Circuits and Systems
Mekelweg 4,
2628 CD Delft
The Netherlands
http://ens.ewi.tudelft.nl/

CAS-2014-00

M.Sc. Thesis

Compressive Sensing for Near-field Source Localization

Keke Hu

Abstract

Near-field source localization is an important aspect in many diverse areas such as acoustics, seismology, to list a few. The planar wave assumption frequently used in far-field source localization is no longer valid when the sources are in the near field. Near-field sources can be localized by solving a joint direction-of-arrival and range estimation problem. The original near-field source localization problem is a multi-dimensional non-linear optimization problem which is computationally intractable. In this thesis we study address two important questions related to near-field source localization: (i) Sparse reconstruction techniques for joint DOA and range estimation using a grid-based model. (ii) Matching the sampling grid for off-grid sources. In the first part of this thesis, we use a grid-based model and by further leveraging the sparsity, we can solve the aforementioned problem efficiently using any of the off-the-shelf ℓ_1-norm optimization solvers. When multiple snapshots are available, we can also exploit the cross-correlations among the symmetric sensors of the array and further reduce the complexity by solving two sparse reconstruction problems of lower dimensions instead of a single sparse reconstruction problem of a higher dimension. In the second part of this thesis, we account scenarios where the true source locations are not on the grid resulting in a grid mismatch. Using the first-order Taylor approximation, we model the grid mismatch as a perturbation around the sampling grid. Based on the grid mismatch model, we propose a bounded sparse and bounded joint sparse recovery algorithms to localize near-field sources.

Compressive Sensing for Near-field Source Localization

Thesis

submitted in partial fulfillment of the
requirements for the degree of

Master of Science

in

Telecommunications

by

Keke Hu
born in Yichang, P. R. China

This work was performed in:

Circuits and Systems Group
Department of Electrical Engineering
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology

Delft University of Technology

Delft University of Technology
Department of
Electrical Engineering

The undersigned hereby certify that they have read and recommend to the Faculty of Electrical Engineering, Mathematics and Computer Science for acceptance a thesis entitled **"Compressive Sensing for Near-field Source Localization"** by **Keke Hu** in partial fulfillment of the requirements for the degree of **Master of Science**.

Dated: January 30th, 2014

Chairman: _____
Prof. dr. ir. Geert Leus

Advisor: _____
ir. Sundeep Prabhakar Chepuri

Committee Members: _____
dr. ir. Richard Heusdens

_____
dr. ir. Radmila Pribić

Abstract

Near-field source localization is an important aspect in many diverse areas such as acoustics, seismology, to list a few. The planar wave assumption frequently used in far-field source localization is no longer valid when the sources are in the near field. Near-field sources can be localized by solving a joint direction-of-arrival and range estimation problem. The original near-field source localization problem is a multi-dimensional non-linear optimization problem which is computationally intractable. In this thesis we study address two important questions related to near-field source localization: (i) Sparse reconstruction techniques for joint DOA and range estimation using a grid-based model. (ii) Matching the sampling grid for off-grid sources. In the first part of this thesis, we use a grid-based model and by further leveraging the sparsity, we can solve the aforementioned problem efficiently using any of the off-the-shelf ℓ_1-norm optimization solvers. When multiple snapshots are available, we can also exploit the cross-correlations among the symmetric sensors of the array and further reduce the complexity by solving two sparse reconstruction problems of lower dimensions instead of a single sparse reconstruction problem of a higher dimension. In the second part of this thesis, we account scenarios where the true source locations are not on the grid resulting in a grid mismatch. Using the first-order Taylor approximation, we model the grid mismatch as a perturbation around the sampling grid. Based on the grid mismatch model, we propose a bounded sparse and bounded joint sparse recovery algorithms to localize near-field sources.

Acknowledgments

I would like to thank the people who have helped and supported me thoughout my project. The progress I made during my thesis project is inseparable from their help and support.

First of all, I would like to express my sincere gratitude to my supervisor ir. Sundeep Prabhakar Chepuri for his continuous support of my thesis study and making me progress during the project. I trouble him a lot, and sometimes I even entangled in a minor question, but he alway provides inspired suggestions immediately with patience. From Sundeep, I learnt how to think critically and how to manage a project. Thank Sundeep, and if without his help this work can never be accomplished.

I am grateful to my professor, Prof. dr. ir. Geert Leus for giving me the opportunity to work in this group and for his time, patient, encouragement, support, and invaluable guidance.

Also, I would like to thank dr. ir. Richard Heusdens and dr. ir. Radmila Pribić for being my committee members. It is my honor.

I would like to thank Sumeet Kumar, Shahzad Sarwar Gishkori, Dyonisius Dony Ariananda and Venkat Roy for their help and advice during my thesis project. And I would like to thank all the member of CAS group, they are enthusiasm, friendly and kindness.

Finally, I would like to thank my family and friends for their support, help and encouragement. These are the source of my strength. And thank Ying He for her understanding and support. Thank my mom, if without her I won't come to this beautiful world. And I miss her very much.

Keke Hu
Delft, The Netherlands
January 30th, 2014

Contents

Notations

x	Single element
$\mathbf{x}$	Vector of element
$\mathbf{1}_N$ ($\mathbf{0}_N$)	Vector of ones (zeros)
$\mathbf{X}$	Matrix
$\mathbf{I}_N$	An identity matrix of size N
$(\bullet)^*$	Complex conjugate
$(\bullet)^T$	Transposition
$(\bullet)^H$	Hermitian transposition (complex conjugate transposition)
$\mathbb{R}$	Real number set notation
$\mathbb{C}$	Complex number set notation
$\mathbb{E}\{\bullet\}$	Expectation operation
$\text{vec}(\bullet)$	Vectorize a matrix
$\text{diag}(\bullet)$	Put an vector on the main matrix
$\odot$	Hadamard product
$\circledcirc$	Khatri-rao product
$\|\mathbf{x}\|_0$	The ℓ_0(-quasi) norm that indicates the number of non-zero entries of the vector $\mathbf{x}$
$\|\mathbf{x}\|_1$	The ℓ_1 norm of an $N \times 1$ vector $\mathbf{x}$ equals to $\sum_{n=1}^{N} \|x_n\|$
$\|\mathbf{x}\|_2$	The ℓ_2 norm of an $N \times 1$ vector $\mathbf{x}$ equals to $\sqrt{\sum_{n=1}^{N} \|x_n\|^2}$
$\|\mathbf{x}\|_{2,1}$	The mixed ℓ_2/ℓ_1 norm of vector $\mathbf{x}$, which can be divided into N sub-blocks $\mathbf{x}_n$, is equivalent to $\sum_{n=1}^{N} \|\mathbf{x}_n\|_2$
$\|\mathbf{X}\|_F$	The Frobenius norm of an $M \times N$ matrix $\mathbf{X}$ equals to $\sqrt{\sum_{m=1}^{M} \sum_{n=1}^{N} \|x_{mn}\|^2}$
$\sigma_{s,k}^2$	Variance of k-th signal
σ_w^2	Variance of additive noise

List of Figures

1 Introduction

In this thesis, we study the near-field source localization problem using sparse recovery techniques. This chapter provides the motivation behind this thesis, literature review, and contributions.

1.1 Literature review

Source localization is an important aspect for target tracking and location-aware services, and has many applications in the field of seismology, acoustics, radar, sonar, and oceanography. In Fig. 1.1, an example of source localization applied to seismic imaging is shown.

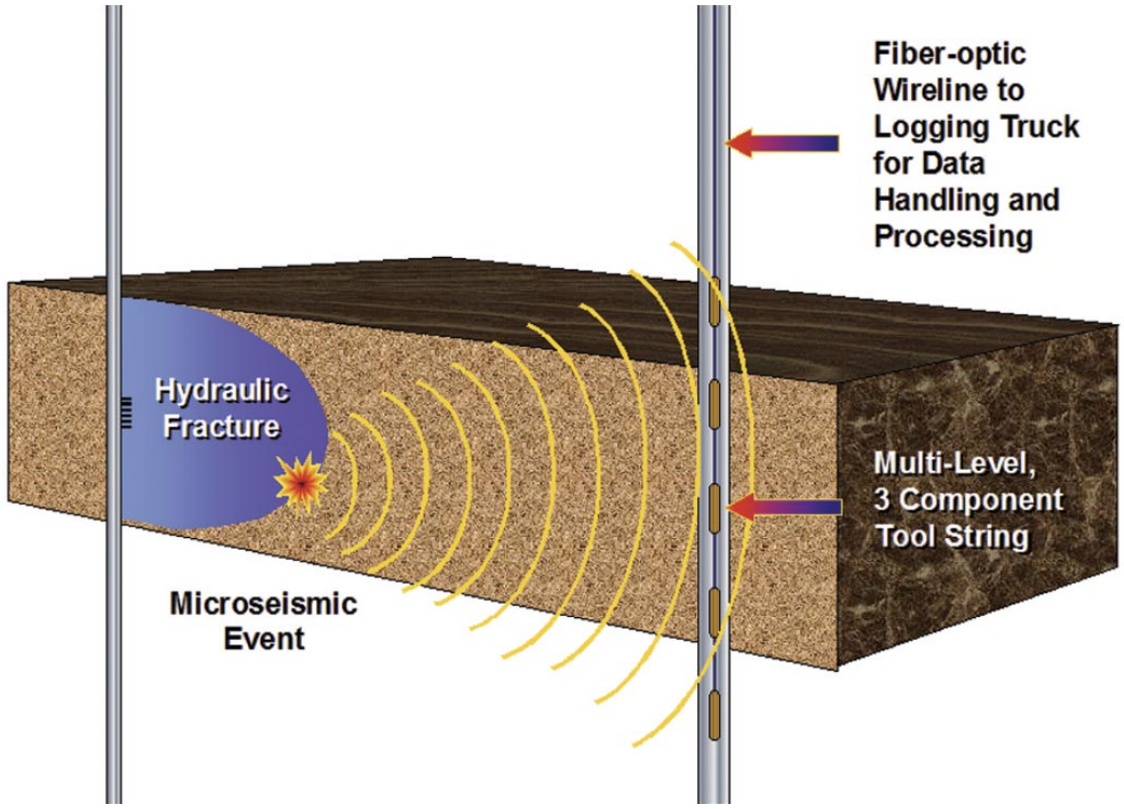


Figure 1.1: Schematic diagram showing an acquisition of micro-seismic data for seismic imaging. Image courtesy of John Logel [1].

Direction-of-arrival (DOA) estimation of narrowband signals is an extensively studied topic [2, 3]. DOA estimation can be categorized into two types, based on the distance between the source and the antenna array: (a) far-field (e.g., $r \gg 2D^2/\lambda$),

and (b) near-field source localization, where r is the range between the source and the phase-reference of the array, D is the array aperture, and λ is the wavelength of the source signal. In far-field source localization, the wavefront of the signal impinging on the array is assumed to be planar [2, 4]. However, the curvature of the wavefront is no longer negligible when sources are located close to the array (i.e., in the near field or Fresnel region) like the scenario shown in Fig. 1.1. Therefore, the algorithms that depend on the planar-wave assumptions for DOA estimation are no longer valid. Such problems often arise in practical term, for example in underwater sources localization by using a vector hydrophone [5], tracking a speaker by using microphones array [6].

Unlike the far-field source localization, near-field localization is traditionally done by a joint DOA and range (distance between the source and the phase-reference of the array) estimation. Traditional approaches to the near-field localization problem extend the techniques like multiple signal classification (MUSIC) to a two-dimensional field. In [7], the MUSIC method used in a two-dimensional (DOA, range) field achieves a high resolution, but, suffers from a high computational load required for searching a two-dimensional range and bearing spectrum. Moreover, the performance of the MUSIC algorithm deteriorates at low SNRs and when the sources are correlated. In [8], the wavefront is assumed to be piece-wise linear, and the uniform linear array (ULA) is divided into several subarrays. The wavefront of the signal impinging on each subarray is then assumed to be planar. By using the method proposed in [2, 4] at each subarray, the location can be estimated after gathering the DOA of each subarray. In [9, 10, 11], instead of using the piece-wise linear approximation, a quadratic approximation (the so-called Fresnel approximation) of the wavefront is made, which makes the wavefront neither planar nor spherical. The phase delay is no longer linear with the position of the antenna element (like in [2]), instead, it varies quadratically with the array position and it is characterized by the azimuth (DOA) and range of the sources (see [9] for more details). However, the array has to still satisfy the Nyquist sampling rate criterion in space, i.e., the spacing between the two adjacent antennas need to less than half a wavelength.

1.2 Motivation

The previous works on near-field localization problem can provide accurate results or relative satisfying resolution but all of them are using the conventional data acquisition methods based on the Nyquist sampling rate theorem [7, 8, 9, 10, 11], which limits their performance. Since in many application the sources can be assumed as point sources and hereby the spectrum of energy versus location can be viewed as sparse with the number of sources is small. Based on this assumption, we try to localize the source locations via a perspective of sparse signal recovery. Generally the sampled measurement used for recovery is less than Nyquist sampling rate, in which case the framework is called compressive sensing (CS) [12]. In [13], the CS framework is used for DOA estimation of far-field sources. The CS framework can provide better resolution than conventional approaches [2, 3, 4] from a lesser number of samples but suffers from the computational load for multiple snapshot case. To avoid the unnecessary computational load, [14, 15] developed a tractable subspace-based ℓ_1-SVD method for multiple

snapshot far-field source localization which can also provide super-high resolution with low complexity. Thus, we wonder whether if the sparse recovery techniques can be applied for the near-field source localization problem also.

Even though the source localization based on the sparse signal reconstruction (SSR) can perform better than the classical source localization methods, its performance is tremendously affected by the choice of the estimation grid. The grid mismatch problem is described in [16], and an error-in-variables (EIV) model is proposed in [17, 18, 19]. In [20], a joint-sparse recovery [21, 22] was used to further exploit the underlying structure with grid mismatch. Using the same concepts, we try to exploit the off-grid compressed sensing for near-field source localization.

1.3 Thesis outline and contributions

In this thesis, we use the CS framework to jointly estimate the DOAs and ranges of multiple narrowband near-field sources. We address two scenarios in this thesis: in the first part, we illustrate how to localize multiple near-field sources using the SSR perspective under assumption that all the desired sources are exactly on the sampling grid; in the second part, we extend the near-field linear model to an EIV model to include the off-grid (all the desired sources need not be on the sampling grid) effect and propose a number of estimators for the same.

Chapter 2: Background

In this chapter, a brief introduction to CS and its applications related to our topic is provided. Firstly, we will introduce some terminologies used in CS. Then, we discuss sparse signal recovery via a ℓ_1-norm regularized least-squares algorithm. Next, we introduce ℓ_1-SVD, which we use when multiple snapshots are available. At last, we introduce the concept of block sparse recovery.

Chapter 3: Signal model

In this chapter, we develop the signal model that we use throughout this thesis. It includes the notation and terminologies that will be used in this thesis. We will also provide spherical wavefront model, quadratic wavefront model (based on the Fresnel approximation).

Chapter 4: Sparse recovery techniques for near-field source localization: Grid-based model

In this chapter, we propose a grid-based model and by further leveraging the sparsity, we solve the aforementioned problem efficiently using ℓ_1-regularized least-squares. When multiple snapshots cases are available, we exploit the cross-correlations among the symmetric sensors to further reduce the complexity by solving two sparse reconstruction problems of lower dimensions instead of a single sparse reconstruction problem of a higher dimension.

Chapter 5: Grid matching for near-field source localization using sparse recovery techniques

In this chapter, we consider a scenario in which the true source locations are not on the sampling grid. By using the Taylor expansion, we model the grid mismatch effect. Based on the grid mismatch model, we propose a sparse recovery and a joint sparse recovery estimators to localize the near-field sources.

Chapter 6: Conclusions

In this chapter, we summarize this thesis and provide some suggestions for future research.

2 Background

The compressive sensing (CS) theory is a technique to reconstruct signals at a rate that could be significantly below the Nyquist rate under the assumption that the desired signals are compressible or sparse [12]. This chapter introduces the concept of CS briefly and along with ℓ_1-norm minimization, ℓ_1-SVD, and block sparse recovery.

2.1 Compressive sensing

Before we discuss the compressive sensing problem, we will introduce the concept of compressive signals. In general, we can say a signal vector $\mathbf{x}$ is a K-sparse vector if it has only K non-zero entries of $\mathbf{x}$. Mathematically, it can be represented as $\|\mathbf{x}\|_0 \leq K$, where $\|\mathbf{x}\|_0$ indicate the number of non-zero entries. We can say a signal $\mathbf{x}$ is compressible if it is K-sparse vector or nearly sparse (with K relatively larger coefficients compared to the rest).

Given a measurement vector $\mathbf{y} \in \mathbb{C}^M$ and a sensing matrix $\mathbf{A} \in \mathbb{C}^{M \times N}$, we want to recovery the original vector $\mathbf{x} \in \mathbb{C}^N$ acquired via a linear measurement process $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w}$, where $\mathbf{w} \in \mathbb{C}^M$ is the additive noise vector. As we known, if we want to recovery the signal $\mathbf{x}$, we normally need to acquire more than N measurements, which means if $M \geq N$ there exists a unique solution of $\mathbf{x}$. When we have fewer measurements, i.e., $M \ll N$, if we know that the desired signal $\mathbf{x}$ is a compressive signal, with a high probability we can still recovery the signal vector $\mathbf{x}$ by solving an optimization problem of form

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x} \in \mathbb{C}^{N \times 1}} \quad \|\mathbf{x}\|_0 \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \epsilon, \tag{2.1}$$

the above framework of reconstructing $\mathbf{x}$ is called CS. The wide matrix $\mathbf{A}$ is often referred to as the sensing matrix. The optimization problem in (2.1) is non-convex due to the ℓ_0-norm cost function and is mathematically intractable. In fact, for a general matrix $\mathbf{A}$ finding an approximate solution of (2.1) is NP-hard. Thus, in practical we do not propose to use this strategy to solve CS problem due to its intractability.

2.2 ℓ_1-minimization algorithm

As we mentioned before, $\|\mathbf{x}\|_0$ is non-convex, and the ℓ_0-norm minimization is an NP-hard problem. In order to transform such a problem into a computational tractable one, it is proposed to relax the ℓ_0-norm cost with the ℓ_1-norm, and hence transforming (2.1) to a convex optimization problem of the form

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x} \in \mathbb{C}^{N \times 1}} \quad \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \epsilon. \tag{2.2}$$

The optimization problem in (2.2) can alternatively be written in form

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbb{C}^{N\times 1}} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \mu\|\mathbf{x}\|_1, \tag{2.3}$$

where parameter μ is used to control tradeoff between the sparsity and the residual norm. In this thesis, we use LASSO, a shrinkage and selection method for linear regression in [23], to solve the ℓ_1 regularized optimization problem in (2.3) .

In fact, it provides accurate solution of $\mathbf{x}$ only under the two conditions that (1) the $\mathbf{x}$ is sufficiently sparse and (2) the sensing matrix $\mathbf{A}$ satisfies an important condition named restricted isometry property (RIP) [24].

In conclusion, the ℓ_1 minimization approach is one of the numerous algorithms applied for CS recovery problem and provides a powerful framework to recovering sparse signals. The advantage of ℓ_1 minimization algorithm is that it provides a provably accurate solution and relaxes the CS problem to a convex optimization problem which is more tractable and can be solved by many existing efficient and accurate solvers.

2.3 ℓ_1-SVD

When multiple snapshot case available, the overcomplete representation is extended and reformulated in the following form

$$\mathbf{y}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{w}(t), \quad t = 1, 2, \ldots, T, \tag{2.4}$$

where $\mathbf{y}(t) \in \mathbb{C}^M$, $\mathbf{x}(t) \in \mathbb{C}^N$ is a K-sparse vector and the positions of non-zero entries will not change at different t, $T \gg K$, and $\mathbf{A} \in \mathbb{C}^{M\times N}$.

One direct approach to reconstruct the sparse signal sequence is to treat each snapshot separately. We separate the problem sample by sample and solve each problem indexed by t as if each problem at index t is a single snapshot problem solved as a ℓ_1-norm minimization problem given by

$$\hat{\mathbf{x}}(t) = \arg\min_{\mathbf{x}(t)\in\mathbb{C}^{N\times 1}} \quad \|\mathbf{y}(t) - \mathbf{A}\mathbf{x}(t)\|_2^2 + \mu\|\mathbf{x}(t)\|_1, \quad t = 1, 2, \ldots, T, \tag{2.5}$$

then we will have a set of T solutions. This algorithm suffers too much computational load and the position of non-zero entries of the estimate may be different at different index t. In Fig. 2.1b, the result of the estimate separately (2.5) is shown. The position of non-zero entries changes for different sample due to the fact that there is no dependence between each sub-problem at different index t.

The ℓ_1-SVD algorithm we will introduce in this section is an efficient regularization framework to find the related columns of $\mathbf{A}$ to the non-zeros entries. It is based on the singular value decomposition (SVD) of the measurement matrix and the idea is to decompose the measurement matrix into signal and noise subspaces.

Stacking (2.4) into matrix $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{W}$, where $\mathbf{X}$ is a row-wise sparse matrix. Consider the singular value decomposition of $\mathbf{Y}$:

$$\mathbf{Y} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H.$$

Since only K rows of $\mathbf{X}$ have non-zero values, then in noiseless case the rank of $\mathbf{Y}$ should be K. Accordingly, the first K columns of matrix $\mathbf{U}$ correspond to the non-zero singular value. While with noise, the $\mathbf{Y}$ has full rank and the largest K singular vectors will correspond to the signal subspace and the rest are related to the noise. Therefore, we can generate a reduced dimension matrix $\mathbf{Y}_{\text{SV}} \in \mathbb{C}^{M\times K}$, which only contains the most of the signal power

$$\mathbf{Y}_{\text{SV}} = \mathbf{U}\mathbf{\Lambda}\mathbf{D}_K = \mathbf{Y}\mathbf{V}\mathbf{D}_K,$$

where $\mathbf{D}_K = [\mathbf{I}_K, \mathbf{0}_K^H]^H \in \mathbb{R}^{K\times T}$, $\mathbf{I}_K$ is a $K\times K$ identity matrix, and $\mathbf{0}_K$ is a $K\times(T-K)$ matrix of all zero value. Here, we multiple both size of equation $\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{W}$ with $\mathbf{V}\mathbf{D}_K$, we obtain

$$\mathbf{Y}_{\text{SV}} = \mathbf{A}\mathbf{X}_{\text{SV}} + \mathbf{W}_{\text{SV}}, \tag{2.6}$$

where $\mathbf{X}_{SV} = \mathbf{X}\mathbf{V}\mathbf{D}_K$ and $\mathbf{W}\mathbf{V}\mathbf{D}_K$. The above equation can be expressed in the vector-form (column by column) as

$$\mathbf{y}_{\text{sv}}(k) = \mathbf{A}\mathbf{x}_{\text{sv}}(k) + \mathbf{w}_{\text{sv}}(k), \text{ for } k = 1, \ldots, K.$$

Here, each column corresponds to a signal subspace singular vector. The reduced data matrix is only spatially sparse, and not in terms of the singular vector index k. In order to take into account this effect, we use a different prior obtained by computing the ℓ_2-norm of the singular values of a particular spatial index of $\mathbf{x}_{\text{sv}}$, i.e., $x_{m,\text{sv}}^{(\ell_2)} = \sqrt{\sum_{k=1}^{K}(x_{m,\text{sv}}(k))^2}$, for $m = 1, \ldots, M$. Note that, the ℓ_2-norm can be computed for all time-samples instead of only the signal subspace singular vectors, however, the former technique adds more computational complexity especially when $T \gg K$. Now, we can find the range by minimizing

$$\|\mathbf{Y}_{\text{sv}} - \mathbf{A}\mathbf{X}_{\text{sv}}\|_F^2 + \mu_{\text{sv}}\|\mathbf{x}_{\text{sv}}^{(\ell_2)}\|_1,$$

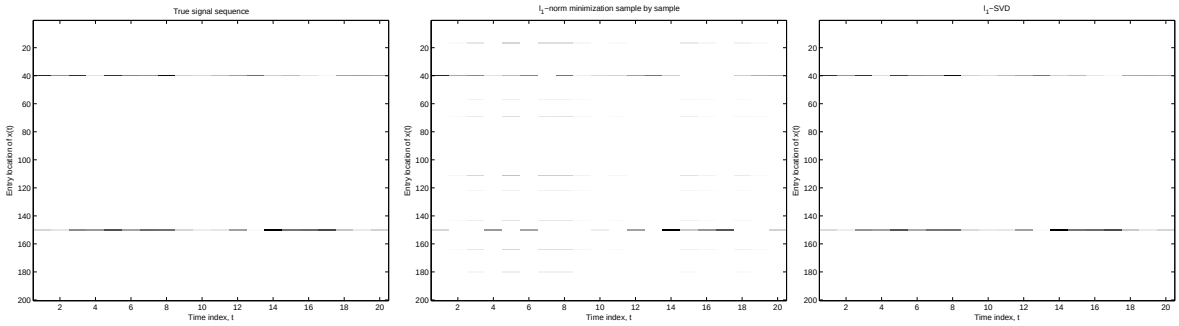
where $\mathbf{x}_{\text{sv}}^{(\ell_2)} = [x_{1,\text{sv}}^{(\ell_2)}, \ldots, x_{M,\text{sv}}^{(\ell_2)}]^T$, and the parameter μ_{sv} controls the spatial sparsity. In Fig 2.1c, we can see that ℓ_1-SVD performance better than the sparse recovery method based on ℓ_1-minimization separately. Since ℓ_1-SVD enforces the row-sparse structure while separate-based method only enforce the sparse elements themselves.

