



M.Sc. Thesis

Acoustic array processing with planar coded cover

Yuan Yuan

Abstract

While the success of improving direction of arrival (DOA) estimation with linear coded covers using a single acoustic vector sensors (AVS) has been established, the extension of this theory to array-based systems remains unexplored. To address this gap, we employ a specially designed coded cover and leverage compressed sensing (CS) and compressed covariance sensing (CCS) methods, extending their application from single AVS systems to an array-based acoustic measurement system. Our results demonstrate that a 14×10 coded cover with 12 PU probes enables accurate localization of 100 sound sources in 3D, even at a signal to noise ratio (SNR) as low as 10 dB, showcasing the scalability and robustness of this approach.

To further enhance localization accuracy, we implement a self-calibration method in the covariance domain to correct phase and gain errors in each receiving channel. Additionally, we combine self-calibration with CCS to improve resolution and reduce side lobes. For geometric mismatch, we first investigate the sparsity-cognizant total least-squares (STLS) with multiple measurement vectors (MMV) variant of the fast iterative shrinkage-thresholding algorithm (FISTA) method. Then a grid-searching strategy is employed to compensate for these mismatches, ensuring better estimation accuracy. Experimental validation confirms that these techniques significantly enhance DOA estimation under non-ideal conditions, contributing to the advancement of acoustic sensing and localization methodologies.

Acoustic array processing with planar coded cover

THESIS

submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE

in

ELECTRICAL ENGINEERING

by

Yuan Yuan
born in Jiangsu, China

This work was performed in:

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



Delft University of Technology

Copyright © 2025 Circuits and Systems Group
All rights reserved.

DELFT UNIVERSITY OF TECHNOLOGY
DEPARTMENT OF
MICROELECTRONICS

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 “**Acoustic array processing with planar coded cover**” by **Yuan Yuan** in partial fulfillment of the requirements for the degree of **Master of Science**.

Dated: My Graduation Date

Chairman:

prof.dr.ir. Geert Leus

Advisor:

dr. Dani Fernandez Comesaña

Committee Members:

dr.ir. J.N. Driessen

Abstract

While the success of improving direction of arrival (DOA) estimation with linear coded covers using a single acoustic vector sensors (AVS) has been established, the extension of this theory to array-based systems remains unexplored. To address this gap, we employ a specially designed coded cover and leverage compressed sensing (CS) and compressed covariance sensing (CCS) methods, extending their application from single AVS systems to an array-based acoustic measurement system. Our results demonstrate that a 14×10 coded cover with 12 PU probes enables accurate localization of 100 sound sources in 3D, even at a signal to noise ratio (SNR) as low as 10 dB, showcasing the scalability and robustness of this approach.

To further enhance localization accuracy, we implement a self-calibration method in the covariance domain to correct phase and gain errors in each receiving channel. Additionally, we combine self-calibration with CCS to improve resolution and reduce side lobes. For geometric mismatch, we first investigate the sparsity-cognizant total least-squares (STLS) with multiple measurement vectors (MMV) variant of the fast iterative shrinkage-thresholding algorithm (FISTA) method. Then a grid-searching strategy is employed to compensate for these mismatches, ensuring better estimation accuracy. Experimental validation confirms that these techniques significantly enhance DOA estimation under non-ideal conditions, contributing to the advancement of acoustic sensing and localization methodologies.

Acknowledgments

I would like to express my sincere gratitude to my university supervisor, Prof. Geert Leus, for giving me the opportunity to undertake this research. His guidance and insightful suggestions have been invaluable throughout this journey. His support has not only helped me navigate challenges but also deepened my understanding of the topic.

I am also deeply grateful to my company supervisor, Dani, for his continuous support and encouragement. His guidance has inspired me to explore new directions while allowing me the freedom to develop my ideas independently. Beyond the technical aspects, he has also helped me integrate smoothly into the company culture, making my experience both enriching and enjoyable.

A special thanks to my colleague Krishna, whose valuable suggestions and insights have broadened my perspective beyond my immediate tasks. His advice has helped me think more critically and holistically about my work.

I would also like to extend my appreciation to all my colleagues at Microflown, who have supported me throughout my time at the company. Their willingness to help and share their knowledge has made my experience both productive and enjoyable.

Most importantly, I want to express my deepest gratitude to my mother and my late father. My mother's unwavering love, strength, and support have been the foundation of my journey. No matter the distance, she has always been there for me, encouraging me through every challenge and celebrating every achievement. My father, though no longer with us, has always been my guiding light. I know he was proud of me, and I believe he is still watching over me, supporting me in his own way. I would not have reached this milestone without them.

It has been an incredible journey, and I feel truly fortunate to have had the support of so many wonderful people along the way.

Yuan Yuan
Delft, The Netherlands
My Graduation Date

Contents

Abstract	v
Acknowledgments	vii
1 Introduction	1
1.1 Acoustic Sound Field Measurements	1
1.1.1 PU probe	1
1.1.2 Array grids	4
1.2 Motivation of the project	5
1.3 Outline	6
2 CCS-Based Localization on RECT-7.5 with a Coded Mask	9
2.1 1D DOA Estimation Using Compressed Sampling	11
2.2 1D DOA Estimation Using Compressive Covariance Sensing	13
2.3 Simulation Results of 1D DOA Estimation with CCS	15
2.4 2D DOA Estimation Using CS method	16
2.5 2D DOA Estimation Using CCS method	23
2.5.1 Structured Recovery Using Two Independent Observations	24
2.5.2 Full Data-Based Recovery	25
2.6 Simulation Results of CCS in 3D source localization	26
2.7 Conclusion	29
3 Self-Calibration of Planar PU Array with Compression	31
3.1 Data Model with Uncalibrated Sensors	32
3.1.1 Self-calibration	32
3.1.2 Ambiguity and Identifiability	33
3.1.3 Self-Calibration Algorithm	37
3.1.4 Simulation Results of Self-calibration Algorithm	38
3.2 Resolution-Enhanced Self-Calibration with CCS Method	45
3.3 Conclusion	45
4 Geometric Mismatch Correction	47
4.1 The Impacts and Categories of Geometric Mismatch	47
4.2 STLS-MMV Method with Perturbation in the Compression Matrix	48
4.2.1 First step to update signal matrix	49
4.3 Simulation Results for STLS with the MMV-FISTA Method	52
4.4 Grid Search Method for Geometric Mismatching with calibrated sensors	63
4.5 Simulation Results of Grid Search Method for Geometric Mismatching with Calibrated Sensors	63
4.6 Self-Calibration integrated Grid Search Method for Geometric Mismatching with uncalibrated sensors	65

4.7	Simulation Results of Self-Calibration integrated Grid Search Method with uncalibrated sensors	67
4.8	Conclusion	68
5	Experimental Results	71
5.1	3D Printing of the Coded Cover and Assembly of the Protective Enclosure	71
5.2	Experimental Setup	73
5.3	Measurement Results	74
	5.3.1 Calibrated Sensors	76
	5.3.2 Uncalibrated Sensors	77
5.4	Conclusion	78
6	Conclusions and Future Directions	81
6.1	Conclusion	81
6.2	Future Directions for Research	81

List of Figures

1.1	Hot-wire anemometer principle-based particle velocity transducer. . . .	2
1.2	PU Regular	2
1.3	Configuration of a single PU probe and one sound source	3
1.4	Measurement model of a PU array	5
1.5	Family of Hand-Held Array with PU Probes	5
1.6	The model of a single 2D AVS with a specially designed coded cover [6]	6
2.1	Measurement model of a PU array with coded cover	9
2.2	Beam pattern of an array with 12 PUs with coded cover	12
2.3	MUSIC Localization of 23 2D-sources by using CS method with 12 PUs at SNR = 20dB	13
2.4	MUSIC Localization of 61 2D-sources by using CCS method with 12 PUs at SNR = 20dB	16
2.5	MUSIC Localization of 41 2D-sources by using CCS method with 12 P/Us at SNR = 20dB	16
2.6	MUSIC Localization of 61 2D-sources by using CCS method with 12 P/Us at SNR = 20dB	17
2.7	3D source localization model	18
2.8	Beam pattern of CS method with the target at $(70^\circ, 140^\circ)$	20
2.9	MUSIC localization of 23 sources with CS method when SNR = 20dB .	20
2.10	The details of localization results from the CS method	21
2.11	Comparison of MUSIC-based localization using the CS method with and without a coded cover at SNR = 20 dB.	22
2.12	Beam pattern of CCS method with the target at $(70^\circ, 140^\circ)$	26
2.13	MUSIC localization of 100 sources with CCS when SNR = 20dB	27
2.15	MUSIC localization of 100 10kHz sources with CCS when SNR = 10dB	27
2.14	MUSIC localization of 100 8kHz sources with CCS when SNR = 20dB	28
3.1	DOA estimation using unperturbed data for 12 unknown sources with 9° spacing.	39
3.2	DOA estimation using unperturbed data for 12 unknown sources with 8° spacing.	40
3.3	DOA estimation using unperturbed data with 33 sources in 3D case . .	41
3.4	DoA estimation using perturbed data for 12 unknown sources with 9° spacing.	42
3.5	DOA estimation with self-calibration using uniformly perturbed data for 34 sources	43
3.6	Top view of DoA result using CCS method with uniformly perturbed data	44
3.7	Recovered DOAs after applying the self-calibration method.	45
3.8	Final DOA estimation after applying the CCS method.	46
4.1	DOA estimation using geometrically perturbed data with 34 sources, including 10 known sources.	47

4.2	Visualization of geometric errors in the x-z plane.	48
4.3	Localization result using STLS with additive perturbation optimization under geometric mismatch in compression matrix, showing accurate source recovery at 72°	53
4.4	CCS localization results with and without STLS calibration, showing improved accuracy with STLS calibration.	53
4.5	Localization result using STLS with additive perturbation optimization under geometric mismatch and small perturbation in compression matrix, showing accurate source recovery at 56°	54
4.6	CCS localization results with and without STLS calibration under small perturbations, showing improved accuracy with STLS calibration. . . .	54
4.7	Localization result using STLS with additive perturbation optimization under geometric mismatch and larger perturbation in compression, showing accurate source recovery at 130°	55
4.8	CCS localization results with and without STLS calibration under larger perturbations, showing improved accuracy with STLS calibration. . . .	55
4.9	Localization result using STLS MMV-FISTA under geometric mismatch with uncalibrated sensors	56
4.10	Localization result with and without STLS correction under geometric mismatch with uncalibrated sensors, showing improved target recovery.	56
4.11	Localization result using STLS, successfully recovering both targets. . . .	57
4.12	CCS localization result after STLS calibration. Due to an inaccurate perturbation matrix, performance degrades, but STLS already provides accurate localization.	57
4.13	Localization result using STLS with added perturbation noise (variance = 0.5). The first target is exactly recovered, while the second peak has a mismatch of 3 degrees.	58
4.14	CCS localization result after STLS with added perturbation noise. Only the first peak is roughly matched, and the second target is missing. . . .	59
4.15	Localization result using STLS under geometric mismatches. Three sources are roughly recovered.	59
4.16	CCS localization result after STLS under geometric mismatches. Only the first source is roughly estimated, while the others are poorly matched. . . .	60
4.17	STLS recovery result for a single source at (20,30).	60
4.18	CCS recovery result without calibration, failing to recover the single target. . . .	61
4.19	STLS MMV-FISTA recovery result for four sources in the 3D case. . . .	61
4.20	CCS recovery result without calibration, showing complete failure to recover the sources.	61
4.21	Recovered result with STLS-MMV-FISTA method after adding noise to the compression matrix (variance 0.1).	62
4.22	Recovery result without calibration, showing poor performance due to the added noise in the compression matrix.	62
4.23	CCS localization result using ideal compression matrix, where all peaks do not correspond to the true target positions.	64

4.24	CCS localization result after applying grid-search calibration, where the targets are exactly matched by the algorithm.	64
4.25	CCS localization result using the ideal compression matrix in the 3D case, where the detected peaks do not align with the true source positions.	65
4.26	CCS localization result after applying grid-search calibration in 3D, showing precise alignment between the detected peaks and the true source positions.	66
4.27	Localization result in 2D space using grid searching and self-calibration under $\text{SNR} = 20$ dB.	67
4.28	Comparison of 3D localization results with and without grid-searching integrated self-calibration under $\text{SNR} = 20$ dB.	68
5.1	The model of coded cover in SolidWorks	72
5.2	3D printed coded cover in two different views	72
5.3	Design of the enclosure box with all the required length indications.	73
5.4	The physical box constructed from wood with a thickness of 1.8 cm.	73
5.5	Inside view of the enclosure with the RECT-7.5 inside.	74
5.6	Experimental setup with the coded cover	75
5.7	Near-field propagation setup modulation	75
5.8	Localization result using the CS method with calibrated sensors.	76
5.9	Localization result using the STLS-MMV-FISTA method with calibrated sensors.	77
5.10	DOA estimation result using self-calibration with grid-searching.	77
5.11	Enhanced DOA estimation result using CS with MUSIC after self-calibration.	78
5.12	DOA estimation result using CS with MUSIC for uncalibrated sensors.	78
5.13	DOA estimation result using STLS for uncalibrated sensors.	79
5.14	DOA estimation result after self-calibration.	79
5.15	Enhanced DOA estimation result using CS with optimized calibration parameters.	79
6.1	Scan&Paint 3D system from Microflow	82

List of Tables

3.1	Error Metrics	40
3.2	Errors in the estimated correction vector compared to the true values in unperturbed data case.	42
3.3	Errors in the estimated correction vector compared to the true values in uniformly perturbed case.	44

Introduction

The capability of capturing all acoustic information at any point within a sound field provides a detailed view of sound propagation [1]. This allows for precise identification and quantification of sound outputs from various sources. By measuring both the amplitude and direction of sound waves, the technology enhances the accuracy of acoustic analysis. It results in more detailed and reliable acoustic data.

1.1 Acoustic Sound Field Measurements

Acoustic sound fields are made up of both pressure variations and particle velocity. While acoustic pressure is a scalar quantity, particle velocity has both a magnitude and a direction, making it a vector quantity.

Microphones, which are the most commonly used devices to measure sound, operate by detecting changes in air pressure. The pressure quantity is measured by the movement of an internal membrane, which responds to pressure fluctuations. As the pressure force acts on the membrane, it moves back and forth, and this motion is then converted into an electrical signal by a transducer.

Acoustic particle velocity sensors [2], which are transducers used to measure particle velocity in air, operate on the same principles as hot-wire anemometers. In this system, a wire is heated by an electrical current and cools when exposed to airflow. As the temperature of the wire changes, its resistance also varies, generating an electrical signal that is proportional to the flow. To measure acoustic particle velocity, two heated wires are placed close to each other, and the temperature difference between them is monitored. This allows for an accurate measurement of particle velocity. Thanks to their compact design, these sensors can be arranged in three orthogonal directions, enabling the characterization of the acoustic particle velocity vector in the sound field. Figure 1.1 provides a close-up view of the particle velocity transducer, which operates on the principles of hot-wire anemometry.

1.1.1 PU probe

Unlike an acoustic vector sensor (AVS), which measures both acoustic pressure and two or three particle velocity channels, the PU Probe is a sound intensity probe that integrates two sensing technologies: a traditional microphone (P) and a Microflow sensor (U) [3]. By simultaneously measuring acoustic particle velocity and sound pressure, it overcomes many of the limitations typically associated with conventional sound intensity PP probes. This allows sound intensity measurements to be performed directly in real-world conditions, without the need for an acoustically treated environment. Due to this flexibility, PU sound intensity probes are ideal for use in a wide variety of

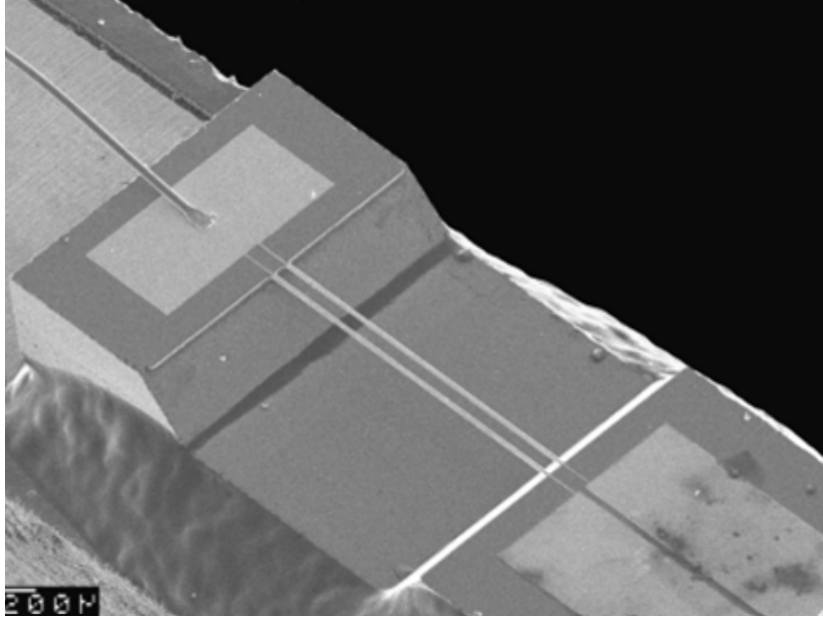


Figure 1.1: Hot-wire anemometer principle-based particle velocity transducer.

challenging settings, such as highly reverberant areas or confined spaces like vehicle interiors.

The Microflowm PU probe provides a distinctive capability to simultaneously capture all acoustic information at any point within a sound field. With its broadband response spanning from 20 Hz to 10 kHz, the probe enables direct measurement of sound pressure, particle velocity, and sound intensity. This comprehensive data collection is achieved without relying on assumptions or introducing computational errors, ensuring exceptional precision and accuracy in acoustic analysis. The appearance of the PU Regular from Microflowm is shown in Figure 1.2.

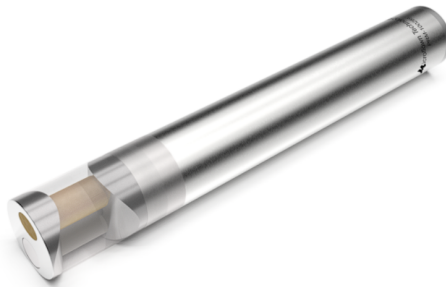
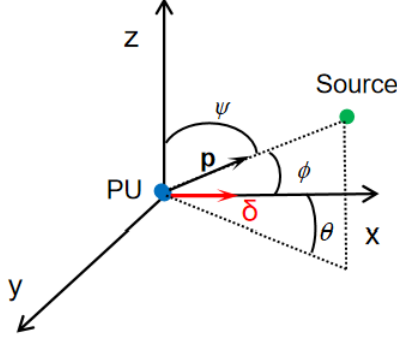
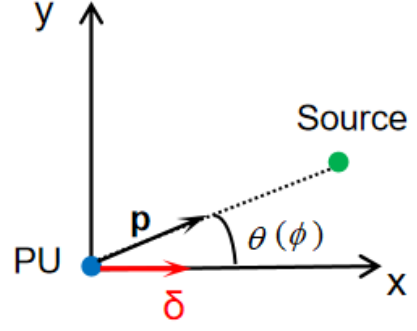


Figure 1.2: PU Regular

When a plane wave impinges on a PU probe, the model can be described using a three-dimensional spherical coordinate system. This system uses azimuth $\theta \in [0, 2\pi)$ and elevation $\psi \in [0, \pi]$, as illustrated in Figure 1.3a.



