Computational resolution limit: a theory towards super-resolution

Ping Liu* and Hai Zhang†

December 12, 2019

Given an image generated by the convolution of point sources with a band-limited function, the deconvolution problem is to reconstruct the source number, positions, and amplitudes. This problem arises from many important applications in imaging and signal processing. It is well-known that it is impossible to resolve the sources when they are close enough in practice. Rayleigh investigated this problem and formulated a resolution limit, also called Rayleigh limit, for the case of two sources with identical amplitudes. On the other hand, many numerical experiments demonstrate that a stable recovery of the sources is possible even if the sources are separated below the Rayleigh limit. This is so-called "super-resolution". In this paper, a new mathematical theory for super-resolution will be developed. The theory will address the issue when one can recover the source number exactly. The key is a new concept "computational resolution limit" which is defined to be the minimum separation distance between the sources such that exact recovery of the source number is possible. This new resolution limit is determined by the signal-to-noise ratio and the sparsity of sources, in addition to the cutoff frequency of the image. Sharp upper bound for this limit is derived, which reveals the importance of the sparsity as well as the signal-to-noise ratio to the recovery problem. The stability for recovering the source positions is further derived when the separation distance is beyond the upper bound. Moreover, a MUSIC-type algorithm is proposed to recover the source number and to verify our theoretical results on the computational resolution limit. Its performance in the super-resolution regime when the sources are separated below the Rayleigh limit is analyzed both theoretically and numerically. The results are based on a multipole expansion method and a novel non-linear approximation theory in Vandermonde space.

*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 partially supported by HK RGC grant GRF 16304517 and GRF 16306318.
1 INTRODUCTION

1.1 PROBLEM SETTING AND BACKGROUND

In numerous imaging and signal processing problems, an image is obtained by convoluting point sources with a band-limited function which is called the point spread function. The problem of recovering the source number, positions and amplitudes is called deconvolution. It has many applications in biology, medical imaging, and astronomy. This paper mainly focuses on recovering the source number and positions from their image under a certain noise level. It’s well-known that it is impossible to resolve the sources when they are sufficiently close in practice. Rayleigh investigated this issue and formulated a resolution limit, which is also called Rayleigh limit, for the case of two sources with identical amplitudes. It is defined to be \( \frac{1.22 \pi}{\Omega} \) for two-dimensional images where \( \Omega \) is the cutoff frequency of the point spread function. The Rayleigh limit is empirical and only applicable to classical instrumental imaging methods. Mathematically, when there is no noise, it is known that one can recover the sources uniquely. Thus, the resolution limit for the deconvolution problem should take into account the noise. In view of this and the limitation of the classical resolution limit, we propose a new concept named computational resolution limit which is defined to be the minimum distance required to recover the source number from the image with a certain noise level.

More precisely, we assume the collection of point sources is a discrete measure

\[
\mu^* = \sum_{j=1}^{n} a_j^* \delta_{y_j^*},
\]

where \( y_1^*, \cdots, y_n^* \) are the locations of point sources and \( a_1^*, \cdots, a_n^* \) the amplitudes. We assume that the amplitudes are real numbers. We denote

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

Let \( f \) be the band-limited point spread function. Throughout the paper, we let \( f(x) = \frac{\sin x}{x} \) for ease of explanation. The corresponding cutoff frequency is one and the Rayleigh limit is \( \pi \). We assume that the sources are located in \([-d, d]\) as in Figure 1.1 with \( d \) of order one. The general cutoff frequency case will be addressed in Section 6 by using a scaling argument. We remark that our results and techniques can be adapted to other band-limited or essentially band-limited functions such as \( (\frac{\sin x}{x})^2 \) and \( e^{-x^2} \).
The noiseless image is the convolution of $\mu^*$ and $f$. We sample the image at evenly spaced points $x_1 = -R, x_2 = -R + h, x_3 = -R + 2h, \cdots, x_N = R$, where $R$ is a large truncation number such that the image outside is negligible, and $h$ is the spacing of the sample points. We employ the Shannon sampling strategy by assuming that $h \leq \pi$. The measurements are

$$Y(x_t) = \mu^* \ast f(x_t) + W(x_t) = \sum_{j=1}^{n} a_j^* f(x_t - y_j^*) + W(x_t), \quad t = 1, \cdots, N$$  \hspace{1cm} (1.1)

where $W(x)$ is band-limited noise. More precisely, $W(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} w(\omega) e^{i\omega x} d\omega$ with $w(\omega) = 0$ for $|\omega| > 1$. We assume the noise level $\|w\|_2 = \sqrt{\int_{-\infty}^{\infty} |w(\omega)|^2 d\omega} \leq \sqrt{2\pi}\sigma$.

We denote

$$Y = (Y(x_1), \cdots, Y(x_N))^T, \quad (\mu \ast f) = (\mu \ast f(x_1), \cdots, \mu \ast f(x_N))^T, \quad W = (W(x_1), \cdots, W(x_N))^T.$$

Then the following estimate holds

$$\sqrt{h}\|W\|_2 \leq \frac{1}{\sqrt{2\pi}}\|w\|_2 \leq \sigma.$$  \hspace{1cm} (1.2)

The deconvolution problem is to recover the source number $n$ and their locations $y_j^*$'s and amplitudes $a_j^*$'s from the measurements in (1.1). The traditional resolution limit problem is concerned with the minimum distance between the two sources such that one can distinguish them from their image. Without noise, by the linear independence of the functions $f(x - y_j^*)$ for different $y_j^*$'s, the locations and amplitudes can be uniquely figured out. What is more interesting is the case with noise as will be addressed in this paper. We shall answer the question when one can recover the source number and positions from their noisy image?

For a discrete measure $\mu$, it is natural to view it as a solution to the deconvolution problem if the image $\mu \ast f$ is very close to $Y$. We thus introduce the following concept of admissible measures.

**Definition 1.1.** For a given a priori noise level $\sigma$, interval size $d$, and total-variation norm bound $M \geq m^*$, we say that $\mu = \sum_{j=1}^{k} a_j \delta_{y_j}$ is a $(d, \sigma, M)$-admissible discrete measure for the image $Y$ only if $\mu$ is supported in $[-d, d]$ such that $\|\mu\|_{TV} = \sum_{j=1}^{k} |a_j| \leq M$ and

$$\sqrt{h}\|(|\mu \ast f| - Y)|_2 \leq \sigma.$$  \hspace{1cm} (1.3)

The set of admissible measures of $Y$ characterizes all possible solutions to the deconvolution problem with the given image $Y$. A good reconstruction algorithm should give an admissible measure. If there exists one admissible measure with less than $n$ sources, then one may say that the image $Y$ is generated by less than $n$ sources, and hence miss the exact source number. On the other hand, if all admissible measures have at least $n$ sources, then one can determine the source number $n$ correctly if one restricts to the sparsest admissible measures. This leads to the following definition of computational resolution limit.

**Definition 1.2.** For an image $Y$ generated by $n$ point sources, the computational resolution limit is defined as the minimum separation distance between the sources beyond which there does not exist any $(d, \sigma, M)$-admissible measure for $Y$ with less than $n$ supports.
The above limit will be determined by the signal-to-noise ratio (SNR) and sparsity of the sources, in addition to the cutoff frequency and a few other a priori information of the sources. We note that it is independent of the algorithms of reconstruction. According to the definition, determining the source number is impossible if the sources are separated below this limit.

In this paper, we mainly focus on estimating the bounds of this computational resolution limit. This gives a quantification of the actual limit of resolution which has been a long-standing problem in signal processing, spectral estimation and the fields related. The idea that the actual resolution limit is determined by SNR was known at a very early age. To our knowledge, as early as 1980, it was pointed out in [37] that, Rayleigh limit is adequate if one relies on the direct observation of the data for the determination of sources, however, not useful if the data are subjected to elaborate processing. In that paper, the authors proposed an extrapolation method and described an application in medical ultrasound which claimed that Rayleigh limit was exceeded in some examples. Apart from the extrapolation methods [36, 37, 44], there were a large group of methods based on statistics and time series, see for instance, the maximum entropy method (MEM) [1, 5, 24] and the maximum likelihood method (MLM) [10, 28]. Besides, the CLEAN method [22] used in astronomy also involves finding a small set of sources that nearly generate the measurements by convoluting with the point spread function. However, the CLEAN method is designed for the case that the brightness distribution contains only a few sources at well separated, small regions [43]. We refer to [25, 26] for the summaries of these early year methods.

To understand the puzzle of resolution limit and the performance of these early year algorithms, Donoho gave the first attempt to unravel the mystery and meanwhile quantified the ill-posedness of the deconvolution problem. In [13], he developed a theory from the optimal recovery point of view to explain the possibility and difficulties of superresolution via sparsity constraint. He considered a grid setting where a discrete measure is supported on the lattice \( \{k\tau : k \in \mathbb{Z}\} \) with spacing \( \tau \), and the available measurement is its low frequency information. He derived bounds for the minimax error of the recovery of a special class of sparse measures. These bounds are given by noise amplification, and increase polynomially with the super-resolution factor which is defined to be the ratio between Rayleigh limit and the grid spacing. His results emphasize the importance of sparsity in superresolution which is embodied in the order of the polynomial. Further discussed in [12], Demanet and Nguyen obtained sharper bounds using estimate of the minimum singular value for the measurement matrix. In [33], using novel extremal functions, Moitra demonstrated a phase transition phenomenon for ill-posedness of the inverse problem when the cutoff frequency is near the inverse of the grid spacing. We remark that all these theoretical results deal with the grid setting where the locations of the sources are given and one need only recover their amplitudes. This approach reduces the nonlinear problem into a linear one. However, the grid setting brings unavoidable model error when the sources are off grids [11].

In addition to the early year algorithms mentioned above, a class of algorithms called subspace method have gained popularity for achieving superresolution. Specific examples including Multiple Signal Classification (MUSIC) [42], Estimation of Signal Parameters via Rotational Invariance Technique (ESPRIT) [39], and Matrix Pencil Method [23]. These algorithms date back to the work of Prony [38]. In [47], a statistical analysis of MUSIC was provided along with the performance limits based on Cramer-Rao bound. In the case of two sources, [45, 46],
gave the characterization of the trade-off between resolution and SNR (signal-to-noise ratio). Moreover, recently in the grid setting, a mathematical theory was developed in [31] to explain the numerical superresolution observed in [32]. Their theory is based on a sharp bound of the minimax error of recovery.

The success of $l_1$ penalty in seismic imaging at early stage [30, 41] and the recent development of the theory of compressive sensing [9, 14] provide new ideas and methods for the deconvolution problem, namely, sparsity promoting algorithms. In [8], it is demonstrated that, in the grid setting, when there is no noise, well-separated sources can be exactly recovered by the sparsity promoting convex optimization. Moreover, if the separation distance is beyond several Rayleigh limits, the measurement matrix may have a good restricted isometry property which guarantees a stable recovery. Spectral compressive sensing [15] and MUSIC method in [17] also employ this idea. In the presence of noise, stability results are further established in [7] for well-separated sources but without the grid constraint. Many interesting results are obtained in this research line, see, for instance, [4, 6, 16, 18, 48]. We remark that in order for most of the sparsity promoting convex optimization to work, it is necessary to assume that the sources are well-separated [16]. As a consequence, these results may not be applicable to the super-resolution regime which we are interested in. In [16, 34], the restriction on the separation distance is relaxed. However, they only deal with positive sources.

1.2 Our Main Results

In this paper, we derived a sharp upper bound for the computational resolution limit we introduced. To achieve the goal, we developed a novel multiple expansion method and reduced the original problem to a non-linear approximation problem in the so-called Vandermonde space. We obtained a sharp bound to the approximation problem. This yields an upper bound for the computational resolution limit in the following explicit form

$$4.7(1 + d)^{2n - 2} \sqrt{\frac{3}{\sigma_{\min}(s^*) m^*_{\min}}} \sigma_{\min}(s^*) m^*_{\min},$$

where $d$ is the a priori estimate of the size of the interval where the sources are located, $n$ is the source number, $\sigma_{\min}(s^*)$ is the minimum singular value of certain multipole matrix and $\sigma_{\min}(s^*) m^*_{\min}$ is the noise-to-signal ratio. The order $2n - 2$ quantitatively indicates the importance of sparsity to the deconvolution problem. The factor $\sigma_{\min}(s^*)$ characterizes the correlation of multipoles and is determined by the point spread function. This upper bound characterizes the ill-posedness of the recovery problem. It also manifests that beyond this upper bound the inverse problem is regularized, and any algorithm finding the sparsest solution in the set of admissible measures can recover the source number. As a consequence, provided that SNR is sufficiently large, one can recover the source number exactly even if they are separated below the Rayleigh limit to achieve the so-called “super-resolution”.

We also considered the stability of recovering the source positions. We showed that when the separation distance exceeds

$$6.24(1 + d)^{2n - 1} \sqrt{\frac{3}{\sigma_{\min}(s^*) m^*_{\min}}} \sigma_{\min}(s^*) m^*_{\min},$$
we can have a stable recovery of the source positions from the admissible measures. It implies that any algorithm searching among the sparsest admissible measures can stably recover the source positions. We also demonstrate the sharpness of this condition.

Finally, we proposed a preliminary MUSIC-type algorithm to recover the source number. Our numerical results show that this algorithm can recover the exact number in the super-resolution regime when the source separation distance is comparable to the upper bound we derived for the computational resolution limit.

1.3 Comment of Super-resolution

The motivation of this work is to understand various results of "super-resolution". While the definition of "super-resolution" varies from fields to fields. Here, we restrict to the deconvolution problem \([1]\), which arises from optical microscopy. We define resolution to be the minimum distance between the sources so that one can resolve their number correctly. It is well-known, since Abbe's work, that there is a fundamental limit on the resolution due to the diffraction of light. This limit has its origin in the physics of waves, which is independent of noise. In our work, all the physics is encoded in the shape of point spread function. On the other hand, from data processing (or computational) point of view, it is well-known that one can recover the source number exactly in the deconvolution problem if there is no noise. As a consequence, one can achieve infinite resolution if there is no noise in the measurement. Therefore, a proper definition of resolution in the deconvolution problem should take into account of SNR, as is discussed in this paper. We characterize an upper bound of the computational resolution limit in terms of SNR and the sparsity of sources. Beyond this upper bound, we can stably recover the source number. We also demonstrate that, as sources separated further, we can stably recover their positions. These results indicate that breaking the Rayleigh limit to achieve the so-called super-resolution by mathematical algorithms is possible for suitable SNR.

On the other hand, our result unveils the fundamental limit of recovering the source number in the presence of noise by using a single snap-shot image of the sources. It indicates that if one needs to resolve the sources beyond this computational resolution limit, there are only physical ways. We here propose two methods for reference. The first is to break the diffraction limit by using various super-lens in the microscopy which can yield a sharper point spread function \([49]\). See also a very promising way to break the diffraction limit by using subwavelength resonators \([2, 3, 29]\). The second way is to use multi-illuminations to generate multi images of sources. Loosely speaking, SIM \([20]\), STED \([21]\) \([27]\), STROM \([40]\) can be viewed in this category. In a forthcoming paper, we shall quantitatively demonstrate that multi-illumination can indeed break the computational resolution limit significantly.

1.4 Organization of this paper

The rest of the paper is organized as follows. In Section 2, we introduce the multipole expansion method and derive the upper bound of the computational resolution limit. In Section 3, we introduce the nonlinear approximation problem in the Vandermonde space. In Section 4, we derive the stability of support recovery from admissible measures. In Section 5, we propose
a MUSIC-type algorithm. In Section 6, we discuss the general frequency case. In Section 7, we show numerical experiments. The conclusion and future work are discussed in Section 8. Finally, technical lemmas and their proofs are given in the appendix.

2 Upper bound of the computational resolution limit

2.1 Multipole expansion

The usual way to solve the deconvolution problem is to consider the following linear problem

\[
\begin{pmatrix}
  Y(x_1) \\
  \vdots \\
  Y(x_N)
\end{pmatrix}
= \begin{pmatrix}
  f(x_1 - y_1) & \cdots & f(x_1 - y_K) \\
  \vdots & \ddots & \vdots \\
  f(x_N - y_1) & \cdots & f(x_N - y_K)
\end{pmatrix}
\begin{pmatrix}
  a_1 \\
  \vdots \\
  a_K
\end{pmatrix}
+ \begin{pmatrix}
  W(x_1) \\
  \vdots \\
  W(x_N)
\end{pmatrix}
\tag{2.1}
\]

where \( y_1, \cdots, y_K \) are given grid points. In this setting, the sources are assumed to be lied on the grid points and one only need to reconstruct their amplitudes. This approach introduces model errors when the sources are off the grid \([11]\). To avoid this issue, we propose a novel multipole expansion method. The key observation is that the measurement \( Y \) has the following multipole expansion:

\[
Y = \sum_{r=0}^{+\infty} c_r^* h_r + W,
\tag{2.2}
\]

where \( c_r^* \)'s are the multipole coefficients and \( h_r \)'s are the multipoles defined as

\[
h_r = \sqrt{2r + 1} (f^{(r)}(x_1), \ldots, f^{(r)}(x_N))^T, r = 0, 1, \ldots.
\]

Here \( \sqrt{2r + 1} \) is a normalization factor. Especially, we call \( h_0 = (f^{(0)}(x_1), \ldots, f^{(0)}(x_N))^T \) the monopole, \( h_1 = \sqrt{3}(f^{(1)}(x_1), \ldots, f^{(1)}(x_N))^T \) the dipole and \( h_2 = \sqrt{5}(f^{(2)}(x_1), \ldots, f^{(2)}(x_N))^T \) the quadrapole. For subsequent derivation, the following estimate is needed:

\[
h \sum_{i=1}^{N} f^{(r)}(x_i)^2 \leq h \sum_{j=-\infty}^{+\infty} f^{(r)}(jh)^2 = \int_{-\infty}^{+\infty} |f^{(r)}|^2 dx = \frac{\pi}{2r + 1}.
\tag{2.3}
\]

By Taylor expansion of \( f(x - y_j^*) \), \( j = 1, \cdots, n \) around origin we have

\[
c_r^* = \sum_{j=1}^{n} a_j^* \frac{(d_j^*)^r}{r! \sqrt{2r + 1}},
\]

where \( d_j^* = 0 - y_j^*, j = 1, \cdots, n \). By \([2,3]\), the 2-norm of each multipole satisfies the following estimate

\[
\sqrt{h} ||h_r||_2 = \sqrt{2r + 1} \sqrt{h \sum_{i=1}^{N} f^{(r)}(x_i)^2} \leq \sqrt{\pi}.
\tag{2.4}
\]

The analysis of the resolution limit is based on the idea that for a certain level of SNR, we can only stably recover a finite number of low-order multipole coefficients from measurements.
We shall show that these partial multipole coefficients set a limit to the resolution. For the purpose, we introduce multipole matrix

\[ H(s) = \left( h_0, h_1, \cdots, h_{s-1} \right). \]

We denote the minimum singular value of \( \sqrt{m}H(s) \) as

\[ \sigma_{\min}(s) = \sigma_{\min}(\sqrt{m}H(s)). \]

We remark that \( \sigma_{\min}(s) \) is determined by \( s \) and the point spread function. Note that \( \sigma_{\min}(\sqrt{m}H(s)) \leq \|\sqrt{m}H(s)\|_2 \leq \sqrt{n} \) by (2.4) for \( n = (1, 0, \cdots, 0)^T \). We have the following estimate

\[ \sigma_{\min}(s) \leq \sqrt{n}, \quad \forall \ s \geq 1. \] (2.5)

2.2 Lower Bound of the Computational Resolution Limit

As a first application of the multipole method introduced in the previous part, we show that to recover the number of sources, the minimum separation distance is at least of the order \( O\left( \frac{\sigma}{\sqrt{m^*}} \right) \).

**Proposition 2.1.** For given \( \sigma > 0 \), integer \( n \geq 2 \), and \( m^* > 0 \), choose \( d \) satisfying

\[ \frac{d}{n-1} = 2 \frac{2^{n-1}}{e} \sqrt{\frac{(2n-2)}{2n-1}} \frac{\sigma}{m^*}. \] (2.6)

Then there exist two measures \( \mu^* = \sum_{j=1}^{n} a_j^* \delta_{y_j^*} \) with \( n \) supports and \( \mu \) with \( n - 1 \) supports, both supported in \([-d, d]\), such that \( \| \mu \|_{TV} \leq \| \mu^* \|_{TV} = m^* \) and

\[ \min_{i \neq j} |y_i^* - y_j^*| = \frac{d}{n-1}. \]

Moreover,

\[ \sigma \leq \| \sqrt{m} [\mu * f] - [\mu^* * f] \|_2. \]

**Proof:** Let \( d_1 = -d, d_2 = d - \frac{d}{n-1}, \cdots, d_{2n-2} = d - \frac{d}{n-1}, d_{2n-1} = d \). We consider the linear system

\[ Aa = 0 \] (2.7)

where

\[ A = \begin{pmatrix} 1 & \frac{d_1}{\sqrt{3}} & \frac{d_2}{\sqrt{3}} & \cdots & \frac{d_{2n-1}}{\sqrt{3}} \\ \frac{d_1}{\sqrt{3}} & 1 & \frac{d_2}{\sqrt{3}} & \cdots & \frac{d_{2n-1}}{\sqrt{3}} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \frac{(d_1)^{2n-3}}{(2n-3)!/\sqrt{4n-5}} & \frac{(d_2)^{2n-3}}{(2n-3)!/\sqrt{4n-5}} & \cdots & \frac{(d_{2n-1})^{2n-3}}{(2n-3)!/\sqrt{4n-5}} \\ \end{pmatrix}, \quad a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_{2n-1} \end{pmatrix}. \]

Since \( A \) is underdetermined, there exists a nonzero \( a = (a_1, \cdots, a_{2n-1}) \) satisfying (2.7). Using the linear independence of the column vectors in the matrix \( A \), we can show that all \( a_j \)'s are
We derive an upper bound for the computational resolution limit introduced in Definition 1.2. Let \( \mu = \sum_{j=1}^{n} a_j \delta_{-d_j} \) and \( \mu = \sum_{j=n+1}^{2n-1} -a_j \delta_{-d_j} \).