2.4 Block sparse recovery

In many scenarios, the non-zero elements of a sparse signal appear as "clusters", i.e., non-zero entries appear in cluster or in blocks. Estimators taking into account such underlying structures can often perform better than the conventional ℓ_1-norm regularized least-squares algorithm.

We firstly introduce the concept of block sparsity. Let us consider a vector $\mathbf{x}$ that can be divided into N sub-blocks, $\mathbf{x} = [\mathbf{x}_1^T, \mathbf{x}_2^T, \ldots, \mathbf{x}_N^T]^T$, out of which only K sub-blocks have non-zero values, then we say that vector $\mathbf{x}$ is a block sparse vector. Define d_n is a set that contains all the indices of the entries of $\mathbf{x}$ that correspond to $\mathbf{x}_n$ and define the set $\mathcal{I} = \{d_1, \ldots, d_N\}$, e.g. suppose $\mathbf{x}$ is of 5 that can be divided into 2 sub-blocks, $\mathbf{x}_1$ and $\mathbf{x}_2$,

$$\mathbf{x} = [\overbrace{x_1, x_2}^{\mathbf{x}_1^T}, \overbrace{x_3, x_4, x_5}^{\mathbf{x}_2^T}]^T$$



(a) (b) (c)

Figure 2.1: Multiple snapshot sparse recovery. $\mathbf{x}(t)$ is a sparse vector and 40-th and 150-th elements are non-zero. SNR = 20 dB and $T = 20$. (a) True signal sequences. (b) ℓ_1-minimization sample by sample. (c) ℓ_1-SVD algorithm.

then $d_1 = \{1, 2\}$ and $d_2 = \{3, 4, 5\}$, and we can say $\mathbf{x}$ is divided via $\mathcal{I} = \{d_1, d_2\}$. Mathematically, a K-block sparse vector $\mathbf{x}$ can be represented as

$$\|x\|_{0,\mathcal{I}} \leq K$$

where

$$\|\mathbf{x}\|_{0,\mathcal{I}} = \sum_{n=1}^{N} \mathbf{I}(\mathbf{x}_n)$$

where $\mathbf{I}(\mathbf{x}_n)$ is an indicator function, i.e. $\mathbf{I}(\mathbf{x}_n) = 1$ if $\mathbf{x}_n$ with non-zero value otherwise it is 0. An example of $\mathbf{I}(\mathbf{x}_n)$ is that $\mathbf{I}(\mathbf{x}_n) = \|\|\mathbf{x}_n\|_2\|_0$.

The block sparse recovery is to reconstruct a K-block sparse vector through measurement vector $\mathbf{y}$ acquired via a linear equation $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w}$. One approach to block sparse recovery is to enforce the block sparsity of $\mathbf{x}$ via a mixed ℓ_2/ℓ_1 minimization. We define the mixed ℓ_2/ℓ_1 norm over the index set $\mathcal{I} = \{d_1, d_2, \ldots, d_N\}$ as

$$\|\mathbf{x}\|_{2,1} = \sum_{n=1}^{N} \|\mathbf{x}_n\|_2.$$

The suggested algorithm is given as

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbb{C}} \quad \|\mathbf{x}\|_{2,1} \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \epsilon, \tag{2.7}$$

which it is equivalent to

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbb{C}^{N\times 1}} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 + \mu\|\mathbf{x}\|_{2,1}. \tag{2.8}$$

To discuss the uniqueness and stability of the block sparse recovery, we would like to introduce the definition of block restricted isometry property [21, 22], which can be view as a extension to the standard RIP. The definition of B-RIP indicates that there exists a unique K-block sparse solution if the sensing matrix $\mathbf{A}$ satisfies the B-RIP of order $2K$. However, we should note that a K-block sparse vector $\mathbf{x}$ is not

necessarily a K-sparse vector. It would have more than K entries with non-zero values. Note that a K-block sparse vector $\mathbf{x}$ is a L-sparse vector in the conventional sense, where L is the sum of the K sub-blocks element numbers (e.g. suppose each sub-block contains 3 elements, therefore, $L = 3K$). Fig. 2.2 shows the result of reconstruction of a block-sparse vector $\mathbf{x}$ by using standard sparse recovery and block sparse recovery. The original sequence $\mathbf{x}$ is a 5-sparse vector but also can be viewed as 1-block sparse vector. The standard sparse recovery fails to recovery $\mathbf{x}$, because the prior knowledge it uses that $\mathbf{x}$ is 5-sparse is not sufficient to determine $\mathbf{x}$ (the 5 non-zero element can be anywhere within the vector not appearing in cluster) and $\sigma_{10} > 1$. While there is a unique block-sparse vector to the problem ($\sigma_{2|\mathcal{I}} < 1$).

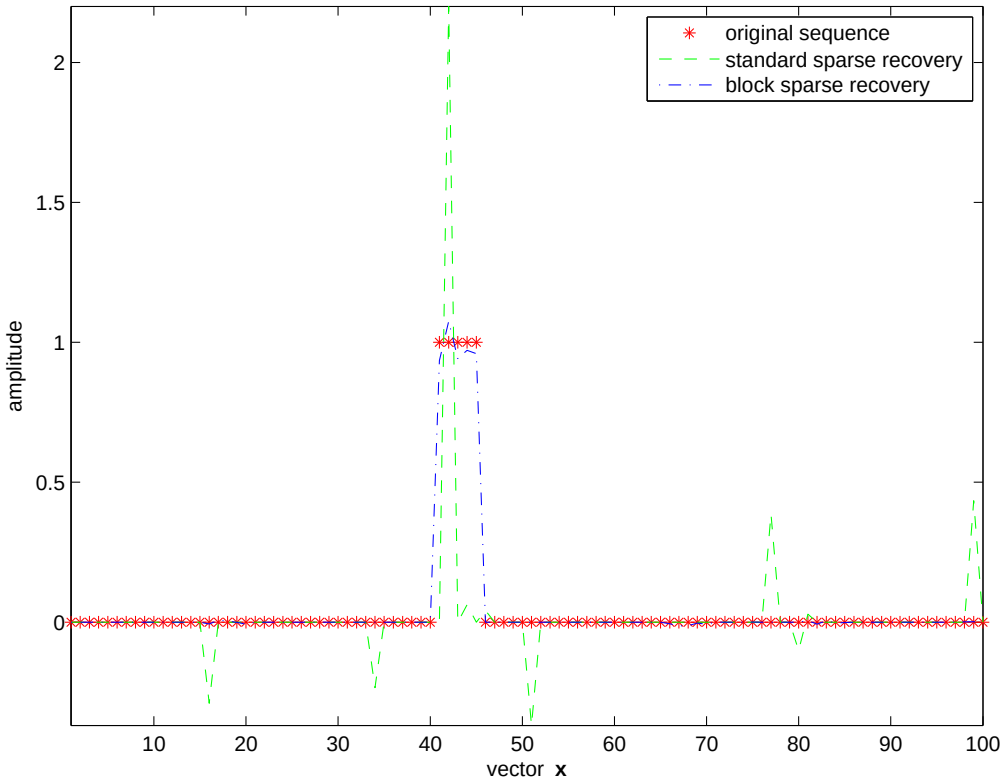


Figure 2.2: Comparison of block sparse recovery with traditional sparse recovery. SNR = 30 dB.

Signal Model

3

In this chapter, the model we use thoughout this thesis will be present. In Fig. 3.1, an illustration of the geometry of source localization: sources $s_k(t)$ impinging on the uniform linear array from location (θ_k, r_k) producing observation outputs $y_m(t)$.

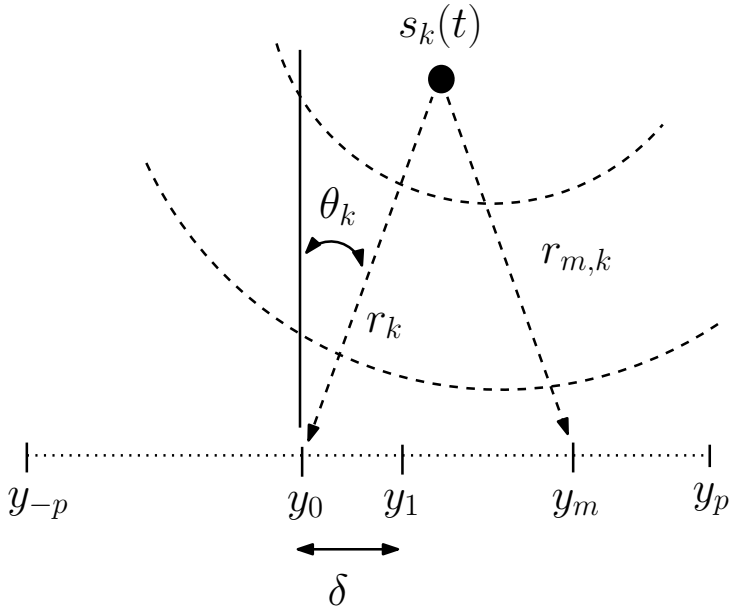


Figure 3.1: A linear array receiving a signal from a near-field point source.

3.1 Spherical wavefront model

Consider K narrowband sources present in the near field impinging on a array of $M = 2p + 1$ sensors as illustrated in Fig. 3.1. Without loss of generality, it is assumed that the phase reference of the array is at the origin, and the sensors are placed at location indices in the range $[-p, p]$. Denoting the spacing between two adjacent sensors as δ, the position of the m-th sensor will be $m\delta$ where $m \in [-p, p]$. Since the sources are present in the near-field region of the ULA, the curvature of the wavefront is no longer negligible. Therefore, the signal received by the m-th sensor at time t can be expressed as

$$y_m(t) = \sum_{k=1}^{K} s_k(t) \exp(j\frac{2\pi}{\lambda}(r_{m,k} - r_k)) + w_m(t), \tag{3.1}$$

where

$$r_{m,k} = \sqrt{r_k + m^2\delta^2 - 2m\delta r_k \sin(\theta_k)} \tag{3.2}$$

represents the distance between the m-th sensor and the k-th source, r_k is the range from the k-th source to the phase reference, $s_k(t)$ is the signal radiated by the k-th source characterized the DOA-range pair (θ_k, r_k), λ denotes the wavelength, and $w_m(t)$ denotes the additive noise.

Stacking the measurements in $\mathbf{y}(t) = [y_{-p}(t), \ldots, y_p(t)]^T \in \mathbb{C}^{M\times 1}$, we get

$$\mathbf{y}(t) = \sum_{k=1}^{K} \mathbf{a}(\theta_k, r_k) s_k(t) + \mathbf{w}(t), \quad \text{for} \quad t = t_1, t_2, \ldots, t_T, \tag{3.3}$$

where t_T denotes the number of snapshots, $\mathbf{a}(\theta_k, r_k) \in \mathbb{C}^{M\times 1}$ is the so-called *steering vector*, and $\mathbf{w}(t) = [w_{-p}(t), \ldots, w_p(t)]^T \in \mathbb{C}^{M\times 1}$ is the noise vector. The non-linear measurement model in (3.3) can be concisely written as

$$\mathbf{y}(t) = \mathbf{A}(\boldsymbol{\theta}, \mathbf{r})\mathbf{s}(t) + \mathbf{w}(t), \tag{3.4}$$

where $\mathbf{A}(\boldsymbol{\theta}, \mathbf{r}) = [\mathbf{a}(\theta_1, r_1), \mathbf{a}(\theta_2, r_2), \ldots, \mathbf{a}(\theta_K, r_K)] \in \mathbb{C}^{M\times K}$ is the array manifold matrix, and $\mathbf{s}(t) = [s_1(t), \ldots, s_K(t)]^T \in \mathbb{C}^{K\times 1}$ is the source vector.

3.2 Fresnel approximation

Using the Taylor series expansion of (3.2), and approximating this up to the second order, we get the so-called *Fresnel* approximation, which is given by

$$r_{m,k} \approx r_k - m\delta \sin\theta_k + m^2\delta^2 \frac{\cos^2\theta_k}{2r_k}.$$

We can now approximate $\tau_{m,k} = \frac{2\pi}{\lambda}(r_{m,k} - r_k)$ as

$$\begin{aligned} \tau_{m,k} &\approx -m\frac{2\pi\delta}{\lambda}\sin(\theta_k) + m^2\frac{\pi\delta^2}{\lambda r_k}\cos^2(\theta_k) \\ &= m\omega_k + m^2\phi_k \end{aligned} \tag{3.5}$$

where we re-parameterize the DOA and range, respectively as

$$\omega_k = -\frac{2\pi\delta}{\lambda}\sin(\theta_k) \quad \text{and} \quad \phi_k = \frac{\pi\delta^2}{\lambda r_k}\cos^2(\theta_k). \tag{3.6}$$

Using the approximation for $\tau_{m,k}$ in (3.1), we get

$$y_m(t) \approx \sum_{k=1}^{K} s_k(t) e^{j(m\omega_k + m^2\phi_k)} + w_m(t). \tag{3.7}$$

Stacking the measurements from all the M sensors, we get

$$\mathbf{y}(t) = \sum_{k=1}^{K} \tilde{\mathbf{a}}(\omega_k, \phi_k) s_k(t) + \mathbf{w}(t), \quad \text{for} \quad t = t_1, t_2, \ldots, t_T, \tag{3.8}$$

where $\tilde{\mathbf{a}}(\omega_k, \phi_k) = [e^{-jp\omega_k} e^{jp^2\phi_k}, \ldots, 1, \ldots, e^{jp\omega_k} e^{jp^2\phi_k}]^T \in \mathbb{C}^{M\times 1}$ is the *modified steering vector*. The output of the ULA can now be written as the following non-linear measurement model

$$\mathbf{y}(t) = \tilde{\mathbf{A}}(\boldsymbol{\omega}, \boldsymbol{\phi})\mathbf{s}(t) + \mathbf{w}(t) \tag{3.9}$$

where $\tilde{\mathbf{A}}(\boldsymbol{\omega}, \boldsymbol{\phi}) = [\tilde{\mathbf{a}}(\omega_1, \phi_1), \tilde{\mathbf{a}}(\omega_2, \phi_2), \ldots, \tilde{\mathbf{a}}(\omega_K, \phi_K)] \in \mathbb{C}^{M\times K}$ is the array manifold, and $\mathbf{s}(t)$ and $\mathbf{w}(t)$ are the source and noise vectors, respectively.

Sparse recovery techniques for near-field source localization: Grid-based model

4

In this chapter, we localize multiple narrowband near-field sources by jointly estimating the DOA and range. Using the sparse representation framework, we form a overcomplete basis constructed using a sampling grid that is related to the possible source locations. By doing so, the original non-linear parameter estimation problem is transformed into a linear ill-posed problem. Assuming the spatial spectrum is sparse, we can localize the sources with super-resolution by solving the well-known ℓ_1-regularized least-squares optimization problem. When multiple time samples are available, we make use of the Fresnel approximation and further assume that the source signals are mutually uncorrelated, which naturally decouples the DOA and range in the correlation domain. This allows us to significantly reduce the complexity, by solving two inverse problems of smaller dimensions one by one, instead of one inverse problem of a higher dimension.

4.1 Grid-based model

4.1.1 Gridding on spherical wavefront model

We provide a framework for localizing multiple near-field sources based on sparse reconstruction techniques. More specifically, we aim to jointly estimate the DOA $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_K]^T$ and the range $\mathbf{r} = [r_1, r_2, \ldots, r_K]^T$. In this section, a single snapshot case is considered, with $T = 1$ in (3.4). The problem in (3.4) as it appears is a non-linear parameter estimation problem, where the matrix $\mathbf{A}(\boldsymbol{\theta}, \mathbf{r})$ depends on the unknown source locations $(\boldsymbol{\theta}, \mathbf{r})$. In order to jointly estimate both the DOA and range using the data model in (3.4), a multi-dimensional non-linear optimization over both $\boldsymbol{\theta}$ and $\mathbf{r}$ is required. This optimization problem is clearly computationally intractable.

Suppose all possible source locations reside in the domain $\theta_k \in [\theta_{\min}, \theta_{\max}]$ and $r_k \in [r_{\min}, r_{\max}]$ for all $k = 1, \ldots, K$. We can then cast the joint DOA-range estimation problem as a sparse reconstruction problem, where we discretize the θ-interval into N_θ and r-interval into N_r bins of resolution $\Delta\theta$ and Δr, respectively. This discretization results in an overcomplete representation of $\mathbf{A}$ in terms of the sampling grid $(\bar{\boldsymbol{\theta}}, \bar{\mathbf{r}})$ that includes all the source locations of interest with $\bar{\boldsymbol{\theta}} = [\bar{\theta}_1, \bar{\theta}_2, \ldots, \bar{\theta}_{N_\theta}]^T$ and $\bar{\mathbf{r}} = [\bar{r}_1, \bar{r}_2, \ldots, \bar{r}_{N_r}]^T$. We construct a matrix with steering vectors corresponding to each potential source location as its columns:

$$\mathbf{A}(\bar{\boldsymbol{\theta}}, \bar{\mathbf{r}}) = [\mathbf{a}(\bar{\theta}_1, \bar{r}_1), \mathbf{a}(\bar{\theta}_1, \bar{r}_2), \ldots, \mathbf{a}(\bar{\theta}_N, \bar{r}_N)] \in \mathbb{C}^{M \times N},$$

where $N = N_\theta N_r$. The matrix $\mathbf{A}$ is now known, and does not depend on the unknown variables $(\boldsymbol{\theta}, \mathbf{r})$. Note that the number of potential source locations N will typically be much greater than the number of sources K or even the number of sensors M.

The signal is now represented by an $N \times 1$ vector $\mathbf{x}(t)$, where every source can be found as a non-zero weight $x_n(t) = s_k(t)$ if source k comes from $(\bar{\theta}_n, \bar{r}_n)$ for some k and is zero otherwise, i.e., the dominant peaks in $\mathbf{x}(t)$ correspond to the true source locations. The discrete grid-based model for a single snapshot is given by

$$\begin{aligned} \mathbf{y} &= \sum_{n=1}^{N} \mathbf{a}(\bar{\theta}_n, \bar{r}_n) x_n + \mathbf{w} \\ &= \mathbf{A}(\bar{\boldsymbol{\theta}}_n, \bar{\mathbf{r}}_n)\mathbf{x} + \mathbf{w}. \end{aligned} \tag{4.1}$$

This model allows us to transform the non-linear parameter estimation problem into a sparse recovery problem based on the central assumption that the vector $\mathbf{x}$ is sparse. An ideal measure for the sparsity of $\mathbf{x}$ is its ℓ_0(-quasi) norm $\|\mathbf{x}\|_0$, and mathematically we must solve for $\arg\min \|\mathbf{x}\|_0$ subject to $\|\mathbf{y} - \mathbf{A}(\bar{\boldsymbol{\theta}}_n, \bar{\mathbf{r}}_n)\|_2^2 \leq \varepsilon$, where the parameter ε controls how much noise we wish to allow. However, this is a mathematically intractable combinatorial problem even for modestly sized problems. Hence, to simplify this problem we use an ℓ_1-norm regularization, which is the traditional best convex surrogate of the ℓ_0(-quasi) norm. The inverse problem can be solved using an ℓ_1-regularized least-squares (LS) methodology which is given by

$$\hat{\mathbf{x}} = \underset{\mathbf{x} \in \mathbb{C}^{N \times 1}}{\arg\min} \quad \|\mathbf{y} - \mathbf{A}(\bar{\boldsymbol{\theta}}, \bar{\mathbf{r}})\mathbf{x}\|_2^2 + \mu \|\mathbf{x}\|_1, \tag{4.2}$$

where μ is the sparsity regulating parameter. This optimization problem can be solved using any of the popular solvers available off-the-shelf (e.g., iterative thresholding, matching pursuit).

4.1.2 Gridding on Fresnel approximation model

Similar as the gridding on the spherical wavefront model, we can construct an overcomplete representation also for $\tilde{\mathbf{A}}$ using the sampling grid $(\bar{\boldsymbol{\theta}}, \bar{\mathbf{r}})$ that includes all possible source locations. This discretization results in a known matrix with steering vectors corresponding to each potential source location as its columns:

$$\tilde{\mathbf{A}}(\bar{\boldsymbol{\omega}}, \bar{\boldsymbol{\phi}}) = [\tilde{\mathbf{a}}(\bar{\omega}_1, \bar{\phi}_1), \tilde{\mathbf{a}}(\bar{\omega}_1, \bar{\phi}_2), \ldots, \tilde{\mathbf{a}}(\bar{\omega}_N, \bar{\phi}_N)] \in \mathbb{C}^{M \times N},$$

where $\bar{\omega}_n = -\frac{2\pi\delta}{\lambda} \sin(\bar{\theta}_n)$ and $\bar{\phi}_n = \frac{\pi\delta^2}{\lambda \bar{r}_n} \cos^2(\bar{\theta}_n)$ for all $n \in \{1, \ldots, N\}$. The discrete grid-based model is finally given by

$$\mathbf{y}(t) = \tilde{\mathbf{A}}(\bar{\boldsymbol{\omega}}, \bar{\boldsymbol{\phi}})\mathbf{x}(t) + \mathbf{w}(t), \tag{4.3}$$

and the corresponding inverse problem for a single snapshot can be solved using an ℓ_1-regularized least-squares optimization problem as earlier. Solving the near-field localization problem using sparse regression with or without Fresnel approximation for a single snapshot incurs the same complexity. Moreover, this approximation can even deteriorate the range estimation (for more details see [9]). However, when multiple snapshots are available, the structure of the Fresnel approximated array manifold matrix allows us to significantly reduce the computational complexity. This is discussed in the next section.

