for the deconvolution problem as well. Finally, we proposed a recursive MUSIC-type algorithm for the number recovery and an augmented MUSIC-algorithm for the support recovery, and analyze their performance both theoretically and numerically. The numerical results also confirm our results on the resolution limit and the phase transition phenomenon. Compare to the closely related work \cite{3,6} which focuses on the support and the amplitude recovery problem, our bounds for the resolution limit to the support recovery problem is more explicit in the special case of a single cluster of spectral lines. Moreover, our result shows that the support recovery error scales as $\text{SRF}^{2n-2} \frac{\sigma}{\Omega^{2n-1}}$ which is an improvement of the scaling $(\text{SRF}^{2n-1} \frac{\sigma}{\Omega^{2n-1}})$ in \cite{6}.

The paper is organized in the following way. Section 2 presents the main results to the LSE problem and Section 3 the consequence for the deconvolution problem. Section 4 provides the main technique of the paper, Vandermonde space approximation. Section 5 proposes the MUSIC-type algorithms for both number recovery and support recovery and numerical experiments. Section 6 gives a conclusion and Section 7 provides the proofs of the main results. Finally the Appendix contains some inequalities.
2 Main results

We present our main results on the resolution limit for the LSE problem in this section. All the results in this section shall be proved in Section 7. We consider the case when the spectral lines \( y_j, j = 1, \cdots, n \) are tightly spaced and form a cluster. To be more specific, we define the interval

\[
I(n, \Omega) := \left[ - \frac{(n-1)\pi}{2\Omega}, \frac{(n-1)\pi}{2\Omega} \right]
\]

and assume that \( y_j \in I(n, \Omega), 1 \leq j \leq n \). Recall that the Raleigh limit is \( \frac{\pi}{\Omega} \). For a discrete measure \( \hat{\mu} = \sum_{j=1}^{k} \hat{a}_j \delta_{\hat{y}_j} \), we can only determine if it is the solution to the LSE problem by comparing the data \([\hat{\mu}]\) it generated with the measurements \( Y \). In this principle, we introduce the following concept of \( \sigma \)-admissible measure.

**Definition 2.1.** Given measurement \( Y \), we say that \( \hat{\mu} = \sum_{j=1}^{k} \hat{a}_j \delta_{\hat{y}_j} \) is a \( \sigma \)-admissible discrete measure of \( Y \) only if

\[
||[\hat{\mu}] - Y||_\infty < \sigma.
\]

The set of \( \sigma \)-admissible measures of \( Y \) characterizes all possible solutions to the LSE with the given measurement \( Y \). A good reconstruction algorithm should give a \( \sigma \)-admissible measure. If there exists one \( \sigma \)-admissible measure with less than \( n \) supports, then one may say that there are less than \( n \) spectral lines, and hence miscalculate the exact number. On the other hand, if all \( \sigma \)-admissible measures have at least \( n \) supports, then one can determine the number \( n \) correctly if one restricts to the sparsest admissible measures. This leads to the following definition of resolution limit to the number recovery problem in LSE.

**Definition 2.2.** For measurement \( Y \) generated by \( n \) spectral lines, the computational resolution limit to the number recovery problem is defined as the minimum separation distance between the spectral lines beyond which there does not exist any \( \sigma \)-admissible measure for \( Y \) with less than \( n \) supports.

Theorem 2.1 gives an upper bound of the resolution limit to recover the spectral number in the LSE problem. With minimum separation distance greater than the upper bound in (2.1), the number recovery problem is regularized, and any algorithm targeting at the sparsest
admissible measures can recover the correct number. Compared with Rayleigh limit $\frac{\pi}{\Omega}$, the upper bound indicates that super-resolution is theoretically possible for suitable SNR.

We next show that the above upper bound is optimal in the sense that recovering the source number is theoretically impossible when sources are separated by $\sqrt{C_{m}}\Omega \frac{(\sigma \Omega)^{n-2}}{n-2}$ for some constant $C$.

**Proposition 2.1.** For given $\sigma > 0$, integer $n \geq 2$ and $m > 0$, choose $\tau$ satisfying

$$\tau = \frac{2}{\pi} \left( \frac{(\pi(n - 1))^{1/2}}{e^{n-1} \tau} \right)^{1/2} \frac{\pi^{1/2}}{m^{1/2}}.$$  \hfill (2.2)

There exist two measures $\mu = \sum_{j=1}^{n} \alpha_{j} \delta_{y_{j}}$ with $n$ supports and $\hat{\mu} = \sum_{j=1}^{n-1} \hat{\alpha}_{j} \delta_{\hat{y}_{j}}$ with $n - 1$ supports such that $||\hat{\mu}||_{TV} \leq ||\mu||_{TV} \leq m$ and

$$\min_{\hat{y}_{j} \neq y_{j}} |\hat{y}_{j} - y_{j}| = \frac{\tau}{\Omega}.$$  

Moreover,

$$||[\hat{\mu}] - [\mu]||_{\infty} < \sigma.$$  

The above result gives a lower bound for the resolution limit to the number recovery problem. We emphasize that similar to parallel results in [3, 6, 26], our lower bound is the worst-case bound, and one may achieve better resolution bound for the case of random noises.

**Example of the upper bound and the lower bound:**

We calculate the upper and the lower bounds of the resolution limit derived above to show that super-resolution can be achieved. We set $n = 2, M = 20, \Omega = 1, \sigma = 10^{-3}$ and the collection of spectral lines is $\mu = \delta_{y_{1}} + \delta_{y_{2}}$.

The noisy measurements are $Y = (Y(x_{1}), \ldots, Y(x_{M}))^{T}$ with the noise $W$ satisfying $||W||_{\infty} < \sigma$. According to Theorem 2.1, the upper bound for the resolution limit to number recovery is

$$2.22e \left( \frac{2\sigma}{m_{\min}} \right)^{1/2} \frac{\pi^{1/2}}{\Omega} = 0.27 \frac{\pi}{\Omega},$$

which is much better than the Rayleigh limit $\frac{\pi}{\Omega}$. From Proposition 2.1, the lower bound to the resolution limit is

$$\frac{\tau}{\Omega} = \frac{2}{\pi} \left( \frac{(\pi(n - 1))^{1/2}}{e^{n-1} \frac{\tau}{\Omega}} \right)^{1/2} \frac{\pi^{1/2}}{m^{1/2}} \frac{\pi}{\Omega} = 0.007 \frac{\pi}{\Omega}.$$  

We now consider the support recovery problem in the LSE. We first introduce the following concept of $\delta$-neighborhood of a discrete measure.

**Definition 2.3.** Let $\mu = \sum_{j=1}^{n} \alpha_{j} \delta_{y_{j}}$ be a discrete measure and let $\delta > 0$ be such that the $n$ intervals $(y_{k} - \delta, y_{k} + \delta), 1 \leq k \leq n$ are pairwise disjoint. We say that $\hat{\mu} = \sum_{j=1}^{n} \hat{\alpha}_{j} \delta_{\hat{y}_{j}}$ is within $\delta$-neighborhood of $\mu$ if each $\hat{y}_{j}$ is contained in one and only one of the $n$ intervals $(y_{k} - \delta, y_{k} + \delta), 1 \leq k \leq n$. 
According to the above definition, a measure in a \( \delta \)-neighbourhood preserves the inner structure of the real spectral lines. For any stable support recovery algorithm, the output should be a measure in some \( \delta \)-neighborhood. Moreover, \( \delta \) should tend to zero as the noise level \( \sigma \) tends to zero. We now introduce the resolution limit for stable support recovery. For ease of exposition, we only consider measures supported in \( I(n, \Omega) \) where \( n \) is the number of supports.

**Definition 2.4.** For measurement \( Y \) generated by \( \mu = \sum_{j=1}^{n} a_j \delta y_j \) which is supported in \( I(n, \Omega) \), the computational resolution limit to the stable support recovery problem is defined as the minimum separation distance between the spectral lines beyond which there exists \( \delta > 0 \) such that any \( \sigma \)-admissible measure for \( Y \) with \( n \) supports in \( I(n, \Omega) \) is within \( \delta \)-neighbourhood of \( \mu \).

To state the results on the resolution limit to stable support recovery, we need to introduce one more concept: super-resolution factor, which is usually utilized to characterize the ill-posedness of the super-resolution problem \([8]\). It is defined as the ratio between Rayleigh limit and the grid scale (in the grid setting) or the minimum separation distance (off-the-grid setting). In our case, since the Rayleigh limit is \( \frac{\pi}{\Omega} \), we define the super-resolution factor as

\[
SRF := \frac{\pi}{\Omega d_{\text{min}}},
\]

where \( d_{\text{min}} = \min_{p \neq j} |y_p - y_j| \). We have the following theorem.

**Theorem 2.2.** Let \( n \geq 2 \), assume that \( \mu = \sum_{j=1}^{n} a_j \delta y_j \) is supported on \( I(n, \Omega) \) and satisfies the separation condition that

\[
\min_{p \neq j} |y_p - y_j| \geq \frac{3.07 \pi e}{\Omega m_{\text{min}}} \left( \frac{2 \sigma}{m_{\text{min}}} \right)^{\frac{1}{2n-1}}. \tag{2.3}
\]

If \( \hat{\mu} = \sum_{j=1}^{n} \hat{a}_j \delta y_j \) supported on \( I(n, \Omega) \) is a \( \sigma \)-admissible measure for the measurement generated by \( \mu \), then \( \hat{\mu} \) is within the \( \frac{d_{\text{min}}}{2} \)-neighborhood of \( \mu \). Moreover, for \( 1 \leq j \leq n \),

\[
|\hat{y}_j - y_j| \leq \frac{C(n)}{\Omega} SRF^{2n-2} \frac{\sigma}{m_{\text{min}}} \tag{2.4}
\]

where

\[
C(n) = (2n + 1)2^{2n-2}e^{2n+1} \pi^{-\frac{1}{2}}.
\]

Theorem 2.2 gives an upper bound for the computational resolution limit to stably recover the supports of spectral lines. It states that when the spectral lines are separate beyond the upper bound in (2.3), we can stably recover their supports and any \( \sigma \)-admissible measure preserves the structure of real one. Moreover, it implies that the inverse problem of recovering the spectral supports can be regularized by the minimum separation condition with separation distance greater than the upper bound of the resolution limit in Theorem 2.2. In that case, any algorithm looking for the sparsest admissible measure can achieve stable recovery. Compared with the Rayleigh limit \( \frac{\pi}{\Omega} \), the upper bound indicates achieving super-resolution is possible under suitable SNR.

We next show that the lower bound for the separation distance to ensure a stable support recovery is of the order \( O(\frac{1}{\Omega} \left( \frac{\sigma}{m} \right)^{\frac{1}{2n-1}}) \), which demonstrates that our upper bound is optimal.
Proposition 2.2. For given $\sigma > 0$, integer $n \geq 2$ and $m > 0$, chose $\tau$ satisfying
\[
\tau = \frac{2}{e^{\left(\frac{\pi (n - \frac{1}{2})}{2 e^{\frac{1}{2e}}\frac{1}{m}}\right)}}. ~ (2.5)
\]
Then there exist two measures $\mu$ and $\hat{\mu}$, supported in $\{-\frac{\tau}{\Omega}, -2\frac{\tau}{\Omega}, -n\frac{\tau}{\Omega}\}$ and $\{\frac{\tau}{\Omega}, 2\frac{\tau}{\Omega}, \ldots, n\frac{\tau}{\Omega}\}$ respectively, such that $||\hat{\mu}||_{TV} \leq ||\mu||_{TV} \leq m$ and
\[
||\hat{\mu} - \mu||_{\infty} < \sigma.
\]

Proposition 2.2 and Theorem 2.2 reveal that the resolution limit of stable support recovery is of the order $O\left(\frac{1}{\text{SNR} \frac{\pi \Omega}{\Omega}}\right)$.

Example of the upper bound and the lower bound:
We calculate the upper and the lower bounds of the resolution limit derived above to show that achieving the super-resolution is theoretically possible. We set $n = 2, M = 20, \Omega = 1, \sigma = 10^{-4}$ and the collection of spectral lines is
\[
\mu = \delta_{y_1} + \delta_{y_2}.
\]
The noisy measurements are
\[
Y = (Y(x_1), \ldots, Y(x_M))^T
\]
with the noise $W$ satisfying $||W||_{\infty} < \sigma$. According to Theorem 2.2, the upper bound for the resolution limit to support recovery is
\[
3.07e^{\left(\frac{2\sigma}{m_{\min}}\right)^\frac{1}{2e^{\pi}}\frac{\pi \Omega}{\Omega}} = 0.49\frac{\pi \Omega}{\Omega},
\]
which is much better than the Rayleigh limit $\frac{\pi \Omega}{\Omega}$. From Proposition 2.2, we calculate the lower bound that
\[
\frac{\tau}{\Omega} = \frac{2}{e\pi^{\left(\frac{\pi (n - \frac{1}{2})}{2 e^{\frac{1}{2e}}\frac{1}{m}}\right)}} = 0.008\frac{\pi \Omega}{\Omega}.
\]

Remark 2.1. We have quantitatively characterized the resolution limit to both the number recovery and the support recovery problems in the LSE by using the SNR and the sparsity of the spectral lines. It shows that the number recovery has a better resolution limit than the support recovery. As a direct consequence, the exact recovery of the number does not guarantee a stable recovery of supports. This is confirmed in our numerical experiments in Section 5.

Remark 2.2. Our results imply that phase transition may occur in both the number and the support recovery problems in the LSE. See Section 5.3 for detail.

3 The resolution limit in the deconvolution problem
As an application of the results in the previous section, we consider the resolution limit in the deconvolution problem which is closely related to the LSE problem in this section. By sampling the image in the frequency domain, see (1.1), we get
\[
Y(x_t) = \mathcal{F}f(x_t) \cdot \mathcal{F}\mu(x_t) + W(x_t) = \mathcal{F}f(x_t)\left(\sum_{j=1}^{n} a_j e^{iy_j x_t}\right) + W(x_t), \quad t = 1, \ldots, M.
\]
where \( Y(x_t) = \mathcal{F} y(x_t) \) and \( W(x_t) = \mathcal{F} e(x_t) \). We assume \( ||W||_\infty < \sigma \) with \( \sigma \) being the noise level.