(a) Configuration of a single PU probe and one sound source in $\mathbb{R}^3$.



(b) Configuration of a single PU probe and one sound source in $\mathbb{R}^2$.

Figure 1.3: Configuration of a single PU probe and one sound source

In this coordinate system, the unit vector in the direction of the incoming wave is defined as:

$$\mathbf{p} = \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} = \begin{pmatrix} \cos(\theta) \sin(\psi) \\ \sin(\theta) \sin(\psi) \\ \cos(\psi) \end{pmatrix}.$$

For the special case of a two-dimensional spherical coordinate system where the elevation angle ψ is $\frac{\pi}{2}$, the unit vector simplifies to:

$$\mathbf{p} = \begin{pmatrix} p_x \\ p_y \end{pmatrix} = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}.$$

This reduction to two dimensions occurs because, with $\psi = \frac{\pi}{2}$, the wavefront is confined to the horizontal plane, making the elevation angle irrelevant.

The orientation of the PU probe δ can be set randomly. Here, we set it along with x-axis. Since a PU probe measures only one component of particle velocity, we can define the angle between $\mathbf{p}$ and δ as ϕ .

From the Euler equation of motion under the assumption of a far-field narrow band source signal [4], the sound pressure $P(\mathbf{r}, t)$ and the acoustic particle velocity $v(\mathbf{r}, t)$ at position $\mathbf{r}$ corresponding to the sensor and time t are linearly related as:

$$\mathbf{v}(\mathbf{r}, t) = -\frac{\mathbf{p}}{\rho c} P(\mathbf{r}, t),$$

where ρ represents the density of the medium, and c is the speed of sound in the medium. The product term ρc represents the characteristic acoustic impedance of the medium.

Due to the linear relationship between sound pressure and acoustic particle velocity, the wave phase delay relative to the reference point, $e^{jk(\mathbf{r} \cdot \mathbf{p})}$, can be factored out. Here, $k = \frac{2\pi}{\lambda}$ is the spatial frequency of the source signal, also called the wave number. This leaves only a difference in scaling by ρc . By adjusting for this scaling, the element-space

measurement data at time instant t is [5]:

$$\mathbf{y}(t) = e^{jk(\mathbf{r} \cdot \mathbf{p})} \begin{bmatrix} 1 \\ \cos(\phi) \end{bmatrix} s(t) + \mathbf{n}(t), \quad (1.1)$$

where $s(t)$ represents the source signal, and $\mathbf{n}(t)$ denotes the measurement noise.

1.1.2 Array grids

While individual sensors, such as the PU probe and an AVS, provide valuable data on acoustic pressure and particle velocity at a single point, scaling this approach to array grids offers enhanced capabilities.

First, we consider a PU based array randomly placed in the 3D space. The array contains the K sensors located at arbitrary positions indicated by the position vectors $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_K$ and N far-field sound sources $S_1, \dots, S_N$. Since the sources are transmitted from the far field, it can be approximated as plane waves. The n th source is located in the direction of the unit vector $\mathbf{p}_n$, which is parameterized by two angles in the 3D case: θ_n (azimuth) and ψ_n (elevation). The orientations of sensors can be arbitrary and are indicated by δ_k . The angle between the n th source and the orientation of the k th sensor is $\phi_{k,n}$. For convenience, we collect all $\phi_{k,n}$ terms corresponding to the same source into a vector. Specifically, we define

$$\boldsymbol{\phi}_n = [\phi_{1,n}, \phi_{2,n}, \dots, \phi_{K,n}], \quad n = 1, \dots, N.$$

The measurement data $\mathbf{y}(t)$ at time t of the measured snapshot corresponding to frequency f (wavelength λ) can be modeled as:

$$\mathbf{y}(t) = [\mathbf{a}(\theta_1, \psi_1) \circ \mathbf{u}(\boldsymbol{\phi}_1) \dots \mathbf{a}(\theta_N, \psi_N) \circ \mathbf{u}(\boldsymbol{\phi}_N)] \mathbf{s}(t) + \mathbf{n}(t), \quad (1.2)$$

where $\mathbf{a}(\theta_n, \psi_n) = \mathbf{a}_{\mathcal{P}}(\theta_n, \psi_n) \otimes \begin{bmatrix} 1 \\ 1 \end{bmatrix} = [e^{jk(\mathbf{r}_1 \cdot \mathbf{p}_n)} \dots e^{jk(\mathbf{r}_K \cdot \mathbf{p}_n)}]^T \otimes \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ represents the extended response vector of an APS array, $\mathbf{u}(\boldsymbol{\phi}_n) = [1 \cos(\phi_{1,n}) \dots 1 \cos(\phi_{K,n})]^T$ is the weighting vector containing directional information of the far-field source with respect to the vector sensor for the n th source. The $\circ$ denotes the Hadamard (element-wise) product. Further, $\mathbf{s}(t) = [s_1(t) \ s_2(t) \ \dots \ s_N(t)]^T$ is the source signal vector and $\mathbf{n}(t)$ is the additive noise vector present in the measurement data. In the further discussion, we refer to this general setup as a PU array, as indicated in Figure 1.4.

When the orientations of all the sensors are aligned in the same direction, the far-field model can be further simplified, since $\phi_{1,n} = \phi_{2,n} = \dots = \phi_{K,n}$ for the same n . We denote this common value as ϕ_n . The modified weighting vector $\mathbf{u}$ can be written as $\mathbf{u}(\phi_n) = [1 \cos(\phi_n)]^T$. Then, the simplified data model can be written as:

$$\mathbf{y}(t) = [\mathbf{a}_{\mathcal{P}}(\theta_1, \psi_1) \otimes \mathbf{u}(\phi_1) \dots \mathbf{a}_{\mathcal{P}}(\theta_N, \psi_N) \otimes \mathbf{u}(\phi_N)] \mathbf{s}(t) + \mathbf{n}(t), \quad (1.3)$$

Multiple sensors can be arranged in a fixed layout. One of the planar array product series from Microflown is the Hand-Held Array. It includes the RECT-7.5, RECT-1.8, and BiHex models. The first two arrays both feature a 3×4 configuration with 12 PU

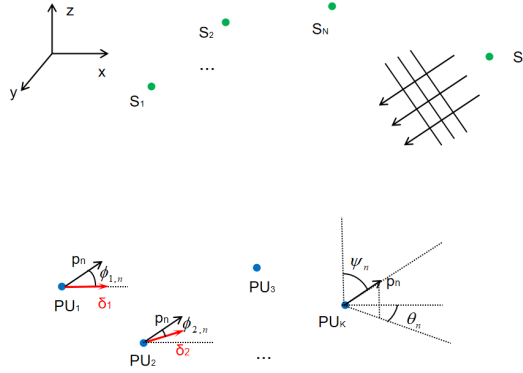


Figure 1.4: Measurement model of a PU array

Mini or PU Match sensors. The key difference is the inter-sensor spacing. As the names indicate, the RECT-7.5 has 75 mm inter-sensor spacing, while the RECT-1.8 has an 18 mm distance. Figures 1.5a and 1.5b show their appearances. Additionally, there is another design optimized for near-field acoustic holography, which features a hexagram layout, as shown in Figure 1.5c. In our work, we focused only on the RECT-7.5, where the orientation of each probe is perpendicular to the array plane. This corresponds to a special case of the general PU array model described earlier, where the sensor positions $\mathbf{r}_k$ lie on a 2D plane and the orientation vectors δ_k are all aligned along the same axis (e.g., the z-axis), orthogonal to the array plane. This structured configuration allows simplification of the general model by fixing both the spatial distribution and sensor orientation.

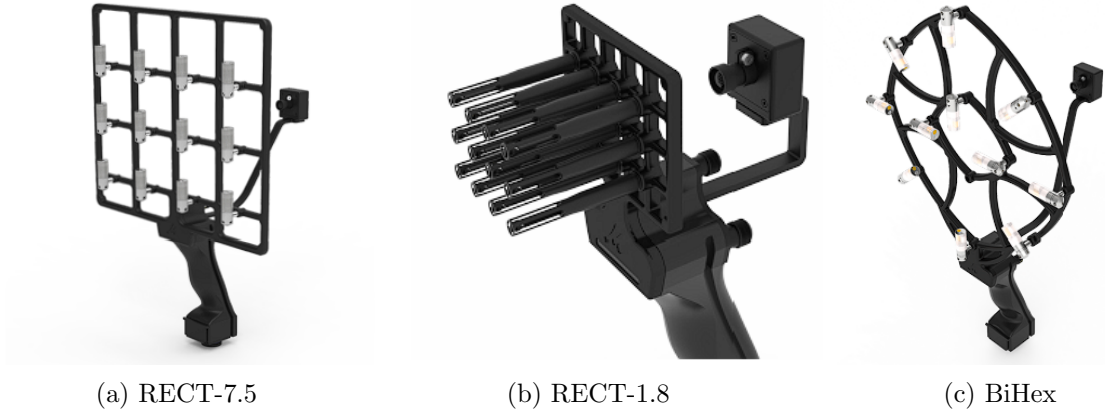


Figure 1.5: Family of Hand-Held Array with PU Probes

1.2 Motivation of the project

Current systems that employ a single AVS with a coded cover [6] have shown improved detection accuracy through compressed sensing (CS) and compressed covariance sensing

(CCS) methods [7]. The model of the considered cover to improve the performance of a 2D AVS is shown in Figure 1.6. As the wave encounters the cover, it passes through each channel independently. At the exit of each channel, we assume the presence of a point source. The sound then propagates from this point source to the position of the AVS in the near field. Notably, the CCS method has extended the ability to detect multiple sound sources with a single AVS. However, there remains a lack of evidence regarding the scalability and performance of these methods in more complex conditions.

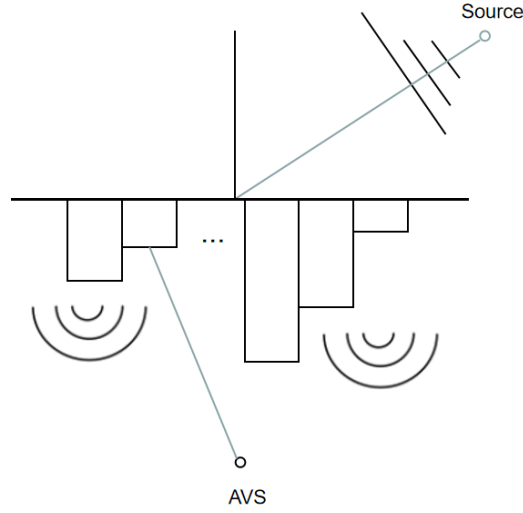


Figure 1.6: The model of a single 2D AVS with a specially designed coded cover [6]

To address these challenges, this project investigates the potential of extending these advantages by employing an array of PU sensors. By leveraging the combined capabilities of multiple sensors, we aim to further enhance DOA estimation accuracy and detection capabilities, surpassing the limitations of single-sensor systems. This investigation seeks to explore how the integration of a PU sensor array can provide significant improvements, offering a more robust solution for complex acoustic environments.

Furthermore, this project investigates calibration within the framework of compressed sensing. Since compressed sensing methods rely on an accurate system model, even small mismatches in the geometric configuration or sensor responses can degrade performance. By integrating calibration techniques into the compressed sensing process, we aim to correct these mismatches and enhance the accuracy of source localization.

1.3 Outline

This thesis explores advanced methods for sound source localization, focusing on the investigation of compressed sensing, self-calibration, and sparsity-cognizant total least-squares (STLS) with multiple measurement vectors (MMV) variant of the fast iterative shrinkage-thresholding algorithm (FISTA) method for improved DOA estimation in

3D. The structure of the thesis is as follows:

Chapter 2: This chapter extends the CCS method from a single AVS system to a planar PU array. The feasibility of 1D DOA estimation is first analyzed, followed by an extension to 2D DOA estimation. Experimental results demonstrate that a 14×10 coded cover enables localization of up to 100 sources in 3D, even in low SNR conditions.

Chapter 3: This chapter introduces a self-calibration method for planar PU arrays with coded cover, addressing per-channel receiving errors. The theoretical identifiability conditions are derived, followed by a joint estimation framework for signal recovery and calibration. An improved approach integrating CCS and self-calibration is proposed to enhance resolution and reduce side lobes.

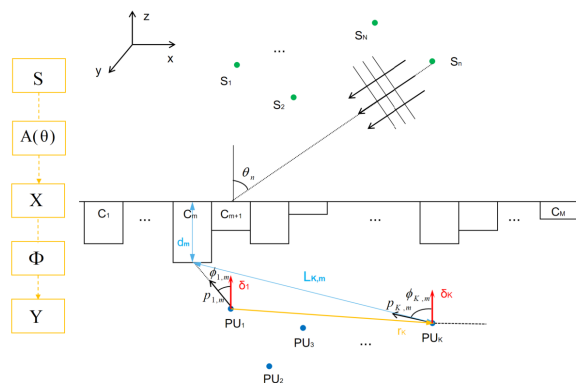
Chapter 4: The feasibility of STLS-MMV-FISTA in compressed sensing is evaluated, revealing its localization limitations due to geometric errors. To address this, a grid-search-based correction is applied, with different calibration strategies for calibrated and uncalibrated sensors. The method is first validated in 2D simulations, and then extended to 3D cases, showing its effectiveness in mitigating geometric mismatches.

Chapter 5: This chapter presents real-world tests using CS, STLS-MMV-FISTA, and self-calibration methods. Despite experimental constraints, results confirm the effectiveness of sparsity-based approaches in side lobe suppression, leading to improved localization accuracy under non-ideal conditions.

Chapter 6: The final chapter summarizes the key findings and highlights future research directions. Potential improvements include the use of irregularly shaped coded covers, alternative 3D sensor arrangements (e.g., spherical arrays), and enhanced self-calibration techniques for handling multiple error types.

2

The total array response with coded cover can be divided into two parts as explained below, and also illustrated in Figure 2.1:



From sources to coded cover

9

$$\mathbf{a}(\theta_i) = \begin{bmatrix} 1 \\ e^{j2\pi\Delta\sin(\theta_i)} \\ \vdots \\ e^{j2\pi(M-1)\Delta\sin(\theta_i)} \end{bmatrix}.$$

Thus the virtual ULA response of all N signals can be stacked as the $M \times N$ matrix:

$$\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_1) \quad \mathbf{a}(\theta_2) \quad \cdots \quad \mathbf{a}(\theta_N)] \quad (2.1)$$

$$= \begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{j2\pi\Delta\sin(\theta_1)} & e^{j2\pi\Delta\sin(\theta_2)} & \cdots & e^{j2\pi\Delta\sin(\theta_N)} \\ \vdots & \vdots & \ddots & \vdots \\ e^{j2\pi(M-1)\Delta\sin(\theta_1)} & e^{j2\pi(M-1)\Delta\sin(\theta_2)} & \cdots & e^{j2\pi(M-1)\Delta\sin(\theta_N)} \end{bmatrix}. \quad (2.2)$$

The signals received at the entrance of M channels can be expressed as:

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t)$$

From coded cover to sensors

Assume the length of the m -th channel is d_m , and the distance between the exit of the m -th channel and the k -th PU sensor is $L_{k,m}$. Since the distance from the coded mask to the PU array is far shorter than 10 times the wavelength, the signal attenuation model from the exit of the channels should be considered in the near-field. The attenuation from the exit of the m -th channel to the k -th PU sensor can be represented as:

$$\beta_{k,m} = \frac{1}{L_{k,m}} e^{-j \frac{2\pi f(L_{k,m} + d_m)}{c}}.$$

We group the attenuation coefficients $\beta_{k,m}$, $k = 1, 2, \dots, K$, that correspond to the same channel m , into a matrix:

$$\mathbf{E}_m = \text{diag}(\beta_{k,m}, k = 1, 2, \dots, K) \otimes \mathbf{I}_{2 \times 2}$$

for later use. Note that we used the Kronecker product here to indicate the linear relation between pressure and the single particle velocity measurement used in the PU probe.

From Section 1.1.2, we have already derived the PU array response of multiple signals. Now we can consider the current situation where there are M sources transmitted from M channels and received by an array of K PU sensors. With the weighting vector $\mathbf{u}(\phi_n)$ defined in section 1.1.2, the compression matrix is therefore written as a $2K \times M$ matrix :

$$\boldsymbol{\Phi} = [\mathbf{E}_1 \mathbf{u}(\phi_1) \quad \mathbf{E}_2 \mathbf{u}(\phi_2) \quad \cdots \quad \mathbf{E}_M \mathbf{u}(\phi_M)] \quad (2.3)$$

$$= \begin{bmatrix} \beta_{1,1} & \beta_{1,2} & \cdots & \beta_{1,M} \\ \beta_{1,1}\cos(\phi_{1,1}) & \beta_{1,2}\cos(\phi_{1,2}) & \cdots & \beta_{1,M}\cos(\phi_{1,M}) \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{K,1} & \beta_{K,2} & \cdots & \beta_{K,M} \\ \beta_{K,1}\cos(\phi_{K,1}) & \beta_{K,2}\cos(\phi_{K,2}) & \cdots & \beta_{K,M}\cos(\phi_{K,M}) \end{bmatrix}. \quad (2.4)$$

As a result, the total array response matrix is

$$\mathbf{B}(\boldsymbol{\theta}) = \Phi \mathbf{A}(\boldsymbol{\theta}) = [\Phi \mathbf{a}(\theta_1) \quad \Phi \mathbf{a}(\theta_2) \quad \cdots \Phi \mathbf{a}(\theta_N)] = [\mathbf{b}(\theta_1) \quad \mathbf{b}(\theta_2) \quad \cdots \mathbf{b}(\theta_N)]. \quad (2.5)$$

The received signals in the PU array can be expressed as:

$$\mathbf{y}(t) = \Phi \mathbf{A}(\boldsymbol{\theta}) \mathbf{s}(t) + \mathbf{n}(t) = \mathbf{B}(\boldsymbol{\theta}) \mathbf{s}(t) + \mathbf{n}(t) \quad (2.6)$$

$$= \Phi \mathbf{x}(t) + \mathbf{n}(t) \in \mathbb{C}^{2K \times 1}, \quad (2.7)$$

where $\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta}) \mathbf{s}(t)$ is the measurement vector of the noiseless measurements at the virtual ULA.

This model equation can be extended to T measurements:

$$\mathbf{Y} = [\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(T)] = \Phi \mathbf{A}(\boldsymbol{\theta}) \mathbf{S} + \mathbf{N} = \mathbf{B}(\boldsymbol{\theta}) \mathbf{S} + \mathbf{N} \quad (2.8)$$

$$= \Phi \mathbf{X} + \mathbf{N} \in \mathbb{C}^{2K \times T}, \quad (2.9)$$

where $\mathbf{S}$, $\mathbf{N}$ and $\mathbf{X}$ are similarly defined as $\mathbf{Y}$.