Then \( d_{\min}^* = \min_{i \neq j} |y_i^* - y_j^*| = \frac{d}{n-1} \). We consider the expansion

\[
[\mu^* * f] - [\mu * f] = \sum_{r=0}^{2n-3} (c_r^* - c_r) h_r + \text{Res},
\]

where \( \text{Res} = \sum_{r=2n-2}^{+\infty} (c_r^* - c_r) h_r \). By (2.7), \( \sum_{r=2n-2}^{2n-3} (c_r^* - c_r) h_r = 0 \). On the other hand,

\[
\sqrt{\pi} \| \text{Res} \|_2 = \sqrt{\pi} \| \sum_{r=2n-2}^{\infty} (c_r^* - c_r) h_r \|_2 \leq \sqrt{\pi} \sum_{r=2n-2}^{\infty} \| (c_r^* - c_r) h_r \|_2 ≤ \sqrt{\pi} \sum_{r=2n-2}^{\infty} \frac{2m^* d^r}{(2n-2)!} \leq \frac{2m^* d^{2n-2}}{(2n-2)!\sqrt{\pi}} (by \ (2.4))
\]

\[
= \frac{2e^d \sqrt{\pi} m^* d^{2n-2}}{(2n-2)!\sqrt{\pi}} \leq \frac{2\sqrt{\pi}(n-1)^{2n-2}(2n-2) 2}{(2n-2)!\sqrt{\pi}} (by \ (2.6)) \leq \sigma. \quad \text{(by Lemma 9.4)}
\]

Thus,

\[
\sqrt{\pi} \| [\mu^* * f] - [\mu * f] \|_2 ≤ \sqrt{\pi} \| \sum_{r=0}^{2n-3} (c_r^* - c_r) h_r \|_2 + \sqrt{\pi} \| \text{Res} \|_2 ≤ \sigma.
\]

2.3 Upper Bound of Computational Resolution Limit

We derive an upper bound for the computational resolution limit introduced in Definition 1.2. For given \( \sigma, M \) and \( d \), we define

\[
s^* := \min \left\{ l \in \mathbb{N} : \frac{d^l}{l! \sqrt{2l+1}} \leq \frac{\sigma}{2e^d \sqrt{\pi} M} \right\},
\]

which can be understood as the maximum number of multipoles one can expect to recover stably with the given SNR.

**Theorem 2.1.** Let \( n \geq 2 \) and let \( \mu^* = \sum_{j=1}^{n} a_j^* \delta_{y_j}^* \) be a measure supported in \([-d, d] \). Assume that the following separation condition is satisfied

\[
\min_{i \neq j} |y_i^* - y_j^*| \geq 4.7(1 + d) \frac{3}{\sigma_{\min}(s^*)^2 m_{\min}^*} \sqrt{\frac{\sigma}{3}},
\]

then any discrete measure \( \mu \) with \( k < n \) supports cannot be a \((d, \sigma, M)\)-admissible measure.
Proof: **Step 1.** We show that \( s^* \geq 2n-1 \). Let \( d_{\min}^* = \min_{i \neq j} |y_i - y_j| \). Then \( d \geq \frac{(n-1)d_{\min}^*}{2} \). We have

\[
\frac{d^{2n-2}}{(2n-2)! \sqrt{n-3}} \geq \frac{(1/2)^{2n-2}(n-1)^{2n-2}(d_{\min}^*)^{2n-2}}{(2n-2)! \sqrt{4n-3}}
\]

\[
\geq \frac{(1/2)^{2n-2}(n-1)^{2n-2}(1 + d)^{2n-2}}{(2n-2)! \sqrt{4n-3}}
\]

\[
\geq (n-1)^{2n-2} \frac{e(2n-2)^{2n-2} (1 + d)^{2n-2}}{3 \sigma} (\text{since } m_{\min}^* = m^*/n \leq M)
\]

\[
\geq e^{2n-2} 2^{-1.175} (1 + d)^{2n-2} \frac{3 \sigma}{\sigma_{\min}(s^*) M/M}
\]

\[
\geq (2\sqrt{2} - 2)^{2n-2} (1 + d)^{2n-2} \frac{3 \sigma}{\sigma_{\min}(s^*) M}
\]

which implies that \( s^* \geq 2n-1 \).

**Step 2.** For \( \mu = \sum_{j=1}^{k} a_j y_j \) with \( k < n \), we have

\[
Y - (\mu * f) = \sum_{r=0}^{+\infty} (c_r - c_r) h_r + \text{Res}, \tag{2.11}
\]

where

\[
c_r^* = \sum_{j=1}^{n} a_j \frac{(d_j)^r}{r! \sqrt{2r + 1}}, \quad c_r = \sum_{j=1}^{k} a_j \frac{(d_j)^r}{r! \sqrt{2r + 1}}, \quad r = 0, 1, \ldots
\]

It follows that

\[
||Y - (\mu * f)||_2 \geq || \sum_{r=0}^{s-1} (c_r^* - c_r) h_r ||_2 - ||W||_2 - ||\text{Res}||_2, \tag{2.12}
\]

where \( \text{Res} = \sum_{r=s}^{+\infty} (c_r^* - c_r) h_r \) is the residual term. Note that \( \sum_{j=1}^{k} |a_j| \leq M, \sum_{j=1}^{n} |a_j^*| = m^* \leq M \), and \( |d_j| \leq d \). We have

\[
\sqrt{n}||\text{Res}||_2 = \sqrt{n} \left| \sum_{r=s}^{+\infty} (c_r^* - c_r) h_r \right|_2 \leq \sqrt{n} \left| \sum_{r=s}^{+\infty} \left( \sum_{j=1}^{n} a_j \frac{(d_j)^r}{r! \sqrt{2r + 1}} - \sum_{p=1}^{k} a_p \frac{d_p^r}{r! \sqrt{2r + 1}} \right) h_r \right|_2
\]

\[
\leq 2M d_s \left( \sum_{r=s}^{+\infty} \frac{d_r^{s-r}}{s! \sqrt{2s + 1}} \right) \sqrt{n} \quad \text{(by (2.4))}
\]

\[
\leq 2e^d \sqrt{\pi} M d_s \quad \text{(since } d^s \leq \sigma \text{)}
\]

**Step 3.** We estimate the term \( \sqrt{n} \left| \sum_{r=0}^{s-1} (c_r^* - c_r) h_r \right|_2 \) in this step. Recall the multipole matrix

\[
H(s^*) = \left[ h_0 \ h_1 \ \cdots \ h_{s-1} \right].
\]
Denote 

\[ b(s^*) = \left( c_0^* - c_0, \cdots, c_{s^* - 1}^* - c_{s^* - 1} \right)^T. \]

Then we have

\[
\sqrt{h} \left| \sum_{r=0}^{s^*-1} (c_r^* - c_r) h_r \right|_2 \geq \sigma_{\min}(\sqrt{h} H(s^*)) \left| b(s^*) \right|_2 \geq \sigma_{\min}(s^*) \left| b(s^*) \right|_2.
\]

where \( \sigma_{\min}(s^*) = \sigma_{\min}(\sqrt{h} H(s^*)) \) is defined in Section 2.1. Note that \( b(s^*) \) can be written in the form \( \tilde{A}a - \tilde{A}^* a^* \) with \( \tilde{A}, \tilde{A}^*, a, a^* \) being defined as in Corollary 3.2. An application of the result therein yields

\[
\left| b(s^*) \right|_2 \geq \left| b(2n-1) \right|_2 \geq \frac{1.15 m_{\min}^* (d_{\min}^*)^{2n-2}}{2^{8n-8}(n-1)(1+d)^{2n-2}}.
\]

Using the separation condition (2.10), we further get

\[
\left| b(s^*) \right|_2 \geq \frac{1.15 m_{\min}^* (d_{\min}^*)^{2n-2}}{2^{8n-8}(n-1)(1+d)^{2n-2}} \geq \frac{3\sigma}{\sigma_{\min}(s^*)}.
\]

It follows that

\[
\sqrt{h} \left| \sum_{r=0}^{s^*-1} (c_r^* - c_r) h_r \right|_2 \geq 3\sigma. \tag{2.14}
\]

**Step 4.** Finally, combining (2.12), (2.13) and (2.14), we have

\[
\sqrt{h} \left| Y - [\mu \ast f] \right|_2 > 3\sigma - 2\sigma = \sigma,
\]

which proves that any discrete measure \( \mu \) with only \( k < n \) supports cannot be a \((d, \sigma, M)\)-admissible measure.

We have derived an upper bound for the computational resolution limit. This upper bound shows that the inverse problem of recovering source number is not ill-posed when the sources are separated beyond this limit. Any algorithm finding the sparsest admissible measure can determine the exact source number. This result highlights the significance of sparsity in the ill-posedness of the inverse problem.

**Example of the upper bound**

We set \( d = 1 \) and consider the following measure in the interval \([-1, 1]\)

\[
\mu^* = \delta_{-0.38} + \delta_{0.38}.
\]

Then \( n = 2, m^* = 2, m_{\min}^* = 1 \). We set the noise level \( \sigma = 5.8 \times 10^{-5} \) and the priori amplitude bound \( M = 2 \). We can calculate that

\[
s^* = \min\{l \in \mathbb{N} : \frac{d}{l!\sqrt{2l + 1}} \leq \frac{\sigma}{2\epsilon d \sqrt{\pi M}}\} = 8.
\]

We sample the image of \( \mu^* \) evenly in \([-100, 100]\) with 101 sample points (this corresponds to letting \( R = 100, h = 2 \)). The measurements are

\[
Y(x_t) = \mu^* \ast f(x_t) + W(x_t), \quad t = 1, \cdots, 101,
\]
Figure 2.1: a: The noisy image.  b: behavior of $\sigma_{\text{min}}(s)$

where $W(x_t)$ is a uniformly distributed random number in $(0, \sigma/\sqrt{2R})$. The upper bound for the computational resolution limit is

$$4.7(1 + d) \cdot \frac{\sigma}{\sigma_{\text{min}}(s^*) \cdot m_{\text{min}}^*} = 0.7597.$$  

The graph of $Y$ is shown in Figure 2.1 a. It is hard to determine visually that there are two sources. However, by Theorem 2.1, only discrete measures with at least two supports can be a $(d, \sigma, M)$-admissible measure.

Discussion of $\sigma_{\text{min}}(s)$

The separation distance in Proposition 2.1 shows that, the lower bound for the computational resolution limit is at least of the order $O(\frac{\sigma}{\sqrt{m^*}})$. The upper bound derived is

$$4.7(1 + d) \cdot \frac{\sigma}{\sigma_{\text{min}}(s^*) \cdot m_{\text{min}}^*}.$$  

The major gap lies in the factor $\sigma_{\text{min}}(s^*)$ which is determined by $s^*$ in (2.9) and the point spread function. We note that this gap is due to the correlation between the multipoles which amplifies the noise to signal ratio $\frac{\sigma}{m_{\text{min}}}$ when one reconstructs the multipole coefficients. In what follows, we estimate this factor $\sigma_{\text{min}}(s)$. Recall that

$$\sigma_{\text{min}}(s) = \sigma_{\text{min}}(\sqrt{hH(s)}) = \sqrt{\lambda_{\text{min}}(hH(s)^TH(s))}.$$  

For $f = \sin\frac{\pi x}{L}$,

$$\hat{f} = \pi \delta_{[-1, 1]}, \quad \hat{f}^{(1)} = \pi (i\omega) \delta_{[-1, 1]}, \quad \hat{f}^{(2)} = \pi (i\omega)^2 \delta_{[-1, 1]}, \quad \ldots$$
where \( \hat{f} = \int_{-\infty}^{+\infty} f e^{-i\omega x} \, dx \) denotes the Fourier transform of \( f \). We have
\[
\int_{-\infty}^{+\infty} f(p) \hat{f}(j) \, dx = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(p) \hat{f}(j) \, d\omega = \frac{\pi}{2} \int_{-1}^{1} (i\omega)^p (-i\omega)^j \, d\omega
\]
\[
= \begin{cases} 
(-1)^{p+j} \frac{\pi}{p+j+1}, & \text{if } p + j \text{ is even}, \\
0, & \text{if } p + j \text{ is odd}.
\end{cases}
\]

Now consider the \((p, j)\)-th entry in \( hH(s)H(s) \), we have
\[
h^{T}h_p h_j \to \sqrt{2p+1} \sqrt{2j+1} \int_{-\infty}^{+\infty} f^{(p)}(x) f^{(j)}(x) \, dx
\]
as \( R \to +\infty \). Thus
\[
h^{T}h_p h_j \approx \mathcal{H}_{p,j}(s) = \begin{cases} 
(-1)^{p+j} \frac{\pi \sqrt{2p+1} \sqrt{2j+1}}{p+j+1}, & \text{if } p + j \text{ is even, } p, j \leq s, \\
0, & \text{if } p + j \text{ is odd, } p, j \leq s.
\end{cases}
\]

The singular values of the matrix \( hH(s)H(s) \) can be calculated using the above approximation. The behavior of \( \sigma_{\min}(s) \) as \( s \) increases is shown in Figure 2.1:b. We see that for \( s > 20 \), \( \sigma_{\min}(s) \) oscillate above \( 10^{-8} \).

### 3 Non-linear Approximation in Vandermonde Space

In this section, we introduce the nonlinear approximation theory which is used to derive the upper bound for the computational resolution limit. We first introduce some definitions. For a given positive integer \( s \) and \( \omega \in \mathbb{R} \), we define the Vandermonde vector
\[ \phi_s(\omega) = (1, \omega, \cdots, \omega^s)^T. \]

We also define the Vandermonde space
\[ W_s = \text{span} \{ \phi_{s}(\omega) : \omega \in \mathbb{R} \}, \]
and \( k \) dimensional Vandermonde subspace
\[ W_{s}^k(\omega_1, \cdots, \omega_k) := \text{span} \{ \phi_{s}(\omega_1), \cdots, \phi_{s}(\omega_k) \}. \]

For a given vector
\[ v = \sum_{j=1}^{k+1} a_{j}^* \phi_{s}(d_{j}^*), \]
we consider the following nonlinear optimization problem in the Vandermonde space
\[
\min_{a_j, d_j \in \mathbb{R}, |d_j| \leq d_1, \cdots, d_k} \left\| \sum_{j=1}^{k} a_j \phi_s(d_j) - v \right\|_2^2. \tag{3.1}
\]

We shall derive a sharp lower bound for this minimization problem. This lower bound shall lead to the upper bound for the computational resolution limit we introduced. The key idea is a volume method to calculate the projection of \( \phi_s(\omega) \) onto the Vandermonde subspace \( W_{s}^k(\omega_1, \cdots, \omega_k) \). We first introduce some preliminary about the volume of parallelootope and Vandermonde matrices.
3.1 Preliminary

Definition 3.1. For \( s \geq k - 1 \), the k-dimensional volume of the parallelootope spanned by the vectors \( \phi_s(\omega_j), \omega_j \in \mathbb{R}, j = 1, \cdots, k \) is \( \sqrt{\det(A^T A)} \) where

\[
A = (\phi_s(\omega_1), \cdots, \phi_s(\omega_k)).
\]

We denote

\[
S_{1n}^j := \{(\tau_1, \cdots, \tau_j) : \tau_p \in \{1, \cdots, n\}, p = 1, \cdots, j \text{ and } \tau_p \neq \tau_q, \text{ for } p \neq q\}.
\]

and

\[
V_n(n) = \begin{pmatrix}
1 & \cdots & 1 \\
d_1 & \cdots & d_n \\
\vdots & \ddots & \vdots \\
d_1 & \cdots & d_n
\end{pmatrix}, \quad V_n(n-1) = \begin{pmatrix}
1 & \cdots & 1 \\
d_1 & \cdots & d_n \\
\vdots & \ddots & \vdots \\
d_1 & \cdots & d_n
\end{pmatrix}.
\]

(3.2)

Lemma 3.1. The matrix \( V_n(n) \) can be reduced to the following form by using elementary column-addition operations

\[
V_n(n)G(1) \cdots G(n-1)DQ(1) \cdots Q(n-1) = \begin{pmatrix}
1 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & 1 \\
v_{(n+1)1} & v_{(n+1)2} & \cdots & v_{(n+1)n}
\end{pmatrix}
\]

where \( G(1), \cdots, G(n-1), Q(1), \cdots, Q(n-1) \) are elementary column-addition matrices, \( D = \text{diag}(1, \frac{1}{d_2-d_1}, \cdots, \prod_{p=1}^{n-1}(d_n-d_p)) \) and

\[
v_{(n+1)j} = (-1)^{n-j} \sum_{(\tau_1, \cdots, \tau_{n+1-j}) \in S_{1n}^j} d_{\tau_1} \cdots d_{\tau_{n+1-j}}.
\]

Proof: See Appendix C.

Lemma 3.2. We have

\[
\sqrt{\frac{\det(V_n(n)^T V_n(n))}{\det(V_n(n-1)^T V_n(n-1))}} = \sqrt{\sum_{j=0}^{n} v_j^2},
\]

(3.4)

where

\[
v_j = \sum_{(\tau_1, \cdots, \tau_j) \in S_{1n}^j} d_{\tau_1} \cdots d_{\tau_j}.
\]

Proof: Note that in Lemma 3.1 all the elementary column-addition matrices have unit determinant. As a result, \( \det(V_n(n)^T V_n(n)) = \det(F^T F) \cdot \frac{1}{(\det D)^2} \), where \( F \) is the final reduced matrix. A direct calculation shows that

\[
\det(F^T F) = \sum_{j=0}^{n} v_j^2.
\]
On the other hand, $V_n(n-1)$ is a standard Vandermonde matrix. Its determinant $\det(V_n(n-1)) = \det(V_n(n-1)^T) = \frac{1}{\det \hat{A}}$. Therefore (3.4) follows.

**Corollary 3.1.** If $|d_j| < d, j = 1, \ldots, n$, for $V_n(n), V_n(n-1)$ in Lemma 3.2 we have

$$\sqrt{\frac{\det(V_n(n)^2 V_n(n))}{\det(V_n(n-1)^2 V_n(n-1))}} \leq (1 + d)^n.$$  

(3.5)

**Proof:** Use (3.4) and the estimate that

$$\sqrt{\sum_{j=0}^n v_j^2} \leq \sum_{j=0}^n |v_j| \leq n \binom{n}{j} d^j = (1 + d)^n.$$

**Lemma 3.3.** Let $V = W^k(\omega_1, \ldots, \omega_k)$ where $\omega_1, \ldots, \omega_k$ are $k$ different real numbers and let $V^\perp$ is the orthogonal complement of $V$ in $\mathbb{R}^k$. Let $P$ be the orthogonal projection onto $V^\perp$, then

$$\|P(v)\|_2 = \sqrt{\frac{\det(A^T \hat{A})}{\det(A^2 A)}},$$

where

$$A = \left( \begin{array}{ccc} \phi_k(\omega_1) & \phi_k(\omega_2) & \cdots & \phi_k(\omega_k) \end{array} \right) \text{ and } \hat{A} = \left( A, v \right).$$

**Proof:** The conclusion follows from the observation that the $k+1$-dimensional volume of the parallelootope spanned by the vectors $\phi_k(\omega_1), \cdots, \phi_k(\omega_k)$ and $v$ can be computed as the product of $\|P(v)\|_2$ and the $k$-dimensional volume of the parallelootope spanned by the $k$ vectors $\phi_k(\omega_1), \cdots, \phi_k(\omega_k)$.