**Assumption 3.1.** There exists \( c_0 > 0 \), such that
\[
|\mathcal{F} f(x)| \geq c_0, \quad \forall x \in [-\Omega, \Omega].
\]

Define
\[
\mathcal{G}(\mu) = (\mathcal{F} f(x_1) \cdot \mathcal{F} \mu(x_1), \ldots, \mathcal{F} f(x_M) \cdot \mathcal{F} \mu(x_M))^T,
\]

**Definition 3.1.** We say that \( \hat{\mu} = \sum_{j=1}^k \hat{a}_j \delta_{\hat{y}_j} \) is a \( \sigma \)-admissible discrete measure of the measurement \( Y \) only if
\[
||\mathcal{G}(\hat{\mu}) - Y||_\infty < \sigma.
\]

Then we have the following result for the resolution limit to number recovery problem.

**Theorem 3.1.** Let \( n \geq 2 \). Under Assumption 3.1, if \( \mu = \sum_{j=1}^n a_j \delta_{y_j} \) is supported on \( I(n, \Omega) \) and satisfies the separation condition that
\[
\min_{p \neq j} |y_p - y_j| \geq 2.22 \pi e \left( \frac{2\sigma}{c_0 m_{\min}} \right)^{\frac{1}{n-2}},
\]
then there does not exist a \( \sigma \)-admissible measure \( \hat{\mu} \) with less than \( n \) supports for the measurement \( Y \) generated by \( \mu \).

Similarly, we have the following result for the resolution limit to the support recovery problem.

**Theorem 3.2.** Let \( n \geq 2 \). Under Assumption 3.1, suppose that \( \mu = \sum_{j=1}^n a_j \delta_{y_j} \) is supported on \( I(n, \Omega) \) and satisfies the separation condition that
\[
\min_{p \neq j} |y_p - y_j| \geq 3.07 \pi e \left( \frac{2\sigma}{c_0 m_{\min}} \right)^{\frac{1}{n-1}}.
\]
If \( \hat{\mu} = \sum_{j=1}^n \hat{a}_j \delta_{\hat{y}_j} \) supported on \( I(n, \Omega) \) is a \( \sigma \)-admissible measure of the measurement generated by \( \mu \), then \( \hat{\mu} \) is within the \( \frac{\sigma}{c_0 m_{\min}} \)-neighborhood of \( \mu \). Moreover, for \( 1 \leq j \leq n \),
\[
|\hat{y}_j - y_j| \leq \frac{C(n)}{\Omega} SR \pi^{2n-2} \frac{\sigma}{c_0 m_{\min}},
\]
where
\[
C(n) = (2n+1)2^{2n-2}e^{2n+1}\pi^{-\frac{1}{2}}.
\]

**Remark 3.1.** The results in Theorem 3.1 and 3.2 improves the upper bounds for the two computational resolution limits derived in [26] by reducing an auxiliary noise amplification factor. This improvement comes from the special data structure in the frequency space, where the measurements are linear combinations of exponentials. In the time-space as considered in [26], the structure in the direct measurements is not clear. One has to solve a linear system first to get the multipole coefficients which have a special structure of being linear combinations of powers. It is in the process of solving the multipole coefficients that we pick up the auxiliary noise amplification factor due to the correlation of different multipoles. Therefore, the better bounds derived here indicate the obvious advantage of solving the deconvolution problem in the frequency space than in the time-space.
4 Main techniques: Vandermonde space approximation

We present the main techniques used in the paper, the approximation theory in Vandermonde space in this section. The theory was first introduced in [26] and was restricted to the case of real vectors. We shall extend the theory to complex vectors. Since most arguments are similar, we shall only highlight the main differences and refer the readers to [26] for more detail.

We first introduce some definitions. For each $\omega \in \mathbb{C}$, we define the Vandermonde-vector
\[ \phi_s(\omega) = (1, \omega, \cdots, \omega^s)^T. \] (4.1)

We also define the Vandermonde space as
\[ W_s = \text{span}\{\phi_s(\omega) : \omega \in \mathbb{C}\}, \]
and $k$ dimensional Vandermonde subspace as
\[ W^k_s(\omega_1, \cdots, \omega_k) = \text{span}\{\phi_s(\omega_1), \cdots, \phi_s(\omega_k)\}. \]

We consider the following approximation problem in $W^k_s(\omega_1, \cdots, \omega_k)$:
\[ \min_{\tilde{a}_j, \tilde{d}_j \in \mathbb{C}, j=1, \cdots, k} \left\| \sum_{j=1}^{k} \tilde{a}_j \phi_s(\tilde{d}_j) - v \right\|_2^2, \] (4.2)

where $v = \sum_{j=1}^{k+1} a_j \phi_s(d_j)$ is given. This is a highly non-linear problem. We aim to derive a lower bound for this minimization problem for a special case that is relevant to our LSE problem in what follows.

4.1 Notations and Preliminary

We first introduce some notations which are used frequently in subsequent sections. Denote $\Gamma$ the unit circle on the complex plane. For $d \in \Gamma$, we define $\text{Arg}(d) \in [-\pi, \pi)$ to be the unique number such that $d = e^{i \text{Arg}(d)}$. For $d_j, d_p \in \Gamma$, we define
\[ \angle(d_j d_p) = |\text{Arg}(d_j) - \text{Arg}(d_p)|. \] (4.3)

For $-\pi \leq \theta < 0$, we define
\[ \Gamma^+(\theta) = \{d_j \mid d_j \in \Gamma, \text{Arg}(d_j) \in [\theta, \theta + \pi)\}. \]

We note that if $d_j, d_p \in \Gamma^+(\theta)$ for some $\theta \in [-\pi, 0)$, then
\[ |d_j - d_p| \geq \frac{2}{\pi} \angle(d_j d_p). \] (4.4)

We denote for integer $n \geq 1$,
\[ \xi(n) = \begin{cases} \left(\frac{n-1}{2}\right)^2, & n \text{ is odd,} \\ \left(\frac{n}{2}\right)! \left(\frac{n-2}{2}\right)!, & n \text{ is even,} \end{cases} \]
\[ \xi(n) = \begin{cases} \frac{1}{2}, & n = 1, \\ \left(\frac{n-1}{2}\right)! \left(\frac{n+1}{2}\right)!, & n \text{ is odd, } n \geq 3, \\ \frac{1}{4}, & n \text{ is even.} \end{cases} \] (4.5)
We also define
\[
\eta_{p,q}(d_1,\ldots,d_p,\hat{d}_1,\ldots,\hat{d}_q) = \begin{bmatrix}
| (d_1 - \hat{d}_1)| \cdots |(d_1 - \hat{d}_q)| \\
| (d_2 - \hat{d}_1)| \cdots |(d_2 - \hat{d}_q)| \\
\vdots \\
| (d_p - \hat{d}_1)| \cdots |(d_p - \hat{d}_q)|
\end{bmatrix}.
\] (4.6)

In this section, we consider complex matrices. For complex matrix \( A \), we denote \( A^* \) its conjugate transpose. A complex vector is viewed as a matrix.

**Lemma 4.1.** For \( s \times k \) matrix \( A \) of rank \( k \) with \( s \geq k \), we have
\[
\min_{a \in \mathbb{C}^k} ||Aa - v||_2 = \sqrt{\frac{\det(V^*V)}{\det(A^*A)}},
\]
where \( V = (A, v) \).

Proof: Since \( V = (A, v) \), we have
\[
V^*V = \begin{pmatrix} A^*A & A^*v \\ v^*A & v^*v \end{pmatrix}.
\]

By column transform we have
\[
\det(V^*V) = \det(A^*A) \det(v^*v - v^*A(A^*A)^{-1}A^*v).
\]

Denote the space spanned by columns of \( A \) as \( S(A) \) and the orthogonal complement of \( S(A) \) as \( S(A)^\perp \). We decompose \( v = v_1 + v_2 \) where \( v_1 \in S(A) \) and \( v_2 \in S(A)^\perp \). We have
\[
\sqrt{\frac{\det(V^*V)}{\det(A^*A)}} = \sqrt{\det(v^*v - v^*A(A^*A)^{-1}A^*v)}
\]
\[
= \sqrt{\det(v_1^*v_1 + v_2^*v_2 - v_1^*A(A^*A)^{-1}A^*v_1)}.
\]

Note that \( A(A^*A)^{-1}A^* \) is the orthogonal projection onto the space \( S(A) \). Therefore
\[
v_1^*A(A^*A)^{-1}A^*v_1 = v_1^*v_1.
\]

It follows that
\[
\sqrt{\frac{\det(V^*V)}{\det(A^*A)}} = \sqrt{v_2^*v_2} = \min_{a \in \mathbb{C}^k} ||Aa - v||_2.
\]

This completes the proof.

We denote for positive integers \( s \) and \( n \),
\[
V_n(s) = (\phi_s(d_1),\ldots,\phi_s(d_n)), \quad V_n(n-1) = (\phi_{n-1}(d_1),\ldots,\phi_{n-1}(d_n)).
\] (4.7)

where \( \phi_s(\omega) \) is the Vandermonde-vector defined in (4.1). We now present several useful lemmas.
Lemma 4.2. For \( d_j \in \Gamma, j = 1, \cdots, n \), we have
\[
\sqrt{\frac{\det(V_n(n)^* V_n(n))}{\det(V_n(n-1)^* V_n(n-1))}} \leq 2^n.
\]

Proof: Similar to Lemma 3.2 in \[26\]. Here we used the fact that \(|d_j| = 1\).

Lemma 4.3. Let \( \theta \in (-\pi,0) \) and let \( d_j \in \Gamma^+(\theta), j = 1, \cdots, n \), be \( n \) different complex numbers. Denote \( \theta_{\min} = \min_{p \neq j} \angle(d_p d_j) \). For the Vandermonde matrix \( V_n(n-1) \) defined in \[4.7\], we have
\[
||V_n(n-1)^{-1}||_{\infty} \leq \frac{\pi^{n-1}}{\zeta(n) \theta_{\min}^{n-1}},
\]
where \( \zeta(n) \) is defined in \[4.3\].

Proof: Using the properties of Vandermonde matrix in \[19\], we have
\[
||V_n(n-1)^{-1}||_{\infty} \leq \max_{1 \leq j \leq n} \Pi_{1 \leq p \leq n, p \neq j} \frac{1 + |d_p|}{|d_j - d_p|}.
\]
It follows that
\[
||V_n(n-1)^{-1}||_{\infty} \leq 2^{n-1} \max_{1 \leq j \leq n} \Pi_{1 \leq p \leq n, p \neq j} \frac{1}{|d_j - d_p|}.
\] (4.8)

On the other hand, note that
\[
\Pi_{1 \leq p \leq n, p \neq j} \frac{1}{|d_j - d_p|} = \Pi_{p < j} \frac{1}{|d_j - d_p|} \Pi_{p > j} \frac{1}{|d_j - d_p|} \leq (\frac{\pi}{2})^{n-1} \frac{1}{(j-1)! \theta_{\min}^{j-1}} \frac{1}{(n-j)! \theta_{\min}^{n-j}} \leq \frac{1}{(j-1)! (n-j)!} \frac{\pi}{(2 \theta_{\min})^{n-1}} \leq \frac{1}{\zeta(n) (2 \theta_{\min})^{n-1}}.
\]
Together with \[4.8\], we get the desired estimate.

Lemma 4.4. For the Vandermonde matrix \( V_n(n-1) \) and \( V_n(s) \) in \[4.7\] with \( s > n - 1 \), the following estimate on their singular values hold:
\[
\frac{1}{\sqrt{n \min_{1 \leq j \leq n} \Pi_{1 \leq p \leq n, p \neq j} |d_j - d_p|}} \leq \frac{1}{||V_n(n-1)^{-1}||_2} \leq \sigma_{\min}(V_n(n-1)) \leq \sigma_{\min}(V_n(s)).
\]

Proof: See \[26\].

Lemma 4.5. Let \( k \geq 1 \). Assume that \( \theta_j \in \mathbb{R}, 1 \leq j \leq k + 1 \) are \( k + 1 \) different real numbers. Let \( \theta_{\min} = \min_{p \neq j} |\theta_p - \theta_j| \). Then we have the following estimate
\[
\min_{\theta_j \in \mathbb{R}, \cdots, \theta_k \in \mathbb{R}} ||\eta_{k+1,k}(\theta_1, \cdots, \theta_k, \hat{\theta}_1, \cdots, \hat{\theta}_k)||_{\infty} \geq \zeta(k)(\theta_{\min})^k,
\]
where \( \eta_{k+1,k}(\theta_1, \cdots, \theta_k, \hat{\theta}_1, \cdots, \hat{\theta}_k) \) is defined in \[4.6\].
Proof: See [26].

Finally, we extend Lemma 4.5 to the complex case.

**Lemma 4.6.** For distinct \( d_1, \ldots, d_{k+1} \in \Gamma^+(\theta) \) with \( \theta \in [-\pi,0) \), if they satisfy the separation condition that \( \min_{p \neq j} \angle(d_p, d_j) = \theta_{\min} \), then for any \( d_1, \ldots, \hat{d}_k \in \Gamma \), we have the following estimate

\[
||\eta_{k+1,k}(d_1, \ldots, d_{k+1}, \hat{d}_1, \ldots, \hat{d}_k)||_{\infty} \geq \xi(k)\left(\frac{2\theta_{\min}}{\pi}\right)^k,
\]

where \( \xi(k) \) is defined in (4.5) and \( \eta_{k+1,k}(d_1, \ldots, d_{k+1}, \hat{d}_1, \ldots, \hat{d}_k) \) is defined in (4.6).