The measurement covariance matrix $\mathbf{R}_y$ can thus be calculated as:

$$\mathbf{R}_y = \Phi \mathbf{A}(\boldsymbol{\theta}) \mathbf{R}_s \mathbf{A}(\boldsymbol{\theta})^H \Phi^H + \mathbf{R}_n = \mathbf{B}(\boldsymbol{\theta}) \mathbf{R}_s \mathbf{B}(\boldsymbol{\theta})^H + \mathbf{R}_n \quad (2.10)$$

$$= \Phi \mathbf{R}_x \Phi^H + \mathbf{R}_n \in \mathbb{C}^{2K \times 2K}, \quad (2.11)$$

where $\mathbf{R}_s = \mathbb{E}\{\mathbf{s}(t)\mathbf{s}(t)^H\}$, $\mathbf{R}_n = \mathbb{E}\{\mathbf{n}(t)\mathbf{n}(t)^H\}$ and $\mathbf{R}_x = \mathbb{E}\{\mathbf{x}(t)\mathbf{x}(t)^H\}$ are the covariance matrices of corresponding signals.

2.1 1D DOA Estimation Using Compressed Sampling

We already observed the difference of the covariance representations in (2.10) and (2.11). We first examined the advantage gained from implementing MUSIC algorithm with the compressive measurement based on (2.10).

Compressive sensing (CS) is a method used to acquire sparse signals and reconstruct them from a small number of measurements [8]. The core concept behind CS is that certain signals can be recovered using far fewer samples than traditional methods [9]-[10]. This is made possible by two key principles: sparsity, meaning the signal can be represented using only a few non-zero elements, and incoherence, which ensures that these signals can be accurately reconstructed from the compressed data.

In our design, the coded cover can be seen as compressive sampling in the spatial domain, compressing or transforming the virtual ULA into an array of fewer sensors.

According to (2.5), a matched filter can be implemented using $\mathbf{w}(\theta) = \mathbf{b}(\theta)$. However, since the compression matrix Φ does not preserve the energy, it is better to normalize the beamformer. This ensures that the noise energy will also not depend on θ . The normalized beam pattern for the model with a coded cover can be expressed as:

$$B = \frac{|\mathbf{b}(\theta)^H \mathbf{b}(\theta_0)|}{|\mathbf{b}(\theta)|^2} = \frac{|\mathbf{a}(\theta)^H \Phi^H \Phi \mathbf{a}(\theta_0)|}{|\mathbf{b}(\theta)|^2}.$$

Figure 2.2 shows the beam patterns of 12 PUs for both ULA and RECT-7.5 arrangements with a coded cover. The number of channels used is 90, with each channel spaced at half-wavelength. The 12 PUs are placed 15 cm away from the coded cover. The length of each channel ranges from 1 cm to 10 cm. For the ULA, the sensor spacing is 7.5 cm. The signal frequency is 10 kHz.

From Figure 2.2, we can see that the highest peaks of the ULA plot and planar plot pointed towards the target. But there is a common issue: due to the presence of multiple high side lobes in the beam pattern, the beamforming algorithm cannot be used effectively to localize multiple sources.

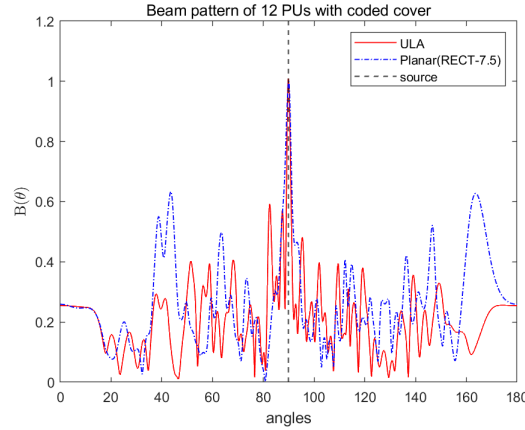


Figure 2.2: Beam pattern of an array with 12 PUs with coded cover

Traditional matched filter-based methods fail to provide satisfactory performance under the given conditions. This necessitates the use of covariance-based DOA estimation methods, such as the MUSIC algorithm and the MVDR algorithm. Since we have a perfect knowledge of the whole propagation model, including the virtual ULA response of the mask $\mathbf{A}$ and the compression matrix Φ , we can leverage these techniques. Research has shown that angle-spectrum-based methods like MUSIC and MVDR perform comparably to CS reconstruction methods [11, 12], while avoiding the high complexity associated with iterative algorithms [13].

MUSIC, in particular, takes advantage of the fundamental property that the signal subspace and the noise subspace are orthogonal, enabling it to perform high-resolution DOA estimation. In contrast to MUSIC, which relies on subspace decomposition, MVDR directly minimizes the interference from all directions except the target direction, thus enhancing DOA estimation while effectively suppressing noise and interference.

We choose to implement the MUSIC algorithm. From (2.10), we perform an eigenvalue decomposition:

$$\mathbf{R}_y = \mathbf{U}_s(\mathbf{\Lambda}_s + \sigma^2 \mathbf{I}_N) \mathbf{U}_s^H + \mathbf{U}_n(\sigma^2 \mathbf{I}_{2K-N}) \mathbf{U}_n^H.$$

Then by scanning $\mathbf{b}(\theta)$, at most $2K - 1$ sources can be localized. The CS MUSIC

spectrum can be expressed as follows:

$$J_{MUSIC}(\theta) = \frac{|\mathbf{U}_n^H \mathbf{b}(\theta)|^2}{|\mathbf{b}(\theta)|^2} = \frac{\mathbf{b}(\theta)^H \mathbf{U}_n \mathbf{U}_n^H \mathbf{b}(\theta)}{\mathbf{b}(\theta)^H \mathbf{b}(\theta)}.$$

The peaks of the MUSIC spectrum correspond to the estimated DOAs.

Compared to the traditional matched filter method, CS can achieve its full potential when combined with MUSIC, as shown in Figure 2.3. It successfully recovers up to $2K - 1$ sources. Specifically, when using 12 PUs, whether arranged in a ULA or a planar array, 23 sources can be accurately localized. This demonstrates that our coded cover enables high-accuracy localization of sources.

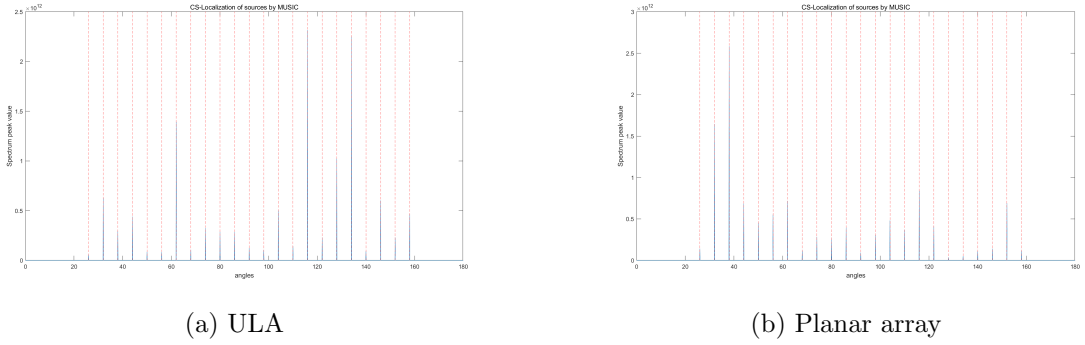


Figure 2.3: MUSIC Localization of 23 2D-sources by using CS method with 12 PUs at SNR = 20dB

2.2 1D DOA Estimation Using Compressive Covariance Sensing

In this part, we investigate the benefit of localizing as much sources as possible with estimating $\mathbf{R}_x$ from $\mathbf{R}_y$ through compressive covariance sensing (CCS) [14]. Specifically, we exploit the setting where the sources are stationary in the spatial domain, which, in this context, is equivalent to assuming uncorrelated sources. Different from the CS method, which requires the original signal to be sparse, CCS focuses on recovering the second-order statistics of a signal from its compressed measurements without imposing any sparsity constraints.

First, we need to obtain the compressed data. Then we recover the estimated $\mathbf{R}_x$ from $\mathbf{R}_y$. The compression stage is done by compressive sampling of the coded mask, as explained in Section 2.1. Following this, we analyze the specific structure of the original covariance matrix of $\mathbf{x}$. In this way, we can handle the over-determined problem by transforming the original under-determined problem.

Since the cover has a virtual ULA structure and sources are uncorrelated, $\mathbf{R}_x$ is a positive semi-definite Hermitian Toeplitz (HT) matrix. It should lie in the subspace formed by the intersection of the set of positive semi-definite matrices and the set of

HT matrices. This subspace can be characterized by the following basis:

$$\mathcal{S} = \{\Sigma_0, \Sigma_1, \dots, \Sigma_{2M-2}\} \quad (2.12)$$

$$= \{\mathbf{I}_M\} \cup \{\mathbf{T}_1, \dots, \mathbf{T}_{M-1}\} \cup \{\tilde{\mathbf{T}}_1, \dots, \tilde{\mathbf{T}}_{M-1}\}, \quad (2.13)$$

where $\mathbf{T}_m$ denotes the HT matrix with all zeros except for the entries on the diagonals $+m$ and $-m$, which have ones, and $\tilde{\mathbf{T}}_m$ represents the HT matrix with all zeros except for the entries on the diagonal $+m$, which have the imaginary unit j , and those on the diagonal $-m$, which have $-j$.

Therefore, $\mathbf{R}_x$ can be represented by the scaled summation of all entries in the subspace, which is:

$$\mathbf{R}_x = \sum_{m=0}^{2M-2} \alpha_m \Sigma_m. \quad (2.14)$$

The number of unknown scalars α_m is decided by the cardinality of the subspace, which is $2M - 1$.

Plugging this into (2.11), we get:

$$\mathbf{R}_y = \sum_{m=0}^{2M-2} \alpha_m \Phi \Sigma_m \Phi^H + \mathbf{R}_n \quad (2.15)$$

$$= \sum_{m=0}^{2M-2} \alpha_m \bar{\Sigma}_m + \mathbf{R}_n. \quad (2.16)$$

Now we can define a set related to $\mathcal{S}$, with all entries being Hermitian, but not Toeplitz:

$$\bar{\mathcal{S}} = \{\bar{\Sigma}_0, \bar{\Sigma}_1, \dots, \bar{\Sigma}_{2M-2}\} \quad \text{with} \quad \bar{\Sigma}_m = \Phi \Sigma_m \Phi^H.$$

If the compression matrix Φ is able to preserve all the second order statistical information of $\mathbf{x}$, it is called an $\mathcal{S}$ -covariance sampler. They are able to preserve the linear independence among matrices in the set of $\bar{\mathcal{S}}$ and knowing α_m is exactly equivalent to knowing the $\mathbf{R}_x$. The fewer the unknowns, the higher the compression ratio achieved by the compressive covariance sensing method.

A necessary condition for this is that (2.16) defines an over-determined system. The number of unknowns is $2M - 1$, while the number of equations is equal to the number of independent real values in $\mathbf{R}_y - \mathbf{R}_n$, which is $(2K)^2 = 4K^2$. Hence, a necessary condition for obtaining an $\mathcal{S}$ -covariance sampler is

$$2M - 1 \leq 4K^2 \iff M \leq (4K^2 + 1)/2 \iff M_{max} = 2K^2.$$

And now we can recover at most $2K^2 - 1$ sources. When 12 PU sensors are used, i.e., $K = 12$, the maximum number of channels becomes 288. This means that up to 287 sources can be recovered in our case. This shows a significant improvement in the number of localizable sources compared to the CS method, which can resolve only up to 23 sources.

To estimate the α_m values, we can use either a maximum-likelihood (ML) approach or a least square method. ML involves high computational costs and requires an accurate statistical characterization of the observations. For these reasons, it is customary to rely on geometrical considerations and project the sample covariance matrix onto the span of $\mathcal{S}$. We consider a simple linear least squares cost based on (2.16):

$$\hat{\alpha} = \min_{\alpha} \left\| \mathbf{R}_y - \mathbf{R}_n - \sum_{m=0}^{2M-2} \alpha_m \bar{\Sigma}_m \right\|_2^2.$$

The solution is described in [7]. First, we need to vectorize $\mathbf{R}_y$, $\mathbf{R}_n$, $\mathbf{R}_x$ and Σ_m as vectors σ_y , σ_n , σ_x and σ_m , respectively. From Equation (2.14), we get the relation between σ_x and σ_m :

$$\sigma_x = \sum_{i=0}^{2M-2} \alpha_i \sigma_i = \mathbf{S} \alpha \quad (2.17)$$

, where $\mathbf{S} = [\sigma_0, \sigma_1, \dots, \sigma_{2M-2}]$, $\alpha = [\alpha_0, \alpha_1, \dots, \alpha_{2M-2}]^T$.

Then, after vectorizing (2.11), we can get:

$$\sigma_y - \sigma_n = (\Phi^* \otimes \Phi) \sigma_x = (\Phi^* \otimes \Phi) \mathbf{S} \alpha. \quad (2.18)$$

Now, $\hat{\alpha}$ can be estimated by solving the above equation with LS: $[(\Phi^* \otimes \Phi) \mathbf{S}]^+ (\sigma_y - \sigma_n)$, where $^+$ denotes the pseudoinverse operation. Finally, we can obtain the estimated

$$\hat{\mathbf{R}}_x = \text{vec}^{-1} \{ \mathbf{S} [(\Phi^* \otimes \Phi) \mathbf{S}]^+ (\sigma_y - \sigma_n) \}.$$

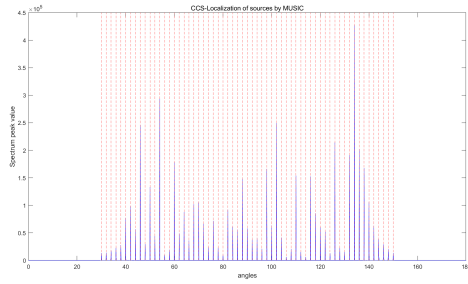
Any ULA-based DOA estimation algorithm can be applied to localize sources with $\hat{\mathbf{R}}_x$. We will focus on MUSIC for this.

2.3 Simulation Results of 1D DOA Estimation with CCS

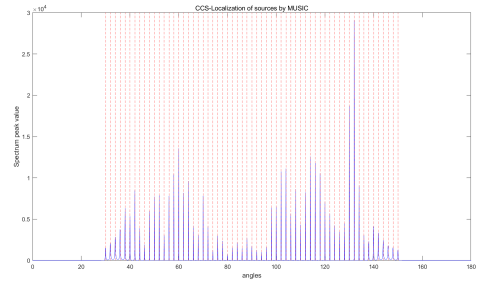
In this section, we will demonstrate the effectiveness of CCS in DOA estimation. First, we need to initialize the parameters. The number of channels in the coded mask is set to 90. The signal frequency is 10 kHz, so the channel spacing is chosen to be 1.7 cm to ensure it is half the wavelength. Twelve PUs are placed 50 cm away from the entry of the coded cover. The length of each channel varies between 1 and 10 cm. For the ULA, the spacing between the sensors is set to 7.5 cm, which is the same as for RECT-7.5. We take 500,000 snapshots for the estimation process.

As shown in Figure 2.4, both the ULA pattern and the RECT-7.5 placement successfully recover 61 sources. This represents a significant improvement over the CS method in terms of source recovery capability.

If all PU probes are replaced with simple P or U sensors, where only one measurement is collected per sensor, the number of channels decreases to 70. All other parameters remain unchanged. As shown in the MUSIC spectrum plots in Figure 2.5, with either a single pressure or particle velocity measurement per sensor, the CCS method can successfully recover 41 sources in all cases. However, achieving the same

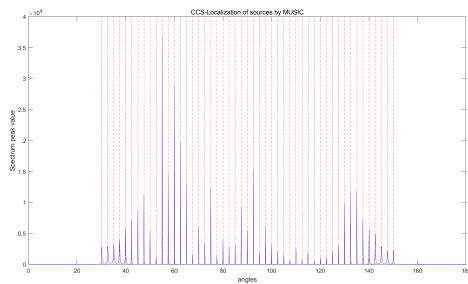


(a) ULA

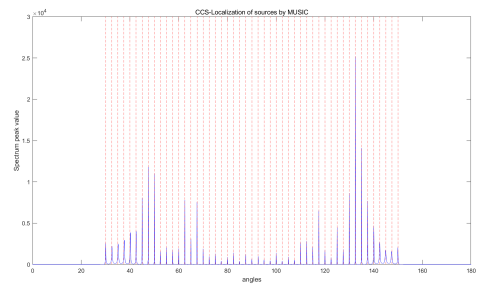


(b) RECT-7.5

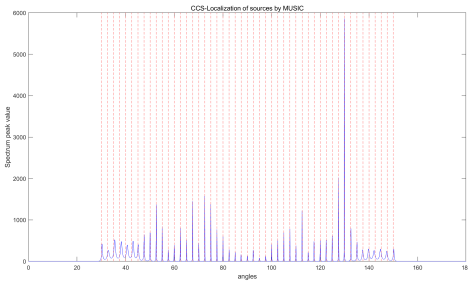
Figure 2.4: MUSIC Localization of 61 2D-sources by using CCS method with 12 PUs at $\text{SNR} = 20\text{dB}$



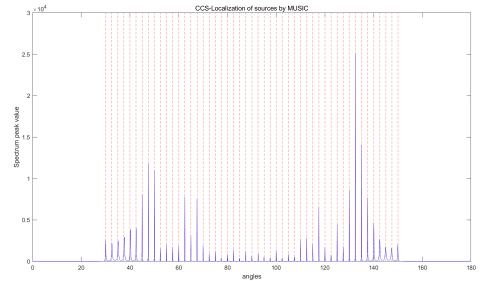
(a) ULA with 12 Pressure sensors



(b) ULA with 12 U sensors



(c) Planar array with 12 Pressure sensors



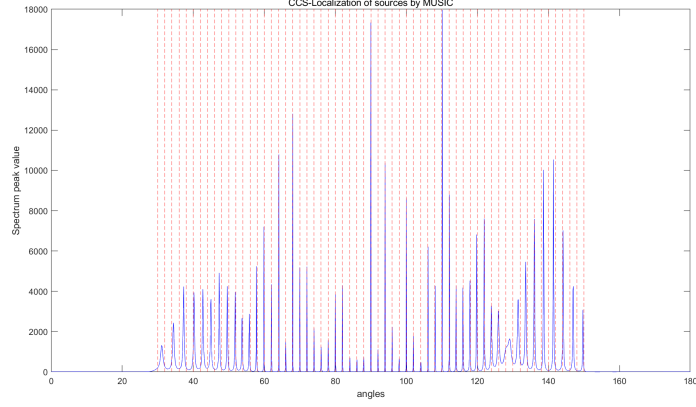
(d) Planar array with 12 U sensors

Figure 2.5: MUSIC Localization of 41 2D-sources by using CCS method with 12 P/Us at $\text{SNR} = 20\text{dB}$

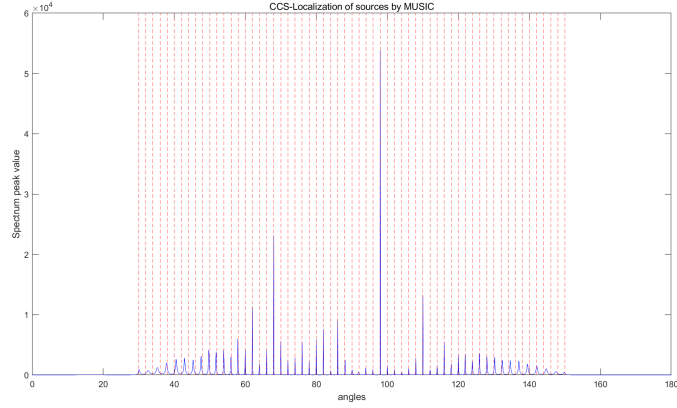
performance as when 12 PUs are used to recover 61 sources is not feasible, as demonstrated by the results in Figure 2.6. It is clear that equivalent performance cannot be obtained with less data.