### 3.2 Lower bound for the approximation in Vandermonde space

In this section we derive a lower bound for the non-linear approximation problem (3.1). We first introduce two notations. We define for integer $k \geq 1$,

$$\zeta(k+1) = \left\{ \begin{array}{ll} \frac{(k+1)!(k+2)!}{(k+2)^2}, & k \text{ odd,} \\ \frac{(k+1)!}{(k+2)}, & k \text{ even,} \end{array} \right\} \text{ and } \xi(k) = \left\{ \begin{array}{ll} \frac{3^k}{4}, & k = 1, \\ \frac{(k+1)!}{(k+2)}, & k \text{ odd, } k \geq 3, \\ \frac{(k+1)!}{(k+2)^2}, & k \text{ even.} \end{array} \right\} \quad (3.6)$$

**Theorem 3.1.** Let $k \geq 1$. Assume that $-d \leq d_1^* < d_2^* < \cdots < d_{k+1}^* \leq d$ and $|a_j^*| \geq m_{\min}^*, j = 1, \cdots, k+1$. Let $d_{\min}^* := \min_{j \neq j} |d_j^* - d_j^*|$. For $q \leq k$, let

$$A(q) = \left( \begin{array}{ccc} 1 & \cdots & 1 \\ d_1 & \cdots & d_q \\ \vdots & \vdots & \vdots \\ d_1^{2k} & \cdots & d_q^{2k} \end{array} \right), \ a(q) = \left( \begin{array}{c} a_1 \\ a_2 \\ \vdots \\ a_q \end{array} \right), \ A^* = \left( \begin{array}{ccc} 1 & \cdots & 1 \\ d_1^* & \cdots & d_{k+1}^* \\ \vdots & \vdots & \vdots \\ (d_1^{*})^{2k} & \cdots & (d_{k+1}^{*})^{2k} \end{array} \right), \ a^* = \left( \begin{array}{c} a_1^* \\ a_2^* \\ \vdots \\ a_{k+1}^* \end{array} \right).$$

15
Then
\[
\min_{a_p, d_p \in \mathbb{R}, \|d_p\| \leq d, p = 1, \ldots, q} \|A(q)a(q) - A^*a^*\|_2 \geq \frac{\zeta(k + 1)\zeta(k)m_{\min}^*(d_{\min}^*)^{2k}}{(1 + d)^{2k}}.
\]

Proof: Step 1. Note that for \(q < k\), we have
\[
\min_{a_p, d_p \in \mathbb{R}, \|d_p\| \leq d, p = 1, \ldots, q} \|A(q)a(q) - A^*a^*\|_2 \geq \min_{a_p, d_p \in \mathbb{R}, \|d_p\| \leq d, p = 1, \ldots, k} \|A(k)a(k) - A^*a^*\|_2.
\]
So we need only to consider the case when \(q = k\). It suffices to show that for any given \(-d \leq d_1 < d_2 < \cdots < d_k \leq d\), the following holds
\[
\min_{a_p, p = 1, \ldots, k} \|A(k)a(k) - A^*a^*\|_2 \geq \frac{\zeta(k + 1)\zeta(k)m_{\min}^*(d_{\min}^*)^{2k}}{(1 + d)^{2k}}.
\]
So we fix \(d_1, \ldots, d_k\) in our subsequent argument.

Step 2. We claim that
\[
\max_{l=0, \ldots, k} \min_{a \in \mathbb{R}^k} \|A_l a - A_l^*a^*\|_2 \geq \frac{\zeta(k + 1)\zeta(k)m_{\min}^*(d_{\min}^*)^{2k}}{(1 + d)^{2k}},
\]
where
\[
A_l = \begin{pmatrix}
  d_1^l & \cdots & d_k^l \\
  d_1^{l+1} & \cdots & d_k^{l+1} \\
  \vdots & \vdots & \vdots \\
  d_1^{l+k} & \cdots & d_k^{l+k}
\end{pmatrix},
\]
\[
A_l^* = \begin{pmatrix}
  (d_1^*)^l & \cdots & (d_k^*)^l \\
  (d_1^*)^{l+1} & \cdots & (d_k^*)^{l+1} \\
  \vdots & \vdots & \vdots \\
  (d_1^*)^{l+k} & \cdots & (d_k^*)^{l+k}
\end{pmatrix}.
\]
The claim will be proved in subsequent two steps.

Step 2.1. For each \(l\), from the decomposition
\[
A_l = \begin{pmatrix}
  d_1^l & \cdots & d_k^l \\
  d_1^{l+1} & \cdots & d_k^{l+1} \\
  \vdots & \vdots & \vdots \\
  d_1^{l+k} & \cdots & d_k^{l+k}
\end{pmatrix} = \begin{pmatrix}
  1 & \cdots & 1 \\
  d_1 & \cdots & d_k \\
  \vdots & \vdots & \vdots \\
  d_1^k & \cdots & d_k^k
\end{pmatrix} \begin{pmatrix}
  d_1^l & 0 & \cdots & 0 \\
  0 & d_2^l & \cdots & 0 \\
  \vdots & \vdots & \vdots & \vdots \\
  0 & 0 & \cdots & d_k^l
\end{pmatrix},
\]
and the similar decomposition for \(A_l^*\), we have
\[
\min_{a \in \mathbb{R}^k} \|A_l a - A_l^*a^*\|_2 \geq \min_{a \in \mathbb{R}^k} \|A_0 a_l - A_0^*a_l^*\|_2, \tag{3.7}
\]
where \(a_l^* = (a_1^*(d_1^*)^l, \cdots, a_{k+1}^*(d_{k+1}^*)^l)^T\). Let \(V\) be the space spanned by column vectors of \(A_0\). Then the dimension of \(V\) is \(k\), and \(V^\perp\), the orthogonal complement of \(V\) in \(\mathbb{R}^{k+1}\) is of dimension one. We let \(v\) be a unit vector in \(V^\perp\) and let \(P_2\) be the orthogonal projection onto \(V^\perp\). Observe that
\[
\min_{a_l \in \mathbb{R}^k} \|A_0 a_l - A_0^*a_l^*\|_2 = \|P_2(A_0^*a_l^*)\|_2 = \left\| \sum_{j=1}^{k+1} a_j^*(d_j^*)^l P_2(\phi_k(d_j^*)) \right\|_2,
\]
where
where $\phi_k(d^*_j) = (1, d^*_j, \cdots, (d^*_j)^k)^T$. We write

$$P_2(\phi_k(d^*_j)) = (-1)^{\sigma(j)}|P_2(\phi_k(d^*_j))|v,$$  \hfill (3.8)

where $\sigma(j) = 0$ or $1$ depends on the direction of $P_2(\phi_k(d^*_j))$. Denote

$$\hat{A}_j = \begin{pmatrix} 1 & \cdots & 1 & 1 \\ d_1 & \cdots & d_k & d^*_j \\ \vdots & \vdots & \vdots & \vdots \\ d^*_{k-1} & \cdots & d^*_{k-1} & \sigma(d^*_{k-1}) \\ d^*_k & \cdots & d^*_k & \sigma(d^*_k) \end{pmatrix}, \quad \hat{A} = \begin{pmatrix} 1 & \cdots & 1 \\ d_1 & \cdots & d_k \\ \vdots & \vdots & \vdots \\ d^*_{k-1} & \cdots & d^*_{k-1} \\ d^*_k & \cdots & d^*_k \end{pmatrix}.$$

We have

$$\left\| \sum_{j=1}^{k+1} a^*_j (d^*_j)^l P_2(\phi_k(d^*_j)) \right\|_2 = \left\| \sum_{j=1}^{k+1} (-1)^{\sigma(j)} a^*_j (d^*_j)^l |P_2(\phi_k(d^*_j))| \right\| \quad \text{(by (3.8))}$$

$$= \frac{1}{\sqrt{\det(A^T_0 A_0)}} \left\| \sum_{j=1}^{k+1} (-1)^{\sigma(j)} a^*_j (d^*_j)^l |\Pi_{1 \leq i < p \leq k} (d_i - d_p) \Pi_{q=1}^k (d^*_j - d_q)| \right\| \quad \text{(by Lemma 3.3)}$$

$$\geq \frac{|\Pi_{1 \leq i < p \leq k} (d_i - d_p)|}{(1 + d)^k \sqrt{\det(A^T_0 A_0)}} \left\{ \sum_{j=1}^{k+1} (-1)^{\sigma(j)} a^*_j (d^*_j)^l |\Pi_{q=1}^k (d^*_j - d_q)| \right\} \quad \text{(by Corollary 3.1)}$$

$$= \frac{1}{(1 + d)^k} \left\| \sum_{j=1}^{k+1} (-1)^{\sigma(j)} a^*_j (d^*_j)^l |\Pi_{q=1}^k (d^*_j - d_q)| \right\|. \quad \text{(3.9)}$$

**Step 2.2.** For $l = 0, 1, \cdots, k$, we let

$$\beta_l = \sum_{j=1}^{k+1} (-1)^{\sigma(j)} a^*_j (d^*_j)^l |\Pi_{q=1}^k (d^*_j - d_q)|,$$

and write the above $k + 1$ equations into the following matrix form

$$B \eta = \beta,$$

where $\beta = (\beta_0, \beta_1, \cdots, \beta_k)^T$ and

$$B = \begin{pmatrix} a^*_1 & a^*_2 & \cdots & a^*_k \\ a^*_1 d^*_1 & a^*_2 d^*_2 & \cdots & a^*_k d^*_k \\ \vdots & \vdots & \vdots & \vdots \\ a^*_1 (d^*_1)^k & a^*_2 (d^*_2)^k & \cdots & a^*_k (d^*_k)^k \end{pmatrix}, \quad \eta = \begin{pmatrix} (-1)^{\sigma(1)} |(d^*_1 - d_1) \cdots (d^*_k - d_k)| \\ (-1)^{\sigma(2)} |(d^*_2 - d_1) \cdots (d^*_2 - d_k)| \\ \vdots \\ (-1)^{\sigma(k+1)} |(d^*_k - d_1) \cdots (d^*_k - d_k)| \end{pmatrix}.$$
By Corollary 9.1, we have

\[ \| \eta \|_\infty = \| B^{-1} \beta \|_\infty \leq \| B^{-1} \|_\infty \| \beta \|_\infty \leq \frac{(1 + d)^k}{\zeta (k + 1) m_{\text{min}}^* (d_{\text{min}}^*)^k \| \beta \|_\infty} \]

Therefore,

\[ \| \beta \|_\infty \geq \frac{\zeta (k + 1) m_{\text{min}}^* (d_{\text{min}}^*)^k}{(1 + d)^k} \| \eta \|_\infty \]

\[ \geq \frac{\zeta (k + 1) \zeta(k) m_{\text{min}}^* (d_{\text{min}}^*)^{2k}}{(1 + d)^k}. \] (by Lemma 3.4)

Recall (3.9), we have

\[ \max_{l=0,\ldots,k} \left\| \sum_{j=1}^{k+1} a_j^*(d_j^*)^l P_2(\phi_k(d_j^*)) \right\|_2 \geq \frac{\zeta (k + 1) \zeta(k) m_{\text{min}}^* (d_{\text{min}}^*)^{2k}}{(1 + d)^k}, \]

or equivalently,

\[ \max_{l=0,\ldots,k} \min_{a_i \in \mathbb{R}^k} \left\| A_0 a_i - A_0^* a_i^* \right\|_2 \geq \frac{\zeta (k + 1) \zeta(k) m_{\text{min}}^* (d_{\text{min}}^*)^{2k}}{(1 + d)^k}. \]

Thus by (3.7),

\[ \max_{l=0,\ldots,k} \min_{a_i \in \mathbb{R}^k} \left\| A_l a - A_l^* a^* \right\|_2 \geq \frac{\zeta (k + 1) \zeta(k) m_{\text{min}}^* (d_{\text{min}}^*)^{2k}}{(1 + d)^k}. \]

Step 3. Finally, observing that

\[ \min_{a(k) \in \mathbb{R}^k} \left\| A(k) a(k) - A^* a^* \right\|_2 \geq \max_{l=0,\ldots,k} \min_{a_i \in \mathbb{R}^k} \left\| A_l a - A_l^* a^* \right\|_2, \]

we have

\[ \min_{a(k) \in \mathbb{R}^k} \left\| A(k) a(k) - A^* a^* \right\|_2 \geq \frac{\zeta (k + 1) \zeta(k) m_{\text{min}}^* (d_{\text{min}}^*)^{2k}}{(1 + d)^k}. \]

This proves the theorem.

We next demonstrate the sharpness of the lower bound above.

**Proposition 3.1.** Let \( k \geq 1 \). There exist \( -d \leq d_1^* < d_2^* < \cdots < d_{k+1}^* \leq d \) and nonzero real numbers \( a_1^*, \ldots, a_{k+1}^* \) with \( \sum_{j=1}^{k+1} |a_j^*| = m^* \) such that

\[ \min_{a_p, d_p \in \mathbb{R}, \sum_{j=1}^{k+1} |a_j^*| = m^*} \left\| A(k) a(k) - A^* a^* \right\|_2 \leq 2 m^* k^2 (d_{\text{min}}^*)^{2k}, \]

where \( A(k), a(k), A^*, a^*, d_{\text{min}}^* \) are defined in Theorem 3.1.
Proof: Let 
\[ \hat{d}_1 = -d, \hat{d}_2 = -d + \frac{d}{k}, \ldots, \hat{d}_{2k} = d - \frac{d}{k}, \hat{d}_{2k+1} = d. \]
There exists nonzero \( \hat{a} \) such that 
\[ Q \hat{a} = \gamma, \]  
where 
\[
Q = \begin{pmatrix}
1 & \cdots & 1 & \cdots & 1 \\
\hat{d}_1 & \frac{1}{(\hat{d}_1)^2} & \frac{1}{(\hat{d}_2)^2} & \cdots & \frac{1}{(\hat{d}_{2k})^2} & \frac{1}{(\hat{d}_{2k+1})^2} \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\left(\hat{d}_1\right)^{2k} & \left(\hat{d}_2\right)^{2k} & \cdots & \left(\hat{d}_{2k}\right)^{2k} & \left(\hat{d}_{2k+1}\right)^{2k} \\
\end{pmatrix}, \quad \hat{a} = \begin{pmatrix}
\hat{a}_1 \\
\hat{a}_2 \\
\vdots \\
\hat{a}_{2k+1} \\
\end{pmatrix}, \quad \gamma = \begin{pmatrix}
0 \\
\vdots \\
0 \\
\sum_{j=1}^{2k+1} \hat{a}_j (\hat{d}_j)^{2k} \\
\end{pmatrix}.
\]
Using the linear independence of the column vectors in the matrix \( Q \), we can show that all \( \hat{a}_j \)'s are nonzero. Without loss of generality, we assume that the first \( (k+1) \)-largest components of \( (\hat{a}_1, \ldots, \hat{a}_{2k+1})^T \) are given by \( \hat{a}_1, \ldots, \hat{a}_{k+1} \). We then define \( d_1^* = \hat{d}_1, \ldots, d_{k+1}^* = \hat{d}_{k+1}, a_1^* = \hat{a}_1, \ldots, a_{k+1}^* = \hat{a}_{k+1} \). \( m^* = \sum_{j=1}^{k+1} |\hat{a}_j|, d_1 = \hat{d}_{k+2}, \ldots, d_k = \hat{d}_{2k+1} \) and \( a_1 = -\hat{a}_{k+2}, \ldots, a_k = -\hat{a}_{2k+1} \).
Note that \( d_{\text{min}}^* = \min_{i \neq j} |d_i^* - d_j^*| \geq \frac{d}{k} \). We have
\[ ||A^* a^* - A(k)a(k)||_2 = ||Q\hat{a}||_2 = ||\gamma||_2 = |\sum_{j=1}^{2k+1} \hat{a}_j (\hat{d}_j)^{2k}| \leq 2m^* d^{2k} \leq 2m^* (kd_{\text{min}}^*)^{2k} = 2m^* k^{2k}(d_{\text{min}}^*)^{2k}. \]
It follows that
\[ \min_{a_p, d_p \in \mathbb{R} |d_p| \leq d, p=1, \ldots, k} ||A(k)a(k) - A^* a^*||_2 \leq 2m^* k^{2k}(d_{\text{min}}^*)^{2k}. \]

Finally, as a consequence of Theorem 3.1, we have the following result.

**Corollary 3.2.** Let \( k \geq 1 \). Assume that \( -d \leq d_1^* < d_2^* < \cdots < d_{k+1}^* \leq d \) and \( a_1^*, \ldots, a_{k+1}^* \) with \( |a_j^*| \geq m_{\text{min}}^*, j = 1, \ldots, k+1 \). For \( q \leq k \), we have
\[ \min_{a_p, d_p \in \mathbb{R} |d_p| \leq d, p=1, \ldots, q} ||\tilde{A}(q)a(q) - \tilde{A}^* a^*||_2 \geq \frac{1.15m^* (d_{\text{min}}^*)^{2k}}{2^{4k} k(1+d)^{2k}}, \]
where \( d_{\text{min}}^* := \min_{i \neq j} |d_i^* - d_j^*| \) and 
\[
\tilde{A}(q) = \begin{pmatrix}
1 & \cdots & 1 \\
\frac{d_1}{\sqrt{3}} & \cdots & \frac{d_q}{\sqrt{3}} \\
\frac{d_1^2}{2\sqrt{5}} & \cdots & \frac{d_q^2}{2\sqrt{5}} \\
\vdots & \ddots & \vdots \\
\frac{d_1^{2k}}{(2k)!\sqrt{4k+1}} & \cdots & \frac{d_q^{2k}}{(2k)!\sqrt{4k+1}} \\
\end{pmatrix}, \quad a(q) = \begin{pmatrix}
a_1 \\
a_2 \\
\vdots \\
a_q \\
\end{pmatrix}, \quad \tilde{A}^* = \begin{pmatrix}
1 & \cdots & 1 \\
\frac{d_1^*}{\sqrt{3}} & \cdots & \frac{d_q^*}{\sqrt{3}} \\
\frac{d_1^{*2}}{2\sqrt{5}} & \cdots & \frac{d_q^{*2}}{2\sqrt{5}} \\
\vdots & \ddots & \vdots \\
\frac{d_1^{*2k}}{(2k)!\sqrt{4k+1}} & \cdots & \frac{d_q^{*2k}}{(2k)!\sqrt{4k+1}} \\
\end{pmatrix}, \quad a^* = \begin{pmatrix}
a_1^* \\
a_2^* \\
\vdots \\
a_{k+1}^* \\
\end{pmatrix}.
\]
Theorem 3.2. Let $k \geq 2$. Assume that $-d \leq d_1^* < d_2^* < \cdots < d_k^* \leq d$ and $|a_j^*| \geq m_{\min}^*, j = 1, \cdots, k$. Let $d_{\min}^* := \min_{i \neq j} |d_i^* - d_j^*|$. Assume that $-d \leq d_1 < d_2 < \cdots < d_k \leq d$ and let

$$
\eta = \begin{pmatrix}
(d_1^* - d_1) \cdots (d_1^* - d_k) \\
(d_2^* - d_1) \cdots (d_2^* - d_k) \\
\vdots \\
(d_k^* - d_1) \cdots (d_k^* - d_k)
\end{pmatrix}.
$$

If

$$
||Aa - A^* a^*||_2 < \sigma,
$$

(3.12)

where

$$
A = \begin{pmatrix}
1 & \cdots & 1 \\
d_1 & \cdots & d_k \\
\vdots & \vdots & \vdots \\
d_1^{2k-1} & \cdots & d_k^{2k-1}
\end{pmatrix},
a = \begin{pmatrix}
a_1 \\
a_2 \\
\vdots \\
a_k
\end{pmatrix},
a^* = \begin{pmatrix}
a_1^* \\
a_2^* \\
\vdots \\
a_k^*
\end{pmatrix}.
$$

(3.13)

Then

$$
||\eta||_\infty < \frac{(1 + d)^{2k-1}}{\zeta(k)(d_{\min}^*)^{k-1}} \frac{\sigma}{m_{\min}^*}.
$$

Proof: Since $||Aa - A^* a^*||_2 < \sigma$, we have

$$
\min_{a \in \mathbb{R}^k} ||Aa - A^* a^*||_2 < \sigma.
$$

(3.14)

Similar to the proof of Theorem 3.1, we consider

$$
\max_{l=0, \cdots, k-1} \min_{a \in \mathbb{R}^k} ||A_l a - A^*_l a^*||_2,
$$
where 

$$A_l = \begin{pmatrix} d_1^l & \cdots & d_k^l \\ d_1^{l+1} & \cdots & d_k^{l+1} \\ \vdots & \vdots & \vdots \\ d_1^{l+k} & \cdots & d_k^{l+k} \end{pmatrix}, \quad A^*_l = \begin{pmatrix} (d_1^*)^l & \cdots & (d_k^*)^l \\ (d_1^{l+1})^* & \cdots & (d_k^{l+1})^* \\ \vdots & \vdots & \vdots \\ (d_1^{l+k})^* & \cdots & (d_k^{l+k})^* \end{pmatrix}. $$

For each \(l\), from the decomposition

$$A_l = \begin{pmatrix} d_1^l & \cdots & d_k^l \\ d_1^{l+1} & \cdots & d_k^{l+1} \\ \vdots & \vdots & \vdots \\ d_1^{l+k} & \cdots & d_k^{l+k} \end{pmatrix} = \begin{pmatrix} 1 & \cdots & 1 \\ d_1 & \cdots & d_k \\ \vdots & \vdots & \vdots \\ d_1 & \cdots & d_k \end{pmatrix} \begin{pmatrix} d_1^l & 0 & \cdots & 0 \\ 0 & d_2^l & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & d_k^l \end{pmatrix},$$

and the similar decomposition for \(A^*_l\), we have

$$\min_{a \in \mathbb{R}^k} \|A_0 a - A^*_0 a^*\|_2 \geq \min_{a \in \mathbb{R}^k} \|A_0 a - A^*_0 a^*\|_2,$$

(3.15)

where \(a^*_j = (d_1^*(d_1^*)^j, \cdots, d_k^*(d_k^*)^j)^T\). Let \(V\) be the space spanned by column vectors of \(A_0\). Then the dimension of \(V\) is \(k\), and \(V^\perp\), the orthogonal complement of \(V\) in \(\mathbb{R}^{k+1}\) is of dimension one. We let \(v\) be a unit vector in \(V^\perp\) and let \(P_2\) be the orthogonal projection onto \(V^\perp\). We observe that

$$\min_{a \in \mathbb{R}^k} \|A_0 a - A^*_0 a^*\|_2 = \|P_2(A_0 a^*)\|_2 = \left\| \sum_{j=1}^k a_j^*(d_j^*)^l P_2(\phi_k(d_j^*)) \right\|_2,$$

where \(\phi_k(d_j^*) = (1, d_j^*, \cdots, (d_j^*)^k)^T\). We write

$$P_2(\phi_k(d_j^*)) = (-1)^{\sigma(j)}|P_2(\phi_k(d_j^*))|v,$$

where \(\sigma(j) = 0\) or \(1\) depends on the direction of \(P_2(\phi_k(d_j^*))\). Similar to Step 2.1 in the proof of Theorem 3.1 we have

$$\left\| \sum_{j=1}^k a_j^*(d_j^*)^l P_2(\phi_k(d_j^*)) \right\|_2 \geq \frac{1}{(1+d)^k} \left\| \sum_{j=1}^k (-1)^{\sigma(j)} a_j^*(d_j^*)^l |\Pi_{q=1}^k (d_j^* - d_q)| \right\|_2, \quad (3.16)$$