Proof: We need to show that

\[
\min_{\hat{d}_1 \in \Gamma, \ldots, \hat{d}_k \in \Gamma} ||\eta_{k+1,k}(d_1, \ldots, d_{k+1}, \hat{d}_1, \ldots, \hat{d}_k)||_{\infty} \geq \xi(k)\left(\frac{2\theta_{\min}}{\pi}\right)^k.
\]  

(4.9)

It is clear that the minimizer to (4.9) exists (may not be unique). Let \( (\hat{d}_1, \ldots, \hat{d}_k) \) be a minimizer to (4.9). Without loss of generality, we may assume that \( \hat{d}_j \in \Gamma^+(\theta), j = 1, \ldots, k \). For otherwise, if \( d_p \notin \Gamma^+(\theta) \) for some \( p \), consider the point \( \hat{d}_p \), which is the symmetry point of \( d_p \) with respect to the straight line connecting \( e^{i\theta} \) and \( e^{i(\theta+\pi)} \). We can check that \( |d_j - \hat{d}_p| \leq |d_j - d_p|, 1 \leq j \leq k + 1 \) and hence

\[
||\eta_{k+1,k}(d_1, \ldots, d_{k+1}, \hat{d}_1, \ldots, \hat{d}_k, \hat{d}_p, \hat{d}_p, \hat{d}_p, \ldots, \hat{d}_p)||_{\infty} \leq ||\eta_{k+1,k}(d_1, \ldots, d_{k+1}, \hat{d}_1, \ldots, \hat{d}_k)||_{\infty}.
\]

This shows that we can choose minimizer \( (\hat{d}_1, \ldots, \hat{d}_k) \) such that \( \hat{d}_j \in \Gamma^+(\theta) \) for \( 1 \leq j \leq k \). Note that \( d_j \in \Gamma^+(\theta) \) for \( 1 \leq j \leq k + 1 \), by (4.4), we have

\[
||\eta_{k+1,k}(d_1, \ldots, d_{k+1}, \hat{d}_1, \ldots, \hat{d}_k)||_{\infty} \geq \left(\frac{2}{\pi}\right)^k \max_{j=1, \ldots, k+1} |\text{Arg}(d_j) - \text{Arg}(\hat{d}_1)| \cdots |\text{Arg}(d_j) - \text{Arg}(\hat{d}_k)| \geq \xi(k)\left(\frac{2\theta_{\min}}{\pi}\right)^k,
\]

where we used Lemma 4.5 for the last inequality above. This completes the proof of the lemma.

### 4.2 Lower bound for space approximation

We derive a lower bound for the non-linear approximation problem (4.2).

**Theorem 4.1.** Let \( k \geq 1 \) and \( \theta \in [-\pi,0) \). Assume that \( d_1, \ldots, d_{k+1} \in \Gamma^+(\theta) \) are \( k + 1 \) distinct points, and \( |a_j| \geq m_{\min}, 1 \leq j \leq k + 1 \). Let \( \theta_{\min} = \min_{p \neq j} \angle(d_p, d_j) \). For \( q \leq k \), let \( \hat{a}(q) = (\hat{a}_1, \ldots, \hat{a}_q)^T, a = (a_1, \ldots, a_{k+1})^T \) and

\[
\hat{A}(q) = (\phi_{2k}(\hat{d}_1), \ldots, \phi_{2k}(\hat{d}_q)), \quad A = (\phi_{2k}(d_1), \ldots, \phi_{2k}(d_{k+1}))
\]

where \( \phi_{2k}(\omega) \) is defined as in (4.7). Then

\[
\min_{\hat{a}_p \in C, d_p \in \Gamma, p=1, \ldots, q} ||\hat{A}(q)\hat{a}(q) - Aa||_2 \geq \frac{\zeta(k+1)\xi(k)m_{\min}\theta_{\min}^{2k}}{\pi^{2k}}.
\]
We rewrite these measurements into a Hankel matrix and the partial measurements are

\[ \hat{X} = \begin{bmatrix} Y(-\Omega) & Y(-\Omega + \frac{1}{2}\Omega) & \cdots & Y(0) \\ Y(-\Omega + \frac{1}{2}\Omega) & Y(-\Omega + \frac{2}{2}\Omega) & \cdots & Y(\frac{1}{2}\Omega) \\ \vdots & \vdots & \ddots & \vdots \\ Y(0) & Y(\frac{1}{2}\Omega) & \cdots & Y(\Omega) \end{bmatrix}. \] (5.1)

We observe that \( \hat{X} \) has the decomposition that

\[ \hat{X} = DAD^T + \Delta, \]

Proof: Using Lemma 4.1, 4.2, 4.6 and 4.3, the proof is similar to the real case in [26].

**Theorem 4.2.** Let \( k \geq 2 \) and \( \theta \in [-\pi, 0) \). Assume that \( d_1, \ldots, d_k \in \Gamma^+(\theta) \) are \( k \) different points and \( |a_j| \geq m_{\text{min}}, 1 \leq j \leq k \). Define \( \theta_{\text{min}} = \min_{p \neq j} \angle(d_p, d_j) \). Assume \( k \) distinct complex numbers \( \hat{d}_1, \ldots, \hat{d}_k \in \Gamma^+(\theta) \) satisfy

\[ ||\hat{A}\hat{a} - Aa||_2 < \sigma, \]

where \( \hat{a} = (\hat{a}_1, \ldots, \hat{a}_k)^T, a = (a_1, \ldots, a_k)^T \) and

\[ \hat{A} = (\phi_{2k-1}(\hat{d}_1), \ldots, \phi_{2k-1}(\hat{d}_k)), \quad A = (\phi_{2k-1}(d_1), \ldots, \phi_{2k-1}(d_k)). \]

Here \( \phi_{2k-1}(\omega) \) is defined as in (4.1). Then for \( \eta_{k,k}(d_1, \ldots, d_k, \hat{d}_1, \ldots, \hat{d}_k) \) defined in (4.6), we have

\[ ||\eta_{k,k}(d_1, \ldots, d_k, \hat{d}_1, \ldots, \hat{d}_k)||_{\infty} < \frac{2^k \pi^{k-1} \sigma}{\zeta(k) \sigma_{\min} m_{\text{min}}}. \]

Proof: Similar to the proof of the parallel result in Vandermonde space approximation for the real case in [26].

### 5 Algorithms and Numerical experiments

In this section, we propose a MUSIC-type algorithm to recover the number of spectral lines and an augmented MUSIC algorithm to recover their supports, and analyze their performance both theoretically and numerically.

#### 5.1 MUSIC-type number recovering Algorithm

We present a MUSIC-type algorithm in this section to recover the number of spectral lines. Instead of considering the full measurements \( Y = (Y(x_1), \ldots, Y(x_M))^T \), we choose partial measurements at \( z_t = x_{(t-1)s+1} \) for \( t = 1, \ldots, 2s+1 \) where \( s \geq n \) and \( r = (M-1) \mod 2s \). For ease of exposition, we assume \( r = \frac{M-1}{2s} \). Thus \( z_t = x_{(t-1)s+1} = -\Omega + \frac{t-1}{s} \Omega \) (since \( x_1 = -\Omega, x_M = \Omega \)), and the partial measurements are

\[ Y(z_t) = \mathcal{F} \mu(z_t) + W(z_t) = \sum_{j=1}^{n} a_j e^{ijz_t} + W(z_t), \quad 1 \leq t \leq 2s+1. \]

We rewrite these measurements into a Hankel matrix

\[ \hat{X} = \begin{bmatrix} Y(-\Omega) & Y(-\Omega + \frac{1}{2}\Omega) & \cdots & Y(0) \\ Y(-\Omega + \frac{1}{2}\Omega) & Y(-\Omega + \frac{2}{2}\Omega) & \cdots & Y(\frac{1}{2}\Omega) \\ \vdots & \vdots & \ddots & \vdots \\ Y(0) & Y(\frac{1}{2}\Omega) & \cdots & Y(\Omega) \end{bmatrix}. \] (5.1)

We observe that \( \hat{X} \) has the decomposition that

\[ \hat{X} = DAD^T + \Delta, \]
where \( A = \text{diag}(e^{-iy_j\Omega} a_1, \ldots, e^{-iy_j\Omega} a_n) \) and \( D = \left( \phi_j(e^{iy_j\Omega}), \ldots, \phi_j(e^{iy_j\Omega}) \right) \) with \( \phi_j(\omega) \) being defined in (4.1) and
\[
\Delta = \begin{pmatrix}
W(-\Omega) & W(-\Omega + \frac{1}{2} \Omega) & \cdots & W(0) \\
W(-\Omega + \frac{1}{2} \Omega) & W(-\Omega + \frac{3}{2} \Omega) & \cdots & W(\frac{1}{2} \Omega) \\
\vdots & \vdots & \ddots & \vdots \\
W(0) & W(\frac{1}{2} \Omega) & \cdots & W(\Omega)
\end{pmatrix}.
\]

We denote the singular value decomposition of \( \hat{X} \) as
\[
\hat{X} = \hat{U} \hat{\Sigma} \hat{U}^*,
\]
where \( \hat{\Sigma} = \text{diag}(\hat{\sigma}_1, \ldots, \hat{\sigma}_n, \hat{\sigma}_{n+1}, \ldots, \hat{\sigma}_{s+1}) \) with the singular values \( \hat{\sigma}_j, 1 \leq j \leq s+1 \), ordered in a decreasing manner. Note that when there is no noise, \( X = DA \Delta^T \). We have the following estimate for the singular values of \( DA \Delta^T \).

**Theorem 5.1.** Let \( n \geq 2 \), \( y_j \in I(n, \Omega), 1 \leq j \leq n \), and let \( U \Sigma U^* \) be the singular value decomposition of the matrix \( DA \Delta^T \). Let \( \Sigma = \text{diag}(\sigma_1, \ldots, \sigma_n, 0, \ldots, 0) \). Then the following estimate holds
\[
\sigma_n \geq \frac{m_{\min}(\zeta(n))^2(\theta_{\min}(\Omega, s))^{2n-2}}{n\pi^{2n-2}},
\]
where \( \zeta(n) \) is defined in (4.3) and
\[
\theta_{\min}(\Omega, s) = \min_{p \neq j} \left| y_p \frac{\Omega}{s} - y_j \frac{\Omega}{s} \right|.
\]

Proof: Recall that \( \sigma_n \) is the minimum non-zero singular value of \( DA \Delta^T \). Let \( \text{ker}(D^T) \) be the kernel space of \( D^T \) and \( \text{ker}^+(D^T) \) be its orthogonal complement, we have
\[
\sigma_n = \min_{\|x\|_2 = 1, x \in \text{ker}^+(D^T)} \|DA \Delta^T x\|_2 \geq \sigma_{\min}(DA)\sigma_n(D) \geq \sigma_{\min}(D)\sigma_{\min}(A)\sigma_{\min}(D).
\]

Since \( s \geq n \), similar to (7.2), there exists \( \theta \in [-\pi, 0) \) such that \( e^{iy_j\theta} \in \Gamma^+(\theta), 1 \leq j \leq n \). By Lemma 4.4 and 4.3, we have
\[
\sigma_{\min}(D) \geq \frac{1}{\sqrt{n}} \left( \frac{\zeta(n)(\theta_{\min}(\Omega, s))^{n-1}}{\pi^{n-1}} \right).
\]
It follows that
\[
\sigma_n \geq \sigma_{\min}(D) \left( \frac{1}{\sqrt{n}} \frac{\zeta(n)(\theta_{\min}(\Omega, s))^{n-1}}{\pi^{n-1}} \right)^2 \geq \frac{m_{\min}(\zeta(n))^2(\theta_{\min}(\Omega, s))^{2n-2}}{n\pi^{2n-2}}.
\]

**Corollary 5.1.** Let \( \mu = \sum_{j=1}^n a_j y_j \) with \( y_j \in I(n, \Omega), 1 \leq j \leq n \). If the following separation condition is satisfied
\[
\min_{p \neq j} |y_p - y_j| > \frac{\pi s}{\Omega} \left( \frac{2n(s+1)}{\zeta(n)} \frac{\sigma}{m_{\min}} \right)^{1/2n-2},
\]
This is (5.2).
then

\[ \hat{\sigma}_n > (s + 1)\sigma, \quad \hat{\sigma}_j \leq (s + 1)\sigma, \quad j = n + 1, \ldots, s + 1. \]

Especially, when \( s = n \), the separation condition

\[ \min_{p \neq j} |y_p - y_j| \geq \frac{2\pi e \left( \frac{7}{\pi} \right)^{\frac{1}{2s-2}} \left( \frac{\sigma}{m_{\min}} \right)^{\frac{1}{2s-2}}}{\Omega}, \tag{5.3} \]

implies that

\[ \hat{\sigma}_n > (n + 1)\sigma, \quad \hat{\sigma}_{n+1} < (n + 1)\sigma. \]

Proof: Since \( ||W||_\infty \leq \sigma \), we have \( ||\Delta||_2 \leq ||\Delta||_F \leq (s + 1)\sigma \). Let

\[ \theta_{\min}(\Omega, s) = \min_{p \neq j} \left| y_p \Omega \frac{s}{s} - y_j \Omega \frac{s}{s} \right| = \frac{\Omega}{s} \min_{p \neq j} \left| y_p - y_j \right|. \]

The separation condition \( \text{(5.2)} \) implies

\[ \theta_{\min}(\Omega, s) \geq \pi \left( \frac{2n(s + 1) - \sigma}{\zeta(n)^2 m_{\min}} \right)^{\frac{1}{2s-2}}. \]

By Theorem \( 5.1 \) we have

\[ \sigma_n \geq \frac{m_{\min} \zeta(n)^2 \theta_{\min}(\Omega, s)^{2n-2}}{n \pi^{2n-2}} > 2(s + 1)\sigma \geq 2||\Delta||_2. \]

By Weyl’s theorem, we have that \( |\hat{\sigma}_n - \sigma_n| \leq ||\Delta||_2 \). Thus,

\[ \hat{\sigma}_n > 2||\Delta||_2 - ||\Delta||_2 \geq (s + 1)\sigma. \]

Similarly, we have for \( j = n + 1, \ldots, s + 1 \),

\[ |\hat{\sigma}_j| \leq ||\Delta||_2 \leq (s + 1)\sigma. \]

When \( s = n \), since

\[ \min_{p \neq j} \left| y_p - y_j \right| \geq \frac{2\pi e \left( \frac{7}{\pi} \right)^{\frac{1}{2s-2}} \left( \frac{\sigma}{m_{\min}} \right)^{\frac{1}{2s-2}}}{\Omega} \left( \frac{n \pi}{\zeta(n)^2} \frac{2n(s + 1) - \sigma}{m_{\min}} \right)^{\frac{1}{2s-2}}, \]

by the above argument, we have

\[ \hat{\sigma}_n > (n + 1)\sigma, \quad \hat{\sigma}_{n+1} < (n + 1)\sigma. \]

This completes the proof of the corollary.