4.2 Two-step estimator with multiple snapshots

When there are multiple measurements available, we can stack (4.1) for the batch of T measurements into a matrix

$$\mathbf{Y} = \mathbf{A}(\bar{\boldsymbol{\theta}}, \bar{\mathbf{r}})\mathbf{X} + \mathbf{W} \tag{4.4}$$

where $\mathbf{Y} = [\mathbf{y}(t_1), \ldots, \mathbf{y}(t_T)] \in \mathbb{C}^{M\times T}$, and matrices $\mathbf{X}$ and $\mathbf{W}$ are defined similarly. An important point to be noted here is that the matrix $\mathbf{X}$ is sparse only spatially, and is generally not sparse in time. A straightforward approach would be to use a joint sparsity promoting ℓ_2/ℓ_1-norm regularization, or an ℓ_1-SVD [14] algorithm to solve the inverse problem. In this paper, we propose to reduce the involved computational complexity of 2D-gridding by solving an inverse problem of smaller dimensions in two-steps. We do this by exploiting the spatial cross-correlation between the symmetric sensors, and the fact that the structure of the Fresnel approximated model naturally decouples the DOA and range.

We now make the following assumptions[1]:

(a1) The source signals are mutually independent and are modeled as independent identically distributed (i.i.d.) complex circular random variables with zero mean and covariance matrix $\mathbb{E}_t\{\mathbf{s}(t)\mathbf{s}^H(t)\} = \mathrm{diag}(\sigma^2_{s,1}, \ldots, \sigma^2_{s,K})$.

(a2) The noise is modeled as a zero-mean spatially white Gaussian process, and it is independent of the source signals. The noise covariance matrix is given by $\mathbb{E}_t\{\mathbf{w}(t)\mathbf{w}^H(t)\} = \sigma^2_w\mathbf{I}$.

Under assumptions (a1) and (a2), the spatial correlation between the m-th and n-th sensor can be written as

$$\begin{aligned} r_y(m,n) &= E_t\{y_m(t)y^*_n(t)\} \\ &= \sum_{k=1}^{K} \sigma^2_{s,k} e^{j(m-n)\omega_k + j(m^2-n^2)\phi_k} + \sigma^2_w\delta(m-n) \end{aligned}$$

where $\delta(.)$ represents the Dirac function, $\mathbb{E}_t\{s_k(t)s^*_k(t)\} = \sigma^2_{s,k}$ denotes the signal power of the k-th source, and σ^2_w is the noise variance. Notice that when $n = -m$ the spatial correlation is independent of the parameter ϕ_k, and we arrive at

$$\begin{aligned} r_y(-m,m) &= E_t\{y_{-m}(t)y^*_m(t)\} \\ &= \sum_{k=1}^{K} \sigma^2_{s,k} e^{-2m\omega_k j} + \sigma^2_w\delta(-2m). \end{aligned} \tag{4.5}$$

This means that by exploiting the cross-correlation between the symmetric sensors, we can transform the original 2D (DOA and range) estimation into a 1D (DOA) estimation. Stacking (4.5) for all the symmetric sensors, we can build a *virtual far-field* scenario:

$$\mathbf{r}_y = \mathbf{A}_\omega(\boldsymbol{\omega})\mathbf{r}_\mathbf{s} + \sigma^2_w\mathbf{e}, \tag{4.6}$$

[1]These assumptions are required only for the two-step estimator discussed in Section 4.2.

where $\mathbf{r}_y = [r_y(0,0), r_y(1,-1), \ldots, r_y(p,-p)]^T \in \mathbb{C}^{L\times 1}$ with $L = p+1$, $\mathbf{r}_s = \text{diag}(\sigma^2_{s,1}, \ldots, \sigma^2_{s,K}) \in \mathbb{C}^{K\times 1}$, and $\mathbf{e} = [1, \mathbf{0}_p^T]^T \in \mathbb{C}^{L\times 1}$, and the corresponding virtual array gain pattern for the k-th source denoted by $\mathbf{a}_\omega(\omega_k)$ can be expressed as $\mathbf{a}_\omega(\omega_k) = [1, e^{-j2\omega_k}, \ldots, e^{-j2p\omega_k}]^T \in \mathbb{C}^{L\times 1}$, with the array manifold

$$\mathbf{A}_\omega(\boldsymbol{\omega}) = [\mathbf{a}_\omega(\omega_1), \mathbf{a}_\omega(\omega_2), \ldots, \mathbf{a}_\omega(\omega_K)] \in \mathbb{C}^{L\times K}.$$

In practice, the vector $\mathbf{r}_y$ containing the statistical correlations is approximated using the measurements from (4.4).

4.2.1 Step-1: DOA estimation

As before, we can construct an overcomplete basis $\mathbf{A}_\omega$ but now with only N_θ columns corresponding to potential source directions of arrival (DOAs) using the sampling grid $\bar{\boldsymbol{\theta}}$, i.e.,

$$\mathbf{A}_\omega(\bar{\boldsymbol{\omega}}) = [\mathbf{a}_\omega(\bar{\omega}_1), \ldots, \mathbf{a}_\omega(\bar{\omega}_{N_\theta})] \in \mathbb{C}^{M\times N_\theta},$$

where $\bar{\omega}_n = -\frac{2\pi\delta}{\lambda}\sin(\bar{\theta}_n)$ for all $n \in \{1, \ldots, N_\theta\}$ as defined earlier. The signal is represented by an $N_\theta \times 1$ vector $\mathbf{u}(t)$, where every source can be found as a non-zero weight $u_n(t) = s_k(t)$ if source k comes from direction $\bar{\theta}_n$ for some k and is zero otherwise, i.e., the dominant peaks in $\mathbf{u}(t)$ correspond to the true source locations. The discrete grid-based model in the correlation domain is then given by

$$\mathbf{r}_y = \mathbf{A}_\omega(\bar{\boldsymbol{\omega}})\mathbf{u} + \sigma_w^2\mathbf{e}. \tag{4.7}$$

Note that the number of potential source DOAs N_θ will typically be much greater than the number of sensors M also in the correlation domain, and the model in (4.7) is still ill-posed. Hence, we solve for the unknown vector $\mathbf{u}$ using an ℓ_1-regularized LS minimization problem which is given by

$$\hat{\mathbf{u}} = \arg\min_{\mathbf{u}} \quad \|\mathbf{r}_y - \mathbf{A}_\omega(\bar{\boldsymbol{\omega}})\mathbf{u}\|_2^2 + \mu_1\|\mathbf{u}\|_1, \tag{4.8}$$

where μ_1 is the sparsity regulating parameter. Alternatively, when the noise variance σ_w^2 is known, the unknown vector $\mathbf{u}$ can be obtained by solving

$$\arg\min_{\mathbf{u}} \quad \|\mathbf{r}_y - \mathbf{A}_\omega(\bar{\boldsymbol{\omega}})\mathbf{u} - \sigma_w^2\mathbf{e}\|_2^2 + \mu_1\|\mathbf{u}\|_1.$$

4.2.2 Step-2: range estimation

Let $\hat{\boldsymbol{\theta}}$ be the estimated DOAs from step-1, and $\hat{K}$ denote the number of DOAs detected (i.e., $\hat{K} = \|\hat{\boldsymbol{u}}\|_0$). We now use the sampling grid $(\hat{\boldsymbol{\theta}}, \bar{\mathbf{r}})$ to form an overcomplete basis $\mathbf{A}(\hat{\boldsymbol{\theta}}, \bar{\mathbf{r}}) \in \mathbb{C}^{M\times \hat{K}N_r}$ to arrive at

$$\mathbf{Y} = \mathbf{A}(\hat{\boldsymbol{\theta}}, \bar{\mathbf{r}})\tilde{\mathbf{X}} + \mathbf{W}, \tag{4.9}$$

where $\tilde{\mathbf{X}}$ is obtained by removing some specific rows of the signal matrix $\mathbf{X}$. In order to solve the inverse problem in (4.9) we use the ℓ_1-SVD algorithm. Note that in step-2

for range estimation we do not use the Fresnel approximation anymore. For the sake of completeness, the ℓ_1-SVD algorithm in [14] is briefly summarized as follows.

Let $\mathbf{Y} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ be the singular value decomposition (SVD) of the data matrix. Keep a reduced $M \times \hat{K}$ matrix $\mathbf{Y}_{\text{sv}} = \mathbf{U}\mathbf{\Sigma}\mathbf{D}_k = \mathbf{Y}\mathbf{V}\mathbf{D}_k$, where $\mathbf{D}_k = [\mathbf{I}_k, \mathbf{0}^T_{\hat{K}\times(T-\hat{K})}]$. The reduced data matrix contains most of the signal power, and forms the basis for the signal subspace. Similarly, let $\tilde{\mathbf{X}}_{\text{sv}} = \tilde{\mathbf{X}}\mathbf{V}\mathbf{D}_k$ and $\mathbf{W}_{\text{sv}} = \mathbf{W}\mathbf{V}\mathbf{D}_k$, to arrive at

$$\mathbf{Y}_{\text{sv}} = \mathbf{A}(\hat{\boldsymbol{\theta}}, \bar{\mathbf{r}})\tilde{\mathbf{X}}_{\text{sv}} + \mathbf{W}_{\text{sv}}, \tag{4.10}$$

which can be expressed in vector form (column by column) as

$$\mathbf{y}_{\text{sv}}(k) = \mathbf{A}(\hat{\boldsymbol{\theta}}, \bar{\mathbf{r}})\tilde{\mathbf{x}}_{\text{sv}}(k) + \mathbf{w}_{\text{sv}}(k), \text{ for } k = 1, \ldots, \hat{K}.$$

Here, each column corresponds to a signal subspace singular vector. The reduced data matrix is only spatially sparse, and not in terms of the singular vector index k. In order to take this effect into account, we use a different prior obtained by computing the ℓ_2-norm of the singular values of a particular spatial index of $\tilde{\mathbf{x}}_{\text{sv}}(k)$, i.e., $\tilde{x}^{(\ell_2)}_{m,\text{sv}} = \sqrt{\sum_{k=1}^{K}(\tilde{x}_{m,\text{sv}}(k))^2}$, for $m \in [-p, p]$. Note that the ℓ_2-norm can be computed for all snapshots instead of only the signal subspace singular vectors, however, the former technique adds more computational complexity especially when $T \gg \hat{K}$ [14]. Now, we can find the range by minimizing

$$\|\mathbf{Y}_{\text{sv}} - \mathbf{A}(\hat{\boldsymbol{\theta}}, \bar{\mathbf{r}})\tilde{\mathbf{X}}_{\text{sv}}\|_F^2 + \mu_{\text{sv}}\|\tilde{\mathbf{x}}^{(\ell_2)}_{\text{sv}}\|_1,$$

where $\tilde{\mathbf{x}}^{(\ell_2)}_{\text{sv}} = [\tilde{x}^{(\ell_2)}_{-p,\text{sv}}, \ldots, \tilde{x}^{(\ell_2)}_{p,\text{sv}}]^T$, and the parameter μ_{sv} controls the spatial sparsity.

Remark 1 (Complexity reduction with multiple snapshots)**.** *Jointly estimating the DOA and range by applying the ℓ_1-SVD algorithm on the model* (4.4) *costs $O((KN_\theta N_r)^3)$. Using the proposed two-step estimator, the complexity is reduced significantly to $O(KN_r^3)$ (reduction by a factor of $O(N_\theta^3)$) with an additional complexity of solving the inverse problem in* (4.8)*, which costs for example, $O(N_\theta \log(N_\theta))$ using the iterative thresholding algorithm [25]. For a typical problem with $N_r = 15$ and $N_\theta = 180$ points on the grid, the complexity reduction is significant.*

Remark 2 (Array geometry)**.** *Any array (uniform or non-uniform) can be used to solve for the variables $(\boldsymbol{\theta}, \mathbf{r})$ based on the optimization problem in* (4.2)*. For the two-step approach, any symmetric array (uniform or non-uniform) can be used.*

4.3 Simulations

We consider a symmetric ULA with $M = 15$ sensors placed such that the inter-sensor spacing $\delta = \lambda/4$ (To avoid aliasing when the virtual far-field model (4.7) in is used), where λ represents the wavelength of the narrowband source signals. For this array, the far-field distance is beyond $2D^2 = 24.5\lambda$, and any sources within the range of 24.5λ from the array will be in the near field. We compare the proposed algorithms with matched-filter beamforming [3] for a single time sample scenario. For the multiple time

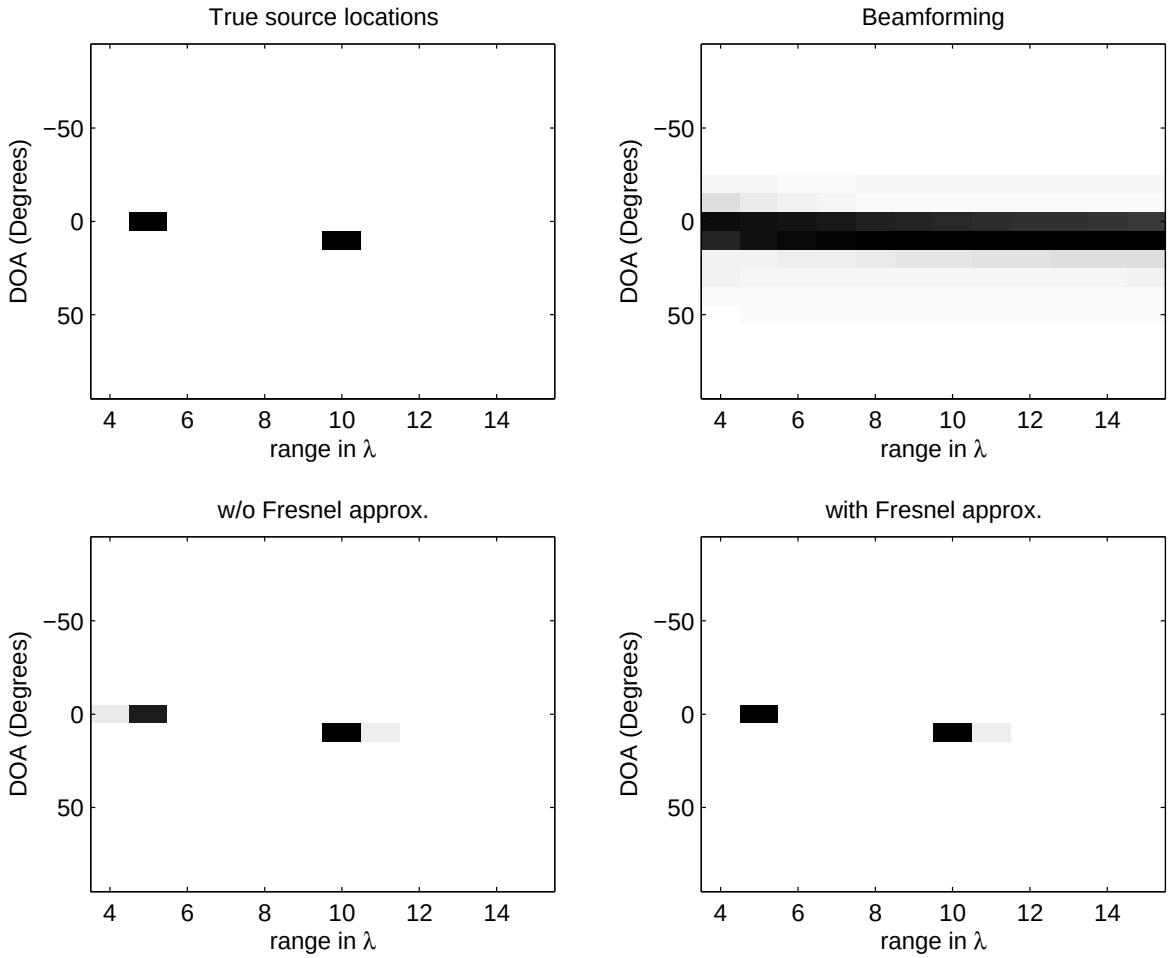


Figure 4.1: Joint DOA and range estimation. Two near-field sources with DOAs: 0° and 10°, and ranges: 5λ and 10λ. SNR $= 20$ dB, and $T = 1$. The sampling grid has a resolution of $\Delta\theta = 10^\circ$ and $\Delta r = 1\lambda$.

sample scenario we compare the performance with the beamforming method [3] and the 2D-MUSIC algorithm [7]. The optimization problems in the proposed algorithms are solved using CVX [26]. The regularization parameter is chosen via cross-validation.

In Fig. 4.1, we illustrate the joint DOA and range estimation obtained by solving the sparse regression problem for a single time sample scenario. We consider two sources at locations $(0^\circ, 5\lambda)$ and $(10^\circ, 10\lambda)$. The SNR is 20 dB with $T = 1$. As can be seen from the plots, high resolution can be achieved by using the sparse modeling framework as compared to the conventional beamforming technique. The performance loss with and without (w/o) Fresnel approximation for the considered scenario is negligible.

In Fig. 4.2, we show the proposed two-step estimator for near-field source localization. We consider four sources are locations $(0^\circ, 5\lambda)$, $(0^\circ, 17\lambda)$ and $(50^\circ, 8\lambda)$. The simulations are provided for SNRs of 10 dB and 30 dB with $T = 200$ snapshots.In the two-step estimator, we solve for the DOAs in the correlation domain in which the range parameters are naturally decoupled from the DOAs (due to the Fresnel approximation).

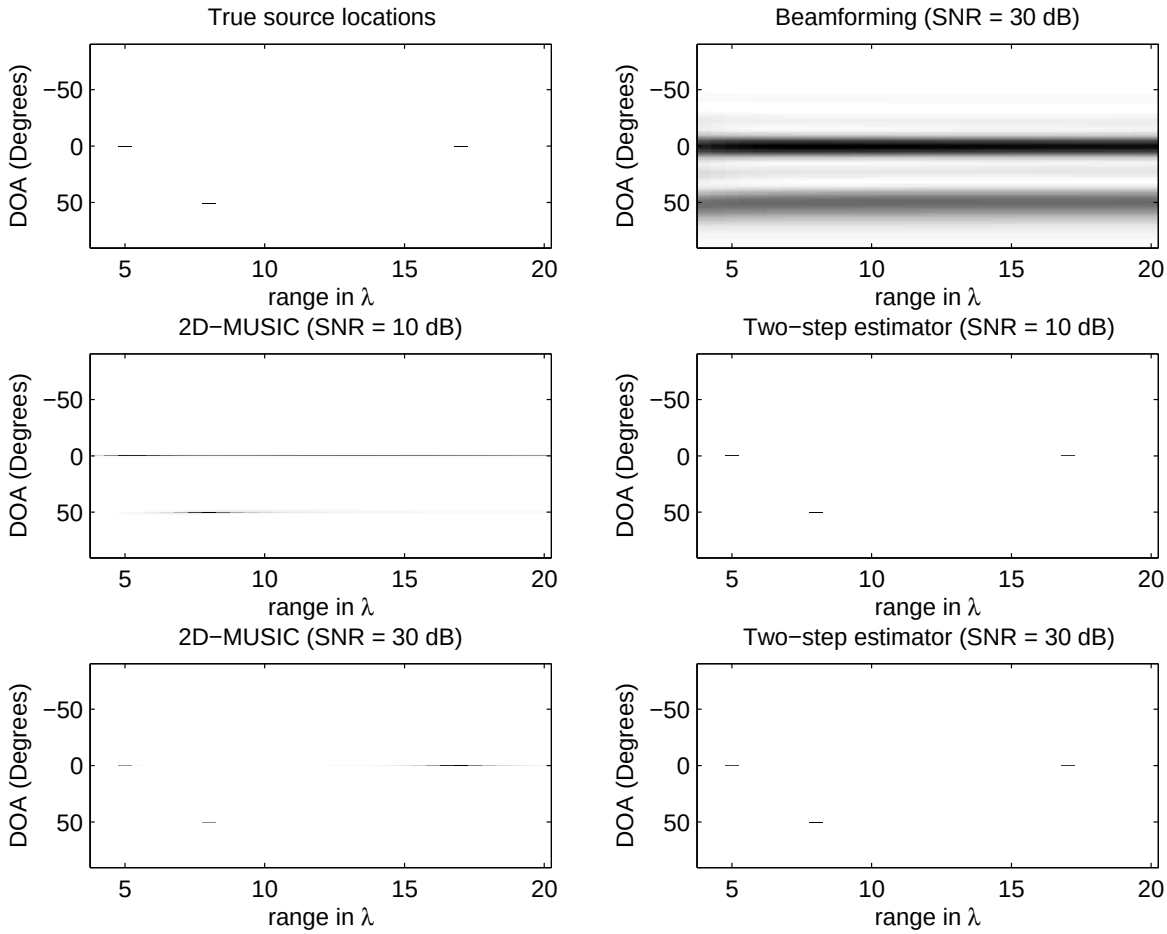


Figure 4.2: Two-step estimator. Three near-field sources at locations: $(0°, 5\lambda)$, $(0°, 17\lambda)$, and $(50°, 8\lambda)$, and $T = 200$. The sampling grid has a resolution of $\Delta\theta = 1°$ and $\Delta r = 0.5\lambda$.

In the second step, we solve for the range using the ℓ_1-SVD algorithm where we use the number of sources as $\hat{K}$ from the step-1. Even though, 2D-MUSIC can achieve high resolution in the DOA domain, its performance deteriorates along the range comain especially when two sources shares the same DOA. The two-step approach would further allow to increase the gridding resolution because of the involved smaller overcomplete dictionary as compared to the dictionary obtained from a 2D grid (cf. (4.1))

Grid matching for near-field source localization using sparse recovery techniques

5

In this chapter, we discuss the grid mismatch effect associated with near-field source localization using sparse recovery techniques. The grid based model presented is the previous chapter is extended to an error-in-variables (EIV) model to account for cases when the true source locations are not on the assumed sampling grid. We propose a number of estimators for correcting such perturbations when both single and multiple time snapshots are available.

5.1 Grid mismatch and its effect

Source localization based on the sparse signal reconstruction (SSR) can provide a high resolution compared with the classical source localization methods, but it should be noted that its performance is tremendously affected by the choice of the sampling grid. In Fig. 5.1, we illustrate a grid mismatch scenario and show how it influences the estimation result, , where the estimator is based on the sparse recovery technique discussed in Chapter 4.