For \(l = 0, \cdots, k - 1\), we let

$$\beta_l = \sum_{j=1}^k (-1)^{\sigma(j)} a_j^*(d_j^*)^l |\Pi_{q=1}^k (d_j^* - d_q)|,$$

and \(\beta = (\beta_0, \cdots, \beta_{k-1})^T\). Similar to Step 2.2 in the proof of Theorem 3.1 we have

$$\|\beta\|_\infty \geq \frac{\zeta(k) m_{\min}^*(d_{\min}^*)^{k-1}}{(1+d)^{k-1}} \|\eta\|_\infty.$$
Recall (3.16), we have
\[
\max_{l=0,\ldots,k-1} \left\| \sum_{j=1}^{k} a_j^* (d_j^*)^l P_2(\phi_k(d_j^*)) \right\|_2 \geq \frac{\zeta(k) m_{\min}^* (d_{\min}^*)^{k-1}}{(1 + d)^{2k-1}} ||\eta||_{\infty},
\]

or equivalently,
\[
\max_{l=0,\ldots,k-1} \min_{\alpha \in \mathbb{R}^k} ||A_0 \alpha_l - A_0^* \alpha_l^*||_2 \geq \frac{\zeta(k) m_{\min}^* (d_{\min}^*)^{k-1}}{(1 + d)^{2k-1}} ||\eta||_{\infty}.
\]

Thus by (3.15),
\[
\max_{l=0,\ldots,k-1} \min_{\alpha \in \mathbb{R}^k} ||A_1 \alpha - A_1^* \alpha^*||_2 \geq \frac{\zeta(k) m_{\min}^* (d_{\min}^*)^{k-1}}{(1 + d)^{2k-1}} ||\eta||_{\infty}.
\]

Recall (3.14) and observe that
\[
\sigma > \min_{\alpha \in \mathbb{R}^k} ||A_\alpha - A_\alpha^* \alpha^*||_2 \geq \max_{l=0,\ldots,k-1} \min_{\alpha \in \mathbb{R}^k} ||A_1 \alpha - A_1^* \alpha^*||_2 \geq \frac{\zeta(k) m_{\min}^* (d_{\min}^*)^{k-1}}{(1 + d)^{2k-1}} ||\eta||_{\infty},
\]

we have
\[
||\eta||_{\infty} < \frac{(1 + d)^{2k-1}}{\zeta(k)(d_{\min}^*)^{k-1} m_{\min}^*} \sigma.
\]

**Corollary 3.3.** Let \( k \geq 2 \). Assume that \(-d \leq d_1^* < d_2^* < \cdots < d_k^* \leq d \) and \( |a_j^*| \geq m_{\min}^*, j = 1, \ldots, k \). Let \(-d \leq d_1 < d_2 < \cdots < d_k \leq d \) and assume that
\[
||\tilde{A} \alpha - \tilde{A}^* \alpha^*||_2 < \sigma,
\]

where
\[
A = \begin{pmatrix}
1 & \cdots & 1
\frac{d_1}{\sqrt{3}} & \cdots & \frac{d_k}{\sqrt{3}}
\frac{d_1^*}{2\sqrt{5}} & \cdots & \frac{d_k^*}{2\sqrt{5}}
\vdots & \ddots & \vdots
\frac{d_1^{k-1}}{(2k-1)!\sqrt{4k-1}} & \cdots & \frac{d_k^{k-1}}{(2k-1)!\sqrt{4k-1}}
\end{pmatrix},
\]
\[
a = \begin{pmatrix}
a_1 \\
a_2 \\
a_3 \\
\vdots \\
a_k
\end{pmatrix}, \quad A^* = \begin{pmatrix}
1 & \cdots & 1
\frac{d_1}{\sqrt{3}} & \cdots & \frac{d_k}{\sqrt{3}}
\frac{d_1^*}{2\sqrt{5}} & \cdots & \frac{d_k^*}{2\sqrt{5}}
\vdots & \ddots & \vdots
\frac{d_1^{k-1}}{(2k-1)!\sqrt{4k-1}} & \cdots & \frac{d_k^{k-1}}{(2k-1)!\sqrt{4k-1}}
\end{pmatrix}, \quad a^* = \begin{pmatrix}
a_1^* \\
a_2^* \\
a_3^* \\
\vdots \\
a_k^*
\end{pmatrix}.
\]

Then
\[
||\eta||_{\infty} < \frac{(2k-1)!\sqrt{4k-1}(1 + d)^{2k-1}}{\zeta(k)(d_{\min}^*)^{k-1}} \frac{\sigma}{m_{\min}^*},
\]

where \( \eta, d_{\min}^* , \zeta(k) \) defined in Theorem 3.2.

Proof: By Theorem 3.2 and a similar argument as in proof of Corollary 3.2.

Finally, we present useful lemma which will be used in Section 4, and whose proof is given in Appendix B.
Lemma 3.4. Let \( k \geq 1 \). Assume that \( d_1^* < d_2^* < \cdots < d_{k+1}^* \) and let \( d_{\min}^* = \min_{i \neq j} |d_i^* - d_j^*| \). Then for real numbers \( d_1 \leq d_2 \leq \cdots \leq d_k \), we have the following estimate
\[
||\eta||_\infty \geq \xi(k)(d_{\min}^*)^k,
\]
where
\[
\eta = \begin{pmatrix}
(d_1^* - d_1) \cdots (d_1^* - d_k) \\
(d_2^* - d_1) \cdots (d_2^* - d_k) \\
\vdots \\
(d_{k+1}^* - d_1) \cdots (d_{k+1}^* - d_k)
\end{pmatrix}. \tag{3.18}
\]

4 Stability Analysis of Support Recovery

In this section, we study the stability of recovering source supports from the admissible measures and quantitatively demonstrate its dependence on the super-resolution factor. In the grid setting, the super-resolution factor can be defined as the ratio between Rayleigh limit and the grid spacing, see for instance \([9]\). In our non-grid setting, recalling that the point spread function is \( \sin x / x \) and the associated Rayleigh limit is \( \pi \), we define the super-resolution factor as
\[
\text{SRF} := \frac{\pi}{d_{\min}^*}.
\]
We have the following stability result.

Theorem 4.1. Let \( n \geq 2 \) and let \( \mu^* = \sum_{j=1}^{n} a_j^* \delta_{y_j^*} \) be a measure such that \(-d \leq y_1^* < y_2^* < \cdots < y_n^* \leq d\). Assume that the following separation condition holds
\[
d_{\min}^* = \min_{i \neq j} |y_i^* - y_j^*| \geq 6.24(1 + d) \sqrt{\frac{3}{\sigma_{\min}(s^*)}} \frac{\sigma}{m_{\min}^*},
\]
where \( s^* \) is defined in \([2,9]\). If \( \mu = \sum_{j=1}^{n} a_j \delta_{y_j} \) is a \((d, \sigma, M)\)-admissible measure such that \(-d \leq y_1 < y_2 < \cdots < y_n \leq d\), then
\[
|y_i - y_i^*| < \frac{d_{\min}^*}{2}, \quad i = 1, \cdots, n.
\]
Moreover,
\[
|y_i - y_i^*| < C(n, d) \text{SRF}^{2n-2} \frac{3\sigma}{\sigma_{\min}(s^*)m_{\min}^*},
\]
where
\[
C(n, d) = \frac{7.73\sqrt{4n-1}}{4}\frac{(2n-1)(4 + 4d)^{2n-1}}{4e\pi^{2n-\frac{1}{2}}}. \tag{2.19}
\]
Proof: **Step 1.** We show that \( s^* \geq 2n \). Since \( d \geq \frac{(n-1)d_{\min}^n}{2} \), we have

\[
\frac{d^{2n-1}}{(2n-1)!\sqrt{4n-1}} \geq \left( \frac{1}{2} \right)^{2n-1}(n-1)^{2n-1}(d_{\min}^n)^{2n-1}
\]

Recalling (2.9), we get \( s^* \geq 2n \).

**Step 2.** Let \( \mu = \sum_{j=1}^{n} a_j \delta_{y_j} \) be a \((d, \sigma, M)\)-admissible measure such that

\[
\sqrt{h}||Y - [\mu \ast f]||_2 \leq \sigma.
\]

Based on the decomposition,

\[
Y - [\mu \ast f] = \sum_{r=0}^{\infty} (c_r^* - c_r)h_r + W,
\]

where \( c_r^* = \sum_{j=1}^{n} a_j \frac{(d_r^*)^r}{r!(2r+1)} \) and \( c_r = \sum_{j=1}^{n} a_j \frac{d_r^r}{r!(2r+1)} \), we have

\[
\sqrt{h}||Y - [\mu \ast f]||_2 \geq \sqrt{h}|| \sum_{r=0}^{s-1} (c_r^* - c_r)h_r ||_2 - \sqrt{h}||W||_2 - \sqrt{h}||Res||_2.
\]

Here Res is the residual term and \( s^* \) is defined in (2.9). By a similar argument as in Step 2 and 3 of the proof of Theorem 2.1 we have

\[
\sqrt{h}||Res|| < \sigma, \text{ and } \sqrt{h} || \sum_{r=0}^{s-1} (c_r^* - c_r)h_r ||_2 \geq ||b(s^*)||_2 \sigma_{\min}(s^*),
\]

where

\[
b(s^*) = (c_0^* - c_0, \ldots, c_{s-1}^* - c_{s-1})^T, \quad \sigma_{\min}(s^*) = \sigma_{\min}(\sqrt{h}H(s^*)),
\]

Here recall that \( H(s^*) = \begin{pmatrix} h_0 & h_1 & \cdots & h_{s-1} \end{pmatrix} \) is the multipole matrix. Therefore, by constraint (4.1) and estimate (4.2)-(4.3), we have

\[
||b(s^*)||_2 < \frac{3\sigma}{\sigma_{\min}(s^*)},
\]

Since by Step 1, \( s^* \geq 2n \), we have \( ||b(2n)||_2 \leq ||b(s^*)||_2 < \frac{3\sigma}{\sigma_{\min}(s^*)} \). Note that \( b(2n) \) can be written in the form \( \tilde{a} \tilde{a}^* \) with \( \tilde{a}, \tilde{a}^* \), \( a, a^* \) being defined as in Corollary 3, Application of the result therein yields

\[
||\eta||_\infty < \frac{(1+d)^{2n-1}}{\zeta(n)(d_{\min}^n)^{n-1}} \sqrt{4n-1}(2n-1)!\frac{3\sigma}{\sigma_{\min}(s^*)} m_{\min}^n,
\]
where
\[
\eta = \left\{ \begin{array}{l}
(d_1^* - d_1) \cdots (d_n^* - d_n) \\
(d_2^* - d_1) \cdots (d_n^* - d_n) \\
\vdots \\
(d_n^* - d_1) \cdots (d_n^* - d_n)
\end{array} \right. \quad .
\]

(4.5)

**Step 3.** By the separation condition and Lemma 9.7 in appendix D, we have
\[
d_{\min}^* \geq 2n^{-1} \sqrt{\frac{(1 + d)^{2n-1} \sqrt{4n-1(2n-1)!}}{\zeta(n)\lambda(n)}} \frac{3}{\sigma_{\min(s^*)} m_{\min}^*},
\]
where
\[
\lambda(n) = \left\{ \begin{array}{l}
1, \\
\xi(n-2), \\
\xi(n-2) = \left\{ \begin{array}{ll}
\frac{7n^5}{4^2}, & n \text{ is odd, } n \geq 5, \\
\frac{2n^3}{4}, & n \text{ is even, } n \geq 4.
\end{array} \right.
\end{array} \right.
\]

Or equivalently,
\[
(d_{\min}^*)^{2n-1} \geq \frac{(1 + d)^{2n-1} \sqrt{4n-1(2n-1)!}}{\zeta(n)\lambda(n)} \frac{3\sigma}{\sigma_{\min(s^*)} m_{\min}^*}.
\]

(4.7)

**Step 4.** We prove the case when \( n = 2 \). We first claim that for each \( d_i^* \), there is one \( d_j \) such that \( |d_i^* - d_j| < \frac{d_{\min}^*}{2} \). By contradiction, assume that
\[
|d_i^* - d_1| \geq \frac{d_{\min}^*}{2}, \quad |d_i^* - d_2| \geq \frac{d_{\min}^*}{2}.
\]
Then, \( ||\eta||_{\infty} \geq |d_i^* - d_1||d_i^* - d_2| \geq \frac{(d_{\min}^*)^2}{4} \). On the other hand, by letting \( n = 2 \) in (4.7), we have
\[
(d_{\min}^*)^3 \geq \frac{(1 + d)^{2n-1} \sqrt{4n-1(2n-1)!}}{\zeta(n)} \frac{3\sigma}{\sigma_{\min(s^*)} m_{\min}^*}.
\]
Therefore,
\[
\frac{(d_{\min}^*)^2}{4} \geq \frac{(1 + d)^{2n-1} \sqrt{4n-1(2n-1)!}}{\zeta(n)} \frac{3\sigma}{\sigma_{\min(s^*)} m_{\min}^*}.
\]
It follows that
\[
||\eta||_{\infty} \geq \frac{(1 + d)^{2n-1} \sqrt{4n-1(2n-1)!}}{d_{\min}^* \zeta(n)} \frac{3\sigma}{\sigma_{\min(s^*)} m_{\min}^*}.
\]
This contradicts (4.4) and hence proves the claim for \( d_i^* \). Similarly, we can prove the claim for \( d_2^* \). As a result, we have
\[
|d_i^* - d_i| < \frac{d_{\min}^*}{2}, \quad i = 1, 2,
\]
(4.8)
which further implies that \( |d_i^* - d_2| \geq \frac{d_{\min}^*}{2} \). On the other hand, by letting \( n = 2 \) in (4.4), we obtain
\[
|d_i^* - d_1| \cdot |d_i^* - d_2| \geq \frac{(1 + d)^{2n-1} \sqrt{4n-1(2n-1)!}}{\zeta(n)(d_{\min}^*)^{n-1}} \frac{3\sigma}{\sigma_{\min(s^*)} m_{\min}^*}.
\]
It follows that
\[
|d_1^* - d_1| < \frac{2(1 + d)^{2n-1}\sqrt{4n - 1}(2n - 1)!}{\zeta(n)} \left( \frac{1}{d_{\min}^*} \right)^{2n-2} \frac{3\sigma}{\sigma_{\min}(s^*)m_{\min}^*}.
\]

In the same fashion, similar estimate holds for \( |d_i^* - d_i| \). Using Lemma 9.8 in appendix D, we have
\[
|d_i^* - d_i| \leq \frac{7.73\sqrt{4n - 1}(2n - 1)(4 + 4d)^{2n-1}}{4e\pi^2} \left( \frac{1}{d_{\min}^*} \right)^{2n-2} \frac{3\sigma}{\sigma_{\min}(s^*)m_{\min}^*}, \quad i = 1, 2.
\]

**Step 5.** We prove the case \( n \geq 3 \) in this step and the next. We claim that for each \( d_i^* \), there is one \( d_j \) satisfies \( |d_i^* - d_j| < \frac{d_{\min}^*}{2} \). We prove the claim by excluding the following two cases.

**Case 1:** There exists \( j_0 \) such that \( |d_i^* - d_{j_0}| \geq \frac{d_{\min}^*}{2} \) for all \( i = 1, \ldots, n \).

Denote
\[
\tilde{\eta} = \begin{pmatrix}
(d_1^* - d_1) \cdots (d_1^* - d_{j_0-1})(d_1^* - d_{j_0+1}) \cdots (d_1^* - d_n) \\
\vdots \\
(d_n^* - d_1) \cdots (d_n^* - d_{j_0-1})(d_n^* - d_{j_0+1}) \cdots (d_n^* - d_n)
\end{pmatrix}.
\]

By Lemma 3.4 we have
\[
||\tilde{\eta}||_\infty \geq \frac{\zeta(n-1)(d_{\min}^*)^{n-1}}{2}.
\]

Note that \( \eta = \text{diag}(d_1^* - d_{j_0}, \ldots, d_n^* - d_{j_0})\tilde{\eta} \). Using the assumption of this case and (4.7), we have
\[
||\eta||_\infty \geq \frac{d_{\min}^*}{2}||\tilde{\eta}||_\infty \geq \frac{\zeta(n-1)}{2} (d_{\min}^*)^n \geq \frac{\zeta(n-2)}{4} (d_{\min}^*)^n \geq \frac{(1 + d)^{2n-1}}{\zeta(n)(d_{\min}^*)^{n-1}} \frac{\sqrt{4n - 1}(2n - 1)3\sigma}{\sigma_{\min}(s^*)m_{\min}^*},
\]
which contradicts to (4.4).

**Case 2:** There exist \( j_1, j_2 \) and \( i_0 \) such that \( |d_{i_0}^* - d_{j_1}| < \frac{d_{\min}^*}{2} \), \( |d_{i_0}^* - d_{j_2}| < \frac{d_{\min}^*}{2} \).

Then for all \( j \neq i_0 \), we have
\[
|(d_j^* - d_{j_1})(d_j^* - d_{j_2})| \geq \frac{(d_{\min}^*)^2}{4}. \tag{4.9}
\]

Denote
\[
\tilde{\eta} = \begin{pmatrix}
(d_1^* - d_1) \cdots (d_1^* - d_{j_1-1})(d_1^* - d_{j_1})(d_1^* - d_{j_1+1}) \cdots (d_1^* - d_n) \\
\vdots \\
(d_n^* - d_1) \cdots (d_n^* - d_{j_1-1})(d_n^* - d_{j_1})(d_n^* - d_{j_1+1}) \cdots (d_n^* - d_n)
\end{pmatrix}.
\]

By Lemma 3.4 (applied to the points \( d_1^* < \cdots < d_{i_0}^* < d_{i_0+1}^* < \cdots < d_n^* \)), we have
\[
||\tilde{\eta}||_\infty \geq \frac{\zeta(n-2)(d_{\min}^*)^{n-2}}{2}. \tag{4.10}
\]
Note that the components of \( \tilde{\eta} \) differ from those of \( \eta \) only by the factors \((d^n_j - d^n_i)(d^n_j - d^n_k)\) for \( j = 1, \cdots, l_0 - 1, l_0 + 1, \cdots, n \). Using (4.10), (4.9) and (4.7), we have

\[
||\eta||_\infty \geq \frac{\zeta(n-2)}{4} (d^*_n)^n \geq \frac{\zeta(n)(d^*_n)^n}{(1 + d)^{2n-1}} \frac{\sqrt{4n-1}(2n-1)!} {3\sigma} \frac{1}{m^*_\min}.
\]

which contradicts to (4.4) and hence proves the claim.

**Step 6.** By the claim in Step 5, we have \(|d^n_i - d^n_j| < \frac{d^*_n}{2} \), \( i = 1, \cdots, n \). It follows that

\[
|d^n_i - d^n_j| \geq \frac{2|i-j|-1}{2} d^*_n.
\]

Thus, for each \( i \in \{1, \cdots, n\} \), we have

\[
\min_{j=1, \cdots, n} |d^n_i - d^n_j| \geq |d^n_i - d^n_j| \cdot (d^*_n)^n - (2i - 3)! (2(n - i) - 1)! (d^*_n)^{n-1} (n - 2)!.
\]

Using (4.4), we further get

\[
|d^n_i - d^n_j| \cdot (d^*_n)^{n-1} (n - 2)! < \frac{(1 + d)^{n-2} \frac{\sqrt{4n-1}(2n-1)!}{3\sigma}}{\zeta(n)(d^*_n)^{n-1}} \frac{1}{m^*_\min^{\min(s^*)}}.
\]

Hence

\[
|d^n_i - d^n_j| < \frac{2^{n-1} (1 + d)^{2n-1} (2n - 1)! \sqrt{4n - 1}}{\zeta(n)(n-2)!} \frac{1}{d^*_n} \frac{3\sigma}{m^*_\min^{\min(s^*)}}.
\]

Finally, using Lemma 9.8 in appendix D, we have

\[
|d^n_i - d^n_j| < \frac{7.73 \sqrt{4n - 1} (2n - 1)! (4 + 4d)^{2n-1}}{4\pi^2} \frac{1}{d^*_n} \frac{3\sigma}{m^*_\min^{\min(s^*)}}.
\]

This proves the theorem by substituting SRF into the above inequality.

The theorem above states that when the sources separated sufficiently, one can stably recover the source positions from admissible measures. With proper separation condition, the inverse problem is not ill-posed in the sense that any algorithms targeting at the sparsest admissible measure can recover the source positions stably. We next demonstrate that the order \( O(\frac{\sqrt{n}}{m^*}) \) is sharp for stability.