Corollary \( 5.1 \) leads to the following number recovery algorithm \textbf{Algorithm 1}, where we also used a decimation strategy implemented in \( \text{[5, 26]} \).

Note that for the above algorithm to work, we need to input the integer \( s \) which requires a priori information on the number of spectral lines to be recovered. This information may not be available in some real applications. To remedy it, we propose a recursive MUSIC-type algorithm \textbf{Algorithm 2} below.
Algorithm 1: MUSIC-type number recovering algorithm

**Input:** Number \( s \), Noise level \( \sigma \)

**Input:** Measurements: \( Y = (Y(x_1), \cdots, Y(x_M))^T \)

1: \( r = (M-1) \mod 2s \), \( Y_{\text{new}} = (Y(x_1), Y(x_{r+1}), \cdots, Y(x_{2sr+1}))^T \);
2: Formulate the \( (s+1) \times (s+1) \) Hankel matrix \( \hat{X} \) from \( Y_{\text{new}} \), and compute the singular value of \( \hat{X} \) as \( \hat{\sigma}_1, \cdots, \hat{\sigma}_{s+1} \) distributed in a decreasing manner;
4: Determine \( n \) by \( \hat{\sigma}_n > (s+1)\sigma \) and \( \hat{\sigma}_j \leq (s+1)\sigma, j = n+1, \cdots, s+1; \)

Return: \( n \)

Algorithm 2: Recursive MUSIC-type number recovering algorithm

**Input:** Noise level \( \sigma \), Measurements: \( Y = (Y(x_1), \cdots, Y(x_M))^T \)

**Input:** \( n_{\text{max}} = 0, n = 1 \)

if \( n > n_{\text{max}} \) then
    \( n_{\text{max}} = n, n = n + 1; \)
    \( r = (M-1) \mod 2n; \)
    \( Y_{\text{new}} = (Y(x_1), Y(x_{r+1}), \cdots, Y(x_{2nr+1}))^T; \)
    Input \( n, \sigma, Y_{\text{new}} \) to Algorithm 1, save the output of Algorithm 1 as \( n; \)
else
    Return \( n \)

We next conduct numerical experiments to demonstrate the efficacy of the recursive MUSIC-type number recovering algorithm.

**Experiment:** We set \( n = 4, \Omega = 1, \sigma = 1 \times 10^{-7} \) and 
\[ \mu = \delta y_1 - \delta y_2 - \delta y_3 + \delta y_4 \]
where \( y_1 = -0.5, y_2 = 0, y_3 = 0.5, y_4 = 1 \). We measure at \( M = 20 \) sample points evenly spaced in \([-\Omega, \Omega]\). The noisy measurements are
\[ Y = (Y(x_1), Y(x_2), \cdots, Y(x_M))^T, \]
where \( Y(x_i) = \sum_{j=1}^n a_j e^{iy_j x_i} + W(x_i) \) with \( W(x_i) \) the noise satisfying \( ||W||_\infty < \sigma \). We apply our recursive MUSIC-type number recovering algorithm and get the recovered number \( n = 4 \).

We then apply the algorithm to spectral lines with different separation distance to find the minimum separation distance required for the success of the algorithm. Precisely, we set \( y_1 = -\tau, y_2 = 0, y_3 = \tau, y_4 = 2\tau \), and recover the number by Algorithm 2 with \( \tau \) varies from 0 to 1. We plot Figure 5.1 which illustrates the number recovered when this minimum separation distance varies. It show that we can exactly recover the source number when they are separated beyond 0.41 (\( \approx 0.13\frac{\pi}{\Omega} \)). In comparison, the upper bound for the resolution limit we derived in Theorem 2.1 is 0.462\( \frac{\pi}{\Omega} \).

5.2 Augmented MUSIC algorithm for support recovery

It is well-known that the MUSIC algorithm for recovering the supports is sensitive to a priori estimate of the model order (the number of spectral lines). With the recursive MUSIC-type
number recovery algorithm at hand, we can employ it to assist the traditional MUSIC algorithm to recover the supports of the spectral lines.

The new augmented MUSIC algorithm consists of three steps. In the first step we use Algorithm 2 to recover the number of the spectral lines. In the second step, we assemble the \((s + 1) \times (s + 1)\) Hankel matrix \(\hat{X}\) from the measurements \(Y = (Y(x_1), Y(x_2), \ldots, Y(x_M))^T\), where \(s = \lfloor \frac{M - 1}{2} \rfloor\), and perform the following singular value decomposition for \(\hat{X}\):

\[
\hat{X} = \hat{U} \hat{\Sigma} \hat{U}^* = [\hat{U}_1 \quad \hat{U}_2] \text{diag}(\hat{\sigma}_1, \hat{\sigma}_2, \ldots, \hat{\sigma}_n, \ldots, \hat{\sigma}_{s+1}) [\hat{U}_1 \quad \hat{U}_2]^*,
\]

where \(\hat{U}_1 = (\hat{U}(1), \ldots, \hat{U}(n))\), \(\hat{U}_2 = (\hat{U}(n + 1), \ldots, \hat{U}(s + 1))\). The orthogonal projection to the space \(\hat{U}_2\) is denoted as \(\hat{P}_2 x = \hat{U}_2 (\hat{U}_2^* x)\). Finally, for a test vector \(\Phi(\omega) = (1, e^{i\omega h}, \ldots, e^{i(s+1)\omega})^T\), we define the MUSIC imaging functional

\[
\hat{R}(\omega) = \frac{||\hat{P}_2 \Phi(\omega)||_2}{||\Phi(\omega)||_2} = \frac{||\hat{U}_2^* \Phi(\omega)||_2}{||\Phi(\omega)||_2},
\]

\[
\hat{J}(\omega) = \frac{||\hat{P}_2 \Phi(\omega)||_2}{||\hat{U}_2^* \Phi(\omega)||_2} = \frac{||\hat{U}_2^* \Phi(\omega)||_2}{||\Phi(\omega)||_2}.
\]

The local minimizers of \(\hat{\omega}_j\) of \(\hat{R}(\omega)\) (or maximizer of \(\hat{J}(\omega)\)) indicate the location of the spectral lines. We summarize the augmented MUSIC algorithm in Algorithm 3 below.

**Algorithm 3: Augmented MUSIC algorithm**

**Input**: Noise level \(\sigma\):

**Input**: Measurements: \(Y = (Y(x_1), \ldots, Y(x_M))^T\) with \(h\) the sampling distance

1: Recover the number \(n\) of spectral lines by Algorithm 2;
2: Let \(s = \lfloor \frac{M - 1}{2} \rfloor\), formulate the \((s + 1) \times (s + 1)\) Hankel matrix \(\hat{X}\) from \(Y\);
3: Compute the singular vector of \(\hat{X}\) as \(\hat{U}(1), \hat{U}(2), \ldots, \hat{U}(s+1)\) and formulate the noise space \(\hat{U}_2 = (\hat{U}(n + 1), \ldots, \hat{U}(s + 1))\);
4: For real number \(\omega\) in the searching region, denote the test vector \(\Phi(\omega) = (1, e^{i\omega h}, \ldots, e^{i(s+1)\omega})^T\). Plot the MUSIC imaging functional \(\hat{J}(\omega) = \frac{||\Phi(\omega)||_2}{||\hat{U}_2^* \Phi(\omega)||_2}\).
We now present several numerical experiments to demonstrate the efficiency of the augmented MUSIC algorithm and to confirm the theoretical results in Section 2.

**Experiment 1:** When the spectral lines are separated beyond the two resolution limits (or their upper bounds) to both the number and the support recovery in Section 2, we can expect an exact recovery of the spectral number and stable recovery of the supports. This is indeed the case in our experiment below.

We set \( n = 4, M = 20, \Omega = 1, \sigma = 10^{-7} \) and the collection of spectral lines is

\[ \mu = \delta_{y_1} - \delta_{y_2} - \delta_{y_3} + \delta_{y_4}, \]

where \( y_1 = -0.5, y_2 = 0, y_3 = 0.5, y_4 = 1 \). We measure at \( M = 20 \) sample points evenly spaced in \([-\Omega, \Omega]\). The noisy measurements are

\[ Y = (Y(x_1), Y(x_2), \cdots, Y(x_M))^T, \]

where \( Y(x_t) = \sum_{j=1}^{n} a_j e^{iy_j x_t} + W(x_t) \) with \( W(x_t) \) being the noise and satisfying \(|W|_\infty < \sigma\). We apply Algorithm 3 to the measurements and we plot the MUSIC imaging function \( \hat{J}(\omega) \), Figure 5.2a. This time the number recovered is 4 and we see that MUSIC algorithm locates the supports very well. The Rayleigh limit is \( \pi \), but we can resolve 4 spectral lines separated by 0.5 with high accuracy.

**Experiment 2:** As is mentioned in Remark 2.1 in Section 2, the resolution limit to the number recovery problem is better than the one for stable support recovery problem. It indicates that the separation distance required for Algorithm 2 for the number recovery is smaller than the one in Algorithm 3 for the support recovery. We demonstrate this by showing in the experiment below that Algorithm 3 may fail to recover the supports even when we can recover the exact number by Algorithm 2.

We set \( n = 4, M = 20, \Omega = 1, \sigma = 10^{-7} \) and

\[ \mu = \delta_{y_1} - \delta_{y_2} - \delta_{y_3} + \delta_{y_4}, \]

where \( y_1 = -0.42, y_2 = 0, y_3 = 0.42, y_4 = 0.84 \). We measure at \( M = 20 \) sample points evenly spaced in \([-\Omega, \Omega]\). The noisy measurements are

\[ Y = (Y(x_1), Y(x_2), \cdots, Y(x_M))^T. \]
This time, we can recover the exact number $n = 4$ by **Algorithm 2**. However, as is seen from Figure 5.2b, the MUSIC algorithm cannot extract all the four spectral lines.

**Experiment 3: Algorithm 2** determines the number by comparing the singular value to the threshold in Corollary 5.1. When this algorithm fails, it means that the $n$-th singular value $\sigma_n$ is buried under the threshold. The corresponding noiseless singular value $\sigma_n$ of the matrix $DAD^T$ in Theorem 5.1 will be comparable to some singular value generated from the noise matrix $\Delta$. This causes the signal vector $\hat{U}(n)$ to be indistinguishable to the noise space $\hat{U}_2$ and results in the failure of MUSIC algorithm to identify the supports. This is evident from the experiment below.

We consider the case when the spectral lines are closely distributed that **Algorithm 2** fails to recover the exact number. We set $n = 4, M = 20, \Omega = 1, \sigma = 10^{-7}$ and

$$\mu = \delta y_1 - \delta y_2 - \delta y_3 + \delta y_4,$$

where $y_1 = -0.39, y_2 = 0, y_3 = 0.39, y_4 = 0.78$. We measure at $M = 20$ sample points evenly spaced in $[-\Omega, \Omega]$. The noisy measurements are

$$Y = (Y(x_1), Y(x_2), \cdots, Y(x_M))^T.$$

The number recovered by **Algorithm 2** is $n = 3$. Figure 5.3a is the imaging function of **Algorithm 3**. It only exhibits three sharp peaks and hence fails to recover the four spectral lines. We now feed the MUSIC algorithm with the exact number $n = 4$, we plot the imaging function Figure 5.3b. It is clear that the four spectral lines cannot be recovered as well.

![Figure 5.3: MUSIC imaging function $\hat{J}(\omega)$](image)

### 5.3 Phase transition

We know from Section 2 that the resolution limit to the number recovery problem is bounded from below and upper by $C_1 \Omega \left( \frac{\sigma}{m_{\min}} \right) \frac{\pi}{d_{\min}}$ and $C_2 \Omega \left( \frac{\sigma}{m_{\min}} \right) \frac{\pi}{d_{\min}}$ respectively for some constants $C_1, C_2$. We shall demonstrate that this implies a phase transition phenomenon for the number recovery. Precisely, recall that the super-resolution factor is $\text{SRF} = \pi \left( \frac{m_{\min}}{\sigma} \right)$ and the signal-to-noise ratio is $\text{SNR} = \frac{m_{\min}}{\sigma}$. From the two bounds for the resolution limit, we can draw the conclusion
that exact number recovery is guaranteed if
\[
\log(SNR) > (2n - 2) \log(SRF) + (2n - 2) \log \frac{C_1}{\pi},
\]
and may fail if
\[
\log(SNR) < (2n - 2) \log(SRF) + (2n - 2) \log \frac{C_2}{\pi}.
\]
As a consequence, we can see that in the parameter space of \(\log SNR - \log SRF\), there exist two lines both with slope \(2n - 2\) such that the number recovery is successful for cases above the first line and unsuccessful for cases below the second. In the intermediate region between the two lines, the number recovery can be either successful or unsuccessful from case to case. This is clearly demonstrated in the numerical experiments below.

We fix \(\Omega = 1\) and consider \(n\) spectral lines equally spaced in \([-\frac{(n-1)\pi}{2},\frac{(n-1)\pi}{2}]\) by \(d_{\min}\) with amplitudes \(a_j\), and the noise level is \(\sigma\). We perform 5000 random experiments (the randomness is in the choice of \((d_{\min}, \sigma, y_j, a_j)\)) to recover the number based on Algorithm 2. Figure 5.4 shows the results for \(n = 2, 4\) and the two lines of slope \(2n - 2\) strictly separate the blue points (successful cases) and red points (unsuccessful cases) respectively, and in-between is the phase transition region.