2.4 2D DOA Estimation Using CS method

We now transition from 2D to 3D sources. To estimate both elevation and azimuth angles, the mask must also be configured in 3D. The form of the coded cover mask



(a) Planar array with 12 Pressure sensors



(b) Planar array with 12 U sensors

Figure 2.6: MUSIC Localization of 61 2D-sources by using CCS method with 12 P/Us at SNR = 20dB

is changed to a planar arrangement of channels, as shown in Figure 2.7. The PU array used behind the planar mask is the RECT-7.5 model, and for all further analysis, $K = 12$ is used. The mask is composed by M channels with J rows and I columns. The localization includes the elevation, which is based on the deviation from the z-axis $\theta \in [0^\circ, 90^\circ)$ and azimuth from the x-axis $\psi \in [0^\circ, 360^\circ)$.

The signal direction is represented by the unit vector $\mathbf{q} = (\sin \theta \cos \psi, \sin \theta \sin \psi, \cos \theta)$. The channel block in the position of (i, j) is also represented by the position vector $\mathbf{m}_{ij} = ((i - 1) \times \delta, (j - 1) \times \delta, 0)$. Here δ is the side length of each square channel entry area. Also, it equals the spacing between channels.

For each source, the coded cover can be viewed as a virtual block array response with a recursive structure:

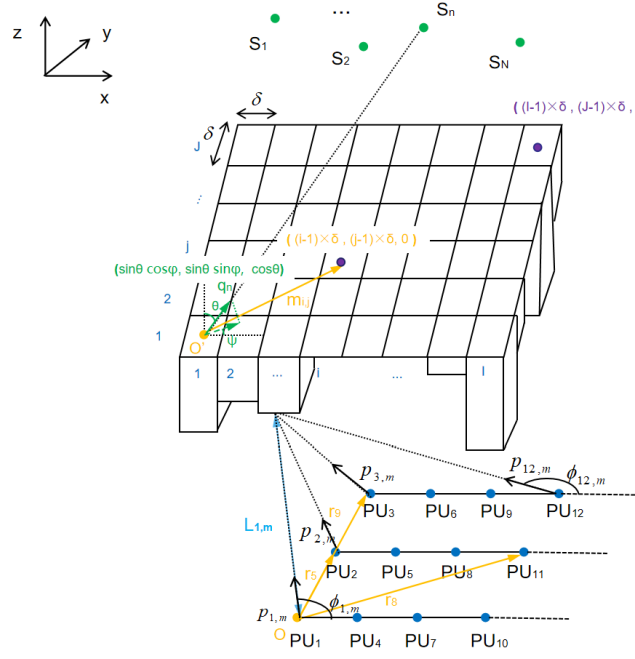


Figure 2.7: 3D source localization model

$$\mathbf{a}(\theta, \psi) = \begin{bmatrix} 1 \\ e^{j2\pi\Delta\sin\theta\cos\psi} \\ \vdots \\ e^{j2\pi(I-1)\Delta\sin\theta\cos\psi} \\ \hline e^{j2\pi\Delta\sin\theta\sin\psi} \\ e^{j2\pi\Delta\sin\theta(\cos\psi+\sin\psi)} \\ \vdots \\ e^{j2\pi\Delta\sin\theta((I-1)\cos\psi+\sin\psi)} \\ \hline \vdots \\ \hline e^{j2\pi\Delta\sin\theta(J-1)\sin\psi} \\ e^{j2\pi\Delta\sin\theta(\cos\psi+(J-1)\sin\psi)} \\ \vdots \\ e^{j2\pi\Delta\sin\theta((I-1)\cos\psi+(J-1)\sin\psi)} \end{bmatrix} = \begin{bmatrix} 1 \\ e^{j2\pi\Delta\sin\theta\cos\psi} \\ \vdots \\ e^{j2\pi(I-1)\Delta\sin\theta\cos\psi} \\ \hline e^{j2\pi\Delta\sin\theta\sin\psi} \cdot 1 \\ e^{j2\pi\Delta\sin\theta\sin\psi} \cdot e^{j2\pi\Delta\sin\theta\cos\psi} \\ \vdots \\ e^{j2\pi\Delta\sin\theta\sin\psi} \cdot e^{j2\pi(I-1)\Delta\sin\theta\cos\psi} \\ \hline \vdots \\ \hline e^{j2\pi(J-1)\Delta\sin\theta\sin\psi} \cdot 1 \\ e^{j2\pi(J-1)\Delta\sin\theta\sin\psi} \cdot e^{j2\pi\Delta\sin\theta\cos\psi} \\ \vdots \\ e^{j2\pi(J-1)\Delta\sin\theta\sin\psi} \cdot e^{j2\pi(I-1)\Delta\sin\theta\cos\psi} \end{bmatrix} \quad (2.19)$$

$$= \underbrace{\begin{bmatrix} 1 \\ e^{j2\pi\Delta\sin\theta\sin\psi} \\ \vdots \\ e^{j2\pi(J-1)\Delta\sin\theta\sin\psi} \end{bmatrix}}_{\mathbf{a}_y(\theta,\psi)} \otimes \underbrace{\begin{bmatrix} 1 \\ e^{j2\pi\Delta\sin\theta\cos\psi} \\ \vdots \\ e^{j2\pi(I-1)\Delta\sin\theta\cos\psi} \end{bmatrix}}_{\mathbf{a}_x(\theta,\psi)}, \quad (2.20)$$

where $\Delta = \frac{\delta f}{c}$ (we set it as 1/2 wavelength to avoid aliasing problem). The order of the terms $\mathbf{a}_y(\theta, \psi)$ and $\mathbf{a}_x(\theta, \psi)$ is not important. Changing the order only affects the stacking order of the shifts for each channel in $\mathbf{a}(\theta, \psi)$.

Now define the virtual array response matrix along with the row and column directions separately, as follows:

$$\mathbf{A}_y(\boldsymbol{\theta}, \boldsymbol{\psi}) = [\mathbf{a}_y(\theta_1, \psi_1) \quad \cdots \quad \mathbf{a}_y(\theta_N, \psi_N)] \in \mathbb{C}^{J \times N},$$

and

$$\mathbf{A}_x(\boldsymbol{\theta}, \boldsymbol{\psi}) = [\mathbf{a}_x(\theta_1, \psi_1) \quad \cdots \quad \mathbf{a}_x(\theta_N, \psi_N)] \in \mathbb{C}^{I \times N}.$$

We can note that both of them are Vandermonde matrices.

The virtual planar array matrix for all sources can thus be written as a column-wise Kronecker product, which is a Khatri-Rao product, denoted as $\odot$:

$$\mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\psi}) = [\mathbf{a}(\theta_1, \psi_1) \quad \cdots \quad \mathbf{a}(\theta_N, \psi_N)] \quad (2.21)$$

$$= [\mathbf{a}_y(\theta_1, \psi_1) \otimes \mathbf{a}_x(\theta_1, \psi_1) \quad \cdots \quad \mathbf{a}_y(\theta_N, \psi_N) \otimes \mathbf{a}_x(\theta_N, \psi_N)] \quad (2.22)$$

$$= \mathbf{A}_y(\boldsymbol{\theta}, \boldsymbol{\psi}) \odot \mathbf{A}_x(\boldsymbol{\theta}, \boldsymbol{\psi}) \in \mathbb{C}^{IJ \times N}. \quad (2.23)$$

The compression matrix Φ can be constructed by calculating the distance from the exit of each channel to each sensor. After obtaining $\mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\psi})$ and Φ , the total array response matrix is:

$$\mathbf{B}(\boldsymbol{\theta}, \boldsymbol{\psi}) = \Phi \mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\psi}) \quad (2.24)$$

$$= [\Phi \mathbf{a}(\theta_1, \psi_1) \quad \Phi \mathbf{a}(\theta_2, \psi_2) \quad \cdots \quad \Phi \mathbf{a}(\theta_N, \psi_N)] \quad (2.25)$$

$$= [\mathbf{b}(\theta_1, \psi_1) \quad \mathbf{b}(\theta_2, \psi_2) \quad \cdots \quad \mathbf{b}(\theta_N, \psi_N)]. \quad (2.26)$$

The received signals in the PU array can be expressed as:

$$\mathbf{y}(t) = \Phi \mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\psi}) \mathbf{s}(t) + \mathbf{n}(t) = \mathbf{B}(\boldsymbol{\theta}, \boldsymbol{\psi}) \mathbf{s}(t) + \mathbf{n}(t) \quad (2.27)$$

$$= \Phi \mathbf{x}(t) + \mathbf{n}(t) \in \mathbb{C}^{2K \times 1}, \quad (2.28)$$

where $\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\psi}) \mathbf{s}(t)$ is the measurement vector of the noiseless measurements at the virtual ULA.

The matched filter beamformer for the CS method in the 2D case is done by scanning both elevation and azimuth together:

$$B_{CS} = \frac{|\mathbf{b}(\theta, \psi)^H \mathbf{b}(\theta_0, \psi_0)|}{|\mathbf{b}(\theta, \psi)|^2} = \frac{|\mathbf{a}(\theta, \psi)^H \Phi^H \Phi \mathbf{a}(\theta_0, \psi_0)|}{|\mathbf{b}(\theta, \psi)|^2},$$

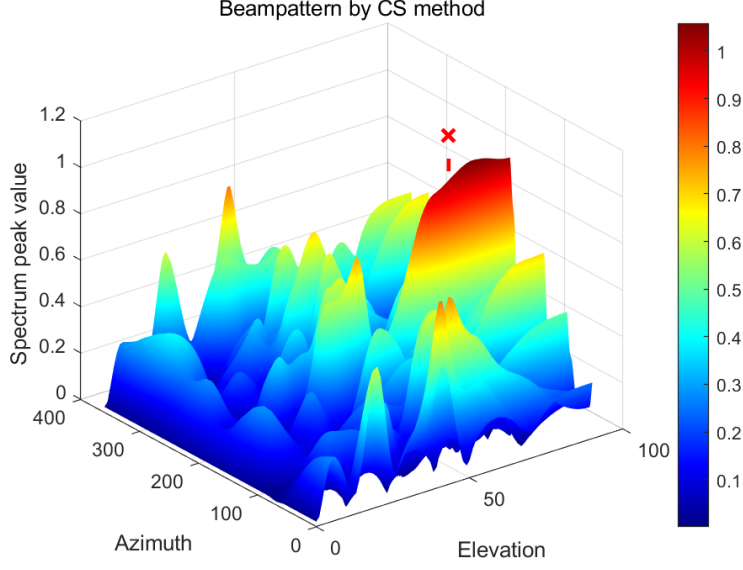


Figure 2.8: Beam pattern of CS method with the target at $(70^\circ, 140^\circ)$

where we assume θ_0 is the elevation of the single source, and the azimuth is localized at the direction of ψ_0 .

We test the case when a 10×14 planar coded cover is used, with 12 PU sensors positioned 15 cm behind the entry plane of the coded cover. The length of the coded cover channels ranges from 1 cm to 10 cm. The target is located at $\theta_0 = 70^\circ$ and $\psi_0 = 140^\circ$. As shown in Figure 2.8, the beamformer fails to accurately focus on the target. The presence of multiple side lobes further complicates localization, especially when additional sources exist. This limitation makes the traditional beamformer unsuitable for the CS method.

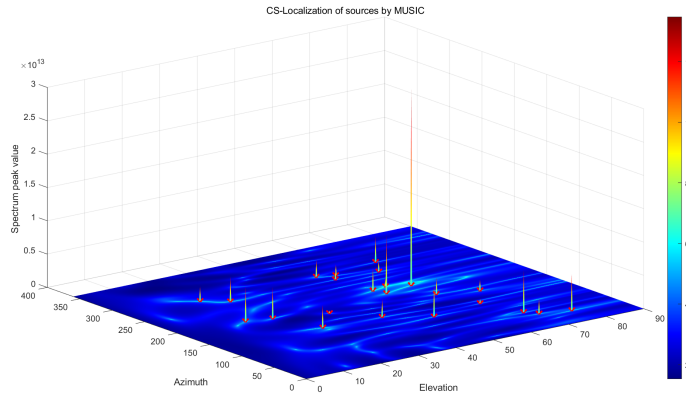
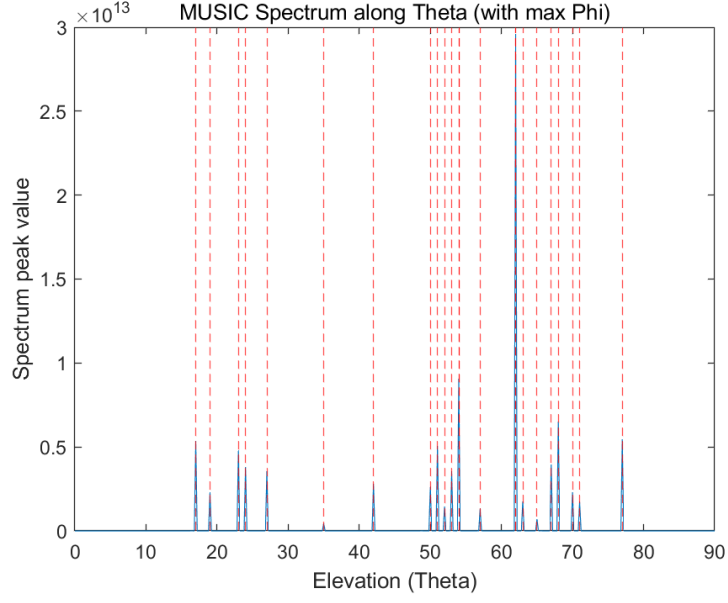
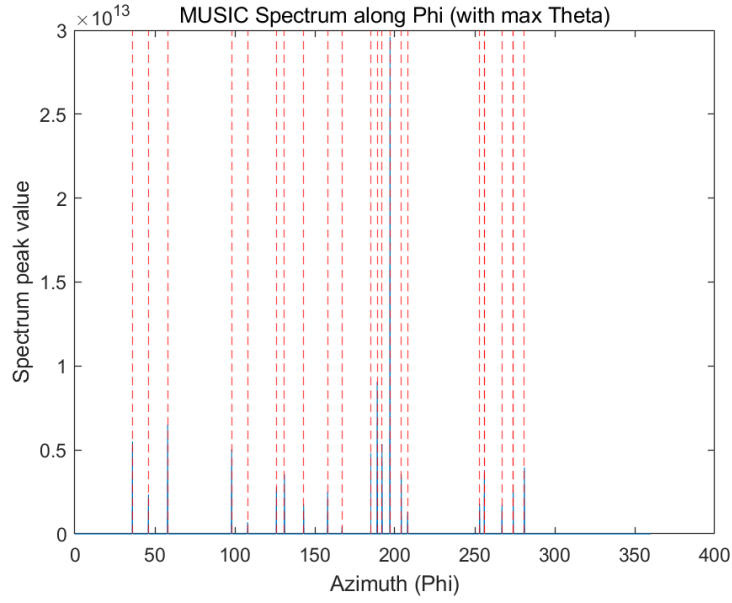


Figure 2.9: MUSIC localization of 23 sources with CS method when $\text{SNR} = 20\text{dB}$

When the MUSIC algorithm is applied, Figure 2.9 shows that 23 sources can be



(a) Elevation aspect



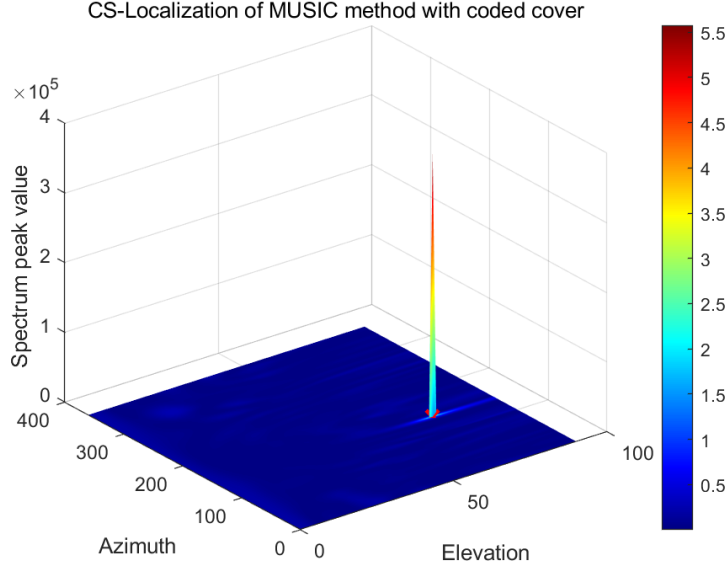
(b) Azimuth aspect

Figure 2.10: The details of localization results from the CS method

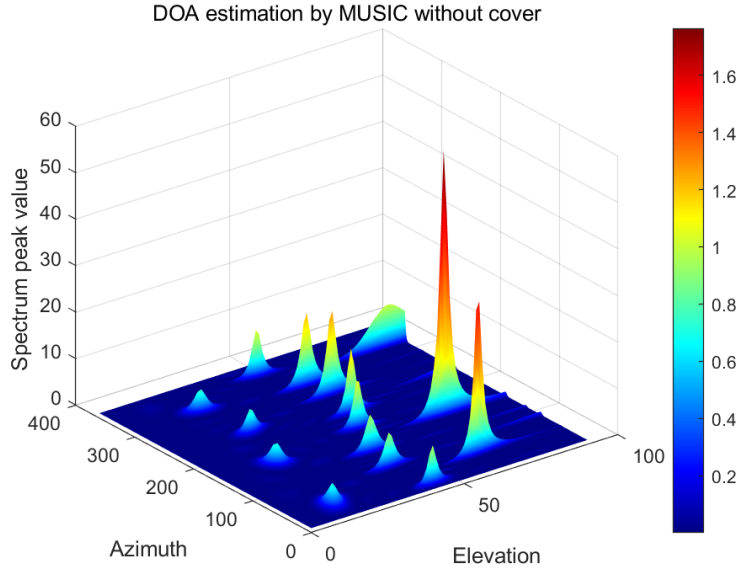
successfully localized. A detailed examination in both the elevation and azimuth dimensions, shown in Figure 2.10, reveals that the localization result is highly accurate. Thus, the CS method demonstrates its full capability in localizing sound sources.