**Proposition 4.1.** For given \( \sigma, n \geq 2 \) and \( m^* > 0 \), choose \( d \) satisfying

\[
\frac{d}{n} = \frac{2}{e^{2n-1}} \sqrt{\frac{2n-1}{e^{4+n}}} \frac{\sqrt{\sigma}}{m^*},
\]

and let \( \tau = \frac{d}{n} \). Then there exist two measures \( \mu^* \) and \( \mu \), supported in \( [-\tau, -2\tau, \cdots, -n\tau] \) and \([\tau, 2\tau, \cdots, n\tau] \), respectively, such that \( ||\mu||_{TV} \leq ||\mu^*||_{TV} = m^* \) and

\[
\sqrt{\bar{h}} \left| \left| \mu * f - \mu^* * f \right| \right|_2 \leq \sigma.
\]
Proof: Let $d_1 = -d = -n\tau, d_2 = -(n-1)\tau, \cdots, d_n = -\tau$, and $d_{n+1} = \tau, \cdots, d_{2n} = n\tau$. We consider the linear system

$$Aa = 0,$$

(4.12)

where

$$A = \begin{pmatrix}
1 & 1 & 1 \\
\frac{d_1}{\sqrt{3}} & \frac{d_2}{\sqrt{3}} & \cdots & \frac{d_n}{\sqrt{3}} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{(d_1)^{2n-2}}{(2n-2)!\sqrt{4n-3}} & \frac{(d_2)^{2n-2}}{(2n-2)!\sqrt{4n-3}} & \cdots & \frac{(d_n)^{2n-2}}{(2n-2)!\sqrt{4n-3}} \\
\end{pmatrix},

a = \begin{pmatrix}
a_1 \\
a_2 \\
\vdots \\
a_{2n}
\end{pmatrix}.
$$

Since $A$ is underdetermined, there exists nonzero $\hat{a} = (\hat{a}_1, \hat{a}_2, \cdots, \hat{a}_{2n})^T$ satisfying (4.12). Moreover, we can show that all $a_j$’s are not zero by using the linear independence of the column vectors in $A$. Without loss of generality, we assume that $\sum_{j=1}^{n} |\hat{a}_j| \geq \sum_{j=n+1}^{2n} |\hat{a}_j|$. By scaling $\hat{a}$, we may assume that $\sum_{j=1}^{n} |\hat{a}_j| = m^*$. We define

$$\mu^* = \sum_{j=1}^{n} \hat{a}_j \delta_{-d_j},\quad \mu = \sum_{j=n+1}^{2n} -\hat{a}_j \delta_{-d_j}.$$

Consider the expansion

$$[\mu^* * f] - [\mu * f] = \sum_{r=0}^{2n-2} (c_r^* - c_r)h_r + \text{Res},$$

(4.13)

where $\text{Res} = \sum_{r=2n-1}^{+\infty} (c_r^* - c_r)h_r$. By (4.12), $\sum_{r=0}^{2n-2} (c_r^* - c_r)h_r = 0$. On the other hand

$$\sqrt{h} \left\Vert \text{Res} \right\Vert_2 = \sqrt{h} \left\Vert \sum_{r=2n-1}^{+\infty} (c_r^* - c_r)h_r \right\Vert_2 \leq \sqrt{h} \sum_{r=2n-1}^{+\infty} \left\Vert (c_r^* - c_r)h_r \right\Vert_2$$

$$= \sqrt{h} \sum_{r=2n-1}^{+\infty} \left\Vert \left( \sum_{j=1}^{n} \hat{a}_j \frac{(d_j)^r}{r!\sqrt{2r+1}} + \sum_{p=n+1}^{2n} \hat{a}_p \frac{(d_p)^r}{r!\sqrt{2r+1}} \right)h_r \right\Vert_2 \\
\leq \sum_{r=2n-1}^{+\infty} \frac{2m^* d^r}{r!\sqrt{2r+1}} \sqrt{h} \left\Vert h_r \right\Vert_2$$

$$< \frac{2m^* d^{2n-1}}{(2n-1)!\sqrt{4n-3}} \left[ \sum_{r=2n-1}^{+\infty} \frac{d^{r-(2n-1)}}{r!(r-(2n-1))!} \right] \sqrt{\pi} \quad \text{by (2.4)}$$

$$= \frac{2e^d \sqrt{\pi} m^* d^{2n-1}}{(2n-1)!\sqrt{4n-3}} \leq \frac{2 \sqrt{\pi} n^{2n-1}(2n-1)}{e(2n-1)!\sqrt{4n-1}} \left( e \right)^{2n-1} \sigma \quad \text{by (4.11)}$$

$$\leq \sigma. \quad \text{by Lemma 9.5}$$

Thus,

$$\sqrt{h} \left\Vert [\mu^* * f] - [\mu * f] \right\Vert_2 \leq \sqrt{h} \left\Vert \sum_{r=0}^{2n-3} (c_r^* - c_r)h_r \right\Vert_2 + \sqrt{h} \left\Vert \text{Res} \right\Vert_2 \leq \sigma.$$

This completes the proof of the proposition.

Proposition 4.1 reveals that the lower bound of the separation distance to ensure a stable recovery of the source positions is at least of the order $O\left( \frac{2n-1}{\sqrt{m^*}} \right)$. This is comparable to the order $O\left( \frac{2n-1}{\sigma_{\min}(S^*)} \alpha_{m_{\min}} \right)$ of the separation condition in Theorem 4.1.
5 A MUSIC-TYPE ALGORITHM

In this section, we propose a preliminary MUSIC-type algorithm to recover the source number. We shall show that the algorithm can recover the source number correctly in the super-resolution regime where the minimum separation distance is comparable to the computational resolution limit.

5.1 RECOVERY OF MULTIPOLe COEFFICIENT

We first consider the recovery of multipole coefficients. For given \( d, \sigma, M \geq m^* \), we define \( s^* \) as in (2.9) and write the measurements as

\[
\begin{pmatrix}
Y(x_1) \\
\vdots \\
Y(x_N)
\end{pmatrix} = \begin{pmatrix}
\begin{array}{c}
f(x_1 - y_1^*) \\
\vdots \\
f(x_N - y_n^*)
\end{array}
\end{pmatrix} \begin{pmatrix}
a_1^* \\
\vdots \\
a_n^* 
\end{pmatrix} + \begin{pmatrix}
W(x_1) \\
\vdots \\
W(x_N)
\end{pmatrix}
\]

\[
= \begin{pmatrix}
h_0 & h_1 & \cdots & h_{s^* - 1}
\end{pmatrix} \begin{pmatrix}
c_0^* \\
c_1^* \\
\vdots \\
c_{s^* - 1}^*
\end{pmatrix} + \begin{pmatrix}
W(x_1) \\
\vdots \\
W(x_N)
\end{pmatrix} + \text{Res},
\]

where \( c_r^* = \sum_{j=1}^{n} a_j^* (d_{j}^r) \frac{r!}{r! \sqrt{2r+1}} \), and \( \text{Res} \) is the residual term of the expansion. We have

\[
\frac{1}{\sqrt{\pi}} \| \text{Res} \|_2 = \frac{1}{\sqrt{\pi}} \left\| \sum_{r=s^*}^{+\infty} \sum_{j=1}^{n} a_j^* (d_{j}^r) \frac{r!}{r! \sqrt{2r+1}} h_r \right\|_2 \leq \sum_{r=s^*}^{+\infty} \sum_{j=1}^{n} \frac{|a_j^*| d^r}{r! \sqrt{2r+1}} \frac{1}{\sqrt{\pi}} \| h_r \|_2
\]

\[
\leq m^* \sum_{r=s^*}^{+\infty} \frac{d^r}{r! \sqrt{2r+1}} \sqrt{\pi} \quad \text{(by (2.4))}
\]

\[
\leq m^* \sqrt{\pi} \frac{d^{s^*}}{s^! \sqrt{2s^* + 1}} \sum_{r=0}^{+\infty} \frac{d^r}{r!} = e^d m^* \sqrt{\pi} d^{s^*} 
\]

\[
\leq \sigma. \quad \text{(by (2.9))}
\]

For the reconstruction, we first find the multipole coefficients by solving the following linear system

\[
\begin{pmatrix}
Y(x_1) \\
\vdots \\
Y(x_N)
\end{pmatrix} = \begin{pmatrix}
h_0 & h_1 & \cdots & h_{s^* - 1}
\end{pmatrix} \begin{pmatrix}
c_0 \\
\vdots \\
c_{s^* - 1}
\end{pmatrix}.
\]

Recall that

\[
H(s^*) = \{ h_0, h_2, \cdots, h_{s^* - 1} \}.
\]

We denote

\[
(c_0^*, c_1^*, \cdots, c_{s^* - 1}^*)^T = \theta^*, \quad (c_0, c_1, \cdots, c_{s^* - 1})^T = \theta.
\]
Then
\[ H(s^*)(\theta - \theta^*) = W + \text{Res} \]
\[ \Rightarrow \sqrt{h}||H(s^*)(\theta - \theta^*)||_2 \leq \sqrt{h}||W||_2 + \sqrt{h}||\text{Res}||_2 \]
\[ \Rightarrow \sqrt{h}||H(s^*)(\theta - \theta^*)||_2 \leq 2\sigma. \quad \text{(by (5.2))} \]

On the other hand
\[ \sqrt{h}||H(s^*)(\theta - \theta^*)||_2 \geq \sigma_{\text{min}}(\sqrt{h}H(s^*))||\theta - \theta^*||_2. \]

Recall that \( \sigma_{\text{min}}(s^*) \) is the minimum singular value of the matrix \( \sqrt{h}H(s^*) \). We have
\[ ||\theta - \theta^*||_2 \leq \frac{2\sigma}{\sigma_{\text{min}}(s^*)}. \tag{5.6} \]

This gives a bound for the matching error of multipole coefficients.

### 5.2 Determine the Exact Source Number

In this section, we determine the source number from the recovered multipole coefficients. Note that for \( n \) sources, we have \( n \) positions and \( n \) amplitudes to recover. But our available data is the \( s^* \) multipole coefficients. So we assume that the source number \( n \leq \frac{s^* - 1}{2} \). Note that in general \( 2n \) multipole coefficients are enough to uniquely determine the \( n \) sources and their amplitudes. We shall only use the first \( s \) multipole coefficients with \( 2n + 1 \leq s \leq s^* \) since the other higher-order multipole coefficients are less reliable. From (5.6), we have the following equations for the first \( s \) multipole coefficients:
\[ c_r = c_r^* + \delta_r, \quad r = 0, \cdots, s - 1, \tag{5.7} \]
where \( \delta_r \) is the perturbation caused by noise. We need to recover the source number \( n \) from (5.7). We rewrite (5.7) as
\[ r!\sqrt{2r + 1}c_r = \sum_{j=1}^{n} a_j^*(d_j^*)^r + r!\sqrt{2r + 1}\delta_r \quad \text{for } r = 0, \cdots, s - 1. \]

For convenience, we suppose \( s \) is odd. We put these coefficients into a data matrix
\[
X = \begin{pmatrix}
    c_0 & \sqrt{3}c_1 & \cdots & (\frac{s-1}{2})!\sqrt{sc_{\frac{s-1}{2}}} \\
    \sqrt{3}c_1 & 2!\sqrt{3}c_2 & \cdots & (\frac{s+1}{2})!\sqrt{s+2c_{\frac{s+1}{2}}} \\
    \vdots & \vdots & \ddots & \vdots \\
    (\frac{s-1}{2})!\sqrt{sc_{\frac{s-1}{2}}} & (\frac{s+1}{2})!\sqrt{s+2c_{\frac{s+1}{2}}} & \cdots & (s-1)!\sqrt{2s-1}c_{s-1}
\end{pmatrix}.
\]

We observe that the data matrix \( X \) has the decomposition that
\[ X = DAD^T + \Delta, \tag{5.8} \]
Theorem 5.1. Let $n \geq 2$ and let $U \Sigma U^*$ be the singular value decomposition of the matrix $DAD^T$. Let $\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_n, 0, \ldots, 0)$, then the following estimate holds

$$\sigma_n \geq \frac{m_{\min}^*(n) \zeta(n)(d_{\min}^*)^{2n-2}}{n(1+d)^{2n-2}},$$

where $\zeta(n)$ is defined in (3.6).

Proof: Note that $\sigma_n$ is the minimum nonzero singular value of the matrix $DAD^T$. Let $S(D^T)$ be the kernel space of $D^T$ and $S^\perp(D^T)$ be its orthogonal complement, we have

$$\sigma_n = \min_{||x||_2=1, x \in S(D^T)} ||DAD^Tx||_2 \geq \sigma_{\min}(DA)\sigma_n(D^T) \geq \sigma_{\min}(D)\sigma_{\min}(A)\sigma_{\min}(D).$$

Since $s \geq 2n + 1$, by Lemma 3.2 and Corollary 3.1 we have

$$\sigma_{\min}(D) \geq \frac{1}{\sqrt{n}} \frac{\zeta(n)(d_{\min}^*)^{n-1}}{(1+d)^{n-1}}.$$

Thus

$$\sigma_n \geq \sigma_{\min}(A)\left(\frac{1}{\sqrt{n}} \frac{\zeta(n)(d_{\min}^*)^{n-1}}{(1+d)^{n-1}}\right)^2.$$

On the other hand, note that $\sigma_{\min}(A) = m_{\min}^*$, we have

$$\sigma_n \geq \frac{m_{\min}^*(n) \zeta(n)(d_{\min}^*)^{2n-2}}{n(1+d)^{2n-2}},$$

Recall that $s \geq 2n + 1$, so $\frac{s}{2} + 1 \geq n + 1$. We denote the singular value decomposition of $X$ as

$$X = \hat{U} \hat{\Sigma} \hat{U}^*,$$

where $\hat{\Sigma} = \text{diag}(\hat{\sigma}_1, \hat{\sigma}_2, \ldots, \hat{\sigma}_n, \hat{\sigma}_{n+1}, \ldots, \hat{\sigma}_{\frac{s}{2}+1})$ with the singular values $\hat{\sigma}_j$ ordered in a decreasing manner.

Note that when there is no noise, $X = DAD^T$. We have the following estimate for the singular values of $DAD^T$.

\begin{equation}
\Delta = \begin{pmatrix}
\delta_0 & \sqrt{3}\delta_1 & \cdots & (\frac{s-1}{2})!\sqrt{s}\delta_{\frac{s}{2}} \\
\sqrt{3}\delta_1 & 2!\sqrt{5}\delta_2 & \cdots & (\frac{s+1}{2})!\sqrt{s+2}\delta_{\frac{s}{2}+1} \\
\vdots & \vdots & \ddots & \vdots \\
(\frac{s-1}{2})!\sqrt{s}\delta_{\frac{s}{2}} & (\frac{s+1}{2})!\sqrt{s+2}\delta_{\frac{s}{2}+1} & \cdots & (s-1)!\sqrt{2s-1}\delta_{s-1}
\end{pmatrix},
\end{equation}

Recall that $s \geq 2n + 1$, so $\frac{s}{2} + 1 \geq n + 1$. We denote the singular value decomposition of $X$ as

$$X = \hat{U} \hat{\Sigma} \hat{U}^*,$$

where $\hat{\Sigma} = \text{diag}(\hat{\sigma}_1, \hat{\sigma}_2, \ldots, \hat{\sigma}_n, \hat{\sigma}_{n+1}, \ldots, \hat{\sigma}_{\frac{s}{2}+1})$ with the singular values $\hat{\sigma}_j$ ordered in a decreasing manner.

Note that when there is no noise, $X = DAD^T$. We have the following estimate for the singular values of $DAD^T$.
which completes the proof of the theorem.

Theorem 5.1 illustrates that the minimum singular value of $DAD^T$ will increase as the minimum separation distance increases. With presence of noise, $X = DAD^T + \Delta$. For a given noise level, by Theorem 5.1, we expect to be able to distinguish singular values from the real sources and from the noise when the sources are separated sufficiently. Precisely, we have following result.

**Corollary 5.1.** Let $n \geq 2$ and let $s^*$ be defined in (2.9). Assume that $2n + 1 \leq s \leq s^*$ and that the minimum separation distance of the measure $\mu^* = \sum_{j=1}^{n} a_j^* \delta y_j^*$ satisfies the following condition

$$\min_{i \neq j} |y_i^* - y_j^*| > (1 + d) \frac{1}{\sqrt{6} \sigma_{\min}(s^*)} \sqrt{\frac{4 \pi n (s-1)! \sqrt{2s-1}}{\zeta(n)^2}}. \quad (5.9)$$

Then the following estimate on the singular values of the data matrix $X$ holds

$$\hat{\sigma}_n > \frac{2\pi (s-1)! \sqrt{2s-1} \sigma}{\sqrt{6} \sigma_{\min}(s^*)}, \quad \hat{\sigma}_j \leq \frac{2\pi (s-1)! \sqrt{2s-1} \sigma}{\sqrt{6} \sigma_{\min}(s^*)}, \quad j = n + 1, \cdots, \frac{s + 1}{2}.$$

Proof: First, by (5.6), we have

$$||\Delta||_2 \leq ||\Delta||_F \leq \sqrt{\left(\sum_{j=1}^{\infty} \frac{1}{j^2}\right)^2 (\frac{2\sigma}{\sigma_{\min}(s^*)})^2}, \quad \text{by} \quad \sum_{j=1}^{\infty} \frac{1}{j^2} = \frac{\pi^2}{6}.$$

By Theorem 5.1 and the separation condition (5.9), we have

$$\sigma_n \geq \frac{m^*_\min \zeta(n)^2 (d^*_\min)^{2n-2}}{n(1+d)^{2n-2}} \geq \frac{4\pi (s-1)! \sqrt{2s-1} \sigma}{\sqrt{6} \sigma_{\min}(s^*)} \geq 2||\Delta||_2.$$

On the other hand, Weyl’s theorem implies that

$$|\hat{\sigma}_n - \sigma_n| \leq ||\Delta||_2.$$

Thus

$$\hat{\sigma}_n > ||\Delta||_2 \geq \frac{2\pi (s-1)! \sqrt{2s-1} \sigma}{\sqrt{6} \sigma_{\min}(s^*)}.$$

In the same fashion, we have that for $j = n + 1, \cdots, \frac{s + 1}{2}$,

$$|\hat{\sigma}_j| \leq ||\Delta||_2 \leq \frac{2\pi (s-1)! \sqrt{2s-1} \sigma}{\sqrt{6} \sigma_{\min}(s^*)}.$$
We note that the minimum separation distance in Corollary 5.1 is comparable to the upper bound we derived in Theorem 2.1. Indeed, for \( s = 2n + 1 \), the required separation distance is

\[
(1 + d)^{-2n-2} \sqrt{\frac{4\pi n(2n)!}{\sqrt{6\zeta(n)^2}}} \leq 4.75(1 + d) \left( \frac{n + 1/2}{\sigma_{\min}(s^*)} \right)^{2n-2} \frac{\sigma}{\sigma_{\min}(s^*)} m_{\min}^*.
\]

In comparison, the upper bound in Theorem 2.1 is \( 4.7(1 + d)^{2n-2} \frac{3}{\sigma_{\min}(s^*)} m_{\min}^* \). Under the separation condition, the lower bound of the first \( n \) singular values \( \frac{2\pi(x-1)(x-3)(2\pi-1)\sigma}{\sqrt{6\sigma_{\min}(s^*)}} \) gives a natural threshold to differentiate the singular values which are generated by the real sources and those by the noise.

In conclusion, our preliminary MUSIC-type algorithm can be summarized as follows. First, we choose a priori estimate of \((d, \sigma, M)\) for the unknown measure, and compute \(s^*\) which is the number of multiple coefficients that can be stably recovered from the given image. Second, we solve a linear system to get the first \( s^* \) multipole coefficients; third, we choose suitable \( s \), rearrange the coefficients to form a data matrix and compute the singular value of the matrix. Finally, we determine the source number by counting the number of singular values that exceed the threshold in Corollary 5.1. Numerical examples of this algorithm will be given in Section 7. We remark that this preliminary algorithm is more of theoretical value than practical one because of the assumption of a priori information such as the number of multiple coefficients to be used in the reconstruction. In a future work, the algorithm will be improved and augmented to be able to detect not only the number of the sources but also their support.

### 6 General Cutoff Frequency Case

In this section, we generalize the theoretical results on the computational resolution limit in the previous sections to the general cutoff frequency case by using a proper scaling argument.

We consider the point spread function \( f_{\Omega}(x) = \frac{1}{\sqrt{2\pi}} \sin(\Omega x) \) with cutoff frequency \( \Omega \). Let \( \mu^* = \sum_{j=1}^{n} a_j^* \delta_{y_j^*} \). We assume the sources \( y_j^* \)'s are located in \([d/\Omega, d/\Omega]\) with \( d = O(1) \). We denote

\[
m_{\min}^* = \min_{j=1,\ldots,n} |a_j^*|, \quad m^* = \|\mu^*\|_{TV} = \sum_{j=1}^{n} |a_j^*|.
\]

The image is measured at evenly spaced sample points \( x_1 = \frac{R}{\Omega}, x_2 = \frac{R}{\Omega} + \frac{h}{\Omega}, \ldots, x_N = \frac{R}{\Omega} \), with spacing \( \frac{h}{\Omega} \leq \frac{\pi}{\Omega} \). The measured data at \( x_t \) is

\[
Y(x_t) = \mu^* \ast f(x_t) + \sqrt{W(x_t)} = \sum_{j=1}^{n} a_j^* f_{\Omega}(x_t - y_j^*) + \sqrt{W(x_t)},
\]

where \( W(x) \) is a band-limited noise with cutoff frequency at \( \Omega \).