We now consider the support recovery. As shown in Section 2, there are lower and upper bounds for the resolution limit, which implies a phase transition for successful support recovery. Similar to the number recovery, we see that in the parameter space of \(\log SNR - \log SRF\), there exist two lines both with slope \(2n - 1\) such that the support recovery is successful for cases above the first line and unsuccessful for cases below the second. In the intermediate region between the two lines, the support recovery can be either successful or unsuccessful.
from case to case. This is clearly demonstrated in the numerical experiments below where
used the augmented MUSIC algorithm.

We fix \( \Omega = 1 \) and consider \( n \) spectral lines equally spaced in \([-\frac{(n-1)\pi}{2}, \frac{(n-1)\pi}{2}]\) by \( d_{\text{min}} \) with amplitudes \( a_j \), and the noise level is \( \sigma \). We perform 5000 random experiments (the randomness is in the choice of \( (d_{\text{min}}, \sigma, y_j, a_j) \) to recover the spectral lines and we summarize the
detail of a single experiment in Algorithm 4. As shown in Figure 5.5, two lines with slope
\( 2n - 1 \) strictly separate the blue points (successful cases) and red points (unsuccessful cases)
separately, and in-between is the phase transition region.

Algorithm 4: A single experiment

```
\textbf{Input:} Sources \( \mu = \sum_{j=1}^{n} a_j \delta_{y_j} \), Noise level \( \sigma \)
\textbf{Input:} Measurements: \( Y = (Y(x_1), \ldots, Y(x_M))^T \)
1: Compute the imaging function \( \hat{J}(\omega) \) by Algorithm 3;
2: Determine the supports, denoted by \( RM \), by spectrum peak search in \( \hat{J}(\omega) \);
3: Successnumber = 0;
\textbf{for each} 1 \leq j \leq n \textbf{do}
    \quad Compute the error for the spectral line \( y_j \): \( e_j := \min_{\hat{y}_l \in RM} |\hat{y}_l - y_j| \);
    \quad The spectral line \( y_j \) is recovered successfully if \( e_j < \min_{p \neq j} |y_p - y_j| / 3 \);
    \quad and \( \text{Successnumber} = \text{Successnumber} + 1 \);
\textbf{if} Successnumber = n \textbf{then}
    \quad Return SUCCESS
\textbf{else}
    \quad Return FAIl
```

6 Conclusion

In this paper, we introduced two resolution limit for the number recovery problem and the
support recovery problem respectively in the LSE. We quantitatively characterized the two
limits by establishing their sharp upper and lower bounds for a cluster of spectral lines which
are bounded by multiple of Rayleigh limit. We developed simple MUSIC-type algorithm for
the number recovery problem with theory guarantee. This algorithm is incorporated to the
usual MUSIC algorithm for support recovery. The numerical experiments showed the effi-
ciency of these algorithms, adding one more example to the many success of applying vari-
ous subspace methods to the super-resolution problems. In addition, phase transition phe-
nomenon in both recovery problems are also demonstrated, confirming our theory and also
revealing the subtle difference between the two.

The results offer a starting point for several interesting topics in the future research. First,
the extension to multiple clusters with a fast and efficient algorithm taking advantage of the
separation of clusters. Second, the extension to multiple-dimensions. And finally, a com-
prehensive comparison, both theoretically and numerically, of various super-resolution algorithms, showing their advantages and disadvantages in various cases.

7 Proofs of main results

7.1 Proof of Theorem 2.1

For ease of explanation, we denote \( d_j = e^{iy_j \frac{2\pi}{M}} \) and \( \hat{d}_j = e^{i\hat{y}_j \frac{2\pi}{M}} \). Recall that \( h = \frac{2\Omega}{M-1} \).

Step 1. We write

\[
M = (2n - 1)r + q, \quad (7.1)
\]

where \( r, q \) are integers with \( r \geq 1 \) and \( 0 \leq q < 2n - 1 \). We first show that for \( y_j \in I(n, \Omega), 1 \leq j \leq n \), there exists some \( \theta \in (-\pi, 0) \) such that

\[
e^{iy_j \frac{2\Omega}{M-1}} \in \Gamma^+(\theta), \quad 1 \leq j \leq n. \quad (7.2)
\]

Indeed, by (7.1) \( \frac{2r}{M-1} \leq \frac{1}{n-1} \). Therefore we have

\[
\max_{p, q \in I(n, \Omega)} |p \frac{2r\Omega}{M-1} - q \frac{2r\Omega}{M-1}| \leq \pi.
\]

This implies that there exists \( \theta \in (-\pi, 0) \) such that (7.2) holds.

Step 2. By Lemma 8.2 we have

\[
\frac{2.22e}{n} \geq \left( \frac{\sqrt{2n - 1}}{\xi(n)\xi(n-1)} \right)^{\frac{1}{2n-2}},
\]
Thus, the separation condition (2.1) implies that

$$\frac{2.22\pi e}{\Omega} \left( \frac{2\sigma}{\min} \right)^{\frac{1}{2n-2}} \geq \frac{\pi n}{\Omega} \left( \frac{\sqrt{2n-1} - \frac{2\sigma}{\min}}{\zeta(n) \zeta(n-1)} \right)^{\frac{1}{2n-2}} \geq \frac{\pi (M-1)}{2r \Omega} \left( \frac{2\sqrt{2n-1} - \sigma}{\zeta(n) \zeta(n-1) \min} \right)^{\frac{1}{2n-2}}.$$ 

Thus, the separation condition (2.1) implies that

$$\theta_{\min} := \min_{p \neq j} \left| \frac{y_p}{M - 1} \right| \geq \frac{2r \Omega}{2.22\pi e} \left( \frac{2\sigma}{\min} \right)^{\frac{1}{2n-2}} \geq \frac{\pi \left( \frac{2\sqrt{2n-1} - \sigma}{\zeta(n) \zeta(n-1) \min} \right)^{\frac{1}{2n-2}}}{M - 1}. \quad (7.3)$$

**Step 3.** For $\hat{\mu} = \sum_{j=1}^k \hat{a}_j \hat{y}_j$ with $k < n$. Observe that that

$$[\hat{\mu}] - [\mu] = \hat{B} \hat{a} - B a,$$

where $\hat{a} = (\hat{a}_1, \cdots, \hat{a}_k)^T$, $a = (a_1, \cdots, a_n)^T$ and

$$\hat{B} = \begin{pmatrix} e^{-i \hat{y}_1 \Omega} & \cdots & e^{-i \hat{y}_n \Omega} \\ e^{-i \hat{y}_1 \Omega + i \hat{y}_1 \Omega} & \cdots & e^{-i \hat{y}_1 \Omega + i \hat{y}_n \Omega} \\ e^{-i \hat{y}_1 \Omega + i \hat{y}_1 (M-2) \Omega} & \cdots & e^{-i \hat{y}_1 \Omega + i \hat{y}_n (M-2) \Omega} \\ \vdots & \ddots & \vdots \\ e^{-i \hat{y}_1 \Omega + i \hat{y}_1 (2n-3) \Omega} & \cdots & e^{-i \hat{y}_1 \Omega + i \hat{y}_n (2n-3) \Omega} \\
 e^{-i \hat{y}_1 \Omega + i \hat{y}_1 (2n-2) \Omega} & \cdots & e^{-i \hat{y}_1 \Omega + i \hat{y}_n (2n-2) \Omega} \end{pmatrix}, B = \begin{pmatrix} e^{-i y_1 \Omega} & \cdots & e^{-i y_n \Omega} \\ e^{-i y_1 \Omega + i y_1 \Omega} & \cdots & e^{-i y_1 \Omega + i y_n \Omega} \\ e^{-i y_1 \Omega + i y_1 (M-2) \Omega} & \cdots & e^{-i y_1 \Omega + i y_n (M-2) \Omega} \\ \vdots & \ddots & \vdots \\ e^{-i y_1 \Omega + i y_1 (2n-3) \Omega} & \cdots & e^{-i y_1 \Omega + i y_n (2n-3) \Omega} \\
 e^{-i y_1 \Omega + i y_1 (2n-2) \Omega} & \cdots & e^{-i y_1 \Omega + i y_n (2n-2) \Omega} \end{pmatrix}.$$

Extracting row vectors from $\hat{B}$ and $B$ respectively, we assemble the following two partial matrices

$$\hat{B}_1 = \begin{pmatrix} e^{-i \hat{y}_1 \Omega} & \cdots & e^{-i \hat{y}_n \Omega} \\ e^{-i \hat{y}_1 \Omega + i \hat{y}_1 r \Omega} & \cdots & e^{-i \hat{y}_1 \Omega + i \hat{y}_n r \Omega} \\ e^{-i \hat{y}_1 \Omega + i \hat{y}_1 (2n-3) r \Omega} & \cdots & e^{-i \hat{y}_1 \Omega + i \hat{y}_n (2n-3) r \Omega} \\
 e^{-i \hat{y}_1 \Omega + i \hat{y}_1 (2n-2) r \Omega} & \cdots & e^{-i \hat{y}_1 \Omega + i \hat{y}_n (2n-2) r \Omega} \end{pmatrix}, B_1 = \begin{pmatrix} e^{-i y_1 \Omega} & \cdots & e^{-i y_n \Omega} \\ e^{-i y_1 \Omega + i y_1 r \Omega} & \cdots & e^{-i y_1 \Omega + i y_n r \Omega} \\ e^{-i y_1 \Omega + i y_1 (2n-3) r \Omega} & \cdots & e^{-i y_1 \Omega + i y_n (2n-3) r \Omega} \\
 e^{-i y_1 \Omega + i y_1 (2n-2) r \Omega} & \cdots & e^{-i y_1 \Omega + i y_n (2n-2) r \Omega} \end{pmatrix}.$$

It is clear that

$$\min_{\hat{a} \in \mathbb{C}^k, \hat{y}_j \in \mathbb{R}, j=1, \ldots, k} ||\hat{B}_1 \hat{a} - B_1 a||_\infty \leq \min_{\hat{a} \in \mathbb{C}^k, \hat{y}_j \in \mathbb{R}, j=1, \ldots, k} ||\hat{B} \hat{a} - B a||_\infty = \min_{\hat{a} \in \mathbb{C}^k, \hat{y}_j \in \mathbb{R}, j=1, \ldots, k} ||[\hat{\mu}] - [\mu]||_\infty. \quad (7.4)$$

**Step 4.** We consider the optimization problem

$$\min_{a \in \mathbb{C}^k, \hat{y}_j \in \mathbb{R}, j=1, \ldots, k} ||\hat{B}_1 a - B_1 a||_2,$$

where $a = (a_1, \cdots, a_n)^T$. Note that

$$\hat{B}_1 = (\phi_{2n-2}(\hat{d}_1), \cdots, \phi_{2n-2}(\hat{d}_k)) \text{diag}(e^{-i \hat{y}_1 \Omega}, \cdots, e^{-i \hat{y}_n \Omega}),$$

$$B_1 = (\phi_{2n-2}(d_1), \cdots, \phi_{2n-2}(d_n)) \text{diag}(e^{-i y_1 \Omega}, \cdots, e^{-i y_n \Omega}). \quad (7.5)$$
where \( \phi_{2n-2}(\omega) \) is defined as in (4.1). We have
\[
\min_{a \in C^k, \gamma \in \mathbb{R}, j = 1, \ldots, k} ||\hat{B}_1 \alpha - B_1 a||_2 = \min_{a \in C^k, j \in \mathbb{R}, j = 1, \ldots, k} ||\hat{D}a - D\tilde{a}||_2,
\]
where \( \tilde{a} = (a_1 e^{-iy_1}, \ldots, a_n e^{-iy_n})^T, \hat{D} = (\phi_{2n-2}(d_1), \ldots, \phi_{2n-2}(d_k)) \) and \( D = (\phi_{2n-2}(d_1), \ldots, \phi_{2n-2}(d_n)) \). Using (7.6), we can apply Theorem 4.1 to get
\[
\min_{a \in C^k, j \in \mathbb{R}, j = 1, \ldots, k} ||\hat{D}a - D\tilde{a}||_2 \geq \frac{m_{\min}(\zeta(n)(n-1)(\theta_{\min})^{2n-2}}{\pi^{2n-2}},
\]
which yields
\[
\min_{a \in C^k, j \in \mathbb{R}, j = 1, \ldots, k} ||\hat{B}_1 \alpha - B_1 a||_2 \geq \frac{m_{\min}(\zeta(n)(n-1)(\theta_{\min})^{2n-2}}{\pi^{2n-2}}.
\]
It follows that
\[
\min_{a \in C^k, j \in \mathbb{R}, j = 1, \ldots, k} ||\hat{B}_1 \alpha - B_1 a||_\infty \geq \frac{1}{\sqrt{2n-1}} \min_{a \in C^k, j \in \mathbb{R}, j = 1, \ldots, k} ||\hat{B}_1 \alpha - B_1 a||_2
\]
\[
\geq \frac{m_{\min}(\zeta(n)(n-1)(\theta_{\min})^{2n-2}}{\sqrt{2n-1}\pi^{2n-2}}
\]
\[
\geq 2\sigma. \quad \text{(by 7.3)} \tag{7.6}
\]

**Step 5.** Combining (7.4) and (7.6), we get
\[
\min_{a \in C^k, j \in \mathbb{R}, j = 1, \ldots, k} ||[\hat{\mu}] - [\mu]||_\infty \geq 2\sigma.
\]
Therefore
\[
||[\hat{\mu}] - Y||_\infty = ||[\hat{\mu}] - [\mu] - W||_\infty
\]
\[
\geq ||[\hat{\mu}] - [\mu]||_\infty - ||W||_\infty \geq ||[\hat{\mu}] - [\mu]||_\infty - \sigma \geq \sigma,
\]
which shows that \( \hat{\mu} \) cannot be a \( \sigma \)-admissible measure. This completes the proof of the theorem.