Further, we validate the effectiveness of the CS with MUSIC method by comparing its localization performance with and without the coded cover. For a fair comparison, the signal frequency in the case without the cover is adjusted to satisfy the half-

wavelength spacing condition. From Figure 2.11, we observe that although the target at $\theta_0 = 70^\circ$ and $\psi_0 = 140^\circ$ can be successfully localized in both scenarios, the result with the coded cover exhibits significantly lower side lobes and a much sharper peak. This further validates the advantage of using a coded cover in enhancing localization accuracy.



(a) MUSIC localization with CS method using coded cover



(b) MUSIC localization with CS method without coded cover

Figure 2.11: Comparison of MUSIC-based localization using the CS method with and without a coded cover at SNR = 20 dB.

2.5 2D DOA Estimation Using CCS method

We have derived the virtual planar array response of the coded cover in Section 2.4, which can be expressed as the Kronecker product of the Vandermonde vector for each source separately. For a single source, the covariance matrix of $\mathbf{x}(\mathbf{t}) = \mathbf{a}(\theta, \psi)\mathbf{s}(\mathbf{t})$ is thus can be decomposed as

$$\begin{aligned} \mathbf{R}_{\mathbf{x}} &= \mathbf{a}(\theta, \psi)R_s\mathbf{a}(\theta, \psi)^H \\ &= \mathbf{a}(\theta, \psi)\mathbf{a}(\theta, \psi)^H R_s \\ &= (\mathbf{a}_y(\theta, \psi) \otimes \mathbf{a}_x(\theta, \psi))(\mathbf{a}_y(\theta, \psi) \otimes \mathbf{a}_x(\theta, \psi))^H R_s \\ &= (\mathbf{a}_y(\theta, \psi)\mathbf{a}_y(\theta, \psi)^H \otimes \mathbf{a}_x(\theta, \psi)\mathbf{a}_x(\theta, \psi)^H)R_s. \end{aligned}$$

Since both $\mathbf{a}_y(\theta, \psi)$ and $\mathbf{a}_x(\theta, \psi)$ exhibit a ULA response structure, the two multipliers in the Kronecker product, $\mathbf{a}_y(\theta, \psi)\mathbf{a}_y(\theta, \psi)^H$ and $\mathbf{a}_x(\theta, \psi)\mathbf{a}_x(\theta, \psi)^H$, also follow an HT structure. Consequently, the subspace defined in (2.12) is no longer a simple HT subspace. Instead, it belongs to a smaller subspace that consists of matrices formed as the Kronecker product of an HT matrix of size $J \times J$ and an HT matrix of size $I \times I$.

For the case with multiple sources: $\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\psi})\mathbf{s}(t)$, assume each source is independent, i.e., the covariance matrix of $\mathbf{s}(t)$ has diagonal structure. Therefore, the covariance matrix of $\mathbf{x}$ can be written as

$$\begin{aligned} \mathbf{R}_{\mathbf{x}} &= \mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\psi})\mathbf{R}_{\mathbf{s}}\mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\psi})^H \\ &= \sum_{n=1}^N \mathbf{a}(\theta_n, \psi_n)\mathbf{a}(\theta_n, \psi_n)^H \sigma_{s_n}^2 \\ &= \sum_{n=1}^N (\mathbf{a}_y(\theta_n, \psi_n)\mathbf{a}_y(\theta_n, \psi_n)^H \otimes \mathbf{a}_x(\theta_n, \psi_n)\mathbf{a}_x(\theta_n, \psi_n)^H) \sigma_{s_n}^2, \end{aligned}$$

where $\sigma_{s_n}^2$ is the unknown signal power of the n th source. It turns out to be a positive semidefinite (PSD) (I, J) , 2-level Toeplitz (2LT) matrix, as defined in [15]. The rank of this matrix is N . According to Theorem 2 in [15], if such a decomposition of the covariance matrix with $N > \min(I, J)$ exists, it can be uniquely found under the weak condition:

$$N \leq I \times J - \max(I, J). \quad (2.29)$$

This provides a theoretical upper bound on the maximum number of sources that can be uniquely recovered using the CCS method.

The subspace of the HT matrix of size $J \times J$ is:

$$\mathcal{S}' = \{\boldsymbol{\Sigma}'_0, \boldsymbol{\Sigma}'_1, \dots, \boldsymbol{\Sigma}'_{2J-2}\} \quad (2.30)$$

$$= \{\mathbf{I}_J\} \cup \{\mathbf{T}'_1, \dots, \mathbf{T}'_{J-1}\} \cup \{\tilde{\mathbf{T}}'_1, \dots, \tilde{\mathbf{T}}'_{J-1}\}. \quad (2.31)$$

And

$$\mathbf{a}_y(\theta_n, \psi_n)\mathbf{a}_y(\theta_n, \psi_n)^H = \sum_{j=0}^{2J-2} \alpha'_j \boldsymbol{\Sigma}'_j.$$

The subspace of the HT matrix of size $I \times I$ is:

$$\mathcal{S}'' = \{\Sigma_0'', \Sigma_1'', \dots, \Sigma_{2I-2}''\} \quad (2.32)$$

$$= \{\mathbf{I}_I\} \cup \{\mathbf{T}_1'', \dots, \mathbf{T}_{I-1}''\} \cup \{\tilde{\mathbf{T}}_1'', \dots, \tilde{\mathbf{T}}_{I-1}''\}. \quad (2.33)$$

And

$$\mathbf{a}_x(\theta_n, \psi_n) \mathbf{a}_x(\theta_n, \psi_n)^H = \sum_{i=0}^{2I-2} \alpha_i'' \Sigma_i''.$$

To estimate the spatial covariance matrix $\mathbf{R}_x$, we consider structured recovery approaches based on matrix decomposition in a specific subspace. Our goal is to express $\mathbf{R}_x$ in terms of a known basis and estimate the unknown coefficients. We propose two different methods: one utilizing independent observations from HT submatrices, and the other relying on full data collection.

2.5.1 Structured Recovery Using Two Independent Observations

The first method estimates $\mathbf{R}_x$ using data collected from two different $\mathbf{R}_y$ s, under the assumption that the signal power is normalized to 1, which will be explained later. The two $\mathbf{R}_y$ s are obtained from the first row and the first column of the coded mask, which can be treated as two separate ULA responses.

By leveraging these two observations, we can separately obtain α_i'' from the first row and α_j' from the first column. Since α_i'' reflects the signal power from the first observation and α_j' reflects the signal power from the second observation, it is not possible to directly extract the individual signal powers from them.

If the signal powers are not normalized to 1, the two observations become inconsistent and cannot be combined to estimate $\mathbf{R}_x$. However, when the signal power is normalized to 1, all necessary unknowns for estimating $\mathbf{R}_x$ are resolved. Specifically, the estimated covariance matrix is given by:

$$\mathbf{R}_x = \sum_{i=0}^{2I-2} \sum_{j=0}^{2J-2} \alpha_i'' \alpha_j' \Sigma_i'' \otimes \Sigma_j'.$$

The advantage of this method is that if only $I + J - 1$ channels are used, the number of resolvable sources is $I \times J - \max(I, J)$. The constraints on I and J also become less strict compared to the second method, allowing more sources to be localized:

$$2I - 1 \leq 4K^2 \implies I \leq \frac{4 \times 12^2 + 1}{2} \implies I_{max} = 288,$$

$$2J - 1 \leq 4K^2 \leq 4K^2 \implies J \leq \frac{4 \times 12^2 + 1}{2} \implies J_{max} = 288.$$

However, in practical scenarios, this method has significant limitations. First, real-world signals rarely have exactly equal power, making the assumption unrealistic. Second, source signals may fluctuate over time, leading to inconsistencies when collecting data in separate measurements. Due to these drawbacks, we mainly focus on the second method in the following analysis.

2.5.2 Full Data-Based Recovery

The second method utilizes all data directly collected in $\mathbf{R}_y$ to estimate $\mathbf{R}_x$. From the deviation above, for a single source s_n , we have:

$$\mathbf{a}(\theta_n, \psi_n) \mathbf{a}(\theta_n, \psi_n)^H \sigma_{s_n}^2 = \sum_{i=0}^{2I-2} \sum_{j=0}^{2J-2} \alpha_i'' \alpha_j' \sigma_{s_n}^2 \boldsymbol{\Sigma}_i'' \otimes \boldsymbol{\Sigma}_j' = \sum_{m=1}^{(2I-1)(2J-1)} \hat{\alpha}_m(\theta_n, \psi_n) \boldsymbol{\Sigma}_m.$$

Here, $\hat{\alpha}_m$ contains all unknown scalar information in the two HT subspaces and the unknown signal power, defined as $\alpha_i'' \alpha_j' \sigma_{s_n}^2$. The basis matrices $\boldsymbol{\Sigma}_m$ form the subspace of $\mathbf{R}_x$:

$$\mathcal{S} = \{\boldsymbol{\Sigma}_m \in \mathbb{C}^{M \times M} | \boldsymbol{\Sigma}_m = \boldsymbol{\Sigma}_i'' \otimes \boldsymbol{\Sigma}_j', \boldsymbol{\Sigma}_i'' \in \mathcal{S}'', \boldsymbol{\Sigma}_j' \in \mathcal{S}'\}. \quad (2.34)$$

The cardinality of $\mathcal{S}$ is $(2J-1)(2I-1)$.

For multiple sources, since each independent source shares the same subspace, the sum of the covariance matrices equals the sum of the scalars for each subspace entry:

$$\begin{aligned} \mathbf{R}_x &= \sum_{n=1}^N \mathbf{a}(\theta_n, \psi_n) \mathbf{a}(\theta_n, \psi_n)^H \sigma_{s_n}^2 \\ &= \sum_{m=1}^{(2I-1)(2J-1)} \sum_{n=1}^N \hat{\alpha}_m(\theta_n, \psi_n) \boldsymbol{\Sigma}_m \\ &= \sum_{m=1}^{(2I-1)(2J-1)} \alpha_m \boldsymbol{\Sigma}_m, \end{aligned}$$

where α_m is the unknown scalar that contains all source information.

To keep the system overdetermined, constraints on I and J must be satisfied:

$$\begin{aligned} (2I-1)(2J-1) &\leq 4K^2 \implies I \leq \frac{2 \times 12^2}{2J-1} + \frac{1}{2} \\ \implies I \times J &\leq \frac{144}{2J-1} + \frac{J}{2} + 144 \implies M_{max} = (I \times J)_{max} = 12 \times 13 = 156. \end{aligned}$$

Since one of the main objectives of CCS is to recover more sources than the CS method, N should be larger than the upper limit of the CS method, which is 23 when 12 PUs are used. This implies that the rank of $\mathbf{R}_x$ will exceed 23, which is greater than the maximum of I and J . Thus, at most $156 - 13$ sources could be localizable in the 3D scenario. Here, we choose $I = 14$ and $J = 10$. According to (2.29), this results in a recovery guarantee of $140 - 14$ sources, which still allows for the recovery of a large number of sources.

If only one acoustic pressure or particle velocity measurement is used in each of 12 sensors, the condition should be changed to

$$(2I-1)(2J-1) \leq K^2 \implies M_{max} = (I \times J)_{max} = 6 \times 7 = 42.$$

So, then at most 35 sources could be localizable in the 2D scenario, which is not large enough.

If a 2D-AVS is used for each sensor, the condition should be changed to

$$(2I - 1)(2J - 1) \leq 9K^2 \implies M_{max} = (I \times J)_{max} = 18 \times 19 = 342.$$

Now, 323 sources could be localizable in the 3D scenario. Considering the trade-off between data processing load and the requirement for the number of identifiable targets in real applications, using a planar AVS array is not necessary in this case. Additionally, when the number of targets is too large, meaning more targets are getting closer, it requires a very high resolution for DOA estimation. This implies that there will be a compromise in improving the number of identifiable sources. So, in our settings, using a PU probe is a suitable choice. Note in this context, that the less measurements we take per sensor, the cheaper the sensor becomes.

2.6 Simulation Results of CCS in 3D source localization

We perform source localization using the estimated covariance matrix of $\mathbf{x}(t)$, rather than directly using $\mathbf{R}_y$ as the CS method does. First, we analyze the beam pattern for the CCS method, which is simply the beam pattern of the coded mask response:

$$B_{CCS} = |\mathbf{a}(\theta, \psi)^H \mathbf{a}(\theta_0, \psi_0)|.$$

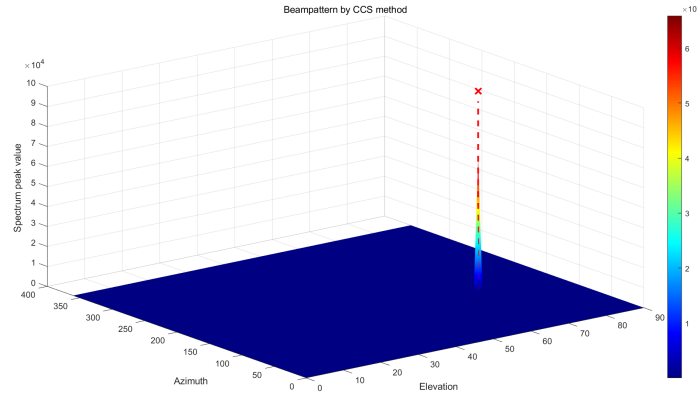
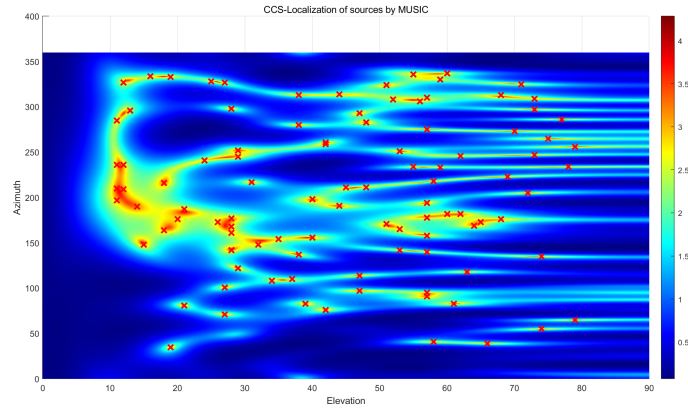


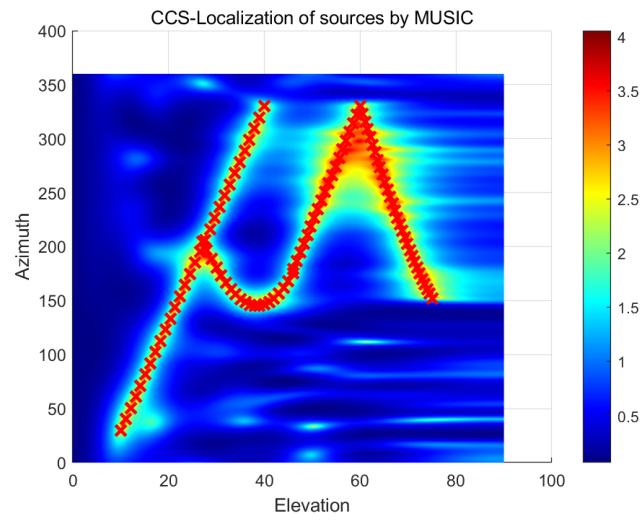
Figure 2.12: Beam pattern of CCS method with the target at $(70^\circ, 140^\circ)$

The parameter settings are identical to those used in the CS method. Figure 2.12 shows that the beamformer of the CCS method can accurately focus on the target. Additionally, no side lobes are present, as there is no compression Φ from the mask that alters the shape of the beam pattern, unlike in the CS method. Therefore, this beamformer is well-suited for the CCS method. However, given the advantages of the MUSIC algorithm, we will focus on MUSIC next.

Figure 2.13 shows that when 100 sources are present, the CCS method combined with the MUSIC algorithm is able to localize them effectively.



(a) 100 random sources



(b) 100 sources with certain pattern

Figure 2.13: MUSIC localization of 100 sources with CCS when $\text{SNR} = 20\text{dB}$

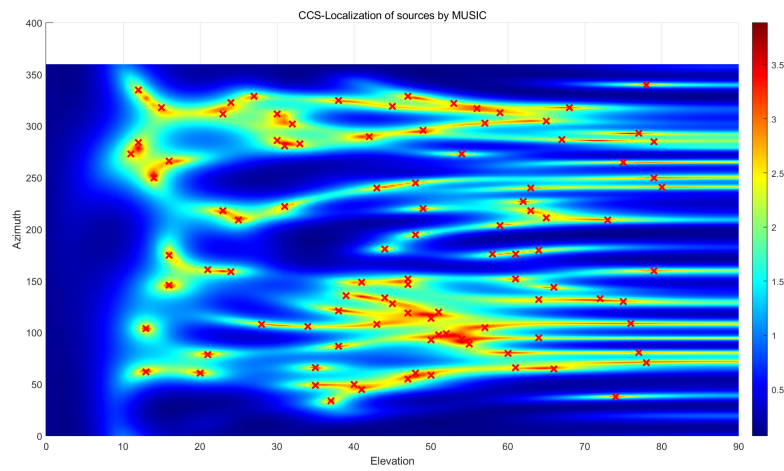
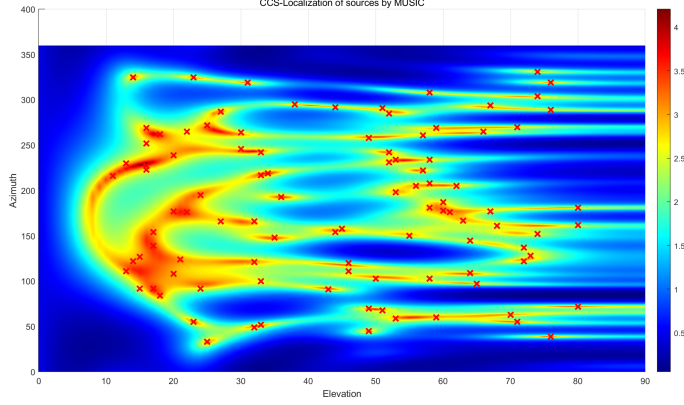
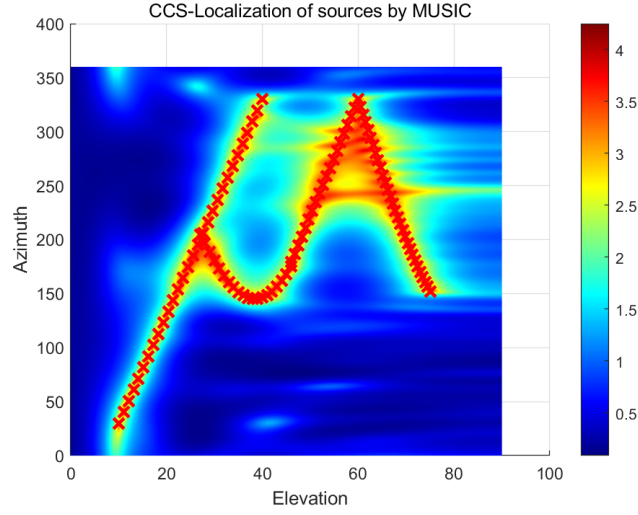


Figure 2.15: MUSIC localization of 100 10kHz sources with CCS when $\text{SNR} = 10\text{dB}$



(a) 100 random sources



(b) 100 sources with certain pattern

Figure 2.14: MUSIC localization of 100 8kHz sources with CCS when $\text{SNR} = 20\text{dB}$

We also examined the operating frequency range of the coded mask. When the signal frequency is reduced from 10 kHz to 8 kHz, the channel spacing is no longer half a wavelength. Figure 2.14 shows that the CCS method can still localize 100 sources effectively. This is attributed to the velocity gain modulation (VGM) of the beam pattern from the array of sensors incorporating acoustic particle velocity measurements [16]. VGM is independent of the frequency of the source signal and the number of sensors in the array, which extends the effective operating frequency range of the array. This indicates that the method is not strictly limited by the half-wavelength spacing condition. Additionally, when the SNR is reduced to 10 dB, Figure 2.15 demonstrates that the algorithm remains robust against noise.