We use the same notation for \( Y, [\mu \ast f_{\Omega}] \) and \( W \) as in Section 1, and we define admissible measures similarly with \( d \) being replaced by \( d/\Omega \) and \( h \) by \( h/\Omega \). Similar to (1.2), we assume that

\[
\sqrt{h}\|W\|_2 \leq \sigma.
\]
Note that

\[ Y - [\mu \ast f_\Omega] = \sum_{r=0}^{\infty} \left( \sum_{j=0}^{n} \frac{d_j^r}{r!} (\frac{d_j^r}{r!})^r \right) - \sum_{j=0}^{k} \frac{d_j^r}{r!} (\frac{d_j^r}{r!})^r \right) J_f + \sqrt{\Omega} W \]

where \( d_j^r/\Omega = 0 - y_j^r \), \( d_j^r/\Omega = 0 - y_j \) and \( J_f = \sqrt{2r + 1} (f^{(r)}(x_1), \ldots, f^{(r)}(x_N))^T \). Since \( f^{(r)}(x) = \sqrt{\Omega} \Omega^f f^{(r)}(\Omega x) \), we have

\[ Y - [\mu \ast f_\Omega] = \sqrt{\Omega} \sum_{r=0}^{\infty} \left( \sum_{j=0}^{n} \frac{d_j^r}{r!} (\frac{d_j^r}{r!})^r \right) - \sum_{j=0}^{k} \frac{d_j^r}{r!} (\frac{d_j^r}{r!})^r \right) h_r + \sqrt{\Omega} W \]

where

\[ h_r = \sqrt{2r + 1} (f^{(r)}(\Omega x_1), \ldots, f^{(r)}(\Omega x_N))^T. \]

Denote \( H(s) = (h_0, \ldots, h_{r-1}) \) and \( \sigma_{\min}(s) = \sigma_{\min}(\sqrt{\Omega} H(s)) \). Note \( H(s) \) is the same as the one defined in Section 2 and so is \( \sigma_{\min}(s) \). As a consequence, the condition for admissible measure

\[ \sqrt{\frac{h}{\Omega}} |[\mu \ast f_\Omega] - Y| \leq \sigma, \]

is equivalent to

\[ \sqrt{h} \left| \sum_{r=0}^{\infty} \left( \sum_{j=0}^{n} \frac{d_j^r}{r!} (\frac{d_j^r}{r!})^r \right) - \sum_{j=0}^{k} \frac{d_j^r}{r!} (\frac{d_j^r}{r!})^r \right) h_r + W \right| \leq \sigma, \]

which is of the same form as the case when \( \Omega = 1 \). As a result, the previous results can be applied readily to our case.

We first show that to be able to recover the number of sources, the minimum separation distance is required to be at least of the order \( O(\frac{1}{\Omega \cdot \sqrt{\Omega} \cdot m^*}) \).

**Proposition 6.1.** For given \( \sigma > 0, n \geq 2, m^* > 0 \), choose \( d \) satisfying

\[ \frac{d}{n-1} = 2^{n-2} e^{(2n-2) \sigma \sqrt{\frac{\sigma}{e^d}}} \]

Then there exist two measures \( \mu^* = \sum_{j=0}^{n} \frac{d_j^r}{r!} (\frac{d_j^r}{r!})^r \) with \( n \) supports and another \( \mu \) with \( n-1 \) supports, both supported in \( [-d/\Omega, d/\Omega] \), such that \( \|\mu\|_{TV} \leq \|\mu^*\|_{TV} = m^* \) and

\[ \min_{\mu \neq \mu^*} |y^r_i - y^r_j| = \frac{d}{(n-1)\Omega}. \]

Moreover,

\[ \sqrt{\frac{h}{\Omega}} \left| [\mu \ast f] - [\mu^* \ast f] \right| \leq \sigma. \]

Next, for given \( \sigma, M, d \), we define \( s^* \) as (2.9). We can derive the following upper bound for the computational resolution limit.
Theorem 6.1. Let \( n \geq 2 \) and let \( \mu^* = \sum_{j=1}^{n} a_j^* \delta_{y_j^*} \) be a measure supported in \([-\frac{d}{\Omega}, \frac{d}{\Omega}]\). Assume that it satisfies the separation condition that

\[
\min_{i \neq j} |y_i^* - y_j^*| \geq \frac{4.7(1 + d)}{\Omega} \sqrt{\frac{3}{\sigma_{\min}(s^*)} \frac{\sigma}{m^*}},
\]

then any discrete measure \( \mu \) with \( k < n \) supports cannot be a \((d, \sigma, M)\)-admissible measure.

Now, let \( d^*_\min = \min_{i \neq j} |y_i^* - y_j^*| \), we define the super-resolution factor

\[
\text{SRF} = \frac{\pi}{\Omega d^*_\min}.
\]

We can prove the following stability result for the support recovery.

Theorem 6.2. Let \( n \geq 2 \) and let \( \mu^* = \sum_{j=1}^{n} a_j^* \delta_{y_j^*} \) with \(-\frac{d}{\Omega} \leq y_1^* < y_2^* < \cdots < y_n^* \leq \frac{d}{\Omega}\) is a \((d, \sigma, M)\)-admissible measure, then we have

\[
|y_i - y_i^*| < \frac{d^*_\min}{2}, \quad i = 1, \cdots, n.
\]

Moreover,

\[
|y_i - y_i^*| \leq C(n,d) \frac{\text{SRF}^{2n-2} \sigma}{\sigma_{\min}(s^*) m^*},
\]

where

\[
C = \frac{7.73 \sqrt{4n-1}(2n-1)(4 + d)^{2n-1}}{4e\pi^{2n-1}}.
\]

The sharpness of the above stability estimate can be seen from the proposition below.

Proposition 6.2. For given \( \sigma, n \geq 2, m^* > 0 \), choose \( d \) satisfying

\[
\frac{d}{n} = \frac{2}{e} \sqrt{\frac{2n-1}{e^{1+d}}} \sqrt{\frac{\sigma}{m^*}}.
\]

Let \( \tau = \frac{d}{n} \). Then there exist two measures \( \mu^* \) and \( \mu \), supported in \((-\tau, -2\tau, \cdots, -n\tau)\) and \((\tau, 2\tau, \cdots, n\tau)\), respectively, such that \( \|\mu\|_{TV} \leq \|\mu^*\|_{TV} = m^* \) and

\[
\sqrt{\frac{h}{\Omega}} \left\| (\mu * f) - (\mu^* * f) \right\|_2 \leq \sigma.
\]
7 NUMERICAL EXPERIMENTS

We perform numerical experiments for our MUSIC-type algorithm in this section.

**Experiment 1:** We consider the recovery of the discrete measure

\[ \mu^* = \delta_{y_1^*} + \delta_{y_2^*} \]

where \( y_1^* = 0.37, \ y_2^* = -0.37 \). Then the source number \( n = 2, \ m^* = 2, \ m_{\min}^* = 1 \). We set the noise level \( \sigma = 7.1 \times 10^{-6}, \ d = 0.5 \) and \( M = 3 \). We let \( R = 100, \ h = 2 \) and sample the image evenly in \([-100, 100]\) with 101 samples points as follows:

\[ Y(x_t) = \mu^* f(x_t) + W(x_t), \quad t = 1, \cdots, 101. \]

where \( W(x_t) \) are uniformly distributed random numbers in \( (0, \frac{\sigma}{\sqrt{2R}}) \).

The measurements are shown in Figure 7.1a. It’s impossible to determine visually that the source number is two. Note that the number of multipole coefficients that we can stably recover is \( s^* = \min \{ l \in \mathbb{N} : \frac{d^l}{l! \sqrt{2l + 1}} \leq \frac{\sigma}{2M \sigma_{\min}} \} = 7. \) We consider the multipole matrix \( \hat{H}(7) = (h_0 \cdots h_6) \) and solve the linear equation

\[ \hat{H}(7) \hat{\theta} = \hat{Y}. \]

We get \( \theta = (2.0, -1.5912 \times 10^{-6}, 0.2738, -2.3621 \times 10^{-5}, 0.0379, -1.9360 \times 10^{-4}, 0.0129)^T \). We use the first 5 multipole coefficients to recover the source number. Consider singular value decomposition of the data matrix

\[ X = \begin{pmatrix} \theta(1) & \sqrt{3}\theta(2) & 2!\sqrt{5}\theta(3) \\ \sqrt{3}\theta(2) & 2!\sqrt{5}\theta(3) & 3!\sqrt{7}\theta(4) \\ 2!\sqrt{5}\theta(3) & 3!\sqrt{7}\theta(4) & 4!\sqrt{9}\theta(5) \end{pmatrix} = \hat{U} \hat{\Sigma} \hat{U}^*. \]

We have \( \hat{\Sigma} = \text{diag}(2.0375, 0.2738, 4.2763 \times 10^{-4}) \) and the threshold derived in Corollary 5.1 is \( \frac{2\pi(s^*-1)!\sqrt{2s^*-1}\sigma}{\sqrt{6\sigma_{\min}(s^*)}} = 0.0227 \). Thus, the first two singular values exceed this threshold and we are
able to recover the source number two exactly by the MUSIC-type algorithm. On the other hand, we note that the minimum separation distance required in Corollary 5.1 is

\[(1 + d)^{2n-2} \sqrt{\frac{4\pi n(s-1)!}{\sqrt{6\zeta(n)^2}}} \sqrt[2n-2]{\frac{\sigma}{\sigma_{\text{min}}(s^*) m_{\text{min}}^*}} = 0.4519,
\]

which is smaller than the actual separation distance \(|y_1^* - y_2^*| = 2 \times 0.37 = 0.74\).

We now investigate the minimum separation distance required in this example beyond which one can figure out the source number by this MUSIC-type algorithm. For the purpose, we draw a graph about the relation between \(\hat{\sigma}^2\) and the separation distance between \(y_1^*\) and \(y_2^*, |y_1^* - y_2^*|\). For simplicity, we only consider \(y_1^* = -y_2^* = |y_1^* - y_2^*|\). The relation is shown by Figure 7.1:b. The threshold derived in Corollary 5.1 is

\[\frac{2\pi (s-1)!}{\sqrt{6\sigma_{\text{min}}(s^*)}} = 0.0227.
\]

As shown by Figure 7.1:b, when sources are separated beyond 0.2, we are able to determine the source number by the MUSIC-type algorithm. To demonstrate the efficiency of the algorithm, we calculate the upper bound for the computational resolution limit in Theorem 2.1, which is equal to

\[4.7(1 + d)^{2n-2} \frac{3}{\sqrt{\sigma_{\text{min}}(s^*)}} m_{\text{min}}^* = 0.1353.
\]

In comparison, the minimum separation distance required for our MUSIC-type algorithm in the above example is 0.2.

**Experiment 2:** We give an example of 3 sources. We consider the recovery of the measure

\[\mu^* = \delta y_1^* + \delta y_2^* + \delta y_3^*\]

where \(y_1^* = 0.4, y_2^* = 0, y_3^* = -0.4\). Then the source number \(n = 3, m^* = 3, m_{\text{min}}^* = 1\). We set \(d = 0.5, \sigma = 1.38 \times 10^{-9}\) and \(M = 4\). We sample the image evenly in \([-100, 100]\) with 101 samples (with \(R = 100, h = 2\)) as follows:

\[Y(x_t) = \mu^* * f(x_t) + W(x_t), \quad t = 1, \ldots, 101,
\]

where \(W(x_t)\)'s are uniformly distributed random numbers in \((0, \sigma/\sqrt{2}\pi)\).

The measurements are shown in Figure 7.2:a. It is impossible to discern visually that the source number is three. Note that the number of multipole coefficients that we can stably recover is \(s^* = \min \{l \in \mathbb{N} : \frac{d^l}{l!\sqrt{2l+1}} \leq \frac{\sigma}{2M\sigma^l}\sqrt{s} \} = 10\). We consider the multipole matrix \(H(10) = (h_0 \cdots h_9)\) and solve the following linear equations

\[H(10)\theta = Y.
\]

We have \(\theta = (3.0, -1.6171 \times 10^{-10}, 0.32, -1.8209 \times 10^{-9}, 0.0512, 1.8293 \times 10^{-7}, 0.0082, 2.1694 \times 10^{-5}, 0.0016, 0.001)^T\). We use the first 7 multipole coefficients to recover the source number.
Figure 7.2: a: The noisy image. b: behavior of the third singular value

Consider the singular value decomposition of the data matrix

\[
X = \begin{pmatrix}
\theta(1) & \sqrt{3}\theta(2) & 2\sqrt{5}\theta(3) & 3\sqrt{7}\theta(4) \\
\sqrt{3}\theta(2) & 2\sqrt{5}\theta(3) & 3\sqrt{7}\theta(4) & 4\sqrt{9}\theta(5) \\
2\sqrt{5}\theta(3) & 3\sqrt{7}\theta(4) & 4\sqrt{9}\theta(5) & 5\sqrt{11}\theta(6) \\
3\sqrt{7}\theta(4) & 4\sqrt{9}\theta(5) & 5\sqrt{11}\theta(6) & 6\sqrt{13}\theta(7)
\end{pmatrix} = \hat{U}\hat{\Sigma}\hat{U}^*.
\]

We have \(\hat{\Sigma} = \text{diag}(3.0343, 0.3282, 0.0169, 1.2608 \times 10^{-5})\) and the threshold in Corollary 5.1 is 0.0017. Thus we can determine exactly the source number 3 by the MUSIC-type algorithm. We next investigate the minimum separation distance required to determine the exact source number. Figure 7.2b illustrates the relation between \(\hat{\sigma}_3\) and the minimum separation distance of \(y_1^*, y_2^*, y_3^*\). For simplicity, we consider \(y_1^* = -y_3^*\) and \(y_2^* = 0\). The threshold in Corollary 5.1 is

\[
\frac{2\pi(s-1)!\sqrt{2s-1}\sigma}{\sqrt{6}\sigma_{\text{min}}(s^*)} = 0.0017.
\]

It is shown in Figure 7.2b that, when the sources separated beyond 0.23, we are able determine the source number by the MUSIC-type algorithm. To show the efficiency of our MUSIC-type algorithm, we calculate the upper bound for the computational resolution limit, which is equal to

\[
4.7(1 + d)^{2n-2} \frac{3\sigma}{\sigma_{\text{min}}(s^*) m_{\text{min}}^*} = 0.2083.
\]

It is comparable to the separation distance required for our MUSIC-type algorithm.

8 Conclusions and Future Work

In this paper, we introduced a new concept computational resolution limit to the recovery of source number in the deconvolution problem and derived sharp upper bound for the
limit. The bound quantitatively demonstrates the dependence on the SNR and sparsity of the sources. Our results reveal that one can achieve super-resolution when the SNR is sufficiently large such that the computational resolution limit is below the Rayleigh limit. It also demonstrates the challenges when the sources are not sparse. We further derived a sharp stability result for recovering source positions when the sources are separated beyond the computational resolution limit. We finally proposed a preliminary MUSIC-type algorithm, based on the multipole expansion method we introduced, to find the source number. Its promising performance in the super-resolution regime when the sources are separated below the Rayleigh limit is analyzed both theoretically and numerically.

The work in this paper opens many research avenues. First, the MUSIC-type algorithm may be improved to use less a priori information and be augmented to be able to determine not only the source number but also the source supports. Second, one may consider the case of multiple clusters of sources. By using multipole expansion for each of the clusters, one may reduce the global deconvolution problem to multiple parallel local problems in which one can have an efficient local solver. Third, the methods developed in the paper can be readily applied to the spectral estimation problem. Fourth, in practice, multiple images of the same distribution of sources under different illuminations are used to enhance resolution in many super-resolution techniques, see for instance: SIM [20], STED [21] and STROM [40]. It is expected that the computational resolution limit in such cases will improve the one for a single image. Finally, the extension of the results to higher-dimensional space is also possible. All these works will be reported in forthcoming papers.

9 APPENDICES

9.1 APPENDIX A: SOME ESTIMATES OF VANDERMONDE MATRICES

In this section, we give some estimates for the Vandermonde matrices that used in this paper.

**Lemma 9.1.** Let \( d_i \neq d_j \) for \( i \neq j, i, j = 1, \cdots, n \). For the Vandermonde matrix

\[
V = \begin{pmatrix}
1 & \cdots & 1 \\
1 & \cdots & d_n \\
\vdots & \cdots & \vdots \\
d_1^{n-1} & \cdots & d_n^{n-1}
\end{pmatrix},
\]

we have the following operator norm estimate of its inverse,

\[
||V^{-1}||_{\infty} \leq \max_{1 \leq s \leq n} \prod_{l \leq p s \neq l} \frac{1 + |d_p|}{|d_l - d_p|}.
\]

Proof: see [19].

**Corollary 9.1.** Let \( d_{\min} = \min_{i \neq j} |d_i - d_j| \) and assume that \( \max_{i=1, \cdots, n} |d_i| \leq d \). Then we have the operator norm estimate of inverse of the Vandermonde matrix \( V \) in (9.1)

\[
||V^{-1}||_{\infty} \leq \frac{(1 + d)^{n-1}}{\zeta(n)(d_{\min})^{n-1}}.
\]
where $\zeta(n)$ is defined in (3.6).

Proof: WLOG, we may arrange $d_i$ such $d_1 < d_2 < \cdots < d_n$. Then, $|d_i - d_p| \geq |i - p|d_{\min}$. By Lemma 9.1 we have

$$||V^{-1}||_{\infty} \leq \max_{1 \leq i \leq n} \Pi_{1 \leq i \neq p \leq n} \frac{1 + |d_p|}{|d_i - d_p|}.$$  

It follows that

$$\max_{1 \leq i \leq n} \Pi_{1 \leq i \neq p \leq n} \frac{1 + |d_p|}{|d_i - d_p|} \leq \max_{1 \leq i \leq n} \Pi_{1 \leq i \neq p \leq n} \frac{1 + d}{|i - p|d_{\min}} \leq \frac{(1 + d)^{n-1}}{\zeta(n)(d_{\min})^{n-1}}.$$  

Lemma 9.2. For the Vandermonde matrices

$$V = \begin{pmatrix} 1 & \cdots & 1 \\ d_1 & \cdots & d_n \\ \vdots & \vdots & \vdots \\ d_1^{n-1} & \cdots & d_n^{n-1} \end{pmatrix}, \quad W = \begin{pmatrix} 1 & \cdots & 1 \\ d_1 & \cdots & d_S \\ \vdots & \vdots & \vdots \\ d_1^S & \cdots & d_S^S \end{pmatrix},$$

assume that $S > n - 1$. Then the following estimate on their singular values holds:

$$\frac{1}{\sqrt{n}} \min_{1 \leq i \leq n} \Pi_{1 \leq i \neq p \leq n} \frac{|d_i - d_p|}{1 + |d_p|} \leq \frac{1}{||V^{-1}||_2} \leq \sigma_{\min}(V) \leq \sigma_{\min}(W).$$

Proof: Because $V$ is invertible, by Lemma 9.1 we have

$$||V^{-1}||_{\infty} \leq \max_{1 \leq i \leq n} \Pi_{1 \leq i \neq p \leq n} \frac{1 + |d_p|}{|d_i - d_p|}.$$  

Then

$$||V^{-1}||_2 \leq \sqrt{n}||V^{-1}||_{\infty} \leq \sqrt{n} \max_{1 \leq i \leq n} \Pi_{1 \leq i \neq p \leq n} \frac{1 + |d_p|}{|d_i - d_p|}.$$  

It follows that

$$\sigma_{\min}(V) \geq \frac{1}{||V^{-1}||_2} \geq \frac{1}{\sqrt{n}} \min_{1 \leq i \leq n} \Pi_{1 \leq i \neq p \leq n} \frac{|d_i - d_p|}{1 + |d_p|}.$$  

Therefore, we have

$$\frac{1}{\sqrt{n}} \min_{1 \leq i \leq n} \Pi_{1 \leq i \neq p \leq n} \frac{|d_i - d_p|}{1 + |d_p|} \leq \sigma_{\min}(V) \leq \sigma_{\min}(W),$$

where the last inequality follows from the observation that

$$\min_{x \in \mathbb{R}^n, ||x||_2 = 1} ||Vx||_2 \leq \min_{x \in \mathbb{R}^n, ||x||_2 = 1} ||Wx||_2.$$
9.2 APPENDIX B: PROOF OF LEMMA 3.4

We prove Lemma 3.4 in this section.

Proof: **Step 1.** Define

\[ \eta_j(d_1, \cdots, d_k) = \Pi_{q=1}^k (d_j^* - d_q), \quad ||\eta_j(d_1, \cdots, d_k)||_\infty = \max_{j=1, \cdots, k+1} ||\eta_j(d_1, \cdots, d_k)||. \]

We only need to prove that

\[ \min_{d_1, \cdots, d_k} ||\eta_j(d_1, \cdots, d_k)||_\infty \geq \xi(k)(d_{\min}^*)^k. \]  

(9.2)

It is easy to verify the result for \( k = 1 \). For \( k \geq 2 \), we argue as follows. It is clear that the minimizer to (9.2) exists (may not be unique) and \( \min_{d_1, \cdots, d_k} ||\eta||_\infty > 0 \). Let \((\hat{d}_1, \cdots, \hat{d}_k)\) be such a minimizer with \( \hat{d}_1 \leq \hat{d}_2 \cdots \leq \hat{d}_k \).