Consider K narrowband sources present in near field impinging on a array comprising of $M = 2p + 1$ sensors. We recall the signal model described in Chapter 3

$$\mathbf{y}(t) = \mathbf{A}(\boldsymbol{\theta}, \mathbf{r})\mathbf{s}(t) + \mathbf{w}(t),$$

where $(\boldsymbol{\theta}, \mathbf{r}) = [(\theta_1, r_1), (\theta_2, r_2), \ldots, (\theta_K, r_K)]$ indicates the unknown source location. Suppose all the possible source location reside in the domain $\theta_k \in [\theta_{\min}, \theta_{\max}]$ and $r_k \in [r_{\min}, r_{\max}]$ for all $k = 1, \ldots, K$. We showed in Chapter 4 that we can discretize the DOA-range domain into $N = N_\theta N_r$ bins to obtain the grid-based model (4.1)

$$\begin{aligned}\mathbf{y} &= \sum_{n=1}^{N} \mathbf{a}(\bar{\theta}_n, \bar{r}_n)x_n + \mathbf{w} \\ &= \mathbf{A}(\bar{\boldsymbol{\theta}}, \bar{\mathbf{r}})\mathbf{x} + \mathbf{w},\end{aligned}$$

where $(\bar{\boldsymbol{\theta}}, \bar{\mathbf{r}}) = [(\bar{\theta}_1, \bar{r}_1), \ldots, (\bar{\theta}_N, \bar{r}_N)]$ denotes the sampling grid.

In the grid-based estimation, we assume that the desired sources lie on top of the sampling grid. However, in practice there is no reason to believe that the true source locations lie on the sampling grid. When the source location (θ_k, r_k) for any k does not lie on the considered sampling grid, the discrete model in (4.1) leads to grid mismatch. Note that the grid mismatch leads to less sparse (or maybe even non-sparse) solutions due to the energy leakage of the desired signals over the adjacent grid points. To illustrate this effect, consider a unit amplitude source present within the near-field region at $(10°, 10\lambda)$ and we try to localize the near-field source by using the classical

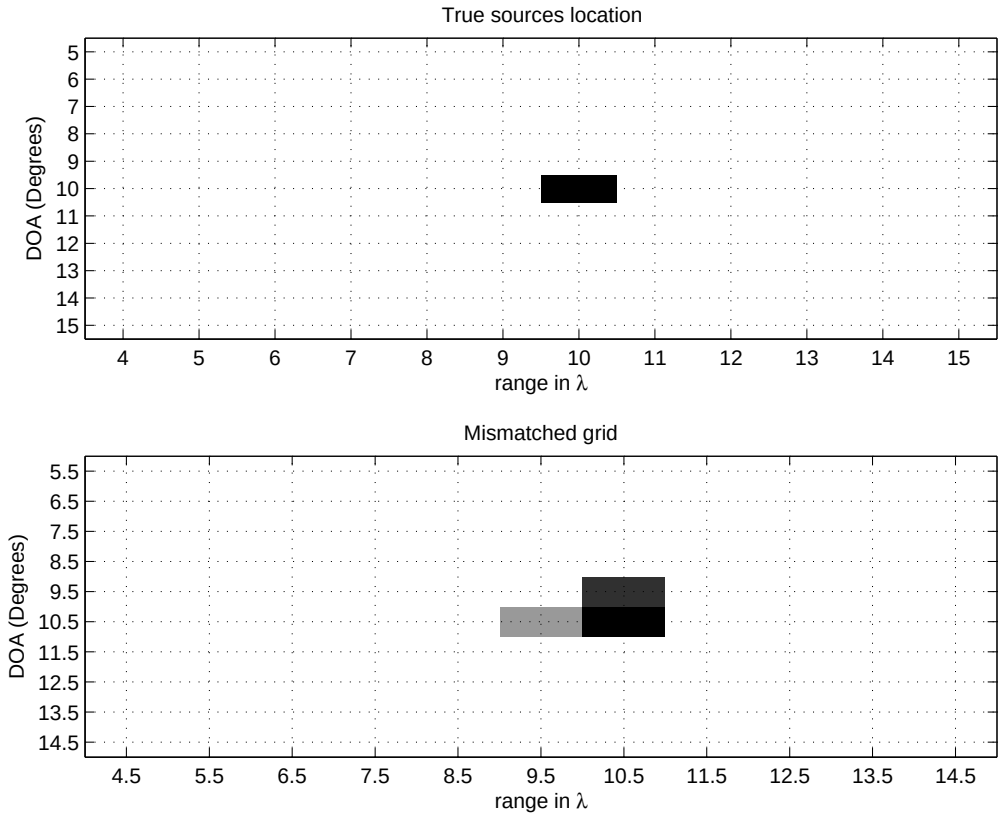


Figure 5.1: Source located at $(10°, 10\lambda)$. The sampling grid has a resolution of $\Delta\theta = 1°$ and $\Delta r = 1\lambda$ within the domain $[5.5°, 14.5°]$ and $[4.5\lambda, 14.5\lambda]$.

grid-based approach, where the sampling grid has the resolution of $\Delta\theta = 1°$ and $\Delta r = 1\lambda$ within the domain $[5.5°, 14.5°]$ and $[4.5\lambda, 14.5\lambda]$ such that the true source locations does not lie on the sampling grid. As seen in Fig. 5.1, the non-zero coefficients that correspond to the true source locations are spilled over the adjacent grid points. On the other hand, increasing the gridding resolution would lead to the columns of $\mathbf{A}$ that are highly correlated (i.e., $\mathbf{A}$ would have high coherence) which makes the sparse recovery difficult. The classical sparse regression does not account for such perturbations due to the grid mismatch.

5.2 Grid mismatch model

The EIV model mentioned in [17, 18, 19] treats the grid mismatch effect as an additive error matrix $\mathbf{E} \in \mathbb{C}^{M\times N}$. Taking the perturbation into account, we can rewrite the signal model in (4.1) as

$$\mathbf{y}(t) = \hat{\mathbf{A}}\mathbf{x}(t) + \mathbf{w}, \quad \text{where} \quad \hat{\mathbf{A}} = \mathbf{A} + \mathbf{E}. \tag{5.1}$$

If a source location (θ_k, r_k) is not on the sampling grid, then there exists a pair

of unknown error parameters $(e_{\bar{\theta}_n}, e_{\bar{r}_n})$ that accounts for the perturbation, $(\theta_k, r_k) = (\bar{\theta}_n + e_{\bar{\theta}_n}, \bar{r}_n + e_{\bar{r}_n})$.

Using a first-order Taylor expansion of the steering vector $\mathbf{a}(\bar{\theta}_n + e_{\bar{\theta}_n}, \bar{r}_n + e_{\bar{r}_n})$, we can approximate it up to the second order as follows

$$\mathbf{a}(\bar{\theta}_n + e_{\bar{\theta}_n}, \bar{r}_n + e_{\bar{r}_n}) \approx \mathbf{a}(\bar{\theta}_n, \bar{r}_n) + \left.\frac{\partial \mathbf{a}(\theta, r)}{\partial \theta}\right|_{\theta=\bar{\theta}_n} e_{\bar{\theta}_n} + \left.\frac{\partial \mathbf{a}(\theta, r)}{\partial r}\right|_{r=\bar{r}_n} e_{\bar{r}_n}. \tag{5.2}$$

The misaligned grid can be corrected by estimating the perturbation $(e_{\bar{\theta}_n}, e_{\bar{r}_n})$. In other words, near-field localization accounting for the grid mismatch is a joint estimation problem in which we jointly estimate $(\bar{\theta}_n, \bar{r}_n)$ as well as $(e_{\bar{\theta}_n}, e_{\bar{r}_n})$.

Using the approximation in (5.2), we can write the measurement vector $\mathbf{y}(t)$ for $T = 1$ as:

$$\begin{aligned}
\mathbf{y}(t) &= \sum_{n=1}^{N} \left(\mathbf{a}(\bar{\theta}_n, \bar{r}_n) + \left.\frac{\partial \mathbf{a}(\theta, r)}{\partial \theta}\right|_{\theta=\bar{\theta}_n} e_{\bar{\theta}_n} + \left.\frac{\partial \mathbf{a}(\theta, r)}{\partial r}\right|_{r=\bar{r}_n} e_{\bar{r}_n} \right) x_n(t) + \mathbf{w} \\
&= \sum_{n=1}^{N} \mathbf{a}(\bar{\theta}_n, \bar{r}_n) x_n(t) + \sum_{i=1}^{N} \left.\frac{\partial \mathbf{a}(\theta, r)}{\partial \theta}\right|_{\theta=\bar{\theta}_n} e_{\bar{\theta}_n} \\
&\quad + \sum_{n=1}^{N} x_n(t) \left.\frac{\partial \mathbf{a}(\theta, r)}{\partial r}\right|_{r=\bar{r}_n} e_{\bar{r}_n} + \mathbf{w},
\end{aligned}$$

and can be concisely written as

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{A}_{\boldsymbol{\theta}} [\mathbf{e}_\theta \odot \mathbf{x}] + \mathbf{A}_{\mathbf{r}} [\mathbf{e}_r \odot \mathbf{x}] + \mathbf{w}, \tag{5.3}$$

where $\mathbf{e}_\theta = \left[e_{\bar{\theta}_1}, \ldots, e_{\bar{\theta}_N}\right]^T \in \mathbb{R}^{N\times 1}$ and $\mathbf{e}_r = \left[e_{\bar{\theta}_1}, \ldots, e_{\bar{\theta}_N}\right]^T \in \mathbb{R}^{N\times 1}$ are the perturbation vectors, and $\mathbf{A}_{\boldsymbol{\theta}} = \left[\left.\frac{\partial \mathbf{a}(\theta,r)}{\partial \theta}\right|_{\theta=\bar{\theta}_1}, \ldots, \left.\frac{\partial \mathbf{a}(\theta,r)}{\partial \theta}\right|_{\theta=\bar{\theta}_N} \right] \in \mathbb{C}^{M\times N}$, matrices $\mathbf{A}_{\mathbf{r}} = \left[\left.\frac{\partial \mathbf{a}(\theta,r)}{\partial r}\right|_{r=\bar{r}_1}, \ldots, \left.\frac{\partial \mathbf{a}(\theta,r)}{\partial r}\right|_{=\bar{r}_N} \right] \in \mathbb{C}^{M\times N}$ contains the first-order derivatives.

Using matrix-vector properties, we have $\mathbf{A}_{\boldsymbol{\theta}} [\mathbf{e}_\theta \odot \mathbf{x}] + \mathbf{A}_{\mathbf{r}} [\mathbf{e}_r \odot \mathbf{x}] = \big(\mathbf{A}_{\boldsymbol{\theta}}\mathbf{E}_\theta + \mathbf{A}_{\mathbf{r}}\mathbf{E}_r\big)\mathbf{x}$, where $\mathbf{E}_\theta = \mathrm{diag}(\mathbf{e}_\theta)$ and $\mathbf{E}_r = \mathrm{diag}(\mathbf{e}_r)$. Recall that the error matrix $\mathbf{E}$ in (5.1) is given by $\mathbf{E} = \mathbf{A}_{\boldsymbol{\theta}}\mathbf{E}_\theta + \mathbf{A}_{\mathbf{r}}\mathbf{E}_r \in \mathbb{C}^{M\times N}$.

Now the problem is briefly stated as follows: given the wide matrix $\boldsymbol{\Phi} = [\mathbf{A}|\mathbf{A}_\theta|\mathbf{A}_r] \in \mathbb{C}^{M\times 3N}$ and the measurement vector $\mathbf{y}$, estimate the unknown parameters $\mathbf{x}$, $\mathbf{e}_\theta$, and $\mathbf{e}_r$.

Remark 3:

(1) Since the perturbations are related to the DOA and range, both $\mathbf{e}_\theta$ and $\mathbf{e}_r$ are real parameters.

(2) When $x_n = 0$ for a certain n, then $e_{\bar{\theta}_n}$ and $e_{\bar{r}_n}$ can take any value without any contributions to the observation vector $\mathbf{y}$. Hence, there is no meaning to recovery such $e_{\bar{\theta}_n}$ and $e_{\bar{r}_n}$. Hence, through out this paper we focus only on recovering the entries of $\mathbf{e}_\theta$ and $\mathbf{e}_r$ related to non-zero entries of $\mathbf{x}$.

Algorithm 1 Alternating minimization

1. **Input:** $\mathbf{A}, \mathbf{A}_{\boldsymbol{\theta}}, \mathbf{A}_{\mathbf{r}}, \epsilon, i_{\max}, \eta$
2. **Output:** $\hat{\mathbf{e}}_\theta, \hat{\mathbf{e}}_r, \hat{\mathbf{x}}$
3. **Initial:** $\mathbf{e}_\theta^{(0)} = \mathbf{0}, \mathbf{e}_r^{(0)} = \mathbf{0}$
4. **for** $i = 0$ to $i_{\max}$
5. $\quad \hat{\mathbf{x}}^{(i)} = \arg\min \|\mathbf{x}\|_1$ s.t. $\|\mathbf{A}\mathbf{x} + \mathbf{A}_{\boldsymbol{\theta}}[\mathbf{e}_\theta^{(i)} \odot \mathbf{x}] + \mathbf{A}_{\mathbf{r}}[\mathbf{e}_r^{(i)} \odot \mathbf{x}] - \mathbf{y}\|_2 \leq \epsilon$
6. $\quad$ **update** $\hat{\mathbf{x}}$
7. $\quad\quad$ **if** $\hat{x}_n^{(i)} \leq \eta$
8. $\quad\quad\quad \hat{x}_n^{(i)} = 0$
9. $\quad\quad$ **end**
10. $\quad \mathbf{l} = |\hat{\mathbf{x}}^{(i)}| ./ \max(|\hat{\mathbf{x}}^{(i)}|)$
11. $\quad \{\mathbf{e}_\theta^{(i+1)}, \mathbf{e}_r^{(i+1)}\} = \arg\min_{\mathbf{e}_\theta, \mathbf{e}_r} \|\mathbf{A}\mathbf{x} + \mathbf{A}_{\boldsymbol{\theta}}[\mathbf{e}_\theta \odot \hat{\mathbf{x}}^{(i)}] + \mathbf{A}_{\mathbf{r}}[\mathbf{e}_r \odot \hat{\mathbf{x}}^{(i)}] - \mathbf{y}\|_2$
$\quad\quad$ s.t. $-0.5\Delta\theta\mathbf{l} < \mathbf{e}_\theta < 0.5\Delta\theta\mathbf{l} \quad -0.5\Delta r\mathbf{l} < \mathbf{e}_r < 0.5\Delta r\mathbf{l}$
12. **end**
13. $\mathcal{I}$ is the set of the support of the non-zero elements of $\hat{\mathbf{x}}^{(i_{\max})}$
14. $\hat{\theta}_k = \bar{\theta}_k + e_{\bar{\theta}_k}$; $\hat{r}_k = \bar{r}_k + e_{\bar{r}_k}$, where $k \in \mathcal{I}$

Table 5.1: Alternation minimization algorithm.

5.3 Sparse recovery for off-grid sources with single snapshot

5.3.1 Bounded sparse recovery

A pair of unknown perturbations $(e_{\bar{\theta}_n}, e_{\bar{r}_n})$ indicates the error between a source location (θ_k, r_k) and the nearest grid point $(\bar{\theta}_n, \bar{r}_n)$ and is always bounded by the size of grid cell by definition. With bounded perturbations $e_{\bar{\theta}_n}$ and $e_{\bar{r}_n}$, we can reconstruct the sparse signal $\mathbf{x}$ in the mismatch model (5.3) by solving an optimization problem of the form [20]

$$\begin{aligned}
[\hat{\mathbf{x}}, \hat{\mathbf{e}}_\theta, \hat{\mathbf{e}}_r] \quad = \quad & \arg \min_{\mathbf{x}\in\mathbb{C}^N, \mathbf{e}_\theta\in\mathbb{R}^N, \mathbf{e}_r\in\mathbb{R}^N} \quad \mu_1\|\mathbf{x}\|_1 + \frac{1}{2}\|\left(\mathbf{A} + \mathbf{A}_{\boldsymbol{\theta}}\mathbf{E}_\theta + \mathbf{A}_{\mathbf{r}}\mathbf{E}_r\right)\mathbf{x} - \mathbf{y}\|_2^2 \\
\text{s.t.} \quad & \mathbf{E}_\theta = \operatorname{diag}(\mathbf{e}_\theta), \quad \mathbf{E}_r = \operatorname{diag}(\mathbf{e}_r) \\
& -0.5\Delta\theta < e_{\bar{\theta}_n} < 0.5\Delta\theta; \\
& -0.5\Delta r < e_{\bar{r}_n} < 0.5\Delta r; \quad \text{for} \quad n = 1, 2, \ldots, N.
\end{aligned} \tag{5.4}$$

The above optimization is non-convex and generally difficult to solve. In [27], an alternating minimization algorithm to solve such non-convex optimization problems was proposed. We adapt the algorithm to suit our problem, and this is summarized in the Table. 5.1. Typically, the threshold η in Algorithm 1 is chosen depending on the noise level, i.e., $\eta \geq \sigma_w$, where σ_w is the standard deviation of the additive noise.

5.3.2 Bounded joint sparse recovery

The model in (5.3) can be expressed as

$$\begin{aligned}\mathbf{y} &= \overbrace{\left[\mathbf{A}\middle|\mathbf{A}_{\boldsymbol{\theta}}\middle|\mathbf{A}_{\mathbf{r}}\right]}^{\boldsymbol{\Phi}} \overbrace{\begin{bmatrix}\mathbf{x}\\ \mathbf{e}_\theta \odot \mathbf{x}\\ \mathbf{e}_r \odot \mathbf{x}\end{bmatrix}}^{\mathbf{q}} + \mathbf{w} \\ &= \boldsymbol{\Phi}\mathbf{q} + \mathbf{w}, \end{aligned} \tag{5.5}$$

where $\boldsymbol{\Phi} \in \mathbb{C}^{M\times 3N}$, and $\mathbf{q} \in \mathbb{C}^{3N\times 1}$ is a $3K$-sparse vector. Recalling remark 3, the vector $\mathbf{q}$ can be grouped into N sub-blocks and each block contain 3 elements, i.e. the j-th sub-block will have elements $\mathbf{q}_j = [x_j, e_{\bar{\theta}_j}x_j, e_{\bar{r}_j}x_j]^T$ for $j = 1, 2, \ldots, N$. Then the vector $\mathbf{q}$ can be viewed as K-block sparse vector since only K sub-blocks have non-zero values. In Chapter 3, we discussed the sparse recovery techniques involving block sparse recovery. We can enforce block sparsity on $\mathbf{q}$ via a mixed ℓ_2/ℓ_1-norm minimization to recover vector $\mathbf{q}$ instead of enforcing the sparsity of $\mathbf{q}$ via ℓ_1-norm minimization [21, 22]. Define the mixed ℓ_2/ℓ_1-norm of $\mathbf{q}$ as

$$\|\mathbf{q}\|_{2,1} = \sum_{n=1}^{N} \sqrt{x_n^2 + (e_{\bar{\theta}_n}x_n)^2 + (e_{\bar{r}_n}x_n)^2}.$$

The optimization problem taking this structure into account can be formulated as

$$\begin{aligned}[\hat{\mathbf{x}}, \hat{\mathbf{e}}_\theta, \hat{\mathbf{e}}_r] \quad = \quad & \arg \min_{\mathbf{x}\in\mathbb{C}^N, \mathbf{e}_\theta\in\mathbb{R}^N, \mathbf{e}_r\in\mathbb{R}^N} \quad \mu_2\|\mathbf{q}\|_{2,1} + \frac{1}{2}\|\boldsymbol{\Phi}\mathbf{q} - \mathbf{y}\|_2^2, \\ & \text{s.t.} \quad \begin{aligned} -0.5\Delta\theta|\mathbf{x}| < \mathbf{q}_{\boldsymbol{\theta}} < 0.5\Delta\theta|\mathbf{x}|, \\ -0.5\Delta r|\mathbf{x}| < \mathbf{q}_{\mathbf{r}} < 0.5\Delta r|\mathbf{x}|, \end{aligned}\end{aligned} \tag{5.6}$$

where $|\mathbf{x}|$ denotes the entry-wise absolute value of vector $\mathbf{x}$, $\mathbf{q}_{\boldsymbol{\theta}} = \mathbf{e}_\theta \odot \mathbf{x}$ and $\mathbf{q}_{\mathbf{r}} = \mathbf{e}_r \odot \mathbf{x}$. The optimization problem in (5.6) is non-convex, and can be relaxed to a convex optimization problem using separating technique, where we assume $\mathbf{x} = \mathbf{x}^+ - \mathbf{x}^-$. The relaxed optimization problem is referred to as the relaxed bounded joint sparse (R-BJS) estimator [27], and is given by

$$\begin{aligned}[\hat{\mathbf{x}}, \hat{\mathbf{e}}_\theta, \hat{\mathbf{e}}_r] = \arg \min_{\mathbf{x}^+, \mathbf{x}^-, \mathbf{q}} \quad & \frac{1}{2}\|\mathbf{A}(\mathbf{x}^+ - \mathbf{x}^-) + \mathbf{A}_{\boldsymbol{\theta}}\mathbf{q}_{\boldsymbol{\theta}} + \mathbf{A}_{\mathbf{r}}\mathbf{q}_{\mathbf{r}} - \mathbf{x}\|_2^2 + \gamma\|\mathbf{q}\|_{2,1} \\ \text{s.t.} \quad & \mathbf{x}^+ \geq \mathbf{0}, \quad \mathbf{x}^- \geq \mathbf{0} \\ & -0.5\Delta\theta(\mathbf{x}^+ + \mathbf{x}^-) < \mathbf{q}_{\boldsymbol{\theta}} < 0.5\Delta\theta(\mathbf{x}^+ + \mathbf{x}^-) \\ & -0.5\Delta r(\mathbf{x}^+ + \mathbf{x}^-) < \mathbf{q}_{\mathbf{r}} < 0.5\Delta r(\mathbf{x}^+ + \mathbf{x}^-)\end{aligned} \tag{5.7}$$

The above convex problem provides an accurate solution only when $\mathbf{q}_{\boldsymbol{\theta}} \odot \mathbf{q}_{\mathbf{r}} = \mathbf{0}$. During simulations it was observed that this non-convex constraint is always satisfied. The R-BJS method is much more efficient than the alternating algorithm. The performance analysis of R-BJS is given in appendix A.4.

Remark 4

The result of R-BJS is an approximate K-block sparse vector, then we choose K columns of matrix $\mathbf{\Phi}$ related to the K sub-blocks with largest ℓ_2-norm value to generated a new sensing matrix $\tilde{\mathbf{\Phi}}$, and substitute into (5.7) to refine the R-BJS results. For example, in single source case, the estimated result is related to the j-th sub-block, where $j = 1, 2, \ldots, N$, hence we pick up the related columns of $\mathbf{\Phi}$ to generate sensing matrix $\tilde{\mathbf{\Phi}}$, then run optimization (5.7) again to find the optimal perturbation parameters.

5.4 Sparse recovery for off-grid sources with multiple snapshots

When multiple snapshots are available, we can stack (3.4) for the batch of T measurements into a matrix

$$\mathbf{Y} = \mathbf{A}(\boldsymbol{\theta}, \mathbf{r})\mathbf{S} + \mathbf{W} \tag{5.8}$$

where $\mathbf{Y} = [\mathbf{y}(1), \ldots, \mathbf{y}(T)] \in \mathbb{C}^{M\times T}$, and matrices $\mathbf{S}$ and $\mathbf{W}$ are defined similarly.