### 7.2 Proof of Proposition 2.1

For \( \gamma = \sum_{j=1}^n a_j \delta_j \),
\[
\mathcal{F}(\gamma)(x) = \sum_{j=1}^n a_j e^{ij \omega j x} = \sum_{j=1}^n a_j \sum_{k=0}^\infty \frac{(i \omega j x)^k}{k!} = \sum_{k=0}^\infty m_k(\omega) \frac{(i \omega j x)^k}{k!}, \tag{7.7}
\]
where \( m_k(\omega) = \sum_{j=1}^n a_j j^k \) (see also (3.4)). Let \( t_1 = -\frac{(n-1)r}{\Omega}, t_2 = -\frac{(n-2)r}{\Omega}, \ldots, t_n = \frac{r}{\Omega}, \ldots, t_{2n-1} = \frac{(n-1)r}{\Omega} \). Considering the linear system that
\[
Aa = 0 \tag{7.8}
\]
Thus for \( \gamma = \sum_{j=1}^{2n-1} a_j \delta_{t_j} \) we have

\[
m_k(\gamma) = 0, k = 0, \ldots, 2n-3, \quad |m_k(\gamma)| \leq 2m(n-1)k^\frac{(r-1)}{\Omega}, \quad k \geq 2n - 2.
\]

Together with Taylor series (7.7), we have for \(|x| \leq \Omega,\)

\[
|\mathcal{F}(\gamma)(x)| \leq \sum_{k \geq 2n-2} 2m(n-1)k^\frac{r}{\Omega}\frac{\Omega^k}{k!} \leq \sum_{k \geq 2n-2} 2m(n-1)k^\frac{r}{\Omega}\frac{\Omega^k}{k!} \leq \frac{2m(n-1)^2 \pi^{2n-2}}{(2n-2)!} \sum_{k=0}^{\infty} \frac{(n-1)^r k^r \Omega^k}{k!} = \frac{2m(n-1)^2 \pi^{2n-2}}{(2n-2)!} e^{\frac{(n-1)^r}{\Omega}} \\
\leq \frac{m}{\sqrt{\pi} (n-1)^2 e^{\frac{(n-1)^r}{\Omega}}} \leq \sigma \quad \text{(by Lemma 8.1)}
\]

Take

\[
\mu = \sum_{j=1}^{n} a_j \delta_{t_j}, \quad \tilde{\mu} = \sum_{j=n+1}^{2n-1} -a_j \delta_{t_j},
\]

then the above estimate yields

\[
||[\tilde{\mu}] - [\mu]||_\infty < \sigma.
\]

### 7.3 Proof of Theorem 2.2

For ease of explanation, we denote \( d_j = e^{iy_j \frac{2\alpha}{M-1}}, \quad \hat{d}_j = e^{iy_j \frac{2\alpha}{M-1}}, \quad \text{and} \quad h = \frac{2\alpha}{M-1}. \)

**Step 1.** Similar to the proof of (7.2), we first write

\[
M = 2nr + q,
\]

where \( r, q \) are integers with \( r \geq 1 \) and \( 0 \leq q < 2n - 1 \). We can show that there exists \( \theta \in [-\pi, 0) \) such that

\[
d_j \in \Gamma^+(\theta), \quad \hat{d}_j \in \Gamma^+(\theta), \quad 1 \leq j, p \leq n.
\]

**Step 2.** Using the separation condition (2.3), Lemma 8.3 and the fact that \( \frac{M-1}{r} \leq 2n+1 \), we can derive that

\[
\theta_{\min} := \min_{p \neq j} |y_p \frac{2r\Omega}{M-1} - y_j \frac{2r\Omega}{M-1}| \geq \pi \left( \frac{8\sqrt{2n}}{\zeta(n)\omega(m_{\min})} \right)^{\frac{1}{M-1}},
\]

26
where
\[\zeta(n) = \begin{cases} 
\left(\frac{n}{2}\right)! \left(\frac{n-2}{2}\right)! & \text{even } n, \\
\left(\frac{n-1}{2}\right)! & \text{odd } n,
\end{cases} \]
\[\lambda(n) = \begin{cases} 
1, & n = 2, \\
\zeta(n-2), & n \geq 3, \\
\frac{1}{\pi}, & n = 3, \\
\frac{2n!}{(2n+1)!}, & n \geq 5, \\
\frac{2n!}{4}, & n \geq 4.
\end{cases} \]

Equivalently, we have
\[
(\theta_{\min})^{2n-1} \geq \frac{\pi^{2n-1} 8\sqrt{2n}}{\zeta(n) \lambda(n)} \frac{\sigma}{m_{\min}}. \tag{7.12}
\]

**Step 3.** Denote \( \hat{a} = (\hat{a}_1, \cdots, \hat{a}_n)^T, a = (a_1, \cdots, a_n)^T, \) and
\[
\hat{B}_1 = \begin{pmatrix}
e^{-i\hat{\gamma}_i \Omega} & \cdots & e^{-i\hat{\gamma}_n \Omega} \\
e^{-i\hat{\gamma}_i \Omega + i\hat{\gamma}_j r h} & \cdots & e^{-i\hat{\gamma}_n \Omega + i\hat{\gamma}_j r h} \\
\vdots & \ddots & \vdots \\
e^{-i\hat{\gamma}_i \Omega + i\hat{\gamma}_j (2n-2) r h} & \cdots & e^{-i\hat{\gamma}_n \Omega + i\hat{\gamma}_j (2n-2) r h} \\
e^{-i\hat{\gamma}_i \Omega + i\hat{\gamma}_j (2n-1) r h} & \cdots & e^{-i\hat{\gamma}_n \Omega + i\hat{\gamma}_j (2n-1) r h}
\end{pmatrix}, B_1 = \begin{pmatrix}
e^{-i\gamma_1 \Omega} & \cdots & e^{-i\gamma_n \Omega} \\
e^{-i\gamma_1 r h} & \cdots & e^{-i\gamma_n r h} \\
\vdots & \ddots & \vdots \\
e^{-i\gamma_1 (2n-2) r h} & \cdots & e^{-i\gamma_n (2n-2) r h} \\
e^{-i\gamma_1 (2n-1) r h} & \cdots & e^{-i\gamma_n (2n-1) r h}
\end{pmatrix}.
\]

Similar to the argument in the step 3 in the proof of Theorem 2.1, we can show that
\[
||\hat{B}_1 \hat{a} - B_1 a||_\infty \leq ||\hat{\mu} - \mu||_\infty.
\]

Since \( \hat{\mu} = \sum_{j=1}^n \hat{a}_j \hat{\gamma}_j \) is a \( \sigma \)-admissible measure, we have
\[
|||\hat{\mu} - \mathbf{Y}|||_\infty < \sigma.
\]

On the other hand,
\[
|||\hat{\mu} - \mathbf{Y}|||_\infty = |||\hat{\mu} - \mu - (\mathbf{W})|||_\infty \geq |||\hat{\mu} - \mu|||_\infty - |||\mathbf{W}|||_\infty.
\]

Therefore \( |||\hat{\mu} - \mu|||_\infty < 2\sigma \). It follows that
\[
||\hat{B}_1 \hat{a} - B_1 a||_\infty < 2\sigma,
\]
whence we get
\[
||\hat{B}_1 \hat{a} - B_1 a||_2 \leq \sqrt{2\pi ||\hat{B}_1 \hat{a} - B_1 a||_\infty} < 2\sqrt{2n\pi} \sigma. \tag{7.13}
\]

**Step 4.** Consider the optimization problem
\[
\min_{a \in \mathbb{C}^n, \gamma_j \in \mathbb{I}(n, \Omega), j = 1, \cdots, n} ||\hat{B}_1 a - B_1 a||_2.
\]

Using the same decomposition as in (7.5) for \( \hat{B}_1 \) and \( B_1 \), we have
\[
\min_{a \in \mathbb{C}^n, \gamma_j \in \mathbb{I}(n, \Omega), j = 1, \cdots, n} ||\hat{B}_1 a - B_1 a||_2 = \min_{a \in \mathbb{C}^n, \gamma_j \in \mathbb{I}(n, \Omega), j = 1, \cdots, n} ||\hat{D} a - D \hat{a}||_2
\]
where \( \hat{a} = (a_1 e^{-i\gamma_1 \Omega}, \cdots, a_n e^{-i\gamma_n \Omega})^T, \hat{D} = (\phi_{2n-1}(\hat{d}_1), \cdots, \phi_{2n-1}(\hat{d}_n)) \) and \( D = (\phi_{2n-1}(d_1), \cdots, \phi_{2n-1}(d_n)) \).

Note that (7.10) holds. We can apply Theorem 4.2 to get
\[
\min_{a \in \mathbb{C}^n, \gamma_j \in \mathbb{I}(n, \Omega), j = 1, \cdots, n} ||\hat{D} a - D \hat{a}||_2 \geq \frac{m_{\min}(\lambda(n))||\eta||_\infty(\theta_{\min})^{n-1}}{2^n \pi^{n-1}},
\]
which completes the proof.
where \( \theta_{\min} = \min_{p \neq j} |y_p^{2r \Omega/M-1} - y_j^{2r \Omega/M-1}| \) and

\[
\eta = \begin{pmatrix}
|\Pi^p_{j=1}(d_1 - \hat{d}_j)| \\
\vdots \\
|\Pi^p_{j=1}(d_n - \hat{d}_j)|
\end{pmatrix}.
\]

(7.14)

It follows that

\[
\min_{\alpha \in \mathbb{C}^n, \hat{\eta}_j \in \mathbb{I}(n, \Omega), j=1, \ldots, n} ||\hat{B}_1 \alpha - B_1 a||_2 \geq \frac{m_{\min} \zeta(n)||\eta||_\infty (\theta_{\min})^{n-1}}{2^n \pi^{n-1}}.
\]

Combining this with (7.13), we get

\[
||\eta||_\infty < \frac{\sqrt{2n} 2^{n+1} \pi^{n-1}}{\zeta(n)(\theta_{\min})^{n-1}} \frac{\sigma}{m_{\min}}.
\]

(7.15)

**Step 5.** Recall that (see (4.3))

\[
\angle(d_p d_j) = \left| y_p^{2r \Omega/M-1} - y_j^{2r \Omega/M-1} \right|.
\]

(7.16)

We prove that

\[
\angle(\hat{d}_j d_j) < \frac{2^{n+1} \pi^{2n-2} \sqrt{2n}}{\zeta(n)(n-2)!} \left( \frac{1}{\theta_{\min}} \right)^{2n-2} \frac{\sigma}{m_{\min}},
\]

(7.17)

for the case \( n = 2 \). We first claim that for each \( d_j \), there is one \( \hat{d}_p \) such that \( \angle(\hat{d}_p d_j) < \frac{\theta_{\min}}{2} \). Otherwise, WLOG, we assume

\[
\angle(\hat{d}_1 d_1) \geq \frac{\theta_{\min}}{2}, \quad \angle(\hat{d}_2 d_1) \geq \frac{\theta_{\min}}{2}.
\]

Using (4.4), the above inequality yields

\[
|(d_1 - \hat{d}_1)(d_1 - \hat{d}_2)| \geq \left( \frac{2}{\pi} \right)^2 \angle(\hat{d}_1 d_1) \angle(\hat{d}_2 d_1) \geq \frac{(2\theta_{\min})^2}{4} = \frac{(\theta_{\min})^2}{\pi^2}.
\]

(7.18)

On the other hand, letting \( n = 2 \) in (7.12), we have

\[
(\theta_{\min})^3 \geq \frac{\pi^{2n-1} 8 \sqrt{2n}}{\zeta(n)} \frac{\sigma}{m_{\min}}.
\]

Therefore

\[
\frac{(\theta_{\min})^2}{\pi^2} \geq \frac{\pi^{n-1} 8 \sqrt{2n}}{\zeta(n) \theta_{\min}} \frac{\sigma}{m_{\min}}.
\]

Recall (7.18), since \( n = 2 \), it follows that

\[
||\eta||_\infty \geq |(d_1 - \hat{d}_1)(d_1 - \hat{d}_2)| \geq \frac{(\theta_{\min})^2}{\pi^2} \geq \frac{\sqrt{2n} 2^{n+1} \pi^{n-1}}{\zeta(n)(\theta_{\min})^{n-1}} \frac{\sigma}{m_{\min}}.
\]
This contradicts (7.15) and hence proves the claim for $d_1$. Similarly, we can prove the claim for $d_2$. As a result, we have
\[ \angle(d_j d_j) < \frac{\theta_{\min}}{2}, \quad j = 1, 2, \]
which further implies that $\angle(d_2 d_1) \geq \frac{\theta_{\min}}{2}$.

On the other hand,
\[ \angle(d_1 d_1) \angle(d_2 d_1) \leq \frac{\pi}{2} |d_1 - d_1||d_2 - d_1| \leq \left( \frac{\pi}{2} \right)^2 ||\eta||_\infty \leq \frac{\sqrt{2n2\pi^{2n-1}}}{\zeta(n)(\theta_{\min})^{n-1}} \frac{\sigma}{m_{\min}}. \]

Therefore
\[ \angle(d_1 d_1) \leq \frac{1}{\angle(d_2 d_1)} \frac{\sqrt{2n2\pi^{2n-1}}}{\zeta(n)(\theta_{\min})^{n-1}} \frac{\sigma}{m_{\min}} \leq \frac{\sqrt{2n2\pi^{2n-1}}}{\zeta(n)(\theta_{\min})^{2n-2}} \frac{\sigma}{m_{\min}}. \]

Similarly, the inequality holds for $\angle(d_2 d_2)$.

**Step 6.** We prove (7.17) for the case when $n \geq 3$ in this step.

**Step 6.1.** We first claim that for each $d_j$, there is one $d_p$ satisfies $\angle(d_j d_p) < \frac{\theta_{\min}}{2}$. We prove the claim by excluding two cases.