**Step 2.** We prove that \( \hat{d}_1, \cdots, \hat{d}_k \in [d_1^*, d_k^*] \). By contradiction, if \( \hat{d}_p < d_1^* \) for some \( p \), then for those \( j \)'s such that \( \eta_j(\hat{d}_1, \cdots, \hat{d}_k) \neq 0 \), we have

\[ ||(d_j^* - \hat{d}_1) \cdots (d_j^* - \hat{d}_{p-1})(d_j^* - \hat{d}_p)(d_j^* - \hat{d}_{p+1}) \cdots (d_j^* - \hat{d}_k)|| > ||(d_j^* - \hat{d}_1) \cdots (d_j^* - \hat{d}_{p-1})(d_j^* - \hat{d}_p^*)(d_j^* - \hat{d}_{p+1}) \cdots (d_j^* - \hat{d}_k)||. \]

i.e.,

\[ ||\eta_j(\hat{d}_1, \cdots, \hat{d}_k)|| > ||\eta_j(\hat{d}_1, \cdots, \hat{d}_{p-1}, d_p^*, \hat{d}_{p+1}, \cdots, \hat{d}_k)||. \]

While for those \( j \)'s such that \( \eta_j(\hat{d}_1, \cdots, \hat{d}_k) = 0 \), we have \( \eta_j(\hat{d}_1, \cdots, \hat{d}_{p-1}, d_p^*, \hat{d}_{p+1}, \hat{d}_k) = 0 \). This contradicts to the assumption that \((\hat{d}_1, \cdots, \hat{d}_k)\) is a minimizer and hence proves that \( \hat{d}_p \geq d_1^* \) for all \( p = 1, \cdots, k \). Similarly, we can prove that \( \hat{d}_p \leq d_{k+1}^* \) for all \( p = 1, \cdots, k \).

**Step 3.** We claim that each interval \([d_p^*, d_{p+1}^*] \) contains only one \( \hat{d}_q \) for some \( 1 \leq q \leq k \). We prove the claim by excluding the following three cases.

**Case 1:** There exist \( j_0, p \) such that \( d_p^* < \hat{d}_{j_0} < \hat{d}_{j_0+1} < d_{p+1}^* \).

Let \( j_1 \) be the integer such that

\[ ||\eta_j(d_1, \cdots, d_k)||_\infty = ||\eta_{j_1}(d_1, \cdots, d_k)|| > 0. \]  

(9.3)

If \( 1 \leq j_1 \leq p \), then for \( \Delta > 0 \) sufficiently small, we have

\[ ||\eta_{j_1}(\hat{d}_1, \cdots, \hat{d}_{j_0} - \Delta, \hat{d}_{j_0+1} + \Delta, \hat{d}_{j_0+2}, \cdots, \hat{d}_k)|| - ||\eta_{j_1}(\hat{d}_1, \cdots, \hat{d}_k)||
\]

\[ = \left| (d_{j_1}^* - \hat{d}_{j_0} + \Delta)(d_{j_1}^* - \hat{d}_{j_0}) - (d_{j_1}^* - \hat{d}_{j_0})(d_{j_1}^* - \hat{d}_{j_0+1}) \right| ||\Pi_{q=1, q \neq j_0, j_0+1}^k (d_{j_1}^* - \hat{d}_q)||
\]

\[ = \left| \Delta(d_{j_0} - \hat{d}_{j_0+1}) - \Delta^2 \right| ||\Pi_{q=1, q \neq j_0, j_0+1}^k (d_{j_1}^* - \hat{d}_q)||
\]

\[ < 0. \]  

(by \( \hat{d}_{j_0} < \hat{d}_{j_0+1} \) and [9.3]

On the other hand, if \( p + 1 \leq j_1 \leq k + 1 \), then in the same fashion, for \( \Delta \) sufficiently small, we have

\[ ||\eta_{j_1}(\hat{d}_1, \cdots, \hat{d}_{j_0} - \Delta, \hat{d}_{j_0+1} + \Delta, \hat{d}_{j_0+2}, \cdots, \hat{d}_k)|| - ||\eta_{j_1}(\hat{d}_1, \cdots, \hat{d}_k)|| < 0. \]
Thus in all cases we have
\[ ||\eta(\hat{d}_1, \ldots, \hat{d}_{j_0} - \Delta, \hat{d}_{j_0+1} + \Delta, \ldots, \hat{d}_k)||_{\infty} < ||\eta(\hat{d}_1, \ldots, \hat{d}_k)||_{\infty}. \]
This contradicts the assumption that \((\hat{d}_1, \ldots, \hat{d}_k)\) is a minimizer.

**Case 2:** There exist \(j_0, p\) such that \(\hat{d}_{j_0} = d_p^*, \ d_p^* < \hat{d}_{j_0+1} < d_{j_0+1}^*\). We still denote \(j_1\) the integer in \(\{1, \ldots, k + 1\}\) satisfying \((9.3)\). Since \(\eta_p = 0, j_1 \neq p\). Let \(\Delta > 0\) be sufficiently small. Similar to \((9.4)\), we can show that in both cases \(1 \leq j_1 \leq p - 1\) and \(p + 1 \leq j_1 \leq k + 1\),
\[ ||\eta_{j_1}(\hat{d}_1, \ldots, \hat{d}_{j_0} - \Delta, \hat{d}_{j_0+1} + \Delta, \hat{d}_{j_0+2}, \ldots, \hat{d}_k)||_{\infty} < ||\eta_{j_1}(\hat{d}_1, \ldots, \hat{d}_k)||_{\infty}. \]
Thus,
\[ ||\eta(\hat{d}_1, \ldots, \hat{d}_{j_0} - \Delta, \hat{d}_{j_0+1} + \Delta, \ldots, \hat{d}_k)||_{\infty} < ||\eta(\hat{d}_1, \ldots, \hat{d}_k)||_{\infty}. \]
This contradicts the assumption that \((\hat{d}_1, \ldots, \hat{d}_k)\) is a minimizer.

**Case 3:** There exist \(j_0, p\) such that \(\hat{d}_{j_0} = d_{j_0}^* = d_{j_0+1}^*\). Denote \(j_1\) the integer in \(\{1, \ldots, k + 1\}\) satisfying \((9.3)\). Since \(\eta_p = 0, j_1 \neq p\). Let \(\Delta > 0\) be sufficiently small, we have for \(1 \leq j_1 \leq p - 1\),
\[ ||\eta_{j_1}(\hat{d}_1, \ldots, \hat{d}_{j_0} - \Delta, \hat{d}_{j_0+1} + \Delta, \hat{d}_{j_0+2}, \ldots, \hat{d}_k)||_{\infty} < ||\eta_{j_1}(\hat{d}_1, \ldots, \hat{d}_k)||_{\infty}. \]
\[ ||\eta_{j_1}(\hat{d}_1, \ldots, \hat{d}_{j_0} - \Delta, \hat{d}_{j_0+1} + \Delta, \hat{d}_{j_0+2}, \ldots, \hat{d}_k)||_{\infty} < ||\eta_{j_1}(\hat{d}_1, \ldots, \hat{d}_k)||_{\infty}. \]
This contradicts the assumption that \((\hat{d}_1, \ldots, \hat{d}_k)\) is a minimizer.

Finally, combining the results in the above three cases proves the claim.

**Step 4.** By the result in Step 3, we have for \(j = 1, \ldots, k\),
\[ d_j^* \leq \hat{d}_j < d_{j+1}^*. \] (9.6)

We then prove the lemma by considering the following two cases.

**Case 1:** For all \(1 \leq j \leq k\), \(\hat{d}_j - d_j^* < \frac{d_{\min}}{2}\).
Clearly \(|d_{k+1}^* - \hat{d}_k| > \frac{d_{\min}}{2}\). Thus
\[ |\prod_{q=1}^{k} (d_{k+1}^* - \hat{d}_q)| = |\prod_{q=1}^{k-1} (d_{k+1}^* - \hat{d}_q)||d_{k+1}^* - \hat{d}_k| \]
\[ \geq |\prod_{q=1}^{k-1} (d_{k+1}^* - \hat{d}_q)||d_{k+1}^* - \hat{d}_k| \quad \text{(by (9.6))} \]
\[ \geq (k - 1)(d_{\min}^{k-1}) \frac{d_{\min}^*}{2} \geq \xi(k)(d_{\min}^*)^k, \]
which implies that $|\eta|_\infty \geq \xi(k)(d_{\min}^*)^k$. 

**Case 2:** There exist some $j$ such that $\hat{d}_j - d_j^* \geq \frac{d_{\min}^*}{2}$.

We let $j_0$ be the smallest integer such that $\hat{d}_{j_0} \geq d_j^* + \frac{d_{\min}^*}{2}$. Then

$$ d_j^* - \hat{d}_{j_0 - 1} \geq \frac{d_{\min}^*}{2}, \quad |d_j^* - \hat{d}_{j_0}| \geq \frac{d_{\min}^*}{2}. \quad (9.7) $$

It follows that

$$ |\Pi_{q=1}^k (d_{j_0}^* - \hat{d}_q)| = |\Pi_{q=1}^k (d_{j_0}^* - \hat{d}_q)| |d_j^* - \hat{d}_{j_0 - 1}| |d_j^* - \hat{d}_{j_0}| |\Pi_{q=j_0+1}^k (d_{j_0}^* - \hat{d}_q)| $$

$$ \geq |\Pi_{q=1}^k (d_{j_0}^* - d_{j_0}^*)| |\Pi_{q=j_0+1}^k (d_{j_0}^* - d_{j_0}^*)| |d_j^* - \hat{d}_{j_0 - 1}| |d_j^* - \hat{d}_{j_0}| \quad \text{(by (9.6))} $$

$$ \geq (j_0 - 2)! |(d_{\min}^*)^{j_0 - 2} (k - j_0)! |d_{\min}^*|^{j_0 - 2} |(d_{\min}^*)^{j_0 - 1}| |d_{\min}^*| $$

$$ \geq (j_0 - 2)! |(k - j_0)! |d_{\min}^*|^{k - 2} \left( \frac{d_{\min}^*}{2} \right)^2 \quad \text{(by (9.7))} $$

$$ = \frac{(j_0 - 2)! (k - j_0)!}{4} (d_{\min}^*)^k. $$

Minimizing $\frac{(j_0 - 2)! (k - j_0)!}{4} (d_{\min}^*)^k$ over $j_0 = 1, \ldots, k$ gives

$$ \min_{j_0=1, \ldots, k} \frac{(j_0 - 2)! (k - j_0)!}{4} (d_{\min}^*)^k \geq \left\{ \begin{array}{ll} \frac{(k - 1)! k!}{(k - 2)!} (d_{\min}^*)^k, & \text{k is odd,} \\ \frac{(k - 2)! 2}{4} (d_{\min}^*)^k, & \text{k is even.} \end{array} \right. \quad (9.8) $$

It follows that $||\eta||_\infty \geq \xi(k)(d_{\min}^*)^k$.

### 9.3 Appendix C: Proof of Lemma 3.1

We prove Lemma 3.1 in this section. Denote

$$ S_{\tau}^l = \{ (\tau_1, \ldots, \tau_j) : \tau_p \in \{l, l + 1, \ldots, n\}, p = 1, \ldots, j, \tau_p \neq \tau_q \text{ for } p \neq q \}. $$

and

$$ D_{\tau}^l = \{ (\tau_1, \ldots, \tau_j) : \tau_p \in \{l, l + 1, \ldots, n\}, p = 1, \ldots, j \}. $$

Let $M_t(d_1, \ldots, d_j)$ be the sum of all monomials of degree $t$ $(t \geq 0)$ in $d_1, \ldots, d_j$, more precisely,

$$ M_t(d_1, \ldots, d_j) = \sum_{(\tau_1, \ldots, \tau_j) \in D_{\tau}^l, \sum_{p=1}^j \tau_p = t} (d_1)^{\tau_1} \cdots (d_j)^{\tau_j}, \quad 0 \leq t \leq n. \quad (9.9) $$

We first note that by a result in [35], the Vandermonde matrix $V_n(n)$ defined in [3.2] can be reduced to the following form by applying a sequence of elementary column-addition matrices $G(1), G(2), \ldots, G(n - 1)$,

$$ W = \begin{pmatrix}
  w_{11} & 0 & 0 & \cdots & 0 \\
  w_{21} & w_{22} & 0 & \cdots & 0 \\
  w_{31} & w_{32} & w_{33} & \cdots & 0 \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  w_{n1} & w_{n2} & w_{n3} & \cdots & w_{nn} \\
  w_{(n+1)1} & w_{(n+1)2} & w_{(n+1)3} & \cdots & w_{(n+1)n}
\end{pmatrix}, \quad (9.10) $$

43
where
\[ w_{ij} = M_{i-j}(d_1, \cdots, d_j) \Pi_{p=1}^{j-1}(d_j - d_p), \quad i \geq j. \]  

(9.11)

After extracting the common factors, we get
\[
V(0) = \begin{pmatrix}
1 & 0 & 0 & \cdots & 0 \\
v_{21}(0) & 1 & 0 & \cdots & 0 \\
v_{31}(0) & v_{32}(0) & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
v_{n1}(0) & v_{n2}(0) & v_{n3}(0) & \cdots & 1 \\
v_{(n+1)1}(0) & v_{(n+1)2}(0) & v_{(n+1)3}(0) & \cdots & v_{(n+1)n}(0)
\end{pmatrix} = WD,
\]

where \( D = \text{diag}(1, \frac{1}{d_2-d_1}, \cdots, \frac{1}{\Pi_{p=1}^{n-1}(d_n-d_p)}) \). We note that \( v_{ij}(0) = M_{i-j}(d_1, \cdots, d_j) \) for \( j \leq i \). Especially \( v_{(n+1)n}(0) = \sum_{p=1}^{n} d_p \).

We now perform further Gaussian eliminations to the matrix \( V(0) \), using only elementary column-addition matrices.

**Step 1:** Eliminate \( v_{n1}(0), \cdots, v_{n(n-1)}(0) \) by elementary column operations to get
\[
V(1) = \begin{pmatrix}
1 & 0 & 0 & \cdots & 0 \\
v_{21}(1) & 1 & 0 & \cdots & 0 \\
v_{31}(1) & v_{32}(1) & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
v_{(n-1)1}(1) & v_{(n-1)2}(1) & v_{(n-1)3}(1) & \cdots & 1 \\
v_{(n+1)1}(1) & v_{(n+1)2}(1) & v_{(n+1)3}(1) & \cdots & v_{(n+1)(n-1)}(1) & v_{(n+1)n}(1)
\end{pmatrix} = V(0)Q(1).
\]

Precisely, we have
\[
v_{ij}(1) = \begin{cases} 
M_{i-j}(d_1, \cdots, d_j), & j < i \leq n - 1, \\
0, & j < i = n, \\
v_{(n+1)j}(0) - v_{(n+1)n}(0)v_{nj}(0), & j \leq n - 1, i = n + 1, \\
v_{(n+1)j}(0), & j = n, i = n + 1.
\end{cases}
\]

**Step 2:** In the same fashion, eliminate \( v_{(n-1)1}(1), v_{(n-1)2}(1), \cdots, v_{(n-1)(n-2)}(1) \) by elementary column operations to get \( V(2) = V(1)Q(2) \). Precisely, we have
\[
v_{ij}(2) = \begin{cases} 
M_{i-j}(d_1, \cdots, d_j), & j < i \leq n - 2, \\
0, & j < i = n - 1, \\
v_{(n+1)j}(1) - v_{(n+1)(n-1)}(1)v_{(n-1)j}(1), & j \leq n - 2, i = n + 1, \\
v_{(n+1)j}(1), & j = n, n - 1, i = n + 1.
\end{cases}
\]

\[
\cdots
\]

**Step t:** Eliminate \( v_{(n-t+1)1}(t-1), v_{(n-t+1)2}(t-1), \cdots, v_{(n-t+1)(n-t)}(t-1) \) by elementary column
operations to get $V(t) = V(t-1)Q(t)$. Precisely, we have

\[
v_{ij}(t) = \begin{cases} 
M_{i-j}(d_1, \ldots, d_j), & j < i \leq n - t, \\
0, & j < i = n, \ldots, n - t + 1, \\
v_{(n+1)j}(t-1) - v_{(n+1)(n-t+1)}(t-1)v_{(n-t+1)j}(t-1), & j \leq n - t, i = n + 1, \\
v_{(n+1)j}(t-1), & j = n, \ldots, n - t + 1, i = n + 1.
\end{cases}
\]  
(9.12)

\[
\begin{pmatrix}
1 & 0 & \cdots & 0 & 0 \\
0 & 1 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 1 & 0 \\
0 & 0 & \cdots & 0 & 1 \\
v_{(n+1)1}(n-1) & v_{(n+1)2}(n-1) & \cdots & v_{(n+1)(n-1)}(n-1) & v_{(n+1)n}(n-1)
\end{pmatrix} = V(n-2)Q(n-1).
\]

Precisely,

\[
v_{ij}(n-1) = \begin{cases} 
0, & j < i = n, n-1, \ldots, 1, \\
v_{(n+1)j}(n-2) - v_{(n+1)2}(n-2)v_{(n+1)j}(n-2), & j = 1, i = n + 1, \\
v_{(n+1)j}(n-2), & j = n, n-1, \ldots, 2, i = n + 1.
\end{cases}
\]

We now present some useful observations. First, by (9.12) we have

\[
v_{ij}(t) = M_{i-j}(d_1, \ldots, d_j), \text{ for } j < i \leq n - t.
\]  
(9.13)

and

\[
v_{(n+1)j}(n-1) = v_{(n+1)j}(t), \text{ for } t = n-j, \ldots, n-2.
\]  
(9.14)

Second, by (9.12) and an induction argument we can calculate that

\[
v_{(n+1)j}(n-1) = v_{(n+1)j}(n-j) \text{ (by (9.14))}
\]

\[
= v_{(n+1)j}(n-j-1) - v_{(n+1)(j+1)}(n-j-1)v_{(j+1)j}(n-j-1) \text{ (by (9.12))}
\]

\[
= v_{(n+1)j}(n-j-1) - v_{(n+1)(j+1)}(n-j-1)v_{(j+1)j}(n-j-1) \text{ (by (9.14))}
\]

\[
= v_{(n+1)j}(n-j-1) - v_{(n+1)(j+1)}(n-1)M_1(d_1, \ldots, d_j) \text{ (by (9.13))}
\]

\[
= v_{(n+1)j}(n-j-2) - v_{(n+1)(j+2)}(n-j-2)v_{(j+2)j}(n-j-2) - v_{(n+1)(j+1)}(n-1)M_1(d_1, \ldots, d_j) \text{ (by (9.12) - (9.14))}
\]

\[
= v_{(n+1)j}(n-j-2) - v_{(n+1)(j+2)}(n-1)M_2(d_1, \ldots, d_j) - v_{(n+1)(j+1)}(n-1)M_1(d_1, \ldots, d_j) \text{ (by (9.13) - (9.14))}
\]

\[
= M_{n-j+1}(d_1, \ldots, d_j) - v_{(n+1)n}(n-1)M_{n-j}(d_1, \ldots, d_j) - v_{(n+1)(n-1)}(n-1)M_{n-j-1}(d_1, \ldots, d_j)
\]

\[
\vdots - v_{(n+1)(j+1)}(n-1)M_1(d_1, \ldots, d_j).
\]

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

\[
v_{(n+1)j}(n-1) = M_{n-j+1}(d_1, \ldots, d_j) - v_{(n+1)n}(n-1)M_{n-j}(d_1, \ldots, d_j) - v_{(n+1)(n-1)}(n-1)M_{n-j-1}(d_1, \ldots, d_j)
\]

\[
\vdots - v_{(n+1)(j+1)}(n-1)M_1(d_1, \ldots, d_j).
\]  
(9.15)

To prove Lemma 3.1 the following result is needed.
Lemma 9.3. The following identity holds for $0 \leq t \leq n - 1$,

$$M_{t+1}(d_1, \ldots, d_{n-t}) - \left( \sum_{p=1}^{n} d_p \right) M_t(d_1, \ldots, d_{n-t}) - (-1)(\sum_{(r_1,r_2) \in S^2_{1k}} d_{r_1} d_{r_2}) M_{t-1}(d_1, \ldots, d_{n-t})$$

$$- \cdots - (-1)^{t-1} \left( \sum_{(r_1, \ldots, r_t) \in S^t_n} d_{r_1} \cdots d_{r_t} \right) M_{1}(d_1, \ldots, d_{n-t}).$$

$$= (-1)^t \sum_{(r_1, \ldots, r_{t+1}) \in S_{1k}^{t+1}} d_{r_1} \cdots d_{r_{t+1}}. \quad (9.16)$$

Proof: We prove by induction. It is clear (9.16) holds for $n = 1$. Suppose that (9.16) holds for $n = k - 1$, we need to prove for the case $n = k$. We argue as follows. For $0 \leq t \leq k - 2$,

$$M_{t+1}(d_1, \ldots, d_{k-t}) - \left( \sum_{p=1}^{k} d_p \right) M_t(d_1, \ldots, d_{k-t}) - (-1)(\sum_{(r_1,r_2) \in S^2_{1k}} d_{r_1} d_{r_2}) M_{t-1}(d_1, \ldots, d_{k-t})$$

$$- \cdots - (-1)^{t-1} \left( \sum_{(r_1, \ldots, r_t) \in S^t_{1k}} d_{r_1} \cdots d_{r_t} \right) M_{1}(d_1, \ldots, d_{k-t}) =: I + II.$$

where

$$I = M_{t+1}(d_1, \ldots, d_{k-t}) - \left( \sum_{p=1}^{k} d_p \right) M_t(d_1, \ldots, d_{k-t}) - \left( \sum_{(r_1,r_2) \in S^2_{1k}} d_{r_1} d_{r_2} \right) M_{t-1}(d_1, \ldots, d_{k-t})$$

$$- \cdots - (-1)^{t-1} \left( \sum_{(r_1, \ldots, r_t) \in S^t_{1k}} d_{r_1} \cdots d_{r_t} \right) M_{1}(d_1, \ldots, d_{k-t})$$

and

$$II = M_{t+1}(d_2, \ldots, d_{k-t}) - \left( \sum_{p=2}^{k} d_p \right) M_t(d_2, \ldots, d_{k-t}) - (-1)(\sum_{(r_1,r_2) \in S^2_{2k}} d_{r_1} d_{r_2}) M_{t-1}(d_2, \ldots, d_{k-t})$$

$$- \cdots - (-1)^{t-1} \left( \sum_{(r_1, \ldots, r_t) \in S^t_{2k}} d_{r_1} \cdots d_{r_t} \right) M_{1}(d_2, \ldots, d_{k-t}).$$