2.7 Conclusion

In this chapter, we extended the CCS method from a single AVS system to an array-based acoustic measurement system. We first analyzed the feasibility of extending DOA estimation for the 1D case. We then extended it to 2D DOA estimation.

Our study demonstrates that using a 14×10 coded cover, it is possible to accurately localize 100 sound sources in three-dimensional space, even under challenging conditions with an SNR as low as 10 dB. This advancement provides valuable insights into the scalability and efficiency of advanced acoustic sensing technologies, paving the way for future improvements in DOA estimation and sound source localization methodologies.

Self-Calibration of Planar PU Array with Compression

3

Based on previous improvements in DOA estimation achieved by applying a coded mask in front of a PU sensor array, we are now considering the impact of perturbations within the system. To enhance source localization accuracy under these conditions, an additional calibration step should be integrated into the current model. Prior researches highlight the significant advantage of combining calibration with DOA estimation, and we aim to develop a similar approach tailored to our specific model.

The configuration of our model includes:

- A 3×4 PU sensor array with a spacing of 7.5 cm,
- A 14×10 planar coded cover with 1.7 cm channel spacing,
- The distance between the coded cover and array is 15 cm and the length of each channel ranges from 1 cm to 10 cm,
- 10 kHz signals are transmitted.

The proposed data model is described as:

$$\mathbf{y}(t) = \mathbf{\Phi} \mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\psi}) \mathbf{s}(t) + \mathbf{n}(t) = \mathbf{B}(\boldsymbol{\theta}, \boldsymbol{\psi}) \mathbf{s}(t) + \mathbf{n}(t) \quad (3.1)$$

$$= \mathbf{\Phi} \mathbf{x}(t) + \mathbf{n}(t) \in \mathbb{C}^{2K \times 1}, \quad (3.2)$$

where $\boldsymbol{\theta}$ collects elevation angles from N sources and $\boldsymbol{\psi}$ denotes the corresponding N azimuth angles. $\mathbf{\Phi} \in \mathbb{C}^{2K \times M}$ represents the compression matrix, with M denoting the number of channels and K the number of sensors. $\mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\psi}) \in \mathbb{C}^{M \times N}$ is the mask response matrix, $\mathbf{s}(t) \in \mathbb{C}^{N \times 1}$ is the signal vector, and N denotes the number of sources. $\mathbf{n}(t) \in \mathbb{C}^{2K \times 1}$ is the noise vector.

The approach presented here allows for accounting for perturbations in individual sensor channels. While numerous studies have addressed sensor calibration, they can generally be categorized into blind and non-blind methods. In the non-blind calibration approaches [17], [18], [19], [20], calibration is performed with the assistance of reference information. On the other hand, blind calibration methods [21], [22], [23], [24], [25] rely on redundant data rather than reference signals to perform the calibration. Although the paper [26] investigated a calibration method without requiring additional received data, it does not consider the compressed sensing framework, and the acoustic sensor used is a ULA AVS rather than a PU planar array.

Our novel contribution to calibration lies in performing calibration within the context of compressed sensing, while simultaneously exploring the sparsity pattern of the signal to enhance localization. The validity of this method is demonstrated under the assumption that only the signal part is perturbed in the received data. However, the method can be extended to scenarios where both signal and noise are affected simultaneously.

3.1 Data Model with Uncalibrated Sensors

When perturbations arise from sensor-related factors, such as uncertainties in analog electronics, sensor elements, or errors in gain, phase, or orientation within the array, these errors can affect the performance of tasks like phase retrieval and DOA estimation [27],[28].

We can express these issues using

$$g_q = \eta_q e^{j\nu_q}$$

to represent the q -th channel's unknown gain and phase mismatch. In our model, there are $Q = 2K$ channels in total. We collect the receiving errors in the diagonal matrix $\text{diag}(\mathbf{g})$ with $\mathbf{g} = [g_1, g_2, \dots, g_Q]^T$. Let us also define the vectors

$$\boldsymbol{\eta} = [\eta_1, \eta_2, \dots, \eta_Q]^T \quad \text{and} \quad \boldsymbol{\nu} = [\nu_1, \nu_2, \dots, \nu_Q]^T.$$

With the assumption that the unknown errors only affect the signal component of the data first, the uncalibrated received signal can be modeled as :

$$\mathbf{y}_0(t) = \text{diag}(\mathbf{g})[\mathbf{B}(\boldsymbol{\theta}, \boldsymbol{\psi})\mathbf{s}(t)] + \mathbf{n}(t) \in \mathbb{C}^{Q \times 1}. \quad (3.3)$$

Since both $\mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\psi})$ and $\mathbf{s}(t)$ are unknown, we construct the DOA finding matrix $\bar{\mathbf{A}}(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\psi}}) \in \mathbb{C}^{M \times G}$ to leverage the sparsity in the signal pattern. Here G is the grid size corresponding to the total number of pixels in the scanning grid. Specifically, the dictionary length for $\bar{\boldsymbol{\theta}}$ is G_1 , and the dictionary length for $\bar{\boldsymbol{\psi}}$ is G_2 . There is an equivalence among these variables: $G = G_1 \times G_2$. The g -th pixel in the scanning area can be represented by $\mathbf{d}_g = (\bar{\theta}_{g_1}, \bar{\psi}_{g_2})$ with index $g = (g_1 - 1) \times G_2 + g_2$.

The data matrix is modified as follows:

$$\mathbf{y}(t) = \Phi \bar{\mathbf{A}}(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\psi}}) \bar{\mathbf{s}}(t) + \mathbf{n}(t) = \bar{\mathbf{B}}(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\psi}}) \bar{\mathbf{s}}(t) + \mathbf{n}(t), \quad (3.4)$$

where $\bar{\mathbf{s}}$ is a length- G vector containing the source signals associated with the corresponding directions in the grid. Note that (3.3) can be expressed using the sparsity-revealing model as well by:

$$\mathbf{y}_0(t) = \text{diag}(\mathbf{g})[\bar{\mathbf{B}}(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\psi}}) \bar{\mathbf{s}}(t)] + \mathbf{n}(t) \in \mathbb{C}^{Q \times 1}. \quad (3.5)$$

Since the signal data is non-zero only in the directions corresponding to the sources, and the grid size is typically much larger than the number of sources, that is $N \ll G$, the vector $\bar{\mathbf{s}}$ exhibits a sparse structure. The non-zero elements indicate the source directions corresponding to the angles in the grid matrix $\bar{\mathbf{A}}$.

3.1.1 Self-calibration

If we collect L snapshots in the element space data matrix $\mathbf{Y}_0$, we have:

$$\mathbf{Y}_0 = \text{diag}(\mathbf{g})[\bar{\mathbf{B}}(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\psi}}) \bar{\mathbf{S}}] + \mathbf{N} \in \mathbb{C}^{Q \times L}. \quad (3.6)$$

Then, the covariance domain model can be written as

$$\mathbf{R}_\circ = \text{diag}(\mathbf{g})(\bar{\mathbf{B}}(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\psi}})\text{diag}(\boldsymbol{\sigma}_s)\bar{\mathbf{B}}^H(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\psi}}))\text{diag}^H(\mathbf{g}) + \mathbf{R}_n. \quad (3.7)$$

The covariance matrix of the uncalibrated receiving data can be represented as $\mathbf{R}_y = \bar{\mathbf{B}}(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\psi}})\text{diag}(\boldsymbol{\sigma}_s)\bar{\mathbf{B}}^H(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\psi}}) + \mathbf{R}_n$. Substituting it in (3.7), we obtain:

$$\mathbf{R}_\circ = \text{diag}(\mathbf{g})(\mathbf{R}_y - \mathbf{R}_n)\text{diag}^H(\mathbf{g}) + \mathbf{R}_n, \quad (3.8)$$

Here, we assume that the source signals $\bar{\mathbf{s}}(t)$ are uncorrelated and have a diagonal covariance matrix $\mathbb{E}\{\bar{\mathbf{s}}(t)\bar{\mathbf{s}}^H(t)\} = \text{diag}(\boldsymbol{\sigma}_s)$, which is not known. Similarly, the noise vector has a diagonal covariance matrix, given by $\mathbb{E}\{\mathbf{n}(t)\mathbf{n}^H(t)\} = \sigma_n \mathbf{I}_Q$.

Let $\mathbf{r}_\circ = \text{vec}(\mathbf{R}_\circ)$, and define the co-array manifold matrix as $\bar{\mathbf{B}}_{\text{co}}(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\psi}}) = \bar{\mathbf{B}}^*(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\psi}}) \odot \bar{\mathbf{B}}(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\psi}})$. The covariance model can then be rewritten as:

$$\mathbf{r}_\circ = \text{diag}(\mathbf{g}^* \otimes \mathbf{g})(\bar{\mathbf{B}}_{\text{co}}(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\psi}})\boldsymbol{\sigma}_s) + \sigma_n \text{vec}(\mathbf{I}). \quad (3.9)$$

In practice, the sample covariance matrix is computed as:

$$\hat{\mathbf{R}}_\circ = \frac{1}{L} \mathbf{Y}_\circ \mathbf{Y}_\circ^H, \quad (3.10)$$

where L is the number of snapshots. For simplicity, we use $\mathbf{R}_\circ$ instead of $\hat{\mathbf{R}}_\circ$ while acknowledging that only an estimate is available.

We aim to jointly estimate the Q complex (i.e., $2Q$ real) receiver gains $\mathbf{g}$ and N directions $(\boldsymbol{\theta}, \boldsymbol{\phi})$ from the received uncalibrated vector $\mathbf{r}_\circ$. Before addressing the estimation problem, we first analyze the feasibility of this approach by investigating the ambiguity and identifiability of the solution.

3.1.2 Ambiguity and Identifiability

We need to determine the identifiability conditions under which a unique solution for both the calibration parameters and the source DOAs can be obtained. The identifiability conditions for APS arrays, based on the element-space model in (3.6) with any geometry placements, have been studied in [28]. In [26], both the element-space model and the co-array data model are examined for APS and AVS ULAs. However, one issue that has not been addressed is that the aforementioned studies primarily focus on the far-field 2D case, where the structure of the array manifold is relatively simple to analyze. In contrast, when considering compression matrix and unknown far-field 3D sources, the problem is getting more complicated.

A fundamental challenge arises because neither $\text{diag}(\mathbf{g})\mathbf{B}(\boldsymbol{\theta}, \boldsymbol{\psi})$ nor $\mathbf{S}$ (or $\boldsymbol{\sigma}_s$) is known a priori, making it impossible to determine them uniquely due to an inherent complex (or real) scaling ambiguity. To resolve this ambiguity, we adopt a calibration reference sensor, located at $\mathbf{p}_1 = 0$. Specifically, we set $g_1 = 1$.

The first condition about how well-posed the problem is, is a necessary but not sufficient condition for a unique solution to exist.

1) The element-space data model: From the element-space data model (3.6), we have $2QL$ nonlinear equations in N unknown 2D DOAs (each source is characterized

by 2 unknown real valued angles), $2(Q - 1)$ unknown calibration parameters, and $2NL$ unknown amplitudes of source signals. Hence, for well-posedness of the calibration problem, we require

$$2QL \geq 2N + 2Q - 2 + 2NL, \quad (3.11)$$

which simplifies to

$$\frac{N + (Q - 1)}{(Q - N)} \leq L. \quad (3.12)$$

This condition is only meaningful when $Q > N$. However, in CCS, the goal is to recover more sources than the number of compressed measurements, i.e., $N > Q$. Therefore, this condition contradicts the fundamental design principle of CCS.

2) The co-array data model: In the covariance domain (3.7), the number of nonlinear equations is determined by the number of parameters required to characterize the covariance matrix. The covariance matrix is completely characterized by $N + 1$ real eigenvalues and $2QN - N^2 + N$ real parameters related to the orthonormal eigenvectors associated with the signal subspace.

As a result, we have $2QN - N^2 + 2N + 1$ nonlinear equations in N unknown 2D DOAs, $2(Q - 1)$ unknown calibration parameters, and N unknown source powers. Hence, for well-posedness, we require

$$2QN - N^2 + 2N + 1 \geq 3N + 2Q - 2. \quad (3.13)$$

This can be simplified to

$$Q \geq \frac{N^2 + N - 3}{2(N - 1)}. \quad (3.14)$$

In our model, we have 24 receiving channels across 12 PUs. After solving (3.14), the maximum N we can get is 46. This allows a maximum recovery of 46 signals.

Due to the non-linear nature of the estimation problem, it is not straightforward to derive the identifiability conditions based on the element-space data model (3.6) or co-array data model (3.7). Therefore, to begin with, we derive sufficient conditions for uniquely estimating $\mathbf{g}$ and $(\boldsymbol{\theta}, \boldsymbol{\psi})$ based on the assumption that $\text{diag}(\mathbf{g})\mathbf{B}(\boldsymbol{\theta}, \boldsymbol{\psi})$ or $(\text{diag}(\mathbf{g}^* \otimes \mathbf{g})\mathbf{B}_{\text{co}}(\boldsymbol{\theta}, \boldsymbol{\psi}))$ is given, with the knowledge that in practice only the column span of it is available from the measurement data.

For convenience in analysis, we explicitly write out the structure of each individual matrix. The total received array manifold is:

$$\mathbf{B}(\boldsymbol{\theta}, \boldsymbol{\psi}) = \boldsymbol{\Phi}\mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\psi}) = [\boldsymbol{\Phi}\mathbf{a}(\theta_1, \psi_1) \quad \boldsymbol{\Phi}\mathbf{a}(\theta_2, \psi_2) \quad \cdots \quad \boldsymbol{\Phi}\mathbf{a}(\theta_N, \psi_N)] \quad (3.15)$$

$$= [\mathbf{b}(\theta_1, \psi_1) \quad \mathbf{b}(\theta_2, \psi_2) \quad \cdots \quad \mathbf{b}(\theta_N, \psi_N)] \quad (3.16)$$

and the known compression matrix $\boldsymbol{\Phi}$:

$$\boldsymbol{\Phi} = [\mathbf{E}_1\mathbf{u}(\phi_1) \quad \mathbf{E}_2\mathbf{u}(\phi_2) \quad \cdots \quad \mathbf{E}_M\mathbf{u}(\phi_M)] \quad (3.17)$$

$$= \begin{bmatrix} \beta_{1,1} & \beta_{1,2} & \cdots & \beta_{1,M} \\ \beta_{1,1}\cos(\phi_{1,1}) & \beta_{1,2}\cos(\phi_{1,2}) & \cdots & \beta_{1,M}\cos(\phi_{1,M}) \\ \vdots & \vdots & \vdots & \vdots \\ \beta_{K,1} & \beta_{K,2} & \cdots & \beta_{K,M} \\ \beta_{K,1}\cos(\phi_{K,1}) & \beta_{K,2}\cos(\phi_{K,2}) & \cdots & \beta_{K,M}\cos(\phi_{K,M}) \end{bmatrix} \quad (3.18)$$

$$= \begin{bmatrix} \frac{1}{L_{1,1}}e^{-jk_b(L_{1,1}+d_1)} & \frac{1}{L_{1,2}}e^{-jk_b(L_{1,2}+d_2)} & \cdots & \frac{1}{L_{1,M}}e^{-jk_b(L_{1,M}+d_M)} \\ \frac{1}{L_{1,1}}e^{-jk_b(L_{1,1}+d_1)}\cos(\phi_{1,1}) & \frac{1}{L_{1,2}}e^{-jk_b(L_{1,2}+d_2)}\cos(\phi_{1,2}) & \cdots & \frac{1}{L_{1,M}}e^{-jk_b(L_{1,M}+d_M)}\cos(\phi_{1,M}) \\ \vdots & \vdots & \vdots & \vdots \\ \frac{1}{L_{K,1}}e^{-jk_b(L_{K,1}+d_K)} & \frac{1}{L_{K,2}}e^{-jk_b(L_{K,2}+d_K)} & \cdots & \frac{1}{L_{K,M}}e^{-jk_b(L_{K,M}+d_K)} \\ \frac{1}{L_{K,1}}e^{-jk_b(L_{K,1}+d_K)}\cos(\phi_{K,1}) & \frac{1}{L_{K,2}}e^{-jk_b(L_{K,2}+d_K)}\cos(\phi_{K,2}) & \cdots & \frac{1}{L_{K,M}}e^{-jk_b(L_{K,M}+d_K)}\cos(\phi_{K,M}) \end{bmatrix} \quad (3.19)$$

$$= \begin{bmatrix} \frac{1}{L_{1,1}}e^{-j2\pi\frac{L_{1,1}+d_1}{\lambda}} & \frac{1}{L_{1,2}}e^{-j2\pi\frac{L_{1,2}+d_2}{\lambda}} & \cdots & \frac{1}{L_{1,M}}e^{-j2\pi\frac{L_{1,M}+d_M}{\lambda}} \\ \frac{l-d_1}{L_{1,1}^2}e^{-j2\pi\frac{L_{1,1}+d_1}{\lambda}} & \frac{l-d_2}{L_{1,2}^2}e^{-j2\pi\frac{L_{1,2}+d_2}{\lambda}} & \cdots & \frac{l-d_M}{L_{1,M}^2}e^{-j2\pi\frac{L_{1,M}+d_M}{\lambda}} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{1}{L_{K,1}}e^{-j2\pi\frac{L_{K,1}+d_1}{\lambda}} & \frac{1}{L_{K,2}}e^{-j2\pi\frac{L_{K,2}+d_2}{\lambda}} & \cdots & \frac{1}{L_{K,M}}e^{-j2\pi\frac{L_{K,M}+d_M}{\lambda}} \\ \frac{l-d_1}{L_{K,1}^2}e^{-j2\pi\frac{L_{K,1}+d_1}{\lambda}} & \frac{l-d_2}{L_{K,2}^2}e^{-j2\pi\frac{L_{K,2}+d_2}{\lambda}} & \cdots & \frac{l-d_M}{L_{K,M}^2}e^{-j2\pi\frac{L_{K,M}+d_M}{\lambda}} \end{bmatrix}. \quad (3.20)$$

Here, l is the distance from the entrance of the coded cover to the PU array, d_m is the length of each channel and $L_{k,m}$ is the distance from the exit of the m th channel to the k th PU, which is decided by the first two values and the position of the sensor:

$$L_{k,m} = \sqrt{p_{k,m}^2(x) + p_{k,m}^2(y) + (l - d_m)^2},$$

where $p_{k,m}(x) = m_m(x) - r_k(x)$, $p_{k,m}(y) = m_m(y) - r_k(y)$.