5.4.1 Step-1: Grid matching in DOA domain

In Chapter 4, we introduced a method to transform the original 2D (DOA and range) estimation into 1D (DOA) estimation via exploiting the cross-correlation between the symmetric sensors of the array based on the Fresnel approximation. The signal model based on the Fresnel approximation is given by (cf.(3.8))

$$\mathbf{y}(t) = \sum_{k=1}^{K} \tilde{\mathbf{a}}(\omega_k, \phi_k)s_k(t) + \mathbf{w}(t), \quad \text{for} \quad t = 1, 2, \ldots, T,$$

$$\text{where} \quad \omega_k = -\frac{2\pi\delta}{\lambda}\sin(\theta_k) \quad \text{and} \quad \phi_k = \frac{\pi\delta^2}{\lambda r_k}\cos^2(\theta_k)$$

The structure of the array manifold based on Fresnel approximation naturally decouples the DOA and range. As in Chapter 4, we have a virtual far-field model (4.6) depending only on DOA, and it is given by

$$\mathbf{r}_y = \mathbf{A}_\omega(\boldsymbol{\omega})\mathbf{r}_\mathbf{s} + \sigma_w^2\mathbf{e}$$

with the array manifold

$$\mathbf{A}_\omega(\boldsymbol{\omega}) = [\mathbf{a}_\omega(\omega_1), \mathbf{a}_\omega(\omega_2), \ldots, \mathbf{a}_\omega(\omega_K)] \in \mathbb{C}^{M\times K},$$

where $\mathbf{r}_y = [r_y(-p, p), \ldots, r_y(0, 0), r_y(1, -1), \ldots, r_y(p, -p)]^T \in \mathbb{C}^{M\times 1}$ with $M = 2p + 1$, $\mathbf{r}_s = [\sigma_{s,1}^2, \ldots, \sigma_{s,K}^2]^T \in \mathbb{C}^{K\times 1}$, and $\mathbf{e} = [1, \mathbf{0}_p^T]^T \in \mathbb{C}^{L\times 1}$.

Based on the virtual far-field model, we can construct an overcomplete basis $\mathbf{A}_\omega$ as we did in Chapter 4 with only N_θ columns corresponding to the potential source DOAs using the sampling grid $\bar{\boldsymbol{\theta}}$, i.e.,

$$\mathbf{A}_\omega(\bar{\boldsymbol{\omega}}) = [\mathbf{a}_\omega(\bar{\omega}_1), \ldots, \mathbf{a}_\omega(\bar{\omega}_{N_\theta})] \in \mathbb{C}^{M\times N_\theta},$$

where $\bar{\omega}_n = -\frac{2\pi\delta}{\lambda}\sin(\bar{\theta}_n)$ for all $n \in \{1, \ldots, N_\theta\}$ is as defined earlier. If an ω_k is not on the sampling grid, then there exists an error parameter $e_{\bar{\omega}_n}$ that accounts perturbation, $\omega_k = \bar{\omega}_n + e_{\bar{\omega}_n}$.

Using a first-order Taylor expansion $\mathbf{a}_\omega(\bar{\omega}_n + e_{\bar{\omega}_n}) = \mathbf{a}_\omega(\bar{\omega}_n) + \frac{\partial \mathbf{a}_\omega}{\partial \omega}\big|_{\omega=\bar{\omega}_n} e_{\bar{\omega}_n}$, we can rewrite (4.6) as

$$\mathbf{r}_y = \left(\mathbf{A}_\omega(\bar{\boldsymbol{\omega}}) + \tilde{\mathbf{A}}_\omega(\bar{\boldsymbol{\omega}})\mathbf{E}_\omega\right)\mathbf{u} + \sigma_w^2\mathbf{e}, \tag{5.9}$$

where $\mathbf{E}_\omega = \text{diag}(e_{\bar{\omega}_1}, e_{\bar{\omega}_2}, \ldots, e_{\bar{\omega}_{N_\theta}}) \in \mathbb{R}^{N_\theta \times N\theta}$ and $\tilde{\mathbf{A}}_\omega(\bar{\boldsymbol{\omega}}) = \left[\frac{\partial \mathbf{a}_\omega}{\partial \omega}\big|_{\omega=\bar{\omega}_1}, \ldots, \frac{\partial \mathbf{a}_\omega}{\partial \omega}\big|_{\omega=\bar{\omega}_{N_\theta}}\right] \in \mathbb{C}^{M\times N_\theta}$.

Choosing a proper constraint box $[-\alpha, \alpha]$ limiting perturbation parameter $e_{\bar{\omega}_n}$, the optimization problem for (4.6) is formulate as

$$\begin{aligned}[\hat{\mathbf{u}}, \hat{\mathbf{p}}_\omega] = \arg\min_{\mathbf{u},\mathbf{p}_\omega} \quad & \frac{1}{2}\|\mathbf{A}_\omega\mathbf{u} + \tilde{\mathbf{A}}_\omega\mathbf{p}_\omega - \mathbf{r}_y\|_2^2 + \mu_3\|\mathbf{z}\|_{2,1} \\ \text{s.t.} \quad & \mathbf{u} \geq \mathbf{0}, \mathbf{z} = [\mathbf{u}^T, \mathbf{p}_\omega^T]^T, \text{ and } -\alpha\mathbf{u} \leq \mathbf{p}_\omega \leq \alpha\mathbf{u},\end{aligned} \tag{5.10}$$

where $\mathbf{p}_\omega = \mathbf{e}_\omega \odot \mathbf{u}$, and $\|\mathbf{z}\|_{2,1} = \sum_{n=1}^{N_\theta}\sqrt{u_n^2 + p_{\omega,n}^2}$.

5.4.1.1 Realigning the DOA estimate

In the first step the matching model we build is under the Fresnel approximation, and the estimated DOA result suffers from the Fresnel approximation. The perturbation generated by the approximation in Fresnel zone influences less in the classical grid-based estimation, and we can still obtain the accurate result, which can be seen in the results presented in Chapter 4. However, for the grid mismatch, the error we estimate is the perturbation on the virtual far-field grid based on the Fresnel approximation. For example, assume we have a near-field source located at $(50°, 5\lambda)$, the estimated DOA based on the Fresnel approximation via algorithms proposed in Chapter 4 (note that the Fresnel zone is related to the aperture of ULA, thus different ULA has different Fresnel zone), will be 48.8067°. This motivates us to realign the estimated result in the Fresnel model to the the estimated result in spherical wavefront model.

Consider a ULA with $M = 2p + 1$ element as illustrated in Fig. 3.1. The distance to each sensor is depicted in (3.2),

$$r_{m,k} = \sqrt{r_k + m^2\delta^2 - 2m\delta r_k \sin(\theta_k)}.$$

In Fresnel approximation, we assume that

$$r_{m,k} \approx r_k - m\delta\sin\theta_k + m^2\delta^2\frac{\cos^2\theta_k}{2r_k}.$$

The range difference $d_{m,k} = r_{m,k} - r_k$ is given by

$$\begin{aligned} d_{m,k} &\approx -m\delta\sin(\theta_k) + m^2\frac{\delta^2}{2r_k}\cos^2(\theta_k) \\ &= m\delta c_{k,1} + m^2\frac{\delta^2}{2}c_{k,2}. \end{aligned}$$

Define $\mathbf{r}^{(k)} = [r_{-p,k}, \ldots, r_{p,k}]^T$ as the distance vector of k-th source to each sensor of the ULA. Stack $d_{m,k}$, we can obtain

$$\mathbf{r}^{(k)} - r_k \mathbf{1} = \hat{\mathbf{d}}_k, \quad k = 1, 2, \ldots, K.$$

We can now build linear equation of the form

$$\mathbf{r}^{(k)} - r_k \mathbf{1} = \mathbf{B}\mathbf{c}_k.$$

where

$$\mathbf{B} = \begin{bmatrix} -\delta p & 0.5\delta^2(-p)^2 \\ \vdots & \vdots \\ \delta p & 0.5\delta^2(p)^2 \end{bmatrix} \in \mathbb{R}^{M \times 2}$$

is a tall matrix and $\mathbf{c}_k = [c_{k,1}, c_{k,2}]^T$. We can then find Fresnel based location $(\hat{\theta}, \hat{r})$ via the true delay vector by using least-squares method. Then we can build a table for us to realign the Fresnel approximated one to the one based on spherical model.

5.4.2 Step-2: grid matching in range domain

The measurement covariance matrix is given by

$$\mathbf{R}_{\mathbf{yy}} = \mathbf{A}(\boldsymbol{\theta}, \mathbf{r}) \mathbf{R}_{ss} \mathbf{A}(\boldsymbol{\theta}, \mathbf{r})^H + \sigma_w^2 \mathbf{I}_M. \tag{5.11}$$

By using the Khatri-rao product, (5.11) can be written as

$$\begin{aligned} \mathbf{p} &= [\mathbf{A}(\boldsymbol{\theta}, \mathbf{r})^* \circledcirc \mathbf{A}(\boldsymbol{\theta}, \mathbf{r})]\mathbf{r_s} + \sigma_w^2 \hat{\mathbf{e}} \\ &= \boldsymbol{\Psi}(\boldsymbol{\theta}, \mathbf{r})\mathbf{r_s} + \sigma_w^2 \hat{\mathbf{e}}, \end{aligned}$$

where $\mathbf{p}$ represents the $\text{vec}(\mathbf{R}_{\mathbf{yy}}) \in \mathbb{C}^{M^2 \times 1}$, $\hat{\mathbf{e}} = \text{vec}(\mathbf{I}_M) \in \mathbb{C}^{M^2 \times 1}$, and

$$\boldsymbol{\Psi}(\boldsymbol{\theta}, \mathbf{r}) = [\boldsymbol{\psi}(\theta_1, r_1), \ldots, \boldsymbol{\psi}(\theta_K, r_K)] \in \mathbb{C}^{M^2 \times K}.$$

Let $\hat{\boldsymbol{\theta}}$ be the DOA from step-1, and $\hat{K}$ (notice that $\hat{K} \leq K$, e.g. two sources may share the same DOA but at different range) denotes the number of different source DOAs detected (i.e., $\hat{K} = \|\hat{\boldsymbol{u}}\|_0$). We now firstly realign $\hat{\boldsymbol{\theta}}$ to $\hat{\bar{\boldsymbol{\theta}}}$ at different range, and then using sampling grid $(\hat{\bar{\boldsymbol{\theta}}}, \bar{\mathbf{r}})$ (in Chapter 3, we have $\hat{K}N_r$ points $\{(\hat{\theta}_1, \bar{r}_1), (\hat{\theta}_1, \bar{r}_2), \ldots\}$, and here the nodes number is not changed but $\{(\hat{\bar{\theta}}_1, \bar{r}_1), (\hat{\bar{\theta}}_2, \bar{r}_2), \ldots\}$) to form an overcomplete basis $\boldsymbol{\Psi}(\hat{\bar{\boldsymbol{\theta}}}, \bar{\mathbf{r}}) \in \mathbb{C}^{M^2 \times \hat{K}N_r}$ as well as the partial derivative matrix $\boldsymbol{\Psi}_r(\hat{\bar{\boldsymbol{\theta}}}, \bar{\mathbf{r}}) \in \mathbb{C}^{M^2 \times \hat{K}N_r}$

$$\boldsymbol{\Psi}_r(\hat{\bar{\boldsymbol{\theta}}}, \bar{\mathbf{r}}) = \left[\frac{\partial \boldsymbol{\psi}(\hat{\bar{\theta}}_1, r)}{\partial r}\bigg|_{r=\bar{r}_1}, \ldots, \frac{\partial \boldsymbol{\psi}(\hat{\bar{\theta}}_1, r)}{\partial r}\bigg|_{r=\bar{r}_{N_r}}, \frac{\partial \boldsymbol{\psi}(\hat{\bar{\theta}}_e, r)}{\partial r}\bigg|_{r=\bar{r}_1}, \ldots, \frac{\partial \boldsymbol{\psi}(\hat{\bar{\theta}}_{\hat{K}}, r)}{\partial r}\bigg|_{r=\bar{r}_{N_r}} \right].$$

And then we generate the grid matching model on range domain

$$\mathbf{p} = \left(\boldsymbol{\Psi}(\hat{\bar{\boldsymbol{\theta}}}, \bar{\mathbf{r}}) + \boldsymbol{\Psi}_r(\hat{\bar{\boldsymbol{\theta}}}, \bar{\mathbf{r}}) \tilde{\mathbf{E}}_r \right) \tilde{\mathbf{r}}_{\mathbf{s}} + \sigma_w^2 \hat{\mathbf{e}}, \tag{5.12}$$

where $\tilde{\mathbf{r}}_{\mathbf{s}} \in \mathbb{C}^{\hat{K}N_r \times 1}$ is a K-sparse vector containing the signal power spectrum. Rewrite (5.12) as

$$\begin{aligned} \mathbf{p} &= \overbrace{\left[\boldsymbol{\Psi}(\hat{\hat{\boldsymbol{\theta}}}, \bar{\mathbf{r}}), \boldsymbol{\Psi}_r(\hat{\hat{\boldsymbol{\theta}}}, \bar{\mathbf{r}})\right]}^{\boldsymbol{\Phi}_r} \overbrace{\begin{bmatrix} \tilde{\mathbf{r}}_{\mathbf{s}} \\ \mathbf{e}_r \odot \tilde{\mathbf{r}}_{\mathbf{s}} \end{bmatrix}}^{\mathbf{u}_r} + \sigma_w^2 \hat{\mathbf{e}} \\ &= \boldsymbol{\Phi}_r \mathbf{u}_r + \sigma_w^2 \hat{\mathbf{e}} \end{aligned} \tag{5.13}$$

where $\boldsymbol{\Phi}_r \in \mathbb{C}^{M^2 \times 2\hat{K}N_r}$, and $\mathbf{u}_r \in \mathbb{C}^{2\hat{K}N_r \times 1}$ is a $2K$-sparse vector, which can be divided into $\hat{K}N_r$ sub-blocks and be viewed as K-block sparse vector. Letting $\tilde{\mathbf{q}}_r = \tilde{\mathbf{e}}_r \odot \tilde{\mathbf{r}}_{\mathbf{s}}$, then we can reconstruct the K-block sparse vector $\tilde{\mathbf{r}}_{\mathbf{s}}$ via the mixed ℓ_2/ℓ_1-norm minimization optimization

$$\begin{aligned} \arg\min_{\tilde{\mathbf{r}}_{\mathbf{s}}, \tilde{\mathbf{q}}_r} \quad & \frac{1}{2}\|\mathbf{p} - \boldsymbol{\Phi}_r \mathbf{u}_r\|_2^2 + \mu_4 \|\mathbf{u}_r\|_{2,1} \\ \text{s.t.} \quad & \tilde{\mathbf{r}}_{\mathbf{s}} \geq \mathbf{0}, \text{ and } -0.5\Delta r \tilde{\mathbf{r}}_{\mathbf{s}} < \tilde{\mathbf{q}}_r < \tilde{\mathbf{r}}_{\mathbf{s}} 0.5\Delta r, \end{aligned} \tag{5.14}$$

where $\|\mathbf{u}_r\|_{2,1} = \sum_{n=1}^{\hat{K}N_r} \sqrt{r_{s,n}^2 + q_{r,n}^2}$.

5.5 Simulations

We consider a symmetric ULA with $M = 15$ sensors and an inter-sensor spacing $\delta = \frac{\lambda}{4}$ where λ represents the wavelength of the narrowband source signals. The near-field region for this array is within the distance $2D^2 = 24.5\lambda$. We compare the results of the proposed agorithms to the grid matching problem and the optimization problems we solved in such estimations is done by CVX [26].

In Fig. 5.2, we show the results of grid matching for near-field source localization when single snapshot is available. We consider a unit amplitude single source at location $(0.3°, 1.28\lambda)$. The sampling grid has a resolution of $\Delta\theta = 1°$ and $\Delta r = 0.1\lambda$, resulting in a maximum DOA and range misalignment of $0.5°$ and 0.05λ, respectively. The SNR is 40 dB. The threshold of alternating algorithm is 0.4. Alternating algorithm performed better than R-BJS form the plot, but R-BJS is much faster and suffer less computational load. In Fig. 5.3, we put all the proposed algorithms together and the performance of the proposed algorithm can be visually compared. Also, it can be seen that refined R-BJS improves the performance of R-BJS.

In Fig. 5.4, we show the results for multiple near-field sources localization via mismatched sensing grid for single snapshot scenario. We consider two sources with a unit amplitude locating at $(0.3°, 1.28\lambda)$ and $(-3.2°, 1.52\lambda)$, using the same sampling grid that was used in Fig 5.3. The SNR is 40 dB. As can be seen that refining the result of R-BJS and improve the performance and provide a much more accurate estimation result. The improvement can be seen much more visually in Fig. 5.5.

In Fig. 5.6, we show the proposed two-step approach to grid matching for near-field source localization when multiple snapshots are available. We consider three near-field sources at locations $(-20.82°, 6.1\lambda)$ and $(10.12°, 10.9\lambda)$. The simulations are provided for SNRs of 0 dB with $T = 200$ snapshots. In the two-step grid matching algorithm, we

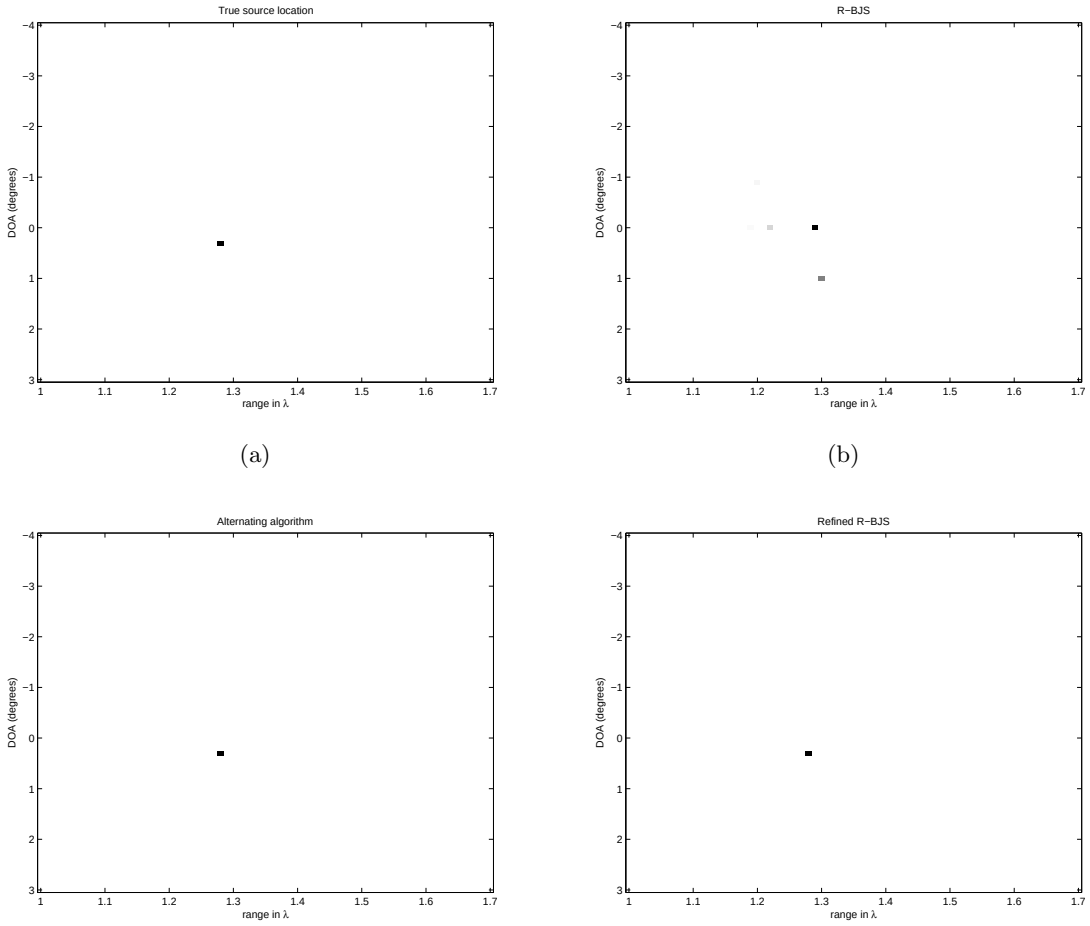
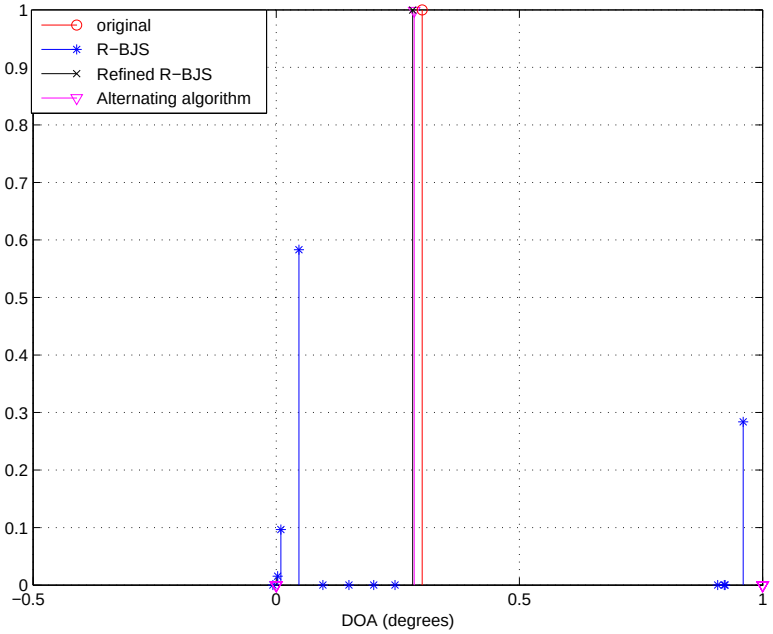
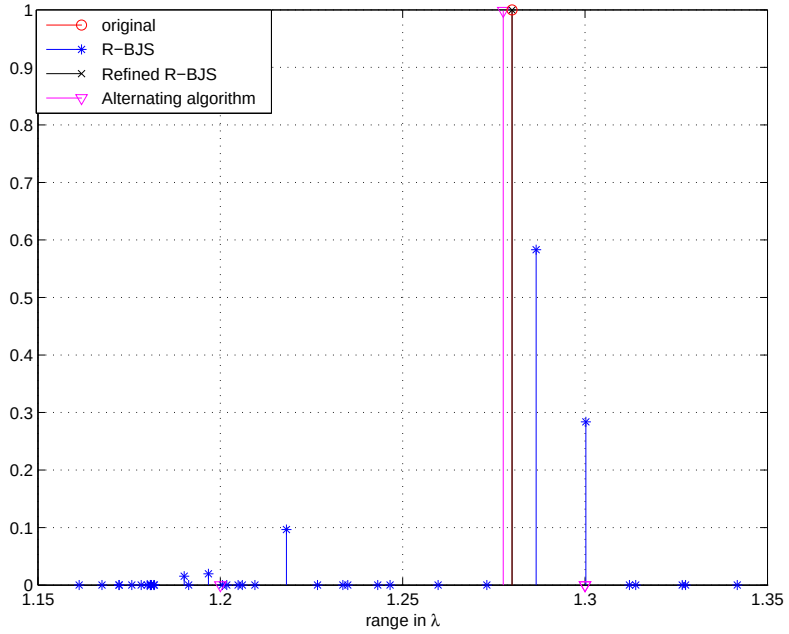


Figure 5.2: Near-field source localization via mismatched sensing grid. Source located at $(0.3°, 1.28\lambda)$. SNR = 40 dB, and T=1. The sampling grid has a resolution of $\Delta\theta = 1°$ and 0.1λ within the domain $[-4°, 3°]$ and $[1\lambda, 1.7\lambda]$. (a) True source location. (b) R-BJS. (c) Alternating minimization algorithm with threshold $\eta = 0.4$. (d) Refined R-BJS.

solve for the Fresnel approximated DOAs in the correlation domain in which the range parameters are naturally decoupled from the DOAs (due to the Fresnel approximation). Due to the Fresnel approximation, in the next step, we check the Fresnel error at different range to the corresponding Fresnel estimated DOAs and realign to $\hat{\hat{\theta}}$ at different range. By solving the optimization problem proposed in grid matching in range domain, we locate the near-field sources. The results have been zoomed in for visual comparison.