**Case 1:** There exists $p_0$ such that $\angle(d_j d_{p_0}) \geq \frac{\theta_{\min}}{2}, j = 1, \ldots, n$.

Denote $\hat{\eta} = \eta_{n-1}(d_1, \ldots, d_n, \hat{d}_1, \ldots, \hat{d}_{p_0-1}, \hat{d}_{p_0+1}, \ldots, \hat{d}_n)$. By lemma 4.6, we have
\[ ||\hat{\eta}||_\infty \geq \zeta(n-1) \left( \frac{2}{\pi} \theta_{\min} \right)^{n-1}. \]

Note that $\eta = \text{diag}(d_1 - d_{p_0}, \ldots, d_n - d_{p_0}) \hat{\eta}$ and that $|d_j - \hat{d}_{p_0}| \geq \frac{\theta_{\min}}{2}$, we have
\[
||\eta||_\infty \geq \frac{1}{2} \left( \frac{2}{\pi} \right) ||\hat{\eta}||_\infty \geq \frac{\zeta(n-1)}{2} \left( \frac{2}{\pi} \right)^n \theta_{\min}^n \geq \frac{\zeta(n-2)}{4} \left( \frac{2}{\pi} \right)^n \theta_{\min}^n \geq \frac{\sqrt{2n2\pi^{2n-1}}}{\zeta(n)(\theta_{\min})^{n-1}} \frac{\sigma}{m_{\min}}, \quad \text{by (7.19)}
\]

which contradicts to (7.15).

**Case 2:** There exist $p_1, p_2$, and $j_0$ such that $\angle(d_{j_0} \hat{d}_{p_1}) < \frac{\theta_{\min}}{2}, \angle(d_{j_0} \hat{d}_{p_2}) < \frac{\theta_{\min}}{2}$.

Then for all $j \neq j_0$, we have $\angle(d_j \hat{d}_{p_1}) \angle(d_j \hat{d}_{p_2}) \geq \frac{\theta_{\min}^2}{4}$ and consequently,
\[ |(d_j - \hat{d}_{p_1})(d_j - \hat{d}_{p_2})| \geq \frac{2}{\pi} \angle(d_j \hat{d}_{p_1}) \angle(d_j \hat{d}_{p_2}) \geq \frac{1}{4} \left( \frac{2}{\pi} \right)^2 \theta_{\min}^2. \]

Denote
\[ \hat{\eta} = \eta_{n-1, n-2}(d_1, \ldots, d_{j_0-1}, d_{j_0+1}, \ldots, d_n, \hat{d}_1, \ldots, \hat{d}_{p_1-1}, \hat{d}_{p_1+1}, \ldots, \hat{d}_{p_2-1}, \hat{d}_{p_2+1}, \ldots, \hat{d}_n). \]

Applying Lemma 4.6 to the $n - 1$ points $d_1, \ldots, d_{j_0-1}, d_{j_0+1}, \ldots, d_n$, we have
\[ ||\hat{\eta}||_\infty \geq \zeta(n-2) \left( \frac{2}{\pi} \theta_{\min} \right)^{n-2}. \]
Note that the components of $\tilde{\eta}$ differ from those of $\eta$ only by the factors $|(d_j - \hat{d}_p)(d_j - d_{p_2})|$ for $j = 1, \ldots, j_0 - 1, j_0 + 1, \ldots, n$. Using (7.21), (7.20) and (7.12), we have

$$
||\eta||_\infty \geq \frac{1}{4}\left(-\frac{2\theta_{\min}}{\pi}\right)^2||\tilde{\eta}\|_\infty
$$

$$
\geq \frac{\xi(n-2)}{4}\left(-\frac{2\theta_{\min}}{\pi}\right)^n \geq \frac{\sqrt{2n^2n^{n-1}}}{\zeta(n)(\theta_{\min})^{n-1}} \frac{\sigma}{m_{\min}},
$$

which contradicts to (7.15) and hence proves the claim.

**Step 6.2.** By the result in step 6.1, we have for $1 \leq p \leq n$, $\angle(d_p \hat{d}_p) < \frac{\theta_{\min}}{2}$. Thus for $p < j$, since $\angle(d_j d_p) \geq (j-p)\theta_{\min}$, we have $\angle(d_j \hat{d}_p) \geq \angle(d_j d_p) - \angle(d_p \hat{d}_p) \geq (j-p-\frac{1}{2})\theta_{\min}$. Similar result also holds for $p > j$. Therefore

$$
\angle(d_j \hat{d}_p) \geq \begin{cases} (j-p-\frac{1}{2})\theta_{\min} & p < j, \\ (p-j-\frac{1}{2})\theta_{\min} & p > j. 
\end{cases} \quad (7.22)
$$

We next show that

$$
|(d_j - \hat{d}_1) \cdots (d_j - \hat{d}_n)| \geq |d_j - \hat{d}_j|(\frac{\theta_{\min}}{\pi})^{\frac{1}{n-1}(n-2)!}, \quad j = 1, 2, \cdots, n. \quad (7.23)
$$

Indeed, for $2 \leq j \leq n-1$, we have

$$
|(d_j - \hat{d}_1) \cdots (d_j - \hat{d}_n)| = |(d_j - \hat{d}_j)|\Pi_{1 < p \leq j-1}|(d_j - \hat{d}_p)|\Pi_{j+1 \leq p \leq n}|(d_j - d_{p_2})|
$$

$$
\geq |(d_j - \hat{d}_j)|\left(\frac{2}{\pi}\right)^{n-1}\Pi_{1 < p \leq j-1}\angle(d_j \hat{d}_p)\Pi_{j+1 \leq p \leq n}\angle(d_j d_p) \quad \text{(by 4.4)}
$$

$$
\geq |(d_j - \hat{d}_j)|\left(\frac{2}{\pi}\right)^{n-1}\left(\Pi_{1 < p \leq j-1}\frac{2(j-p)-1}{2}\theta_{\min}\right)\left(\Pi_{j+1 \leq p \leq n}\frac{2(p-j)-1}{2}\theta_{\min}\right) \quad \text{(by 7.22)}
$$

$$
= |d_j - \hat{d}_j|(\frac{2\theta_{\min}}{\pi})^{n-1}(2j-3)!!(2(n-j)-1)!!
$$

$$
\geq |d_j - \hat{d}_j|(\frac{\theta_{\min}}{\pi})^{n-1}(n-2)! \quad \text{(since (2j-3)!!(2(n-j)-1)!! \geq (n-2)!)}
$$

Similarly, we can prove (7.23) for $j = 1$ and $j = n$. Combining (7.23) and (7.15), we further get

$$
|d_j - \hat{d}_j|(\frac{\theta_{\min}}{\pi})^{n-1}(n-2)! < \frac{\sqrt{2n^2(n^{n-1})}}{\zeta(n)(\theta_{\min})^{n-1}} \frac{\sigma}{m_{\min}}, \quad j = 1, 2, \cdots, n.
$$

Equivalently, we have

$$
|\hat{d}_j - d_j| < \frac{\sqrt{2n^2n^{n-1}}}{\zeta(n)(n-2)!\theta_{\min}^{n-2}} \frac{\sigma}{m_{\min}}, \quad j = 1, 2, \cdots, n.
$$

It follows that for $j = 1, \cdots, n$,

$$
\angle(\hat{d}_j d_j) \leq \frac{\pi}{2}|\hat{d}_j - d_j| < \frac{\sqrt{2n^2n^{n-1}}}{\zeta(n)(n-2)!\theta_{\min}^{n-2}} \frac{\sigma}{m_{\min}}. \quad (7.24)
$$
Step 7. We analyze $|\hat{y}_j - y_j|$ in this step. Note that
\[
|\frac{2r\Omega}{M-1} - \frac{2r\Omega}{M-1}| = \angle(d_j d_j), \quad \theta_{\text{min}} = \frac{2r\Omega}{M-1} d_{\text{min}},
\]
where $d_{\text{min}} = \min_{p \neq j} |y_p - y_j|$. Recall that $\angle(\hat{d}_j d_j) < \frac{\theta_{\text{min}}}{2}$, $1 \leq j \leq n$ (see Step 6.2), we have
\[
|\hat{y}_j - y_j| = \frac{M-1}{2r\Omega} |\hat{y}_j - \frac{2r\Omega}{M-1} - y_j| = \frac{M-1}{2r\Omega} \angle(\hat{d}_j d_j) < \frac{M-1 \theta_{\text{min}}}{2r\Omega} = \frac{d_{\text{min}}}{2}.
\]
Thus $\hat{\mu}$ is within the $\frac{d_{\text{min}}}{2}$-neighborhood of $\mu$. On the other hand, by (7.24), we have
\[
|\hat{y}_j - y_j| < \frac{(M-1)}{2r\Omega} \sqrt{2n^2 \pi^{n-1}} \left(\frac{\sqrt{2n^2 \pi^{n-1}}}{\zeta(n)(n-2)!}\right)^{2n-2} \frac{\sigma}{m_{\text{min}}}, \quad j = 1, \ldots, n.
\]
Together with (7.25), we have
\[
\frac{\sigma}{m_{\text{min}}} \leq C(n) \frac{\pi}{\omega^{\frac{2n-2}{2}}} \left(\frac{2n+1}{\omega}\right)^{2n-2} \frac{\sigma}{m_{\text{min}}}, \quad \text{by Lemma 8.4}
\]
where $C(n) = (2n+1)^2 e^{2n+1} \pi^{\frac{1}{2}}$. This completes the proof.

7.4 Proof of Proposition 2.2

Proof: Let $t_1 = -\frac{\tau}{\Omega}$, $t_2 = -\frac{(n-1)\tau}{\Omega}$, \ldots, $t_n = -\frac{\tau}{\Omega}$, $t_{n+1} = \frac{\tau}{\Omega}$, \ldots, $t_{2n} = \frac{\tau}{\Omega}$. Considering the linear system that
\[
Aa = 0
\]
where $A = (\phi_{2n-2}(t_1), \cdots, \phi_{2n-2}(t_{2n}))$ with $\phi_{2n-2}(\omega)$ defined in (4.1). Since $A$ is underdetermined, we have non-zero $a = (a_1, \cdots, a_{2n})$ satisfying the above equation. By the linear independence of the column vectors in the matrix $A$, we can show that all $a_i$’s are not zero. WLOG, we can suppose $\sum_{j=1}^{n} |a_j| \geq \sum_{j=n+1}^{2n} |a_j|$. By scaling $a$, we may also assume that $\sum_{j=1}^{n} |a_j| = m$. Then we calculate that, for $k \geq 2n-1$,
\[
|\sum_{j=1}^{2n} a_j t_j^k| \leq \sum_{j=1}^{2n} |a_j| \left(\frac{mT}{\Omega}\right)^k \leq 2mn^k \left(\frac{T}{\Omega}\right)^k.
\]
Thus for $\gamma = \sum_{j=1}^{2n} a_j \delta_{t_j}$, recall Taylor series (7.7), we have
\[
m_k(\gamma) = 0, k = 0, \cdots, 2n-2, \quad |m_k(\gamma)| \leq 2mn^k \left(\frac{T}{\Omega}\right)^k, k \geq 2n-1.
\]
Therefore, for $|x| \leq \Omega$,

$$\left| \mathcal{F}(\gamma)(x) \right| \leq \sum_{k \geq 2n-1} 2mn^k \frac{\frac{1\!}{\!\Omega}}{k!} k^{|x|^k} \leq \sum_{k \geq 2n-1} 2mn^k \frac{\frac{1\!}{\!\Omega}}{k!} k^{|x|^k},$$

$$\leq \frac{2mn^{2n-1} \tau^{2n-1}}{(2n-1)!} \sum_{k=0}^{\infty} \frac{(n\tau)^k}{k!} = \frac{2mn^{2n-1} \tau^{2n-1}}{(2n-1)!} e^{n\tau} \leq \sqrt{\pi(n - \frac{1}{2})} \left( \frac{e}{2} \right)^{2n-1} \leq \frac{me^{n\tau+1}}{\sqrt{\pi(n - \frac{1}{2})} \left( \frac{e}{2} \right)^{2n-1}} \leq \frac{1}{2} \leq \sigma. \quad \text{(by Lemma 8.1)}$$

Take

$$\mu = \sum_{j=1}^{n} a_j \delta_{t_j}, \quad \hat{\mu} = \sum_{j=n+1}^{2n} -a_j \delta_{t_j},$$

then the above estimate yields $\| [\hat{\mu}] - [\mu] \|_\infty < \sigma$ as desired.

8 Appendix

In this Appendix, we present some inequalities that are used in this paper. We first recall the Stirling approximation of factorial that for $n \geq 1$,

$$\sqrt{2\pi n^{n+\frac{1}{2}}} e^{-n} \leq n! \leq e^{n+\frac{1}{2}} e^{-n}, \quad (8.1)$$

which will be used frequently in subsequent derivation.