A direct calculation shows that

$$I = \left[d_1 M_t(d_1, \ldots, d_{k-t})\right] - \left[d_1 M_{t-1}(d_1, \ldots, d_{k-t}) + d_1 \left( \sum_{p=2}^{k} d_p \right) M_{t-1}(d_1, \ldots, d_{k-t})\right]$$

$$- (-1) \left[d_1 \left( \sum_{p=2}^{k} d_p \right) M_{t-1}(d_1, \ldots, d_{k-t}) + d_1 \left( \sum_{(r_1,r_2) \in S^2_{2k}} d_{r_1} d_{r_2} \right) M_{t-2}(d_1, \ldots, d_{k-t})\right] - \cdots$$

$$- (-1)^{t-1} \left[d_1 \left( \sum_{(r_1, \ldots, r_{t+1}) \in S^t_{1k}} d_{r_1} \cdots d_{r_{t+1}} \right) M_0(d_1, \ldots, d_{k-t}) + d_1 \left( \sum_{(r_1, \ldots, r_t) \in S^t_{2k}} d_{r_1} \cdots d_{r_t} \right) M_{t-1}(d_1, \ldots, d_{k-t})\right]$$

$$= (-1)^t d_1 \sum_{(r_1, \ldots, r_{t+1}) \in S^t_{1k}} d_{r_1} \cdots d_{r_{t+1}}. \quad (9.17)$$
On the other hand, by the assumption for \( n = k - 1 \), we have

\[
II = (-1)^t \sum_{(r_1, \ldots, r_{t+1}) \in S_{2k}^{t+1}} d_{r_1} \cdots d_{r_{t+1}}, \quad 0 \leq t \leq k - 2.
\] (9.18)

Therefore for \( 0 \leq t \leq k - 2 \),

\[
I + II = (-1)^t d_1 \sum_{(r_1, \ldots, r_{t+1}) \in S_{2k}^{t+1}} d_{r_1} \cdots d_{r_t} + (-1)^t \sum_{(r_1, \ldots, r_{t+1}) \in S_{2k}^{t+1}} d_{r_1} \cdots d_{r_{t+1}}
\]

and hence \( 9.16 \) holds. Finally, for \( t = k - 1 \), using the decomposition

\[
\sum_{(r_1, \ldots, r_q) \in S_{2k}^{q}} d_{r_1} \cdots d_{r_q} = d_1 \left( \sum_{(r_1, r_{q-1}) \in S_{2k}^{q-1}} d_{r_1} \cdots d_{r_{q-1}} \right) + \sum_{(r_1, \ldots, r_q) \in S_{2k}^{q}} d_{r_1} \cdots d_{r_q}, \quad q = 1, 2, \cdots, k - 1,
\]

we have

\[
M_k(d_1) - \left( \sum_{p=1}^{k} d_p \right) M_{k-1}(d_1) - (-1) \left( \sum_{(r_1, r_2) \in S_{2k}^{k-1}} d_{r_1} \cdots d_{r_2} \right) M_{k-2}(d_1)
\]

\[
- \cdots - (-1)^{k-2} \left( \sum_{(r_1, \ldots, r_{k-1}) \in S_{2k}^{k-1}} d_{r_1} \cdots d_{r_{k-1}} \right) M_1(d_1)
\]

\[
= [M_k(d_1) - d_1 M_{k-1}(d_1)] - \left( \sum_{p=2}^{k} d_p \right) M_{k-1}(d_1) - d_1 \left( \sum_{(r_1, r_2) \in S_{2k}^{k-1}} d_{r_1} \cdots d_{r_2} \right) M_{k-2}(d_1)
\]

\[
- \cdots - (-1)^{k-2} \left( \sum_{(r_1, \ldots, r_{k-1}) \in S_{2k}^{k-1}} d_{r_1} \cdots d_{r_{k-1}} \right) M_1(d_1)
\]

\[
= (-1)^{k-1} d_1 d_2 \cdots d_k.
\]

This completes the proof for the case \( n = k \) and hence the proof of the lemma.

Finally, we prove Lemma 3.1 we need only to show that

\[
v_{(n+1)}(n-1) = (-1)^{n-j} \sum_{(r_1, \ldots, r_{n+1-j}) \in S_{2n}^{n+1-j}} d_{r_1} \cdots d_{r_{n+1-j}}.
\] (9.19)

Recall (9.15), we have

\[
v_{(n+1)}(n-1) = M_{n-j}(d_1, \ldots, d_j) - v_{(n+1)}(n-1)M_{n-j}(d_1, \ldots, d_j) - v_{(n+1)}(n-1)(n-1)M_{n-j-1}(d_1, \ldots, d_j)
\]

\[
- \cdots - v_{(n+1)}(j+1)(n-1)M_1(d_1, \ldots, d_j).
\] (9.20)

For \( j = n \), it is clear that \( v_{(n+1)}(n-1) = \sum_{p=1}^{n} d_p \). By (9.20),

\[
v_{(n+1)}(n-1) = \sum_{p=1}^{n} d_p - v_{(n+1)}(n-1)M_1(d_1, \ldots, d_{n-1})
\]

\[
= M_2(d_1, \ldots, d_{n-1}) - \left( \sum_{p=1}^{n} d_p \right) M_1(d_1, \ldots, d_{n-1}) = - \sum_{(r_1, r_2) \in S_{2n}} d_{r_1} d_{r_2}. \quad \text{(by 9.16)}
\]

Continuing the procedure (using (9.20) and (9.16)), we can show that (9.19) holds for \( j = n, n-1, \ldots, 1 \). This completes the proof of the lemma.
9.4 Appendix D: Some Inequalities

In this section, we present some inequalities that are used in this paper. We first recall the following Stirling approximation of factorial

$$\sqrt{2\pi n^{n+\frac{1}{2}}} e^{-n} \leq n! \leq e^{n+\frac{1}{2}} e^{-n}, \quad (9.21)$$

which will be used frequently in subsequent derivation.

**Lemma 9.4.** For $n \geq 2$,

$$2\sqrt{\pi (n-1)}^{2n-2(2n-2)} \frac{(2\pi (n-1))^{2n-2}}{(2n-2)! \sqrt{4n-3}} \frac{2}{e} \leq 1.$$

Proof: By (9.21), for $n \geq 2$,

$$2\sqrt{\pi (n-1)}^{2n-2(2n-2)} \frac{(2\pi (n-1))^{2n-2}}{(2n-2)! \sqrt{4n-3}} \frac{2}{e} \leq \frac{2\sqrt{\pi (n-1)}^{2n-2(2n-2)}}{\sqrt{2\pi (2n-2)^{2n-2} + \frac{1}{2}} e^{-(2n-2)\sqrt{4n-3}}} \frac{2}{e} \leq 1.$$

**Lemma 9.5.** For $n \geq 2$,

$$\frac{2\sqrt{\pi n^{2n-1}(2n-1)}}{e(2n-1)! \sqrt{4n-1}} \frac{2}{e} \leq 1.$$

Proof: By (9.21), for $n \geq 2$,

$$\frac{2\sqrt{\pi n^{2n-1}(2n-1)}}{e(2n-1)! \sqrt{4n-1}} \frac{2}{e} \leq \frac{2\sqrt{\pi n^{2n-1}(2n-1)}}{e\sqrt{2\pi (2n-1)^{2n-1 + \frac{1}{2}} e^{-(2n-1)\sqrt{4n-1}}} \frac{2}{e} \leq 1."

**Lemma 9.6.** For $k = 1, 2, \cdots$, the following estimate holds

$$\frac{\zeta(k+1)\zeta(k)}{(2k)! \sqrt{4k+1}} \geq \frac{1.15}{2^{4k} k},$$

where $\zeta(k+1), \zeta(k)$ are defined in (3.4).

Proof: For $k = 1, 2, 3$, the inequality holds. For odd $k \geq 4$, by (9.21),

$$\frac{\zeta(k+1)\zeta(k)}{(2k)! \sqrt{4k+1}} \geq \frac{4\pi^2 (\frac{k+1}{2})! (\frac{k-1}{2})! (\frac{k-3}{2})!}{4(2k)! \sqrt{4k+1}} \geq \frac{4\pi^2 (\frac{k+1}{2})! (\frac{k-1}{2})! (\frac{k-3}{2})!}{4(2k)! \sqrt{4k+1}} \geq \frac{4\pi^2 (\frac{k+1}{2})! (\frac{k-1}{2})! (\frac{k-3}{2})!}{4(2k)! \sqrt{4k+1}} \geq \frac{1.15}{2^{4k} k} \quad \text{because} \quad k \geq 4"
For even \( k \geq 4 \),
\[
\zeta(k+1)\xi(k) = \frac{(\frac{k}{2})^{2k}(\frac{1}{2})^{2}}{(2k)\sqrt{4k+1}} \geq \frac{4\pi^2(\frac{k}{2})^{k+1}(\frac{1}{2})^{k-1}e^{-2k+2}}{4(2k)\sqrt{4k+1}}
\]
\[
= \frac{\epsilon n^2(\frac{1}{2})^{2k}}{2^{2k}\sqrt{2k\sqrt{4k+1}}} \geq \frac{1}{2^{4k}k8\sqrt{2}\sqrt{4k+1}} \geq \frac{1.15}{2^{4k}k} \quad \text{(because } k \geq 4) \]

**Lemma 9.7.** For \( n = 2, 3, \ldots \)
\[
2^n \sqrt{\frac{4(1+d)2^{n-1}\sqrt{4n-1}(2n-1)!}{\zeta(n)\lambda(n)}} \leq 6.24(1+d),
\]
(9.22)
where \( \zeta(n) \) is defined in (3.6) and
\[
\lambda(n) = \begin{cases} 
\frac{1}{2}, & n = 2, \\
\xi(n-2), & n \geq 3 \end{cases}
\]
\[
\xi(n-2) = \begin{cases} 
\frac{1}{2}, & n = 3, \\
\left(\frac{n-3}{2}\right)\left(\frac{n-4}{2}\right)^{n-3}, & n \text{ is odd, } n \geq 5, \\
\left(\frac{n-4}{2}\right)^{n} \xi(n), & n \text{ is even, } n \geq 4. 
\end{cases}
\]

Proof: For \( n = 2, 3, 4, 5 \), the inequality holds. For even \( n \geq 6 \), by [9.21],
\[
\frac{4\sqrt{4n-1}(2n-1)!}{\zeta(n)\lambda(n)} = \frac{16\sqrt{4n-1}(2n-1)!}{(\frac{1}{2})(\frac{n-2}{2})(\frac{n-4}{2})^{2}} \leq \frac{16\sqrt{4n-1}e^{2n-\frac{3}{4}}e^{-(2n-1)}}{(2\pi)^{2}(\frac{n}{2})^{n-\frac{1}{2}}(\frac{n-2}{2})^{n-\frac{3}{2}}(\frac{n-4}{2})^{n-3}e^{-(2n-3)}}
\]
\[
= \frac{\epsilon n^2(\frac{1}{2})^{2n-1}}{4\pi^2} \left(\frac{n}{2}\right)^{n-\frac{3}{2}}(n-2)^{n-\frac{3}{2}}(n-4)^{n-3} \leq \frac{e\sqrt{4n-1}(2n-1)^{\frac{3}{2}}2^{n-1}}{4\pi^2} \leq (6.24)^{2n-1}.
\]

For odd \( n \geq 7 \),
\[
\frac{4\sqrt{4n-1}(2n-1)!}{\zeta(n)\lambda(n)} = \frac{16\sqrt{4n-1}(2n-1)!}{(\frac{n-1}{2})^{2}(\frac{n-3}{2})!(\frac{n-5}{2})^{2}} \leq \frac{16\sqrt{4n-1}e^{2n-\frac{3}{4}}e^{-(2n-1)}}{(2\pi)^{2}(\frac{n-1}{2})^{n}(\frac{n-3}{2})^{\frac{n}{2}}(\frac{n-5}{2})^{\frac{n}{2}}e^{-(2n-5)}}
\]
\[
= \frac{\epsilon n^2(\frac{1}{2})^{2n-1}}{4\pi^2} \left(\frac{n}{2}\right)^{n-\frac{3}{2}}(n-2)^{n-\frac{3}{2}}(n-4)^{n-3} \leq \frac{e\sqrt{4n-1}(2n-1)^{\frac{3}{2}}2^{n-1}}{4\pi^2} \leq (6.24)^{2n-1}.
\]

It follows that for \( n \geq 6 \)
\[
2^n \sqrt{\frac{4(1+d)2^{n-1}\sqrt{4n-1}(2n-1)!}{\zeta(n)\lambda(n)}} \leq 6.24(1+d).
\]
Lemma 9.8. For $n = 2, 3, 4, \cdots$

$$\frac{2^{n-1}(2n-1)!\sqrt{4n-1}}{\zeta(n)(n-2)!} \leq \frac{7.73\sqrt{4n-1}(2n-1)4^{2n-1}}{4\pi^2},$$

where $\zeta(n)$ is defined in (3.6).

Proof: For $n = 2$, the inequality holds. For odd $n \geq 3$, by Stirling approximation formula (9.21),

$$\frac{2^{n-1}(2n-1)!\sqrt{4n-1}}{\zeta(n)(n-2)!} = \frac{2^{n-1}(2n-1)!\sqrt{4n-1}}{(\frac{n}{2})!^2(n-2)!} \leq \frac{2^{n-1}\sqrt{4n-1}e(2n-1)2^{n-\frac{1}{2}}e^{-2n-1}}{(\sqrt{2\pi})^3(\frac{n}{2})!^2(n-2)!} \leq \frac{7.73\sqrt{4n-1}(2n-1)4^{2n-1}}{4\pi^2} \cdot \frac{1}{(n-1)^{n-2}\pi^2}.$$  \(\text{since } n \geq 3\)

For even $n \geq 4$,

$$\frac{2^{n-1}(2n-1)!\sqrt{4n-1}}{\zeta(n)(n-2)!} = \frac{2^{n-1}(2n-1)!\sqrt{4n-1}}{(\frac{n}{2})!^2(n-2)!} \leq \frac{2^{n-1}\sqrt{4n-1}e(2n-1)2^{n-\frac{1}{2}}e^{-2n-3}}{(\sqrt{2\pi})^3(\frac{n}{2})!^2(n-2)!} \leq \frac{7.73\sqrt{4n-1}(2n-1)4^{2n-1}}{4\pi^2} \cdot \frac{1}{n^{n-2}(n-2)^{n-2}\pi^2}.$$

REFERENCES

[1] J.G. Ables. Maximum entropy spectral analysis. *Astronomy and Astrophysics Supplement Series*, 15:383, 1974.

[2] Habib Ammari and Hai Zhang. A mathematical theory of super-resolution by using a system of sub-wavelength helmholtz resonators. *Communications in Mathematical Physics*, 337(1):379–428, 2015.

[3] Habib Ammari and Hai Zhang. Super-resolution in high-contrast media. *Proc. R. Soc. A*, 471, 2015.

[4] Brett Bernstein and Carlos Fernandez-Granda. Deconvolution of point sources: a sampling theorem and robustness guarantees. *Communications on Pure and Applied Mathematics*, 72(6):1152–1230, 2019.

[5] J.P. Burg. Maximum entropy spectral analysis. *Proceedings of 37th Meeting, Society of Exploration Geophysics*(Oklahoma City, OK), Oct. 1967.

[6] Jian-Feng Cai, Tianming Wang, and Ke Wei. Fast and provable algorithms for spectrally sparse signal reconstruction via low-rank hankel matrix completion. *Applied and Computational Harmonic Analysis*, 46(1):94–121, 2019.

[7] Emmanuel J. Candès and Carlos Fernandez-Granda. Super-resolution from noisy data. *Journal of Fourier Analysis and Applications*, 19(6):1229–1254, 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] Emmanuel J. Candès, Justin Romberg, and Terence Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on information theory*, 52(2):489–509, 2006.

[10] Jack Capon. High-resolution frequency-wavenumber spectrum analysis. *Proceedings of the IEEE*, 57(8):1408–1418, 1969.

[11] 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.

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

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

[14] David L. Donoho et al. Compressed sensing. *IEEE Transactions on information theory*, 52(4):1289–1306, 2006.

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

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

[17] Albert C. Fannjiang and Wenjing Liao. Coherence pattern–guided compressive sensing with unresolved grids. *SIAM Journal on Imaging Sciences*, 5(1):179–202, 2012.

[18] Carlos Fernandez-Granda. Support detection in super-resolution. *arXiv preprint arXiv:1302.3921*, 2013.

[19] Walter Gautschi. On inverses of vandermonde and confluent vandermonde matrices. *Numerische Mathematik*, 4(1):117–123, 1962.

[20] Rainer Heintzmann and Thomas Huser. Super-resolution structured illumination microscopy. *Chemical reviews*, 117(23):13890–13908, 2017.

[21] Stefan W. Hell and Jan Wichmann. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Optics letters*, 19(11):780–782, 1994.

[22] J.A. Högbom. Aperture synthesis with a non-regular distribution of interferometer baselines. *Astronomy and Astrophysics Supplement Series*, 15:417, 1974.
[23] Yingbo Hua and Tapan K. Sarkar. Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38(5):814–824, 1990.

[24] Satoshi Kawata, Keiichiroh Minami, and Shigeo Minami. Superresolution of fourier transform spectroscopy data by the maximum entropy method. *Applied optics*, 22(22):3593–3598, 1983.

[25] S.M. Kay. Modern spectral estimation: Theory and application. 1988. *Englewood Cliffs, NJ*, 1999.

[26] Steven M Kay and Stanley Lawrence Marple. Spectrum analysis a modern perspective. *Proceedings of the IEEE*, 69(11):1380–1419, 1981.

[27] Thomas A. Klar and Stefan W. Hell. Subdiffraction resolution in far-field fluorescence microscopy. *Optics letters*, 24(14):954–956, 1999.

[28] R.T. Lacoss. Autoregressive and maximum likelihood spectral analysis methods. In *Aspects of Signal Processing With Emphasis on Underwater Acoustics, Part 2*, pages 591–615. Springer, 1977.

[29] Geoffroy Lerosey, Julien de Rosny, Arnaud Tourin, and Mathias Fink. Focusing beyond the diffraction limit with far-field time reversal. 315(5815):1120–1122, 2007.

[30] Shlomo Levy and Peter K Fullagar. Reconstruction of a sparse spike train from a portion of its spectrum and application to high-resolution deconvolution. *Geophysics*, 46(9):1235–1243, 1981.

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

[32] Wenjing Liao and Albert C. Fannjiang. Music for single-snapshot spectral estimation: Stability and super-resolution. *Applied and Computational Harmonic Analysis*, 40(1):33–67, 2016.

[33] 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.

[34] Veniamin I. Morgenshtern and Emmanuel J. Candes. Super-resolution of positive sources: The discrete setup. *SIAM Journal on Imaging Sciences*, 9(1):412–444, 2016.

[35] Halil Oruç and George M. Phillips. Explicit factorization of the vandermonde matrix. *Linear Algebra and its Applications*, 315(1-3):113–123, 2000.

[36] Athanasios Papoulis. A new algorithm in spectral analysis and band-limited extrapolation. *IEEE Transactions on Circuits and systems*, 22(9):735–742, 1975.
[37] Athanasios Papoulis and Christodoulos Chamzas. Improvement of range resolution by spectral extrapolation. *Ultrasonic Imaging*, 1(2):121–135, 1979.

[38] R. Prony. Essai expérimental et analytique. *J. de l’Ecole Polytechnique (Paris)*, 1(2):24–76, 1795.

[39] Richard Roy and Thomas Kailath. Esprit-estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on acoustics, speech, and signal processing*, 37(7):984–995, 1989.

[40] Michael J. Rust, Mark Bates, and Xiaowei Zhuang. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (storm). *Nature methods*, 3(10):793, 2006.

[41] Fadil Santosa and William W Symes. Linear inversion of band-limited reflection seismograms. *SIAM Journal on Scientific and Statistical Computing*, 7(4):1307–1330, 1986.

[42] Ralph Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Transactions on antennas and propagation*, 34(3):276–280, 1986.

[43] U.J. Schwarz. Mathematical-statistical description of the iterative beam removing technique (method clean). *Astronomy and Astrophysics*, 65:345, 1978.

[44] P.J. Sementilli, Bobby R. Hunt, and M.S. Nadar. Analysis of the limit to superresolution in incoherent imaging. *JOSA A*, 10(11):2265–2276, 1993.

[45] Morteza Shahram and Peyman Milanfar. Imaging below the diffraction limit: a statistical analysis. *IEEE Transactions on image processing*, 13(5):677–689, 2004.

[46] Morteza Shahram and Peyman Milanfar. On the resolvability of sinusoids with nearby frequencies in the presence of noise. *IEEE Transactions on Signal Processing*, 53(7):2579–2588, 2005.

[47] Petre Stoica and Arye Nehorai. Music, maximum likelihood, and cramer-rao bound. *IEEE Transactions on Acoustics, speech, and signal processing*, 37(5):720–741, 1989.

[48] Gongguo Tang, Badri Narayan Bhaskar, and Benjamin Recht. Near minimax line spectral estimation. *IEEE Transactions on Information Theory*, 61(1):499–512, 2014.

[49] Xiang Zhang and Zhaowei Liu. Superlenses to overcome the diffraction limit. *Nature Materials*, 7, 2008.