(a) DOA estimation result.



(b) Range estimation result.

Figure 5.3: DOA and range estimation via mismatched sensing grid. Source located at $(0.3°, 1.28\lambda)$. SNR = 40 dB, and T=1. The sampling grid has a resolution of $\Delta\theta = 1°$ and 0.1λ within the domain $[-4°, 3°]$ and $[1\lambda, 1.7\lambda]$.

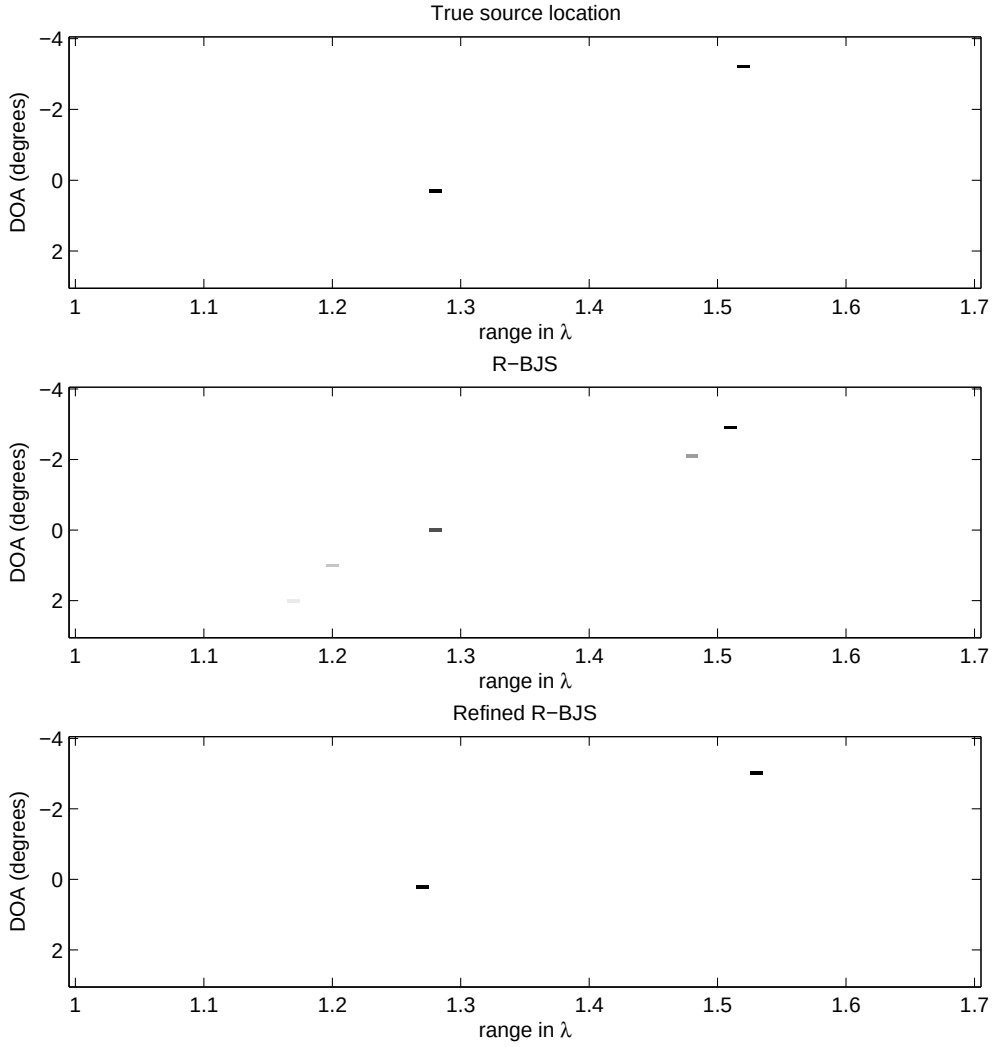
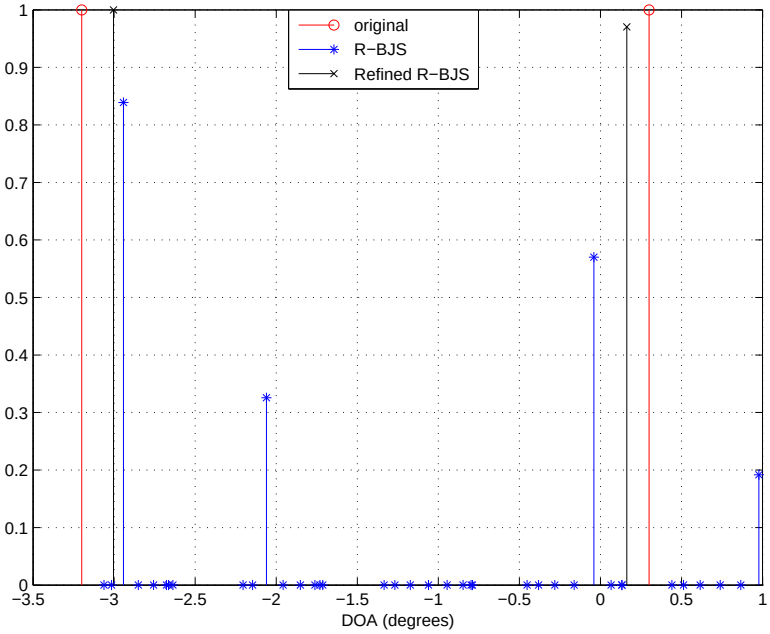
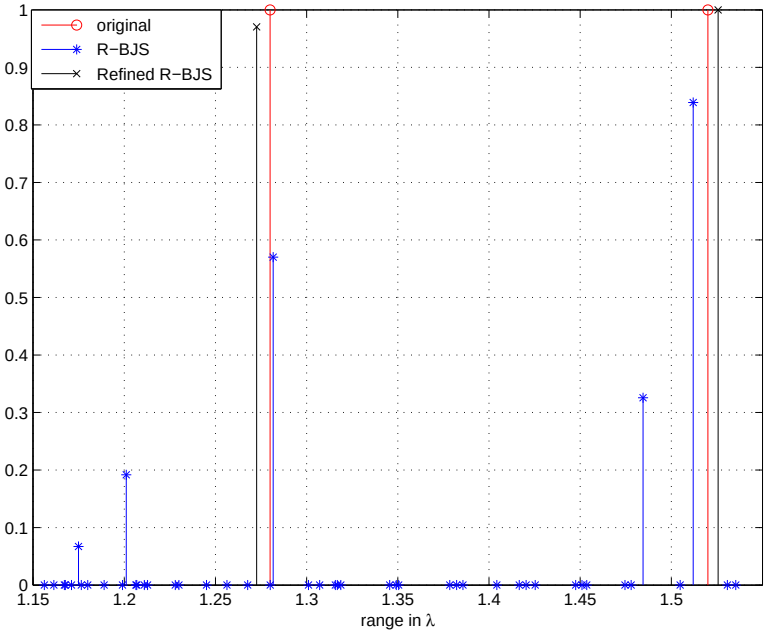


Figure 5.4: Multiple source localization via mismatched sensing grid. Near-field sources with DOAs: 0.3° and -3.2°, and ranges: 1.28λ and 1.52λ. SNR = 40 dB, and $T = 1$. The sampling grid has a resolution of $\Delta\theta = 1^\circ$ and 0.1λ within the domain $[-4^\circ, 3^\circ]$ and $[1\lambda, 1.7\lambda]$.

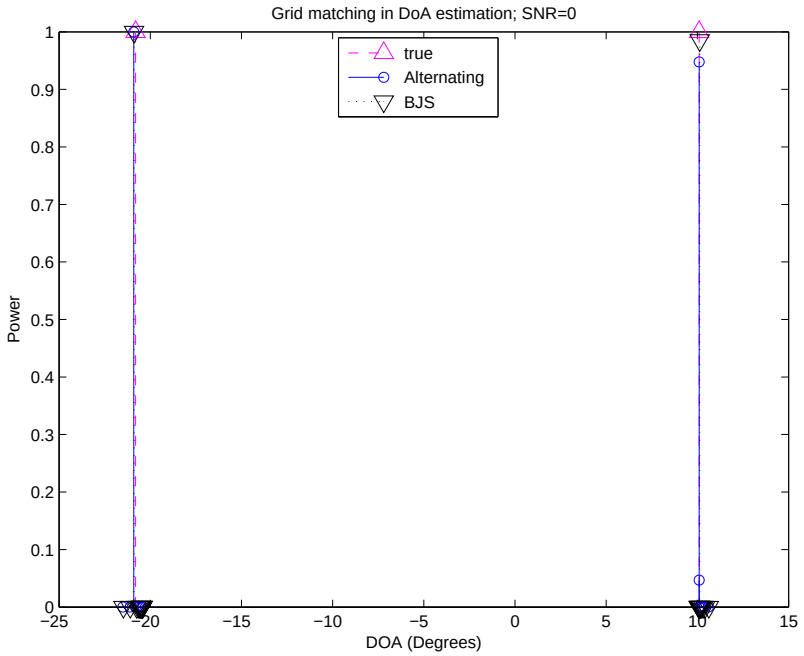


(a) DOA estimation result.

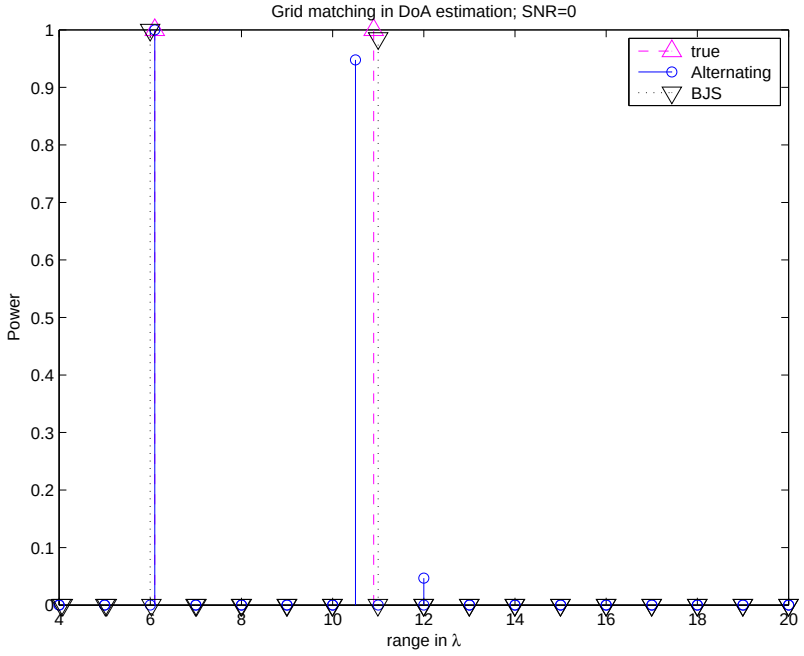


(b) Range estimation result.

Figure 5.5: DOA and range estimation for multiple sources via mismatched sensing grid. Near-field sources with DOAs: $0.3°$ and $-3.2°$, and ranges: 1.28λ and 1.52λ. SNR $= 40$ dB, and $T = 1$. The sampling grid has a resolution of $\Delta\theta = 1°$ and 0.1λ within the domain $[-4°, 3°]$ and $[1\lambda, 1.7\lambda]$.



(a) DOA estimation



(b) range estimation

Figure 5.6: Grid matching via Two-step algorithm. Two near-field source at locations: $(-20.82°, 6.1\lambda)$ and $(10.12°, 10.9\lambda)$. SNR $= 0$ dB and $T = 200$. The sampling grid has a resolution of $\Delta\theta = 1°$ and $\Delta r = 1\lambda$ within domain $[-90°, 90°]$ and $[4\lambda, 20\lambda]$.

6 Conclusions and future work

In this chapter, we summarize this thesis and provide some suggestions for the future research.

6.1 Conclusions

In this thesis, we have considered the problem of localization of near-field point sources. The classical near-field source localization problem is a non-linear (joint DOA and range) parameter estimation problem. The frequently used planar-wave assumption is no more valid for near-field sources as the wavefronts are spherical. Using the sparse representation framework, we transform the original non-linear problem into a linear ill-posed inverse problem. Based on the assumption that the spatial spectrum (i.e., the number of point sources) is sparse, we can localize the sources with a high resolution by solving an ℓ_1-regularized sparse regression.

In this thesis, we address two closely related problems in near-field source localization via sparse recovery:

1. Localize near-field sources with perfect sampling grid.

2. Localize near-field sources with mismatched sampling grid.

For the first situation, sparse recovery can localize near-field sources with single snapshot and provide a relatively accurate result. Additionally, when multiple snapshots are available and the sources are uncorrelated, the DOA and range parameters are naturally decoupled in the correlation domain. This enables us to solve two smaller dimension sparse regression problems instead of one higher dimension sparse regression problem which leads to a significant complexity reduction. Compared with 2D-MUSIC, the two-step estimator via sparse recovery method can even work in low-SNR situation regimes with high-resolution.

Grid mismatch can affect the estimation result tremendously. In this part, extended the grid-based model into EIV model to account for the grid mismatch. We have proposed two estimators based on sparse recovery and block sparse recovery for grid matching. When multiple snapshots are available, we extend the two-step estimator that we proposed in Chapter 3 to account for perturbations also. In the two-step estimator for grid matching, since in the step-1 DOA estimation we use the Fresnel approximation, the estimation result has errors both from the Fresnel approximation as well as perturbation error. Thus, we realign the DOA so as to overcome the effect of Fresnel approximation. The realigned DOA is now used in step-2 in which the range is estimated.

6.2 Suggestion for future research

Sparse recovery is now widely used in source localization but most of them are on far-field model. So we can do lots of work on this aspect. And here is some direction that can be possible to extend in the application of CS in near-field source localization:

- Automatic method to generate the induced parameter μ in Lasso. As we did in this thesis, the recovery result also relies on the choice of regularization parameter, and till now we haven't find a way to automatic generate the suitable regularization parameter, and have to find manually. Find a algorithm that can automatically generate suitable inducing parameter will improve the near-field source localization efficiently and in other field also.
- Wideband near-field source localization. Find an algorithm by using sparse recovery method to localize the wideband near-field sources with relative high performance and computational advantage.
- Using circle array to localize near-field sources. ULA is the simplest array used in array processing, and it can only localize source within domain $[-90°, 90\circ]$. Circle array can estimate omni-direction.
- In this research, we discretize the searching domain to localize the near-field source. E. J. Candes provides a super high resolution algorithm using sparse recovery but no need to discretize the searching domain. And we can avoid the effect of grid mismatch. We test this method and it could work in near-field localization but it need to find a way to achieve a super high resolution as it works in far-field domain.

Performance analysis on sparse recovery

A

A.1 The RIP and the uniqueness of sparse recovery

RIP is an important condition that provide guarantee to recovery approximated sparse vector.

Definition 1. *A matrix* $\mathbf{A}$ *satisfies the restricted isometry property (RIP) of order* K *if there exists a* σ_K[28] $\in (0,1)$ *such that*

$$(1-\sigma_K)\|\mathbf{x}\|_2^2 \leq \|\mathbf{A}\mathbf{x}\|_2^2 \leq (1+\sigma_K)\|\mathbf{x}\|_2^2 \tag{A.1}$$

holds for all $\mathbf{x} \in \Sigma_K$. *Denotes that* $\Sigma_K = \{\mathbf{x} : \|\mathbf{x}\|_0 \leq K\}$ *as the set of all* K*-sparse vector.*

The definition of RIP indicates that there exists a unique solution for the optimization problem in (2.3) if the sensing matrix $\mathbf{A}$ satisfies the RIP of order $2K$. And this can be proved as follow:

Proof. If a matrix $\mathbf{A}$ satisfies the RIP condition of order $2K$, it means $\exists \quad \sigma_{2K} \in (0,1)$ such that

$$(1-\sigma_K)\|\mathbf{h}\|_2^2 \leq \|\mathbf{A}\mathbf{h}\|_2^2 \leq (1+\sigma_K)\|\mathbf{h}\|_2^2$$

holds for all $\mathbf{h} \in \Sigma_{2K}$. $\Rightarrow \|\mathbf{A}\mathbf{h}\|_2^2 > 0$.

Assume $\exists \quad \mathbf{x}_1, \mathbf{x}_2 \in \Sigma_K$ satisfying $bfy = \mathbf{A}\mathbf{x} + \mathbf{w}$ and $\mathbf{x}_1 \neq \mathbf{x}_2$, thus we obtain

$$\mathbf{A}(\mathbf{x}_1 - \mathbf{x}_2) = 0.$$

Let $\mathbf{h} = \mathbf{x}_1 - \mathbf{x}_2$ and $\mathbf{h} \in \Sigma_{2K}$ due to $\mathbf{x}_1, \mathbf{x}_2 \in \Sigma_K$, which contradicts the fact that

$$\|\mathbf{A}\mathbf{h}\|_2^2 > 0.$$

Therefore, $\mathbf{x}_1 = \mathbf{x}_2$ and the uniqueness has been proven. □

Additionally, [24] mentioned that the solution of (2.2) is that of (2.1) if the RIP constants of $\mathbf{A}$ satisfies $\sigma_{2K} < \sqrt{2} - 1$. More details and applications on RIP can be seen in [24, 28, 29, 30, 31]. The induced parameter μ is unknown in general. The general approach to choosing μ, cross-validation, is mentioned in [32, 33, 34] as well as other approaches.

A.2 Performance of block sparse recovery

Suppose that a measurement of a ULA are corrupted by bounded noise so that

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w}$$

where $\|\mathbf{w}\|_2 \leq \epsilon$. Using the matlab tool CVX [26] to recovery $\mathbf{x}$ via the optimization problem (2.7)

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbb{C}} \quad \|\mathbf{x}\|_{2,1} \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \epsilon. \tag{A.2}$$

The B-RIP can be defined as

Definition 2. *A matrix* $\mathbf{A}$ *satisfies the block restricted isometry property (B-RIP) of order* K *if there exists a constant* $\sigma_{K|\mathcal{I}} \in (0,1)$ *such that*

$$(1-\sigma_{K|\mathcal{I}})\|\mathbf{x}\|_2^2 \leq \|\mathbf{A}\mathbf{x}\|_2^2 \leq (1+\sigma_{K|\mathcal{I}})\|\mathbf{x}\|_2^2 \tag{A.3}$$

holds for all K*-block sparse vectors* $\mathbf{x}$ *which can be divided into* N *sub-blocks,* $\mathbf{x} = [\mathbf{x}_1, \ldots, \mathbf{x}_N]$*, via* $\mathcal{I} = \{d_1, \ldots, d_N\}$*, where* $d_n = \{i : x_i \in \mathbf{x}_n\}$ *for* $n = 1, 2, \ldots, N$ *and* i *is the entry index of* $\mathbf{x}$*.*

Then the performance bound of (2.7) can be derived in the following theorem.

Theorem 1. *[21] Assume that a matrix* $\mathbf{\Phi} \in \mathbb{C}^{M\times 3N}$ *satisfies the B-RIP of order* $2K$ *with* $\sigma_{2K} < \sqrt{2}-1$ *and let* $\mathbf{y} = \mathbf{\Phi}\mathbf{q} + \mathbf{w}$ *where* $\|\mathbf{w}\|_2 \leq \epsilon$*. Then the solution* $\hat{\mathbf{q}}$ *of optimization* (2.7) *obeys*

$$\|\hat{\mathbf{q}} - \mathbf{q}\|_2 \leq \frac{c_0}{\sqrt{3}} \frac{\|\mathbf{q} - \mathbf{q}_{(K)}\|_{2,1}}{\sqrt{K}} + \frac{c_1}{\sqrt{3}}\epsilon \tag{A.4}$$

where $\mathbf{q}_{(K)}$ *indicates the best* K*-block sparse approximation to* $\mathbf{q}$*, such that* $\mathbf{q}_{(K)}$ *is* K*-block sparse and minimizes* $\|\mathbf{q} - \mathbf{d}\|_{2,1}$ *over all* K*-block sparse vectors* $\mathbf{d}$ *and both* c_0 *and* c_1 *will be given in lemma 1.*

The proof of above theorem can be found in [21].

A.3 Unbounded joint sparse recovery and its performance

In order to recovery the sparse signal in the mismatch model (5.3), it can be refer to the given optimization problem with unbounded perturbation error $e_{\bar{\theta}_n}$ and $e_{\bar{r}_n}$, the optimization problem [20] can be formulated as

$$\arg\min_{\mathbf{x}\in\mathbb{C}^N, \mathbf{e}_\theta\in\mathbb{R}^N, \mathbf{e}_r\in\mathbb{R}^N} \mu_5\|\mathbf{x}\|_1 + \|\left(\mathbf{A} + \mathbf{A}_{\boldsymbol{\theta}}\mathbf{E}_\theta + \mathbf{A}_{\mathbf{r}}\mathbf{E}_r\right)\mathbf{x} - \mathbf{y}\|_2^2 \tag{A.5}$$

with $\mathbf{E}_\theta = \mathrm{diag}(\mathbf{e}_\theta)$ and $\mathbf{E}_r = \mathrm{diag}(\mathbf{e}_r)$. As we known, the model in (5.3) can be expressed as

$$\begin{aligned} \mathbf{y} &= \overbrace{\left[\mathbf{A}\middle|\mathbf{A}_{\boldsymbol{\theta}}\middle|\mathbf{A}_{\mathbf{r}}\right]}^{\mathbf{\Phi}} \overbrace{\begin{bmatrix} \mathbf{x} \\ \mathbf{e}_\theta \odot \mathbf{x} \\ \mathbf{e}_r \odot \mathbf{x} \end{bmatrix}}^{\mathbf{q}} + \mathbf{w} \\ &= \mathbf{\Phi}\mathbf{q} + \mathbf{w}, \end{aligned} \tag{A.6}$$

where $\mathbf{\Phi} \in \mathbb{C}^{M\times 3N}$, and $\mathbf{q} \in \mathbb{C}^{3N\times 1}$ is a $3K$-sparse vector. Then the optimization (A.5) can be written in a relaxed joint sparse recovery approach

$$\arg \min_{\mathbf{x}\in\mathbb{C}^N, \mathbf{e}_\theta\in\mathbb{R}^N, \mathbf{e}_r\in\mathbb{R}^N} \mu_6\|\mathbf{q}\|_{2,1} + \|\mathbf{y} - \mathbf{\Phi}\mathbf{q}\|_2^2, \tag{A.7}$$

where

$$\|\mathbf{q}\|_{2,1} = \sum_{n=1}^{N} \sqrt{x_n^2 + (e_{\bar{\theta}_n} x_n)^2 + (e_{\bar{r}_n} x_n)^2}.$$

The performance bound for the optimization (A.7) can be derived based on the theorem 1 in the following lemma and the proof will be provided later.