**Lemma 8.1.** For $n \geq 2$, we have

$$\frac{(n-1)^{2n-2}}{(2n-2)!} \leq \frac{1}{2\sqrt{\pi(n-1)}} \left( \frac{e}{2} \right)^{2n-2}, \quad \frac{n^{2n-1}}{(2n-1)!} \leq \frac{e}{2\sqrt{\pi(n - \frac{1}{2})}} \left( \frac{e}{2} \right)^{2n-1}. \quad (8.1)$$

**Proof:** By (8.1),

$$\frac{(n-1)^{2n-2}}{(2n-2)!} \leq \frac{(n-1)^{2n-2}}{\sqrt{2\pi(2n-2)^{2n-2} e^{-(2n-2)}}} = \frac{1}{2\sqrt{\pi(n-1)}} \left( \frac{e}{2} \right)^{2n-2},$$

$$\frac{n^{2n-1}}{(2n-1)!} \leq \frac{n^{2n-1}}{\sqrt{2\pi(2n-1)^{2n-1} e^{-(2n-1)}}} \leq \frac{1}{2\sqrt{\pi(n - \frac{1}{2})}} \left( \frac{e}{2} \right)^{2n-1} \frac{n^{2n-1}}{(2n-1)^{2n-1}} \leq \frac{e}{2\sqrt{\pi(n - \frac{1}{2})}} \left( \frac{e}{2} \right)^{2n-1}. \quad \text{(8.1)}$$

**Lemma 8.2.** For $n \geq 2$, we have

$$\left( \frac{\sqrt{2n-1}}{\zeta(n) \zeta(n-1)} \right)^{\frac{1}{n}} \leq \frac{2.22e}{n},$$

where $\zeta(n)$ and $\zeta(n-1)$ defined in (4.5).
Proof: For $n = 2, 3, 4$, it is easy to check that the above inequality holds. Using (8.1), we have for odd $n \geq 5$,

$$
\zeta(n)\zeta(n-1) = \left(\frac{n-1}{2}\right)^2 \left(\frac{n-3}{2}\right)^2 \geq \pi^2 \left(\frac{n-1}{2}\right)^2 \left(\frac{n-3}{2}\right)^2 e^{-(2n-4)}
$$

$\quad = n^{2n-2} \frac{\pi^2 (2n-1)!}{2^2 n^{2n-2}} \left(\frac{n-1}{2}\right)^{n-1} \left(\frac{n-3}{2}\right)^{n-2} e^{-(2n-4)}$

$\quad = \pi^2 e^{2} \left(\frac{n}{2}\right)^{2n-2} \left(\frac{n-1}{n}\right)^n \left(\frac{n-3}{n}\right)^{n-2} \geq \left(\frac{\pi}{e}\right)^2 \left(\frac{n}{2}\right)^{2n-2},$

and for even $n \geq 6$,

$$
\zeta(n)\zeta(n-1) = \left(\frac{n}{2}\right)! \left(\frac{n-2}{2}\right)! \left(\frac{n-4}{2}\right)! \geq \pi^2 \left(\frac{n}{2}\right)^{n+1} \left(\frac{n-2}{2}\right)^{n-1} \left(\frac{n-4}{2}\right)^{n-2} e^{-(2n-4)}
$$

$\quad = n^{2n-2} \frac{\pi^2 (n+1)!}{2^2 n^{2n-2}} \left(\frac{n-1}{2}\right)^{n-1} \left(\frac{n-3}{2}\right)^{n-2} e^{-(2n-4)}$

$\quad = \pi^2 e^{2} \left(\frac{n}{2}\right)^{2n-2} \frac{n+1}{n} \left(\frac{n-1}{n}\right)^n \left(\frac{n-3}{n}\right)^{n-2} \geq \left(\frac{\pi}{e}\right)^2 \left(\frac{n}{2}\right)^{2n-2}.$

Therefore, for all $n \geq 5$,

$$
\left(\frac{\sqrt{2n-1}}{\zeta(n)\zeta(n-1)}\right)^{\frac{1}{n-1}} \leq 2e \left(\frac{\sqrt{2n-1} e^2}{\pi^2}\right)^{\frac{1}{n-1}} \leq 2.22e.
$$

**Lemma 8.3.** For $n \geq 2$, we have

$$
\left(\frac{4\sqrt{2n}}{\zeta(n)\lambda(n)}\right)^{\frac{n}{n+1}} \leq \frac{3.07e}{n+\frac{3}{2},}
$$

where $\eta(n), \lambda(n), \xi(n-2)$ are as defined in (7.3).

Proof: For $n = 2, 3, 4, 5$, the inequality holds. By the Stirling approximation (8.1), we have for even $n \geq 6$,

$$
\zeta(n)\lambda(n) = \zeta(n)\zeta(n-2) \leq \left(\frac{n}{2}\right)! \left(\frac{n-2}{2}\right)! \left(\frac{n-4}{2}\right)! \geq \pi^2 \left(\frac{n}{2}\right)^{n+1} \left(\frac{n-2}{2}\right)^{n-1} \left(\frac{n-4}{2}\right)^{n-2} e^{-(2n-5)}
$$

$\quad = (n+\frac{1}{2})^{2n-1} \pi^2 (2n-1)! \left(\frac{n-1}{2}\right)^{n-1} \left(\frac{n-3}{2}\right)^{n-3} e^{-(2n-5)}$

$\quad \geq (n+\frac{1}{2})^{2n-1} \frac{4\pi^2}{e^2 (n+\frac{1}{2})^2},$

and for odd $n \geq 7$,

$$
\zeta(n)\lambda(n) = \zeta(n)\zeta(n-2) \leq \left(\frac{n}{2}\right)! \left(\frac{n-2}{2}\right)! \left(\frac{n-4}{2}\right)! \geq \pi^2 \left(\frac{n}{2}\right)^{n+1} \left(\frac{n-2}{2}\right)^{n-1} \left(\frac{n-4}{2}\right)^{n-2} e^{-(2n-5)}
$$

$\quad = (n+\frac{1}{2})^{2n-1} \pi^2 (2n-1)! \left(\frac{n-1}{2}\right)^{n-1} \left(\frac{n-3}{2}\right)^{n-3} e^{-(2n-5)}$

$\quad \geq (n+\frac{1}{2})^{2n-1} \frac{4\pi^2}{e^2 (n+\frac{1}{2})^2}.$

33
Therefore, for all \( n \geq 6, \)
\[
\left( \frac{4\sqrt{2n}}{\zeta(n)\lambda(n)} \right)^{\frac{1}{n-1}} \leq \frac{2e}{n + \frac{1}{2}} \left( \frac{e^{2(n + \frac{1}{2})^2\sqrt{2n}}}{\pi^2} \right)^{\frac{1}{n-1}} \leq 3.07e^{\frac{2}{n + \frac{1}{2}}}.
\]

**Lemma 8.4.** For \( n \geq 2, \) we have
\[
\frac{(2n + 1)^{2n-\frac{1}{2}}}{\zeta(n)(n-2)!} \leq (2n + 1)^{2n-3} e^{2n+1} n^{-\frac{1}{2}}.
\]
where \( \zeta(n) \) is defined in (4.5).

Proof: By the Stirling approximation formula (8.1), when \( n \) is odd and \( n \geq 3, \) we have
\[
\frac{(2n + 1)^{2n-\frac{1}{2}}}{\zeta(n)(n-2)!} = \frac{(2n + 1)^{2n-\frac{1}{2}}}{(n-1)!2(n-2)!} \leq \frac{(2n + 1)^{2n-\frac{1}{2}}}{(\sqrt{2\pi})^3 (n-1)^{n-1} e^{(n-2)^2} (n-2)^{n-2} e^{-2(n-2)^2} e^{(2n-2)^2} e^{2n+1}}
\]
\[
\leq \frac{n^{\frac{1}{2}} e^{2n}}{(e \sqrt{2\pi})^3 (n-1)^{n-1} e^{(n-2)^2} (n-2)^{n-2} e^{-2(n-2)^2} e^{2n+1}} \leq \frac{(2n + 1)^{2n-3} e^{2n+1}}{(e \sqrt{2\pi})^3}.
\]

When \( n \) is even and \( n \geq 4, \) we have
\[
\frac{(2n + 1)^{2n-\frac{1}{2}}}{\zeta(n)(n-2)!} = \frac{(2n + 1)^{2n-\frac{1}{2}}}{(n-1)!2(n-2)!} \leq \frac{(2n + 1)^{2n-\frac{1}{2}}}{(\sqrt{2\pi})^3 (n-1)^{n-1} e^{(n-2)^2} (n-2)^{n-2} e^{-2(n-2)^2} e^{2n+1}}
\]
\[
\leq \frac{n^{\frac{1}{2}} e^{2n}}{(e \sqrt{2\pi})^3 (n-1)^{n-1} e^{(n-2)^2} (n-2)^{n-2} e^{-2(n-2)^2} e^{2n+1}} \leq \frac{(2n + 1)^{2n-3} e^{2n+1}}{(e \sqrt{2\pi})^3}.
\]

For \( n = 2, \) we have
\[
\frac{(2n + 1)^{2n-\frac{1}{2}}}{\zeta(n)(n-2)!} \leq \frac{2^4 e^2}{1} \leq \frac{(2n + 1)^{2n-3} e^{2n+1}}{(e \sqrt{2\pi})^3}.
\]

**References**

[1] Hirotogu Akaike. Information theory and an extension of the maximum likelihood principle. In *Selected papers of hirotogu akaike*, pages 199–213. Springer, 1998.

[2] Hirotugu Akaike. A new look at the statistical model identification. In *Selected Papers of Hirotugu Akaike*, pages 215–222. Springer, 1974.

[3] Andrey Akinshin, Dmitry Batenkov, and Yosef Yomdin. Accuracy of spike-train fourier reconstruction for colliding nodes. In *2015 International Conference on Sampling Theory and Applications (SampTA)*, pages 617–621. IEEE, 2015.

[4] Jean-Marc Azais, Yohann De Castro, and Fabrice Gamboa. Spike detection from inaccurate samplings. *Applied and Computational Harmonic Analysis*, 38(2):177–195, 2015.
[5] Dmitry Batenkov. Stability and super-resolution of generalized spike recovery. *Applied and Computational Harmonic Analysis*, 45(2):299–323, 2018.

[6] Dmitry Batenkov, Gil Goldman, and Yosef Yomdin. Super-resolution of near-colliding point sources. *arXiv preprint arXiv:1904.09186*, 2019.

[7] Badri Narayan Bhaskar, Gongguo Tang, and Benjamin Recht. Atomic norm denoising with applications to line spectral estimation. *IEEE Transactions on Signal Processing*, 61(23):5987–5999, 2013.

[8] Emmanuel J. Candès and Carlos Fernandez-Granda. Towards a mathematical theory of super-resolution. *Communications on Pure and Applied Mathematics*, 67(6):906–956, 2014.

[9] Daniel H. Chae, Parastoo Sadeghi, and Rodney A. Kennedy. Effects of basis-mismatch in compressive sampling of continuous sinusoidal signals. In *2010 2nd International Conference on Future Computer and Communication*, volume 2, pages V2–739. IEEE, 2010.

[10] Venkat Chandrasekaran, Benjamin Recht, Pablo A Parrilo, and Alan S Willsky. The convex geometry of linear inverse problems. *Foundations of Computational mathematics*, 12(6):805–849, 2012.

[11] Yuejie Chi and Maxime Ferreira Da Costa. Harnessing sparsity over the continuum: Atomic norm minimization for super resolution. *arXiv preprint arXiv:1904.04283*, 2019.

[12] Yuejie Chi, Louis L. Scharf, Ali Pezeshki, and A. Robert Calderbank. Sensitivity to basis mismatch in compressed sensing. *IEEE Transactions on Signal Processing*, 59(5):2182–2195, 2011.

[13] Laurent Demanet and Nam Nguyen. The recoverability limit for superresolution via sparsity. *arXiv preprint arXiv:1502.01385*, 2015.

[14] Quentin Denoyelle, Vincent Duval, and Gabriel Peyré. Support recovery for sparse super-resolution of positive measures. *Journal of Fourier Analysis and Applications*, 23(5):1153–1194, 2017.

[15] David L. Donoho. Superresolution via sparsity constraints. *SIAM journal on mathematical analysis*, 23(5):1309–1331, 1992.

[16] Marco F. Duarte and Richard G. Baraniuk. Spectral compressive sensing. *Applied and Computational Harmonic Analysis*, 35(1):111–129, 2013.

[17] Vincent Duval and Gabriel Peyré. Exact support recovery for sparse spikes deconvolution. *Foundations of Computational Mathematics*, 15(5):1315–1355, 2015.

[18] Albert C. Fannjiang and Wenjing Liao. Mismatch and resolution in compressive imaging. In *Wavelets and Sparsity XIV*, volume 8138, page 81380Y. International Society for Optics and Photonics, 2011.
[19] Walter Gautschi. On inverses of vandermonde and confluent vandermonde matrices. *Numerische Mathematik*, 4(1):117–123, 1962.

[20] Thomas Lundgaard Hansen, Bernard Henri Fleury, and Bhaskar D Rao. Superfast line spectral estimation. *IEEE Transactions on Signal Processing*, 66(10):2511–2526, 2018.

[21] Zhaoshui He, Andrzej Cichocki, Shengli Xie, and Kyuwan Choi. Detecting the number of clusters in n-way probabilistic clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(11):2006–2021, 2010.

[22] Jonathan William Helland. *Atomic Norm Algorithms for Blind Spectral Super-resolution Problems*. PhD thesis, Colorado School of Mines. Arthur Lakes Library, 2019.

[23] Qiuwei Li and Gongguo Tang. Approximate support recovery of atomic line spectral estimation: A tale of resolution and precision. *Applied and Computational Harmonic Analysis*, 2018.

[24] Weilin Li and Wenjing Liao. Stable super-resolution limit and smallest singular value of restricted fourier matrices. 2018.

[25] Weilin Li, Wenjing Liao, and Albert Fannjiang. Super-resolution limit of the esprit algorithm. *arXiv preprint arXiv:1905.03782*, 2019.

[26] Ping Liu and Hai Zhang. Computational resolution limit: a theory towards super-resolution. *arXiv preprint arXiv:1912.05430*, 2019.

[27] Dmitry Malioutov, Müjdat Cetin, and Alan S Willsky. A sparse signal reconstruction perspective for source localization with sensor arrays. *IEEE transactions on signal processing*, 53(8):3010–3022, 2005.

[28] Ankur Moitra. Super-resolution, extremal functions and the condition number of vandermonde matrices. In *Proceedings of the Forty-seventh Annual ACM Symposium on Theory of Computing*, STOC ’15, pages 821–830, New York, NY, USA, 2015. ACM.

[29] Athanasios Papoulis and Christodoulos Chamzas. Improvement of range resolution by spectral extrapolation. *Ultrasonic Imaging*, 1(2):121–135, 1979.

[30] Clarice Poon and Gabriel Peyré. Multidimensional sparse super-resolution. *SIAM Journal on Mathematical Analysis*, 51(1):1–135, 1979.

[31] Gideon Schwarz et al. Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464, 1978.

[32] Gongguo Tang. Resolution limits for atomic decompositions via markov-bernstein type inequalities. In *2015 International Conference on Sampling Theory and Applications (SampTA)*, pages 548–552. IEEE, 2015.

[33] Gongguo Tang, Badri Narayan Bhaskar, and Benjamin Recht. Near minimax line spectral estimation. *IEEE Transactions on Information Theory*, 61(1):499–512, 2014.
[34] Gongguo Tang, Badri Narayan Bhaskar, Parikshit Shah, and Benjamin Recht. Compressed sensing off the grid. *IEEE transactions on information theory*, 59(11):7465–7490, 2013.

[35] Dehui Yang, Gongguo Tang, and Michael B Wakin. Super-resolution of complex exponentials from modulations with unknown waveforms. *IEEE Transactions on Information Theory*, 62(10):5809–5830, 2016.