The structure of the virtual planar array response is

$$\mathbf{a}(\theta, \psi) = \begin{bmatrix} 1 \\ e^{j2\pi\Delta\sin\theta\cos\psi} \\ \vdots \\ e^{j2\pi(I-1)\Delta\sin\theta\cos\psi} \\ \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} \\ e^{j2\pi\Delta\sin\theta\sin\psi} \cdot 1 \\ e^{j2\pi\Delta\sin\theta\sin\psi} \cdot e^{j2\pi\Delta\sin\theta\cos\psi} \\ \vdots \\ e^{j2\pi\Delta\sin\theta\sin\psi} \cdot e^{j2\pi(I-1)\Delta\sin\theta\cos\psi} \\ \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} \\ \vdots \\ \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} \\ e^{j2\pi(J-1)\Delta\sin\theta\sin\psi} \cdot 1 \\ e^{j2\pi(J-1)\Delta\sin\theta\sin\psi} \cdot e^{j2\pi\Delta\sin\theta\cos\psi} \\ \vdots \\ e^{j2\pi(J-1)\Delta\sin\theta\sin\psi} \cdot e^{j2\pi(I-1)\Delta\sin\theta\cos\psi} \end{bmatrix}.$$

Through the analysis of their product, we see that it is difficult to extract individual DOA information from $\mathbf{B}(\boldsymbol{\theta}, \boldsymbol{\psi})$ directly because of the existence of the compression matrix. The DOA information is embedded in the phase of the elements of $\mathbf{B}(\boldsymbol{\theta}, \boldsymbol{\psi})$, but we can not extract single DOA information simply from the phase of $\mathbf{B}(\boldsymbol{\theta}, \boldsymbol{\psi})$. Due to this nonlinearity, the condition discussed below is necessary but not sufficient for the uniqueness of DOA estimation with the assumption of $\text{diag}(\mathbf{g})\mathbf{B}(\boldsymbol{\theta}, \boldsymbol{\psi})$ or $(\text{diag}(\mathbf{g}^* \otimes \mathbf{g})\mathbf{B}_{\text{co}}(\boldsymbol{\theta}, \boldsymbol{\psi}))$ is given.

1) The element-space data model: For deriving the sufficient conditions, let us define the phase of $\text{diag}(\mathbf{g})\mathbf{B}(\boldsymbol{\theta}, \boldsymbol{\psi})$ as

$$\rho_q(n) = \frac{1}{2\pi} \text{angle}(g_q[\mathbf{B}(\boldsymbol{\theta}, \boldsymbol{\psi})]_{qn}) = \frac{1}{2\pi} \text{angle}(b_q(\theta_n, \psi_n)) + \nu_q, \quad (3.21)$$

for $q = 1, \dots, Q$ and $n = 1, \dots, N$. Further, if we collect the above equations compactly with the introduction of the phase vector $\boldsymbol{\rho}_n = [\rho_1(n), \rho_2(n), \dots, \rho_Q(n)]^T$, we have

$$\boldsymbol{\rho}_n = [\mathbf{I}_Q \quad \mathbf{I}_Q] \begin{bmatrix} \frac{1}{2\pi} \text{angle}(\mathbf{b}(\theta_n, \psi_n)) \\ \mathbf{d} \end{bmatrix}. \quad (3.22)$$

Since even though $\nu_1 = 0$, the angle of $[\mathbf{B}(\boldsymbol{\theta}, \boldsymbol{\psi})]_{qn}$ is generally non-zero, there are Q equations in $Q + 1$ unknowns.

When $N = 1$, it is possible to solve (3.22) with one more channel's phase error.

When $N \geq 2$, we can use the vector difference to eliminate the phase error:

$$\boldsymbol{\rho}_n - \boldsymbol{\rho}_1 = \frac{1}{2\pi} \text{angle}(\mathbf{b}(\theta_n, \psi_n)) - \frac{1}{2\pi} \text{angle}(\mathbf{b}(\theta_1, \psi_1)) \quad (3.23)$$

$$= \frac{1}{2\pi} \text{angle}(\mathbf{b}(\theta_n, \psi_n) \oslash \mathbf{b}(\theta_1, \psi_1)), \quad (3.24)$$

where $\oslash$ means element-wise division.

Now, there are $Q(N - 1)$ linearly independent equations in $2N$ unknown angles. Since $Q(N - 1) \geq 2N$ can be easily satisfied, the system is typically overdetermined.

2) The co-array data model: First, the phase of $\text{diag}(\mathbf{g}^* \otimes \mathbf{g})\mathbf{B}_{\text{co}}(\boldsymbol{\theta}, \boldsymbol{\psi})$ is

$$\rho_{pq}(n) = \frac{1}{2\pi} \text{angle}(g_p^* g_q([\mathbf{B}^*(\boldsymbol{\theta}, \boldsymbol{\psi})]_{pn} \odot [\mathbf{B}(\boldsymbol{\theta}, \boldsymbol{\psi})]_{qn})) \quad (3.25)$$

$$= \left(\frac{1}{2\pi} \text{angle}(b_p(\theta_n, \psi_n)) - \frac{1}{2\pi} \text{angle}(b_q(\theta_n, \psi_n)) \right) - (\nu_p - \nu_q) \quad (3.26)$$

$$= \left(\frac{1}{2\pi} \text{angle}\left(\frac{b_p(\theta_n, \psi_n)}{b_q(\theta_n, \psi_n)}\right) \right) - (\nu_p - \nu_q), \quad (3.27)$$

for $p, q = 1, \dots, Q$, with $p \neq q$ and $n = 1, \dots, N$.

Defining $\boldsymbol{\rho}_n = [\rho_{12}(n), \rho_{13}(n), \dots, \rho_{Q1}(n), \dots, \rho_{Q(Q-1)}(n)]$, we can compactly express (3.27) as:

$$\boldsymbol{\rho}_n = [\mathbf{D} \quad -\mathbf{D}] \begin{bmatrix} \frac{1}{2\pi} \text{angle}(\mathbf{b}(\theta_n, \psi_n)) \\ \mathbf{d} \end{bmatrix}, \quad (3.28)$$

where $\mathbf{D} \in \mathcal{R}^{(Q^2-Q) \times Q}$ is the difference matrix with only three types of entries: $\{-1, 0, 1\}$. The rank of $\mathbf{D}$ is always $Q - 1$. Multiplying both sides of the (3.28) by the pseudoinverse of $\mathbf{D}$ yields the following system of equations:

$$\mathbf{D}^+ \boldsymbol{\rho}_n = \frac{1}{2\pi} \text{angle}(\mathbf{b}(\theta_n, \psi_n)) - \boldsymbol{\nu}.$$

There are Q equations but $Q + 1$ unknowns when $N = 1$. To determine two unknown angles, at least two phase errors must be known.

If $N \geq 2$, we can eliminate the phase errors $\mathbf{d}$ by difference again:

$$\boldsymbol{\rho}_n - \boldsymbol{\rho}_1 = \mathbf{D} \left(\frac{1}{2\pi} \text{angle}(\mathbf{b}(\theta_n, \psi_n)) - \frac{1}{2\pi} \text{angle}(\mathbf{b}(\theta_1, \psi_1)) \right) \quad (3.29)$$

$$= \mathbf{D} \left(\frac{1}{2\pi} \text{angle}(\mathbf{b}(\theta_n, \psi_n) \oslash \mathbf{b}(\theta_1, \psi_1)) \right), \quad n = 2, \dots, N \quad (3.30)$$

Compared to (3.23) we derived in the element space model, we found that the two equations have a similar structure and both form an overdetermined system. But in the covariance domain, we have more equations, so in the covariance domain, we can obtain more constraints on the solution, leading to potentially improved estimation accuracy.

Since there are $Q \times (Q - 1)$ sub-equations with 4 unknowns, if $Q \geq 3$, the system can potentially be solved without known sources. However, due to the strong nonlinearity of directional information, i.e., (θ, ψ) , in the function $\mathbf{b}(\theta, \psi)$, knowing some pairs of source DOAs can help identify the remaining DOAs more accurately. For example, if (θ_1, ψ_1) is unknown and the solution obtained from other equations is inaccurate, then solving,

$$\frac{1}{2\pi} \text{angle}(\mathbf{b}(\theta_n, \psi_n)) = \mathbf{D}^+(\boldsymbol{\rho}_n - \boldsymbol{\rho}_1) + \frac{1}{2\pi} \text{angle}(\mathbf{b}(\theta_1, \psi_1))$$

would lead to a more inaccurate solution for (θ_n, ψ_n) .

In summary, in 2D DOA estimation with compressed sensing, we need at least one reference channel and possibly some known source directions for uniquely estimating $\mathbf{g}$ and (θ, ψ) .

3.1.3 Self-Calibration Algorithm

From the previous derivation, we know that applying the DOA estimation method in the covariance domain still preserves the compressed sensing framework and provides more accurate results. From the model equation defined in (3.9), we can introduce the calibration matrix $\text{diag}(\mathbf{c}^* \otimes \mathbf{c}) = \text{diag}^{-1}(\mathbf{g}^* \otimes \mathbf{g})$ to correct receiving data through:

$$\text{diag}(\mathbf{c}^* \otimes \mathbf{c})(\mathbf{r}_0 - \sigma_n \text{vec}(\mathbf{I}_Q)) = \text{diag}(\mathbf{r}_0 - \sigma_n \text{vec}(\mathbf{I}_Q))(\mathbf{c}^* \otimes \mathbf{c}) = \bar{\mathbf{B}}_{co}(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\psi}})\boldsymbol{\sigma}_s. \quad (3.31)$$

To simplify notation, we can write $(\mathbf{c}^* \otimes \mathbf{c})$ as $\text{vec}(\mathbf{C})$ with $\mathbf{C} = \mathbf{c}\mathbf{c}^H$. Then (3.31) can be written as

$$\underbrace{\left[\text{diag}(\mathbf{r}_0 - \sigma_n \text{vec}(\mathbf{I}_Q)) \quad -\bar{\mathbf{B}}_{co}(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\psi}}) \right]}_{\mathbf{G}_{co}} \underbrace{\begin{bmatrix} \text{vec}(\mathbf{C}) \\ \boldsymbol{\sigma}_s \end{bmatrix}}_{\boldsymbol{\gamma}_{co}} = \mathbf{0}.$$

The above problem transforms the original non-linear problem into a linear one. But if we further explore the rank one structure of $\mathbf{C}$ and the sparsity of $\boldsymbol{\sigma}_s$, the problem can be further reformulated as:

$$\min_{\mathbf{C}, \boldsymbol{\sigma}_s} \|\mathbf{G}_{co}\boldsymbol{\gamma}_{co}\|_2^2 + \lambda \|\boldsymbol{\sigma}_s\|_0 \quad (3.32)$$

$$\text{s.t.} \quad \mathbf{C} = \mathbf{c}\mathbf{c}^H \quad (3.33)$$

$$\boldsymbol{\sigma}_s \succeq \mathbf{0} \quad (3.34)$$

$$\mathbf{c}(1) = \mathbf{1} \quad (3.35)$$

$$\boldsymbol{\sigma}_s(k) = 1, k \in \mathcal{K} \quad (3.36)$$

where $\sigma_s(k) = 1, k \in \mathcal{K}$ represents the constraint with the knowledge of the known source indices collected in $\mathcal{K}$.

To transform the optimization into a convex problem, we relax it into an SDP problem:

$$\min_{\mathbf{C}, \sigma_s} \quad \|\mathbf{G}_{co}\gamma_{co}\|_2^2 + \lambda \|\sigma_s\|_1 \quad (3.37)$$

$$\text{s.t.} \quad \begin{bmatrix} \mathbf{C} & \mathbf{c} \\ \mathbf{c}^H & 1 \end{bmatrix} \succeq 0 \quad (3.38)$$

$$\sigma_s \succeq \mathbf{0} \quad (3.39)$$

$$\mathbf{c}(1) = 1 \quad (3.40)$$

$$\sigma_s(k) = 1, k \in \mathcal{K} \quad (3.41)$$

For the choice of the regularization parameter λ , we can use any standard method adopted in sparse signal recovery [29]. In practice, for the finite snapshot scenario, $\mathbf{C}$ obtained after solving the above problem might not be rank one and the closest estimates of the calibration parameters can be obtained from the first dominant singular vector of $\mathbf{C}$.

Remark 1: If the phase and gain errors affect both the signal and noise components, they will influence the total received signal. Such distortions may arise from imperfections in sensor analog elements, front-end electronics, or other hardware non-idealities.

In this case, the corresponding convex optimization problem should be modified accordingly as below:

$$\underbrace{\begin{bmatrix} \text{diag}(\mathbf{r}_0) & -\bar{\mathbf{B}}_{co}(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\psi}}) \end{bmatrix}}_{\mathbf{G}_{co}} \underbrace{\begin{bmatrix} \text{vec}(\mathbf{C}) \\ \sigma_s \end{bmatrix}}_{\gamma_{co}} = \sigma_n \text{vec}(\mathbf{I}_Q)$$

$$\min_{\mathbf{C}, \sigma_s} \quad \|\mathbf{G}_{co}\gamma_{co} - \mathbf{r}_n\|_2^2 + \lambda \|\sigma_s\|_1 \quad (3.42)$$

$$\text{s.t.} \quad \begin{bmatrix} \mathbf{C} & \mathbf{c} \\ \mathbf{c}^H & 1 \end{bmatrix} \succeq 0 \quad (3.43)$$

$$\sigma_s \succeq \mathbf{0} \quad (3.44)$$

$$\mathbf{c}(1) = 1 \quad (3.45)$$

$$\sigma_s(k) = 1, k \in \mathcal{K} \quad (3.46)$$

However, in our case, we will further focus on perturbations in the compression matrix Φ . Therefore, the initially proposed problem formulation remains more suitable for our study.

3.1.4 Simulation Results of Self-calibration Algorithm

In this section, we first verify the accuracy of the proposed method in recovering DOAs with unperturbed simulated data. Then, we evaluate the performance of the proposed method for the joint estimation of the DOAs and calibration parameters using a designed experimental model. The test environment is set at SNR of 20 dB. In the 2D case, the one-dimensional scanning angle ranges from 20° to 160° with a resolution of 1° . In the 3D case, the elevation angle θ is uniformly sampled from 0° to 60° with a

resolution of 2° , while the azimuth angle ψ is uniformly sampled from 0° to 350° with a resolution of 10° . The narrowband far-field sources are randomly placed on the grid.

Without loss of generality, we select the first channel of the first PU in the array as the reference channel, with a fixed gain of 1 and a phase of 0° .

Without perturbations in the received data To validate the proposed method, we first evaluate its performance using unperturbed data before considering perturbations. This approach allows us to establish a baseline and ensure that the method can work for unperturbed data as well. By confirming its effectiveness in a perturbation-free scenario, we can better assess the impact of perturbations in later experiments.

The algorithm is first tested in the 2D case using a linear array consisting of 4 PUs with equal sensor spacing of 7.5 cm. A linear coded cover with 14 channels is used, the linear mask design matches that of the planar coded cover. The front of the coded cover is positioned 15 cm away from the array. The 12 sources are distributed at uniformly spaced angles between 40° and 140° with an interval of 9° .

Figure 3.1 shows the localization result for 12 unknown sources. The peaks match the real targets accurately. This demonstrates the initial feasibility of the self-calibration algorithm in the covariance domain.

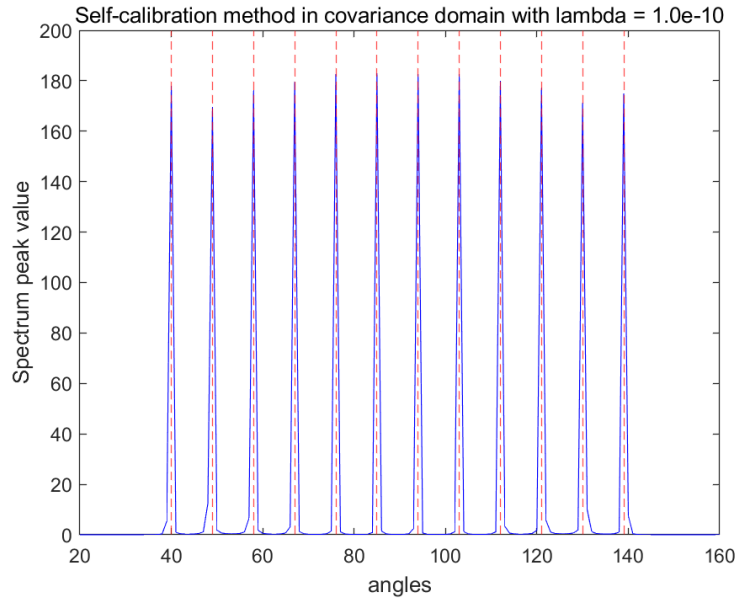


Figure 3.1: DOA estimation using unperturbed data for 12 unknown sources with 9° spacing.

However, when the spacing between sources is reduced, the recovery ability is limited. For example, when the target spacing is 8° with the same number of sources, the result in Figure 3.2 shows a limitation due to resolution. This also suggests that, in the 3D case, the scanning of the elevation angle would be further constrained.