Lemma 1. *Assume that a matrix $\mathbf{\Phi} \in \mathbb{C}^{M\times 3N}$ satisfies the B-RIP of order $2K$ with $\sigma_{2K} < \sqrt{2} - 1$ and let $\mathbf{y} = \mathbf{\Phi}\mathbf{q} + \mathbf{w}$ where $\|\mathbf{w}\|_2 \le \epsilon$. Then the optimal solution $\hat{\mathbf{q}}$ of the optimization* (A.7)*, where $\hat{\mathbf{q}} = [\hat{\mathbf{x}}; \hat{\mathbf{e}}_\theta \odot \hat{\mathbf{x}}; \hat{\mathbf{e}}_r \odot \hat{\mathbf{x}}]$, satisfies*

$$\|\hat{\mathbf{x}} - \mathbf{x}\|_2 \le C_1\|\mathbf{q} - \mathbf{q}_{(K)}\|_{2,1} + C_2\epsilon \tag{A.8}$$

$$\|(\hat{\mathbf{e}}_\theta - \mathbf{e}_\theta) \odot \hat{\mathbf{x}}\|_2 \le C_3\|\mathbf{q} - \mathbf{q}_{(K)}\|_{2,1} + C_4\epsilon \tag{A.9}$$

$$\|(\hat{\mathbf{e}}_r - \mathbf{e}_r) \odot \hat{\mathbf{x}}\|_2 \le C_5\|\mathbf{q} - \mathbf{q}_{(K)}\|_{2,1} + C_6\epsilon \tag{A.10}$$

where $\mathbf{q}_{(K)}$ indicates the best K-block sparse approximation to $\mathbf{q}$ and

$$C_1 = \frac{c_0}{c_2}(1 - \sigma_{2K}); \quad C_2 = \frac{c_1(1-\sigma_{2K}) + 2\sqrt{2}}{c_2}$$

$$C_3 = (1 + \frac{1-\sigma_{2K}}{\|\mathbf{A_r}\|_2})^{-1}(\frac{1}{\|\mathbf{A_r}\|_2} + \frac{\sqrt{2}}{1-\sigma_{2K}})c_3$$

$$C_4 = (1 + \frac{1-\sigma_{2K}}{\|\mathbf{A_r}\|_2})^{-1}(\frac{1}{\|\mathbf{A_r}\|_2} + \frac{\sqrt{2}}{1-\sigma_{2K}})(2 + c_3)$$

$$C_5 = (1 + \frac{1-\sigma_{2K}}{\|\mathbf{A_\theta}\|_2})^{-1}(\frac{1}{\|\mathbf{A_\theta}\|_2} + \frac{\sqrt{2}}{1-\sigma_{2K}})c_3$$

$$C_6 = (1 + \frac{1-\sigma_{2K}}{\|\mathbf{A_\theta}\|_2})^{-1}(\frac{1}{\|\mathbf{A_\theta}\|_2} + \frac{\sqrt{2}}{1-\sigma_{2K}})(2 + c_3)$$

$$c_0 = 2\sqrt{3}\frac{1-\sigma_{2K}}{1-(1+\sqrt{2})\sigma_{2K}}k^{-\frac{1}{2}}$$

$$c_1 = 4\sqrt{3}\frac{\sqrt{1+\sigma_{2K}}}{1-(1+\sqrt{2})\sigma_{2K}}$$

$$c_2 = (1 - \|\mathbf{\Delta}_\theta\|_2 - \|\mathbf{\Delta}_r\|_2)(1-\sigma_{2K}) - \sqrt{2}\|\mathbf{\Phi}\|_2(1 + \|\mathbf{\Delta}_\theta\|_2 + \|\mathbf{\Delta}_r\|_2)$$

$$c_3 = 2 + \|\mathbf{\Phi}\|_2(1 + \|\mathbf{\Delta}_\theta\|_2 + \|\mathbf{\Delta}_r\|_2)$$

with $\mathbf{\Delta}_\theta = diag(\mathbf{e}_\theta)$ and $\mathbf{\Delta}_r = diag(\mathbf{e}_r)$.

Proof. In theorem 1, we have

$$\|\hat{\mathbf{q}} - \mathbf{q}\|_2 \le \frac{c_0}{\sqrt{3}}\frac{\|\mathbf{q} - \mathbf{q}_{(K)}\|_{2,1}}{\sqrt{K}} + \frac{c_1}{\sqrt{3}}\epsilon. \tag{A.11}$$

According to Cauchy-Schwarz inequalities, we obtain

$$\begin{aligned}\|\hat{\mathbf{q}}-\mathbf{q}\|_2 &= \left\|\begin{bmatrix}\mathbf{x}-\hat{\mathbf{x}}\\ \mathbf{e}_\theta\odot\mathbf{x}-\hat{\mathbf{e}}_\theta\odot\hat{\mathbf{x}}\\ \mathbf{e}_r\odot\mathbf{x}-\hat{\mathbf{e}}_r\odot\hat{\mathbf{x}}\end{bmatrix}\right\|_2\\ &\geq \frac{1}{\sqrt{3}}\big(\|\hat{\mathbf{q}}-\mathbf{q}\|_2+\|\mathbf{e}_\theta\odot\mathbf{x}-\hat{\mathbf{e}}_\theta\odot\hat{\mathbf{x}}\|_2+\|\mathbf{e}_r\odot\mathbf{x}-\hat{\mathbf{e}}_r\odot\hat{\mathbf{x}}\|_2\big)\end{aligned}$$

And we know that

$$\begin{aligned}\|\mathbf{e}_\theta\odot\mathbf{x}-\hat{\mathbf{e}}_\theta\odot\hat{\mathbf{x}}\|_2 &= \|\mathbf{e}_\theta\odot(\mathbf{x}-\hat{\mathbf{x}})-\hat{\mathbf{x}}\odot(\mathbf{e}_\theta-\hat{\mathbf{e}}_\theta)\|_2\\ \|\mathbf{e}_r\odot\mathbf{x}-\hat{\mathbf{e}}_r\odot\hat{\mathbf{x}}\|_2 &= \|\mathbf{e}_r\odot(\mathbf{x}-\hat{\mathbf{x}})-\hat{\mathbf{x}}\odot(\mathbf{e}_r-\hat{\mathbf{e}}_r)\|_2.\end{aligned}$$

Defining $\mathbf{d}_p=\mathbf{x}-\hat{\mathbf{x}}$, $\mathbf{d}_\theta=\mathbf{e}_\theta-\hat{\mathbf{e}}_\theta$, $\mathbf{d}_r=\mathbf{e}_r-\hat{\mathbf{e}}_r$ and $\epsilon_0=\|\mathbf{q}-\mathbf{q}_{(K)}\|_{2,1}$. Using the inequality in (A.11)

$$\begin{aligned}&\|\mathbf{d}_p\|_2+\|\mathbf{e}_\theta\odot\mathbf{d}_p+\hat{\mathbf{x}}\odot\mathbf{d}_\theta\|_2+\|\mathbf{e}_r\odot\mathbf{d}_p+\hat{\mathbf{x}}\odot\mathbf{d}_r\|_2\leq c_0\epsilon_0+c_1\epsilon\\ \Rightarrow\ &\|\mathbf{d}_p\|_2(1-\|\boldsymbol{\Theta}\|_2-\|\mathbf{R}\|_2)\leq c_0\epsilon_0+c_1\epsilon+\|\hat{\mathbf{x}}\odot\mathbf{d}_\theta\|_2+\|\hat{\mathbf{x}}\odot\mathbf{d}_r\|_2\\ &\leq c_0\epsilon_0+c_1\epsilon+\sqrt{2}\left\|\begin{bmatrix}\hat{\mathbf{x}}\odot\mathbf{d}_\theta\\ \hat{\mathbf{x}}\odot\mathbf{d}_r\end{bmatrix}\right\|_2 \qquad \text{(A.12)}\end{aligned}$$

Using B-RIP on $[\mathbf{A}_{\boldsymbol{\theta}},\mathbf{A}_{\mathbf{r}}]$

$$\begin{aligned}\sqrt{1-\sigma_K([\mathbf{A}_{\boldsymbol{\theta}},\mathbf{A}_{\mathbf{r}}])}\left\|\begin{bmatrix}\hat{\mathbf{x}}\odot\mathbf{d}_\theta\\ \hat{\mathbf{x}}\odot\mathbf{d}_r\end{bmatrix}\right\|_2 &\leq \left\|[\mathbf{A}_{\boldsymbol{\theta}},\mathbf{A}_{\mathbf{theta}}]\begin{bmatrix}\hat{\mathbf{x}}\odot\mathbf{d}_\theta\\ \hat{\mathbf{x}}\odot\mathbf{d}_r\end{bmatrix}\right\|_2\\ &\leq \left\|\boldsymbol{\Phi}\begin{bmatrix}\mathbf{0}\\ \hat{\mathbf{x}}\odot\mathbf{d}_\theta\\ \hat{\mathbf{x}}\odot\mathbf{d}_r\end{bmatrix}\right\|_2\\ &\leq \|\boldsymbol{\Phi}\mathbf{q}\|_2-\left\|\boldsymbol{\Phi}\begin{bmatrix}\mathbf{d}_p\\ \mathbf{e}_\theta\odot\mathbf{d}_p\\ \mathbf{e}_r\odot\mathbf{d}_p\end{bmatrix}\right\|_2\\ &\leq 2\epsilon+\|\boldsymbol{\Phi}\|_2\|\hat{\mathbf{d}}_p\|_2(1+\|\boldsymbol{\Theta}\|_2+\|\mathbf{R}\|_2) \qquad \text{(A.13)}\end{aligned}$$

It is known that $\sigma_K([\mathbf{A}_{\boldsymbol{\theta}},\mathbf{A}_{\mathbf{r}}])<\sigma_{2K}$ and from (A.12) and (A.13) we obtain

$$\|\hat{\mathbf{x}}-\mathbf{x}\|_2\leq C_1\|\mathbf{q}-\mathbf{q}_{(K)}\|_{2,1}+C_2\epsilon. \qquad \text{(A.14)}$$

Using B-RIP on $\mathbf{A}_{\boldsymbol{\theta}}$

$$\begin{aligned}\sqrt{1-\sigma_k(\mathbf{A}_{\boldsymbol{\theta}})}\|\hat{\mathbf{x}}\odot\mathbf{d}_\theta\|_2 &\leq \left\|\boldsymbol{\Phi}\begin{bmatrix}\mathbf{0}\\ \hat{\mathbf{x}}\odot\mathbf{d}_\theta\\ \mathbf{0}\end{bmatrix}\right\|_2\\ &\leq 2\epsilon+\|\boldsymbol{\Phi}\|_2\|\hat{\mathbf{d}}_p\|_2(1+\|\boldsymbol{\Theta}\|_2+\|\mathbf{R}\|_2)\\ &\quad+\|\boldsymbol{\Phi}\|_2\|\hat{\mathbf{x}}\odot\mathbf{d}_r\|_2. \qquad \text{(A.15)}\end{aligned}$$

Combining (A.12), (A.13) and (A.15), we obtain

$$\|(\hat{\mathbf{e}}_\theta-\mathbf{e}_\theta)\odot\hat{\mathbf{x}}\|_2\leq C_3\|\mathbf{q}-\mathbf{q}_{(K)}\|_{2,1}+C_4\epsilon \qquad \text{(A.16)}$$

In the same way we can get

$$\|(\hat{\mathbf{e}}_r - \mathbf{e}_r) \odot \hat{\mathbf{x}}\|_2 \leq C_5 \|\mathbf{q} - \mathbf{q}_{(K)}\|_{2,1} + C_6\epsilon \tag{A.17}$$

Proof has been done. □

A.4 Performance analysis of bounded sparse recovery

In order to analyze the performance of the the constrained block sparse recovery, we recall the definition of B-RIP mentioned in Chapter 2. The sensing matrix $\mathbf{\Phi} \in \mathbb{C}^{M\times 3N}$ is to said obey B-RIP if there exist a constant σ_K [1] such that for every K-block sparse vector $\mathbf{q} \in \mathbb{C}^{3N}$

$$(1 - \sigma_K)\|\mathbf{q}\|_2^2 \leq \|\mathbf{\Phi}\mathbf{q}\|_2^2 \leq (1 + \sigma_K)\|\mathbf{q}\|_2^2. \tag{A.18}$$

Using this definition, the performance bound for a constrained sparse recovery can be obtained by using the techniques mentioned in [21, 29, 27]. The performance bound for (5.4) is given in the following lemma.

Lemma 2. *Assume that the B-RIP constant* $\sigma_{2K} < \left(\sqrt{2c_4} + 1\right)^{-1}$ *and* $\|\mathbf{w}\|_2 \leq \epsilon$. *Then the optimal solution* $(\hat{\mathbf{x}}, \hat{\mathbf{e}}_\theta, \hat{\mathbf{e}}_r)$ *to the optimization problem* (5.4) *satisfies*

$$\begin{aligned}
\|\hat{\mathbf{x}} - \mathbf{x}\|_2 &\leq C_7\|\mathbf{x} - \mathbf{x}^{(K)}\|_1 + C_8\epsilon && \text{(A.19)}\\
\|(\hat{\mathbf{e}}_\theta - \mathbf{e}_\theta) \odot \hat{\mathbf{x}}\|_2 &\leq C_9\|\mathbf{x} - \mathbf{x}^{(K)}\|_1 + C_{10}\epsilon && \text{(A.20)}\\
\|(\hat{\mathbf{e}}_r - \mathbf{e}_r) \odot \hat{\mathbf{x}}\|_2 &\leq C_{11}\|\mathbf{x} - \mathbf{x}^{(K)}\|_1 + C_{12}\epsilon && \text{(A.21)}
\end{aligned}$$

where $\mathbf{x}^{(K)}$ *is the best approximation to* $\mathbf{x}$ *with* K *non-zeros entries, such that* $\mathbf{x}^{(K)}$ *is* K*-sparse vector and minimizes* $\|\mathbf{x} - \mathbf{d}\|_1$ *over all* K*-sparse vectors* $\mathbf{d}$, *and*

$$\begin{aligned}
C_7 &= \left\{\frac{2(1-\sigma_{2K})}{1-(\sqrt{2c_4}+1)\sigma_{2K}}\frac{2\sqrt{k}}{1-\sigma_{2K}}\left[\frac{\sqrt{c_4}}{\sqrt{k}} + 0.5\Delta\theta + 0.5\Delta r\right] + \frac{\sqrt{2}}{\sqrt{k}}\right\}\\
C_8 &= \frac{2(1-\sigma_{2K})}{1-(\sqrt{2c_4}+1)\sigma_{2K}}\frac{2\sqrt{1+\sigma_{2K}}}{1-\sigma_{2K}}\\
C_9 &= \left(1 + \frac{\sqrt{1-\sigma_{2K}}}{\|\mathbf{A}_\mathbf{r}\|_2}\right)^{-1}\left(\frac{1}{\|\mathbf{A}_\mathbf{r}\|_2} + \frac{\sqrt{2}}{\sqrt{1-\sigma_{2K}}}\right)\sqrt{c_4}\|\mathbf{\Phi}\|_2 C_7\\
C_{10} &= \left(1 + \frac{\sqrt{1-\sigma_{2K}}}{\|\mathbf{A}_\mathbf{r}\|_2}\right)^{-1}\left(\frac{1}{\|\mathbf{A}_\mathbf{r}\|_2} + \frac{\sqrt{2}}{\sqrt{1-\sigma_{2K}}}\right)(2 + \sqrt{c_4}\|\mathbf{\Phi}\|_2 C_8)\\
C_{11} &= \left(1 + \frac{\sqrt{1-\sigma_{2K}}}{\|\mathbf{A}_\theta\|_2}\right)^{-1}\left(\frac{1}{\|\mathbf{A}_\theta\|_2} + \frac{\sqrt{2}}{\sqrt{1-\sigma_{2K}}}\right)\sqrt{c_4}\|\mathbf{\Phi}\|_2 C_7\\
C_{12} &= \left(1 + \frac{\sqrt{1-\sigma_{2K}}}{\|\mathbf{A}_\theta\|_2}\right)^{-1}\left(\frac{1}{\|\mathbf{A}_\theta\|_2} + \frac{\sqrt{2}}{\sqrt{1-\sigma_{2K}}}\right)(2 + \sqrt{c_4}\|\mathbf{\Phi}\|_2 C_8)\\
c_4 &= 1 + 0.25\Delta\theta^2 + 0.25\Delta r^2
\end{aligned}$$

[1] In the rest thesis, we use σ_K for the block-RIP constant instead of $\sigma_{K|\mathcal{I}}$ used in Chapter 2 unless we declare

In the above lemma, the performance bound for the box-constrained sparse recovery relies on the noise level and resolution of the sampling grid. The proof is given below:

Proof. As we known, the optimization (5.4) is equivalent to optimization

$$\min_{\mathbf{x}\in\mathbb{C}^N,\mathbf{e}_\theta\in\mathbb{R}^N,\mathbf{e}_r\in\mathbb{R}^N} \quad \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \|\mathbf{y}-\mathbf{\Phi}\mathbf{q}\| \leq \epsilon. \tag{A.22}$$

Denote by $\hat{\mathbf{x}} = \mathbf{x} + \mathbf{h}$ the solution to (5.4) or (A.22). And we decompose $\mathbf{h}$ as $\mathbf{h} = \sum_{i=0}^{l-1} \mathbf{h}_{\mathcal{I}_i}$, where $\mathbf{h}_{\mathcal{I}_i}$ is the restriction of $\mathbf{h}$ to the set $\mathcal{I}_i$ which consists of K entries. We define $\mathcal{I}_0$ as the set of indices for which $\mathbf{x}$ is the K largest entries (since the noise $\mathbf{x}$ is not K-sparse but approximately). Chosen such that the norm of $\mathbf{h}_{\mathcal{I}_1}$ over $\mathcal{I}_1$ is the largest, the norm over $\mathcal{I}_2$ is the second largest and so on. To find the bound of $\mathbf{h}$, we need use the inequalities below

$$\|\mathbf{h}\|_2 = \|\mathbf{h}_{\mathcal{I}_0\cup\mathcal{I}_1} + \mathbf{h}_{(\mathcal{I}_0\cup\mathcal{I}_1)^c}\|_2 \leq \|\mathbf{h}_{\mathcal{I}_0\cup\mathcal{I}_1}\|_2 + \|\mathbf{h}_{(\mathcal{I}_0\cup\mathcal{I}_1)^c}\|_2. \tag{A.23}$$

- Bound on $\|\mathbf{h}_{(\mathcal{I}_0\cup\mathcal{I}_1)^c}\|_2$
  According to Cauchy inequalities, we develop that

$$\|\mathbf{h}_{(\mathcal{I}_0\cup\mathcal{I}_1)^c}\|_2 \leq \frac{1}{\sqrt{K}}\sum_{i=1}^{l-1}\|\mathbf{h}_{\mathcal{I}_i}\|_1 = \frac{1}{\sqrt{K}}\|\mathbf{h}_{\mathcal{I}_0^c}\|_1 \tag{A.24}$$

$$\|\mathbf{h}_{\mathcal{I}_0^c}\|_1 \leq \|\mathbf{h}_{\mathcal{I}_0}\|_1 \leq \sqrt{K}\|\mathbf{h}_{\mathcal{I}_0}\|_2 \tag{A.25}$$

  And according to the optimization (A.22) and $\mathcal{I}_i \cap \mathcal{I}_j = \varnothing$ where $i \neq j$, we obtained

$$\begin{aligned}
\|\mathbf{x}\|_1 &\geq \|\hat{\mathbf{x}}\|_1 = \|\mathbf{x}_{\mathcal{I}_0} + \mathbf{h}_{\mathcal{I}_0}\|_1 + \|\mathbf{x}_{\mathcal{I}_0^c} + \mathbf{h}_{\mathcal{I}_0^c}\|_1 \\
&\geq \|\mathbf{x}_{\mathcal{I}_0}\|_1 - \|\mathbf{h}_{\mathcal{I}_0}\|_1 + \|\mathbf{h}_{\mathcal{I}_0^c}\|_1 - \|\mathbf{x}_{\mathcal{I}_0^c}\|_1 \\
\|\mathbf{x}_{\mathcal{I}_0} + \mathbf{x}_{\mathcal{I}_0^c}\|_1 &\geq \|\mathbf{x}_{\mathcal{I}_0}\|_1 - \|\mathbf{h}_{\mathcal{I}_0}\|_1 + \|\mathbf{h}_{\mathcal{I}_0^c}\|_1 - \|\mathbf{x}_{\mathcal{I}_0^c}\|_1 \\
\Rightarrow \quad \|\mathbf{h}_{\mathcal{I}_0^c}\|_1 &\leq 2\|\mathbf{x}_{\mathcal{I}_0^c}\|_1 + \|\mathbf{h}_{\mathcal{I}_0}\|_1
\end{aligned}$$

  Substitute into (A.24) and invoke the the fact that $\mathbf{x} = \mathbf{x}_{\mathcal{I}_0^c} + \mathbf{x}_{\mathcal{I}_0}$ and (A.25), we conclude that

$$\|\mathbf{h}_{(\mathcal{I}_0\cup\mathcal{I}_1)^c}\|_2 \leq \frac{2}{\sqrt{K}}\|\mathbf{x}-\mathbf{x}_{(K)}\|_1 + \|\mathbf{h}_{\mathcal{I}_0}\|_2 \leq \frac{2}{\sqrt{K}}e_0 + \|\mathbf{h}_{(\mathcal{I}_0\cup\mathcal{I}_1)^c}\|_2 \tag{A.26}$$

  where $e_0 = \|\mathbf{x}-\mathbf{x}_{(K)}\|_1$.

- Bound on $\|\mathbf{h}_{\mathcal{I}_0\cup\mathcal{I}_1}\|_2$
  To bound $\|\mathbf{h}_{\mathcal{I}_0\cup\mathcal{I}_1}\|_2$, we invoke the relationship with $\|\mathbf{f}_{\mathcal{I}_0\cup\mathcal{I}_1}\|_2$. Define $\mathbf{f} = \hat{\mathbf{q}} - \mathbf{q}$, and $\mathbf{f}_{\mathcal{I}_i}$ indicates the K blocks sparse vector. According to the fact that $\mathbf{\Phi}\mathbf{f}_{\mathcal{I}_0\cup\mathcal{I}_1} = \mathbf{\Phi}\mathbf{f} - \sum_{i=2}^{l-1}\mathbf{\Phi}\mathbf{f}_{\mathcal{I}_i}$, we obtain that

$$\|\mathbf{\Phi}\mathbf{f}_{\mathcal{I}_0\cup\mathcal{I}_1}\|_2^2 = \langle\mathbf{\Phi}\mathbf{f}_{\mathcal{I}_0\cup\mathcal{I}_1}, \mathbf{\Phi}\mathbf{f}\rangle - \sum_{i=2}^{l-1}\langle\mathbf{\Phi}\mathbf{f}_{\mathcal{I}_0\cup\mathcal{I}_1}, \mathbf{\Phi}\mathbf{f}_{\mathcal{I}_i}\rangle \tag{A.27}$$

According to the definition of B-RIP, we obtained that

$$|\langle \mathbf{\Phi f}_{\mathcal{I}_0\cup\mathcal{I}_1}, \mathbf{\Phi f}\rangle| \leq \|\mathbf{\Phi f}_{\mathcal{I}_0\cup\mathcal{I}_1}\|_2 \|\mathbf{\Phi f}\|_2 \leq \sqrt{1+\sigma_{2K}}\|\mathbf{f}_{\mathcal{I}_0\cup\mathcal{I}_1}\|_2 \|\mathbf{\Phi f}\|_2 \tag{A.28}$$

Since $\mathbf{f} = \hat{\mathbf{q}} - \mathbf{q}$ and both $\hat{\mathbf{q}}$ and $\mathbf{q}$ are feasible

$$\|\mathbf{\Phi f}\|_2 = \|\mathbf{\Phi}(\hat{\mathbf{q}} - \mathbf{q})\| \leq \|\mathbf{\Phi}\hat{\mathbf{q}} - \mathbf{x}\|_2 + \|\mathbf{\Phi q} - \mathbf{x}\|_2 \leq 2\epsilon \tag{A.29}$$

Then,

$$|\langle \mathbf{\Phi f}_{\mathcal{I}_0\cup\mathcal{I}_1}, \mathbf{\Phi f}\rangle| \leq 2\epsilon\sqrt{1+\sigma_{2K}}\|\mathbf{f}_{\mathcal{I}_0\cup\mathcal{I}_1}\|_2 \tag{A.30}$$

According to the Lemma 2.1 in [29], we have

$$\begin{aligned} |\sum_{i=2}^{l-1}\langle \mathbf{\Phi f}_{\mathcal{I}_0\cup\mathcal{I}_1}, \mathbf{\Phi f}_{\mathcal{I}_i}\rangle| &\leq \sum_{i=2}^{l-1}\sigma_{2K}\left(\|\mathbf{f}_{\mathcal{I}_0}\|_2 + \|\mathbf{f}_{\mathcal{I}_1}\|_2\right)\|\mathbf{f}_{\mathcal{I}_i}\|_2 \\ &\leq \sqrt{2}\sigma_{2K}\|\mathbf{f}_{\mathcal{I}_0\cup\mathcal{I}_1}\|_2 \sum_{i=2}^{l-1}\|\mathbf{f}_{\mathcal{I}_i}\|_2 \end{aligned} \tag{A.31}$$

Notice that

$$\begin{aligned} \mathbf{f} &= \begin{bmatrix} \mathbf{h} \\ \mathbf{e}_\theta \odot \mathbf{h} + \hat{\mathbf{x}} \odot \mathbf{d}_\theta \\ \mathbf{e}_r \odot \mathbf{h} + \hat{\mathbf{x}} \odot \mathbf{d}_r \end{bmatrix} \\ \|\mathbf{f}_{\mathcal{I}_i}\|_2 &= \left\| \begin{bmatrix} \mathbf{h}_{\mathcal{I}_i} \\ \mathbf{e}_\theta \odot \mathbf{h}_{\mathcal{I}_i} \\ \mathbf{e}_r \odot \mathbf{h}_{\mathcal{I}_i} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \hat{\mathbf{x}}_{\mathcal{I}_i} \odot \mathbf{d}_\theta \\ \hat{\mathbf{x}}_{\mathcal{I}_i} \odot \mathbf{d}_r \end{bmatrix} \right\|_2 \\ &\leq \sqrt{c_4}\|\mathbf{h}_{\mathcal{I}_i}\|_2 + 2(0.5\Delta\theta + 0.5\Delta r)\|\hat{\mathbf{x}}_{\mathcal{I}_i}\|_2 \\ \text{where} \quad c_4 &= 1 + 0.25\Delta\theta^2 + 0.25\Delta r^2. \end{aligned} \tag{A.32}$$

Meanwhile

$$\sum_{i=2}^{l-1}\|\hat{\mathbf{x}}_{\mathcal{I}_i}\|_2 \leq \sum_{i=2}^{l-1}\|\hat{\mathbf{x}}_{\mathcal{I}_i}\|_1 = \|\hat{\mathbf{x}}_{(\mathcal{I}_0\cup\mathcal{I}_1)^c}\|_1 \leq e_0 \tag{A.33}$$

Substitue (A.32) and (A.33) into (A.31), we obtain

$$\begin{aligned} |\sum_{i=2}^{l-1}\langle \mathbf{\Phi f}_{\mathcal{I}_0\cup\mathcal{I}_1}, \mathbf{\Phi f}_{\mathcal{I}_i}\rangle| &\leq \sqrt{2}\sigma_{2K}\|\mathbf{f}_{\mathcal{I}_0\cup\mathcal{I}_1}\|_2(\sqrt{c_4}\|\mathbf{h}_{(\mathcal{I}_0\cup\mathcal{I}_1)^c}\|_2 \\ &\quad +2(0.5\Delta\theta + 0.5\Delta r)\|e_0) \end{aligned} \tag{A.34}$$

Substitute (A.30) and (A.34) into (A.27), and invoke the B-RIP we obtain

$$\begin{aligned} (1-\sigma_{2K})\|\mathbf{f}_{\mathcal{I}_0\cup\mathcal{I}_1}\|_2^2 &\leq \|\mathbf{\Phi f}_{\mathcal{I}_0\cup\mathcal{I}_1}\|_2^2 \\ &\leq \|\mathbf{f}_{\mathcal{I}_0\cup\mathcal{I}_1}\|_2 \Bigg\{ 2\epsilon\sqrt{1+\sigma_{2K}} + \sigma_{2K}\sqrt{2c_4}\|\mathbf{h}_{\mathcal{I}_0\cup\mathcal{I}_1}\|_2 \\ &\quad +2\sqrt{2}\left[\frac{\sqrt{c_4}}{\sqrt{K}} + 0.5\Delta\theta + 0.5\Delta r\right] \Bigg\} \end{aligned} \tag{A.35}$$

and as we known that $\|\mathbf{h}_{\mathcal{I}_0\cup\mathcal{I}_1}\|_2 \leq \|\mathbf{f}_{\mathcal{I}_0\cup\mathcal{I}_1}\|_2$, we obtain the bound of $\|\mathbf{h}_{\mathcal{I}_0\cup\mathcal{I}_1}\|_2$ as below

$$\begin{aligned}\|\mathbf{h}_{\mathcal{I}_0\cup\mathcal{I}_1}\|_2 \;\leq\; & \frac{2(1-\sigma_{2K})}{1-(\sqrt{2c_4}+1)\sigma_{2K}}\left\{\frac{2\sqrt{1+\sigma_{2K}}}{1-\sigma_{2K}}\epsilon\right. \\ & \left.+\frac{2\sqrt{K}}{1-\sigma_{2K}}\left[\frac{\sqrt{c_4}}{\sqrt{K}}+0.5\Delta\theta+0.5\Delta r\right]e_0\right\} \qquad \text{(A.36)}\end{aligned}$$

- Bound on $\|\mathbf{h}\|_2$
  Substitute the bound of $\|\mathbf{h}_{\mathcal{I}_0\cup\mathcal{I}_1}\|_2$ and $\|\mathbf{h}_{(\mathcal{I}_0\cup\mathcal{I}_1)^c}\|_2$ into (A.23), we got

$$\|\mathbf{h}\|_2 = \|\hat{\mathbf{x}}-\mathbf{x}\|_2 \leq C_7\|\mathbf{x}-\mathbf{x}_{(K)}\|_1 + C_8\epsilon \qquad \text{(A.37)}$$

- Bound on $\|(\hat{\mathbf{e}}_\theta-\mathbf{e}_\theta)\odot\hat{\mathbf{x}}\|_2$ and $\|(\hat{\mathbf{e}}_r-\mathbf{e}_r)\odot\hat{\mathbf{x}}\|_2$
  By using the similar method we used in Appendix A.2, we could get

$$\begin{aligned}\|(\hat{\mathbf{e}}_\theta-\mathbf{e}_\theta)\odot\hat{\mathbf{x}}\|_2 &\leq C_9\|\mathbf{x}-\mathbf{x}_{(K)}\|_1 + C_{10}\epsilon \qquad \text{(A.38)}\\ \|(\hat{\mathbf{e}}_r-\mathbf{e}_r)\odot\hat{\mathbf{x}}\|_2 &\leq C_{11}\|\mathbf{x}-\mathbf{x}_{(K)}\|_1 + C_{12}\epsilon \qquad \text{(A.39)}\end{aligned}$$

The last, to make sure all the constant will be larger than zero, $\frac{2(1-\sigma_{2K})}{1-(\sqrt{2c_4}+1)\sigma_{2K}}$ should be larger than zero, thus, $\sigma_{2K} \leq \left(\sqrt{2c_4}+1\right)^{-1}$. □

According to lemma 2, SNR and grid offset error (GoE) are the two important factor that influence the performance of the sparse recovery algorithm in the mismatched situation. Here, we analyze performance of RBJS in different SNR and GoE, respectively by invoking Monte-Carlo simulation for 100 runs for each parameter that is used for comparison. The normalized-root-mean-squared-error (NRMSE) of the estimate is used as a measure of performance and is given by

$$E_\theta = \sqrt{\mathbb{E}\{|\theta-\hat{\theta}|^2\}}/\Delta\theta \times 100\% \qquad \text{(A.40)}$$

$$E_r = \sqrt{\mathbb{E}\{|r-\hat{r}|^2\}}/\Delta r \times 100\% \qquad \text{(A.41)}$$

where θ and r are the true location parameters of the sources and $\hat{\theta}$ and $\hat{r}$ are the estimated results. We calculate the NRMSE for different signal-to-noise ratios (SNRs) and grid offset error (GoE), which is the ratio of perturbation to the size of grid cell in percentage. The SNR is ranging from 20 to 65 dB. The perturbation parameters $\mathbf{e}_\theta$ ranges from 0.1° to 0.45° and $\mathbf{e}_r$ ranges from -0.045λ to -0.005λ. Define the grid offset error in DOA domain GoE_θ as

$$\text{GoE}_\theta = e_\theta/\Delta\theta \times 100\%$$

and the grid offset error in range domain GoE_r as

$$\text{GoE}_r = e_r/\Delta r \times 100\%,$$

e.g., if the size of grid cell is $1^\circ \times 0.1\lambda$, then the GoE_θ and GoE_r of a pair of perturbation, $e_\theta = 0.1^\circ$ and $e_r = 0.005\lambda$, are 10% and 5% respectively. In Fig.A.1, the source is

located at fixed location $(0.3°, 1.28\lambda)$, and shows the performance of the R-BJS for different SNR. In the R-BJS algorithm, the NRMSE is decreasing as SNR increasing. In Fig.A.2, the SNR is fixed in 40 dB and shows the performance with different GoE_θ and GoE_r. The performance of R-BJS becomes worse as the GoE increasing and the results agree with lemma 2.

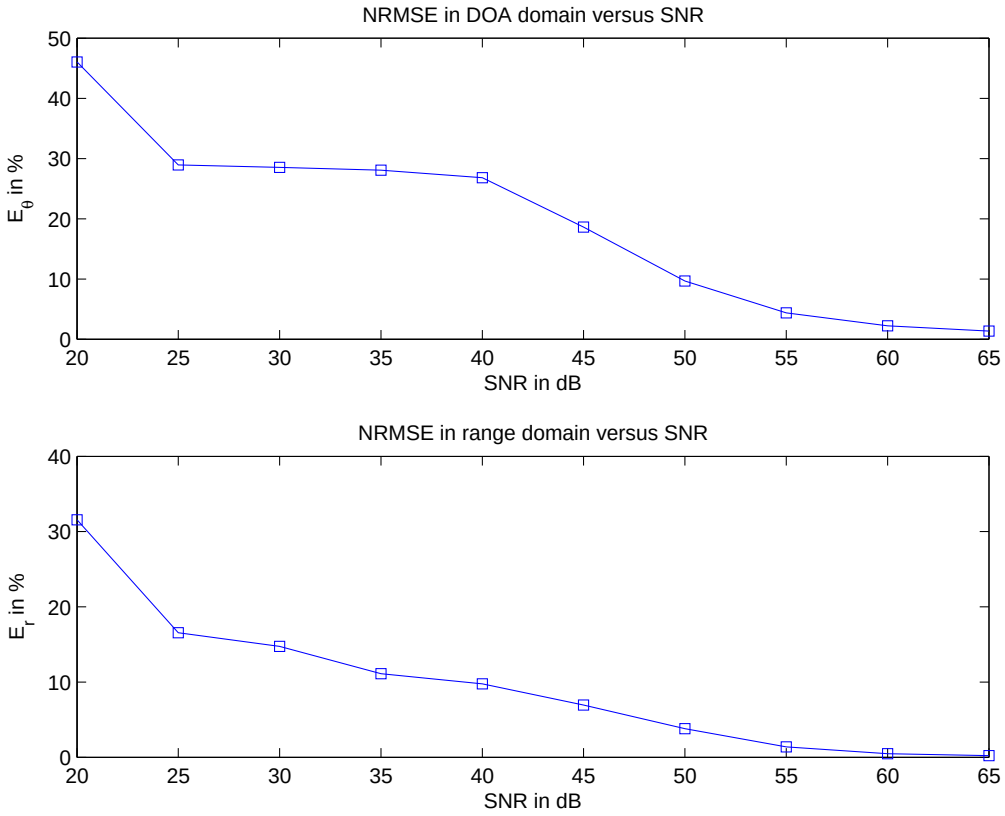


Figure A.1: Performance of R-BJS versus SNR. A fixed single unit source located at $(0.3°, 1.28\lambda)$ with different SNR. SNR ranges from 20 dB to 65 dB. The sampling grid has a resolution of $\Delta\theta = 1°$ and 0.1λ within the domain $[-4°, 3°]$ and $[1\lambda, 1.7\lambda]$.

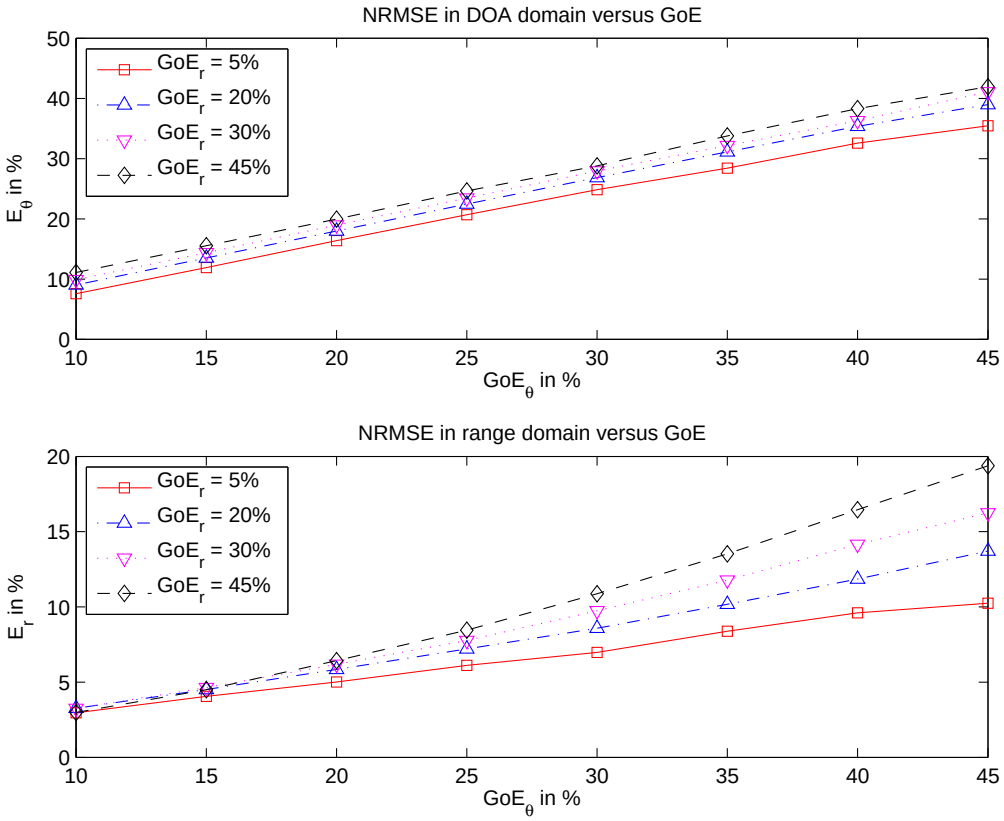


Figure A.2: Performance of R-BJS versus GoE. A unit amplitude single source located close to the grid cell $(0°, 1.3\lambda)$ with different perturbation parameters, $\mathbf{e}_\theta$ ranges from $0.1°$ to $0.45°$ and $\mathbf{e}_r$ ranges from -0.045λ to -0.005λ. SNR is 40 dB and $T = 1$. The sampling grid has a resolution of $\Delta\theta = 1°$ and 0.1λ within the domain $[-4°, 3°]$ and $[1\lambda, 1.7\lambda]$.

Bibliography

[1] Christopher R. Clarkson, "Integration of rate-transient and microseismic analysis for unconventional gas reservoirs: Where reservoir engineering meets geophysics," http://csegrecorder.com/articles/view/integration-of-rate-transient-and-microseismic-analysis, 2011.

[2] Petre Stoica and Kenneth C. Sharman, "Maximum likelihood methods for direction-of-arrival estimation," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 38, no. 7, pp. 1132–1143, July 1990.

[3] Hamid Krim and Mats Viberg, "Two decades of array signal processing research: the parametric approach," *Signal Processing Magazine, IEEE*, vol. 13, no. 4, pp. 67–94, 1996.

[4] Ralph O. Schmidt, "Multiple emitter location and signal parameter estimation," *Antennas and Propagation, IEEE Transactions on*, vol. 34, no. 3, pp. 276–280, March 1986.

[5] Petre Tichavský, Kainam Thomas Wong, and Michael D. Zoltowski, "Near-field/far-field azimuth and elevation angle estimation using a single vector hydrophone," *Signal Processing, IEEE Transactions on*, vol. 49, no. 11, pp. 2498–2510, November 2001.

[6] Güner Arslan and F. Ayhan Sakarya, "A unified neural-network-based speaker localization technique," *Neural Networks, IEEE Transactions on*, vol. 11, no. 4, pp. 997–1002, July 2000.

[7] Yung-Dar Huang and Mourad Barkat, "Near-field multiple source localization by passive sensor array," *Antennas and Propagation, IEEE Transactions on*, vol. 39, no. 7, pp. 968–974, July 1991.

[8] Ning Lv, Ju Liu, Qian Wang, and Jun Du, "An efficient second-order method of near-field source localization and signal recovery," in *Communication Technology (ICCT), 2011 IEEE 13th International Conference on*, September 2011, pp. 780–784.

[9] A. L. Swindlehurst and T. Kailath, "Passive direction-of-arrival and range estimation for near-field sources," in *Spectrum Estimation and Modeling, 1988., Fourth Annual ASSP Workshop on*, 1988, pp. 123–128.

[10] Karim Abed-Meraim, Yingbo Hua, and Adel Belouchrani, "Second-order near-field source localization: algorithm and performance analysis," in *Signals, Systems and Computers, 1996. Conference Record of the Thirtieth Asilomar Conference on*, November 1996, vol. 1, pp. 723–727.

[11] Yuntao Wu, Yi Dong, and Guisheng Liao, "Joint estimation both range and doa of near-field source," *Journal of Electronics (China)*, vol. 21, no. 2, pp. 104–109, March 2004.

[12] Richard G. Baraniuk, "Compressive sensing," *IEEE Signal Processing Magazine*, pp. 118–120,124, June 2007.

[13] B. Anila Satheesh, B. Deepa, Sbuhadra Bhai, and S Anjona Devi, "Compressive sensing for array signal processing," in *India Conference (INDICON), 2012 Annual IEEE*, December 2012, pp. 555 – 560.

[14] Dmitry Malioutov, Mujdat Cetin, and Alan S.Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *Signal Processing, IEEE Transactions on*, vol. 53, no. 8, August 2005.

[15] Fulai Liu, Lu Peng, M. Wei, P. Chen, and S. Guo, "An improved l1-svd algorithm based on noise subspace for doa estimation," *Progress In Electromagnetics Research C*, vol. 29, pp. 109–122, 2012.

[16] Yuejie Chi, Louis L. Scharf, Ali Pezeshki, and A. Robert Calderbank, "Sensitivity to basis mismatch in compressed sensing," *Signal Processing, IEEE Transactions on*, vol. 59, no. 5, pp. 2182–2195, May 2011.

[17] Hao Zhu, Geert Leus, and Georgio B. Giannakis, "Sparsity-cognizant total least-squares for perturbed compressive sampling," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2002–2016, May 2011.

[18] Rakshith Jagannath, Geert Leus, and Radmila Pribić, "Grid matching for sparse signal recovery in compressive sensing," in *Radar Conference (EuRAD), 2012 9th European*, Oct 2012, pp. 111–114.

[19] Rakshith Jagannath and K. V.S Hari, "Block sparse estimator for grid matching in single snapshot doa estimation," *Signal Processing Letters, IEEE*, vol. 20, no. 11, pp. 1038–1041, November 2013.

[20] Tan Zhao, Peng Yang, and Arye Nehorai, "Joint-sparse recovery in compressed sensing with dictionary mismatch," in *Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), 2013 IEEE 5th International Workshop on*, December 2013, pp. 248–251.

[21] Yonina C. Eldar and Moshe Mishali, "Robust recovery of signals from a structured union of subspaces," *Information Theory, IEEE Transactions on*, vol. 55, no. 11, pp. 5302 – 5316, November 2009.

[22] JunHong Lin and Song Li, "Block sparse recovery via mixed ℓ_2/ℓ_1 minimizaton," *Acta Mathematica Sinica, English series*, vol. 29, no. 7, pp. 1401–1412, July 2013.

[23] Robert Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B*, vol. 58, no. 1, pp. 267–288, 1996.

[24] Emmanuel J. Candes and Terence Tao, "Decoding by linear programming," *Information Theory, IEEE Transactions on*, vol. 51, no. 12, pp. 4203–4215, December 2005.

[25] Arian Maleki and David L. Donoho, "Optimally tuned iterative reconstruction algorithms for compressed sensing," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 4, no. 2, pp. 330–341, 2010.

[26] [Online], "CVX: Matlab software for disciplined convex programming, version 2.0 beta," http://cvxr.com/cvx, Sep. 2012.

[27] Zai Yang, Cishen Zhang, and Lihua Xie, "Robustly stable signal recovery in compressed sensing with structured matrix perturbation," *Signal Processing, IEEE Transactions on*, vol. 60, no. 9, pp. 4658–4671, September 2012.

[28] Bubacarr Bah, "Restricted isometry constant in compressed sensing," http://www.maths.ed.ac.uk/pg/thesis/bah.pdf, 2011.

[29] Emmanuel J Candès, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 9, pp. 589–59, February 2008.

[30] Emmanuel J Candes, Justin K Romberg, and Terence Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on pure and applied mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.

[31] Fan Yang, Shengqian Wang, and Chengzhi Deng, "Compressive sensing of image reconstruction using multi-wavelet transforms," in *Intelligent Computing and Intelligent Systems (ICIS), 2010 IEEE International Conference on*, 2010, vol. 1, pp. 702 – 705.

[32] Yonina C. Eldar, "Generalized sure for exponential families: Applications to regularization," *Signal Processing, IEEE Transactions on*, vol. 57, no. 2, pp. 471–481, 2009.

[33] Jerome Friedman, Trevor Hastie, and Rob Tibshirani, "Regularization paths for generalized linear models via coordinate descent," *Journal of statistical software*, vol. 33, no. 1, pp. 1–22, 2010.

[34] Gene H. Golub, Michael Heath, and Grace Wahba, "Generalized cross validation as a method for choosing a good ridge parameter," *Technometrics*, vol. 21, no. 2, pp. 215–223, 1979.