With the planar coded cover placed in front of the Rect-7.5 array, as previously set, we randomly placed 33 sources in a 3D space. This scenario simulates the case where the source directions are completely unknown and no prior information is available

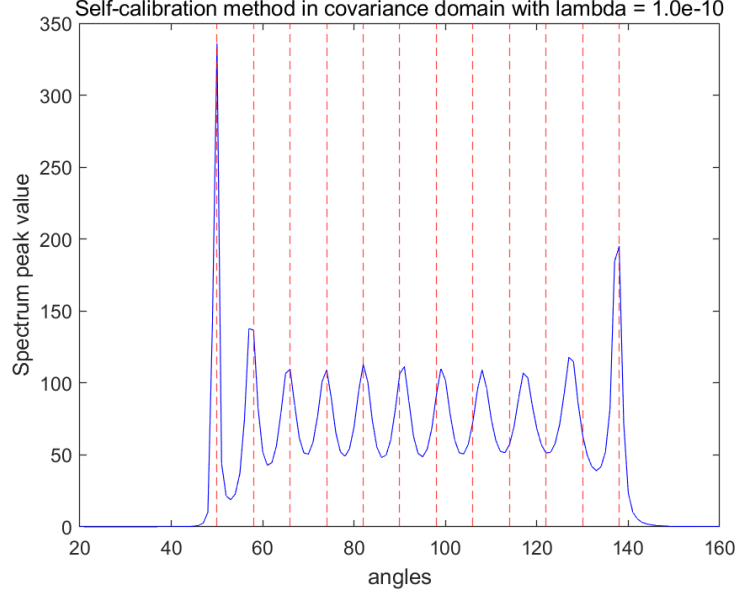


Figure 3.2: DOA estimation using unperturbed data for 12 unknown sources with 8° spacing.

about them. The results for this case are presented in Figure 3.3a, where it can be observed that the estimated peaks closely correspond to the actual target directions. However, there are small side peaks. These side peaks, while not significantly affecting the estimation accuracy, indicate that the algorithm might have some sensitivity to noise or ambiguity about the DOA localizations.

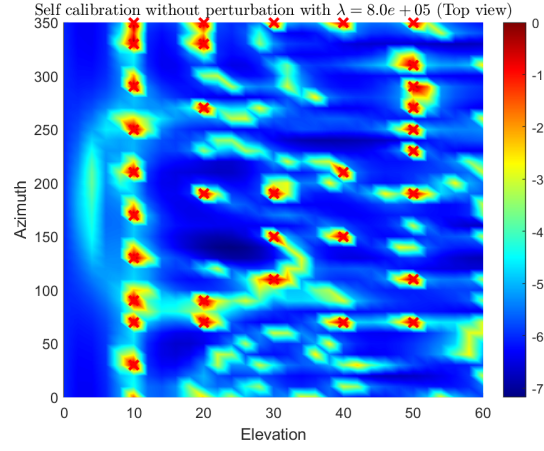
Next, we introduce one known source to examine its effect on estimation accuracy, with results shown in Figure 3.3b. The addition of a known source improves localization by reducing side peaks and refining the estimation of unknown sources. As shown in Figure 3.3c, introducing two more known sources further enhances accuracy, with a greater reduction in side peaks. This demonstrates that incorporating prior knowledge about known sources helps mitigate uncertainties, improving both the precision and robustness of the source localization.

From another perspective, to investigate the influence of prior knowledge, we compare the optimized vector $\mathbf{c}$ with the true $\mathbf{g}^{-1}$. Three types of error are used to evaluate the accuracy of the reconstructed calibration parameters, as shown in Table 3.1.

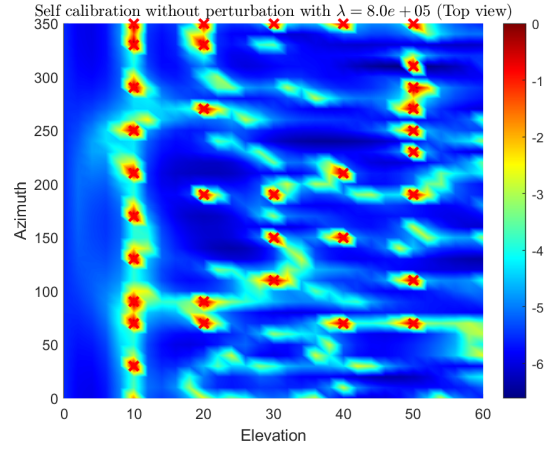
Error Type	Formula
Relative Error (L2 norm)	$\frac{\ \mathbf{g}^{-1} - \mathbf{c}\ _2}{\ \mathbf{g}^{-1}\ _2}$
Mean Gain Error	$\frac{1}{2k} \sum_{i=1}^{2k} \left \mathbf{g}_i^{-1} - \mathbf{c}_i \right $
Mean Phase Error	$\frac{1}{2k} \sum_{i=1}^{2k} \left \angle \mathbf{g}_i^{-1} - \angle \mathbf{c}_i \right $

Table 3.1: Error Metrics

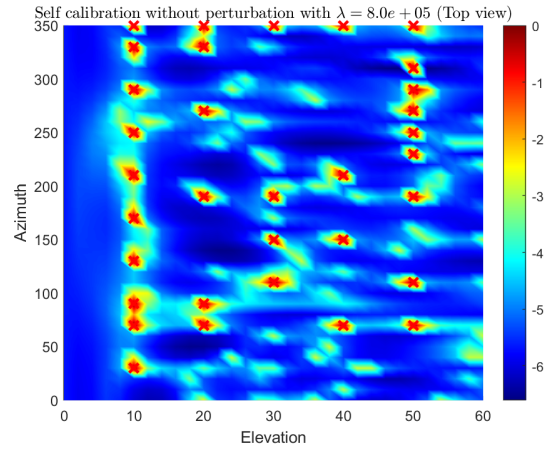
The results are shown in Table 3.2. As the number of known sources increases, all of the errors decrease, except for the mean phase error in the case of 2 known sources. This also implies recovering exact phase value is more challenging since there might be



(a) 0 known source



(b) 1 known source



(c) 2 known sources

Figure 3.3: DOA estimation using unperturbed data with 33 sources in 3D case

an issue of phase ambiguity.

Degree of prior knowledge	Relative Error	Mean Gain Error	Mean Phase Error
0 known source	0.0163	0.0148	0.0031
1 known source	0.0126	0.0109	0.0029
2 known sources	0.0118	0.0098	0.0032
10 known sources	0.0110	0.0094	0.0025

Table 3.2: Errors in the estimated correction vector compared to the true values in unperturbed data case.

In conclusion, the proposed method has been shown to work effectively for unperturbed data. Furthermore, it has been demonstrated that incorporating additional prior knowledge leads to improvements in localization accuracy. Specifically, the results indicate that as the number of known sources increases, the estimation accuracy enhances, further validating the robustness of the method.

Perturbation from gain and phase errors We investigate the performance of the algorithm under perturbed data conditions. We manually introduce errors in the sensor-received signals through a diagonal matrix $\mathbf{g}$, which represents gain and phase errors. The gain errors are drawn from a uniform distribution within the range $[-3, 3]$ dB, while the phase errors are selected from a uniform distribution over $[-20^\circ, 20^\circ]$.

For the 2D case with a linear array, Figure 3.4 shows the result for 12 unknown sources with equal spacing of 9° . The peaks are still very clear even with randomly perturbed data. This demonstrates the efficiency of our algorithm. It establishes the foundation for feasibility with perturbed data in the 3D case.

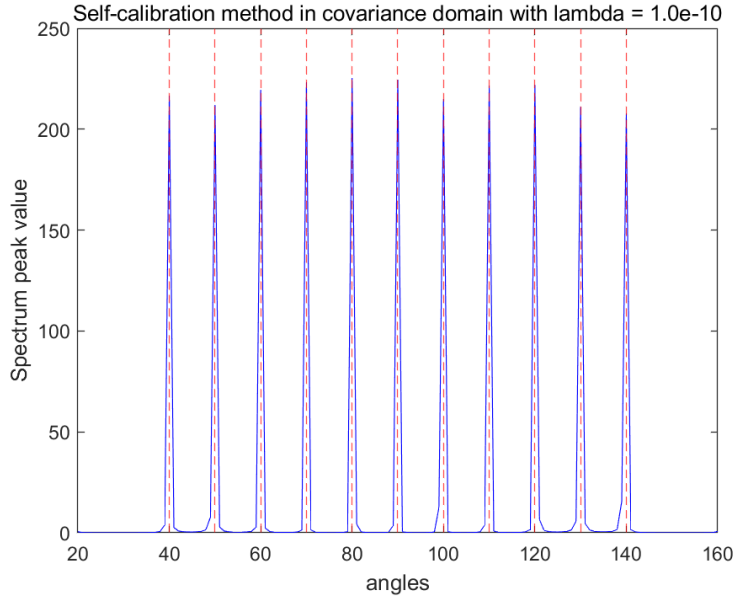
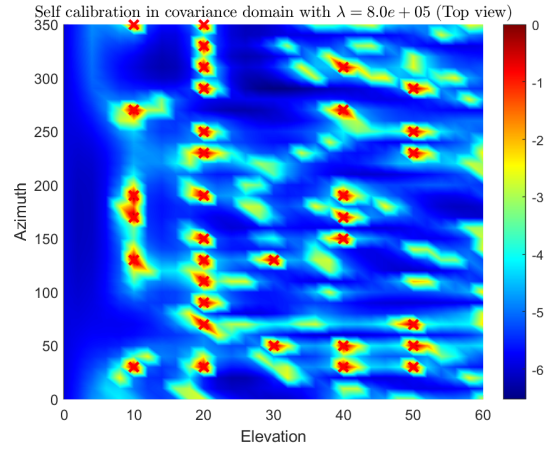
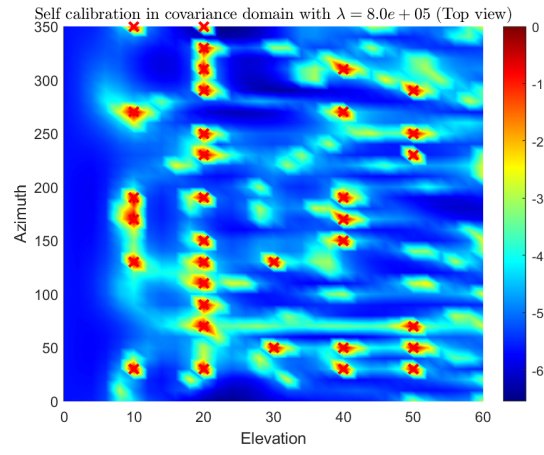


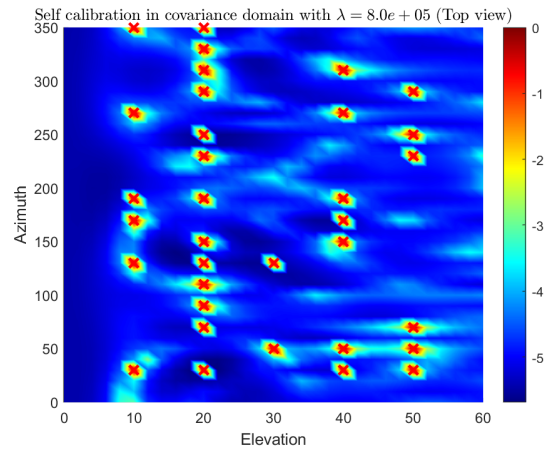
Figure 3.4: DoA estimation using perturbed data for 12 unknown sources with 9° spacing.



(a) 0 known source



(b) 2 known sources



(c) 10 known sources

Figure 3.5: DOA estimation with self-calibration using uniformly perturbed data for 34 sources

Figure 3.5 shows the top view of the estimation results for varying numbers of known sources in 3D case: (a) 0 known source, (b) 2 known sources, and (c) 10 known sources. In comparison, Figure 3.6 shows the corresponding localization result with the classical CCS method. We can see that even without prior knowledge, the localization result

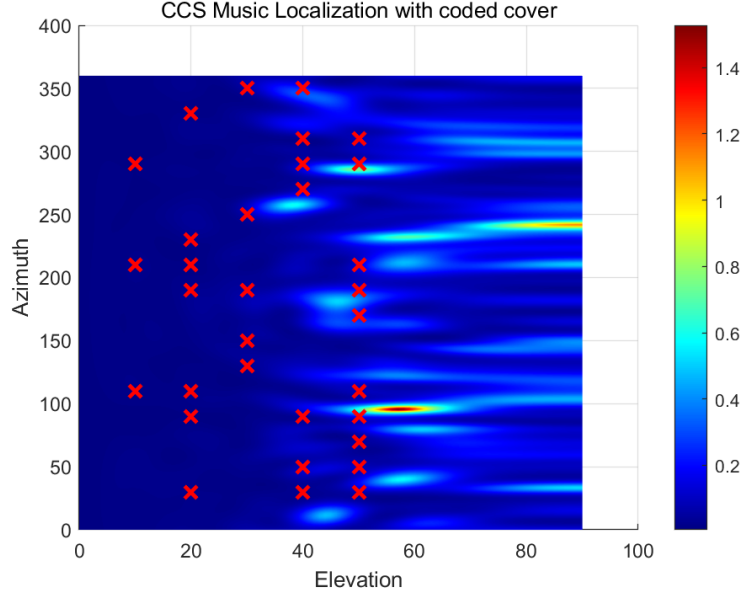


Figure 3.6: Top view of DoA result using CCS method with uniformly perturbed data

from the self-calibration method is already good enough. Without exact knowledge of the receiving compression matrix, the classical CCS method can not work well. Further, taking more known sources into consideration in the self-calibration method, the peaks are getting sharper. The effect of increasing the number of known sources on the localization accuracy and reduction of side lobes is clearly visible.

Table 3.3 shows that with 10 known sources, the calibration matrix becomes closer to the ground truth, which explains why the DOA estimation result is significantly better compared to the other cases.

Degree of prior knowledge	Relative Error	Mean Gain Error	Mean Phase Error
0 known source	0.0269	0.0263	0.0031
2 known sources	0.0107	0.0099	0.0025
10 known sources	0.0096	0.0090	0.0021

Table 3.3: Errors in the estimated correction vector compared to the true values in uniformly perturbed case.

Remark 2: The results shown above can be improved by increasing the scanning angular resolution. However, this also leads to a higher computational load. Moreover, there is a limit to the computational load that MATLAB can handle. Therefore, I present the results with a moderate resolution that balances accuracy and feasibility.

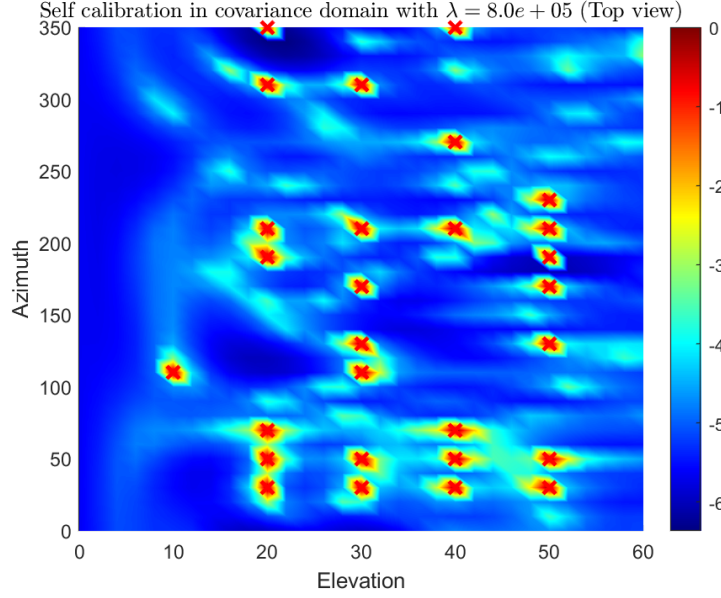


Figure 3.7: Recovered DOAs after applying the self-calibration method.

3.2 Resolution-Enhanced Self-Calibration with CCS Method

From the results of the self-calibration method with compression, we observed that the estimated DOAs are still not sharp enough, and side lobes persist despite the presence of multiple known sources. To address this issue, we propose an improved method. First, the self-calibration method is applied to estimate the correction matrix $\mathbf{C}$ based on the covariance of the received data. Then, the CCS method is performed on the corrected data to achieve higher resolution.

To test the feasibility of this method, the same setting is chosen as the case described in Section 3.1.4. We consider 28 sources, including 2 known sources. After applying the self-calibration method in the covariance domain, the recovered DOAs are shown in Figure 3.7. Then, the obtained correction matrix is used to adjust the data. The result after applying the CCS method is shown in Figure 3.8. It can be observed that the resolution is significantly improved after CCS, demonstrating the feasibility of the proposed method.

3.3 Conclusion

In this chapter, we explored a self-calibration approach for a planar PU array within the framework of compressed sensing. This method integrates calibration and signal recovery by leveraging the inherent sparsity of the signals, leading to improved DOA estimation accuracy in the presence of per-channel receiving errors.

We first analyzed the ambiguity and identifiability of the problem, establishing the theoretical conditions for self-calibration in compressed sensing. A methodology was then introduced to jointly estimate the correction matrix and recover the DOAs, en-

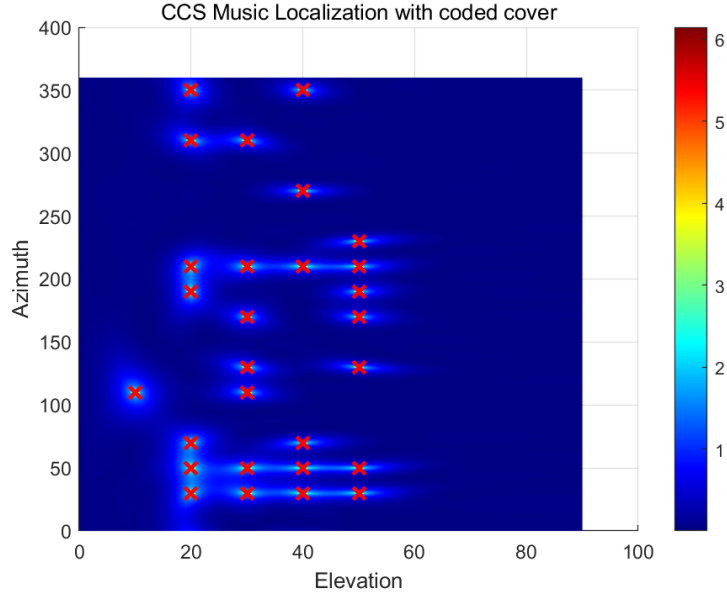


Figure 3.8: Final DOA estimation after applying the CCS method.

sureing improved robustness against calibration errors. Simulation results demonstrated the effectiveness of this approach, though challenges such as insufficient DOA resolution and residual side lobes remained.

To further enhance resolution, we proposed an improved method by integrating self-calibration with the CCS approach. The correction matrix was estimated from the received data covariance, followed by CCS-based signal processing on the corrected data. This method significantly improved resolution and reduced side lobes, demonstrating its potential for more accurate source localization under practical conditions.

Geometric Mismatch Correction

4

In the previous chapter, we addressed sensor-related calibration issues, such as phase and amplitude mismatches. However, accurate DOA estimation also depends on the precise geometric configuration of the sensing system. Small deviations in the geometry, such as sensor positions or array rotations, can significantly distort the compression model and degrade performance. In this chapter, we examine the impact of geometric mismatches and propose a method for compensating these distortions.

4.1 The Impacts and Categories of Geometric Mismatch

To illustrate the severity of geometric errors, we consider a case where the actual distance between the planar PU array and the front of the coded mask is 16 cm, instead of the nominal value of 15 cm. Figure 4.1 presents the DOA estimation results using both the CCS method and a self-calibration method with 34 sources, including 10 known ones. Despite the use of known sources, neither method yields satisfactory localization results, highlighting the need for a robust geometric correction strategy.

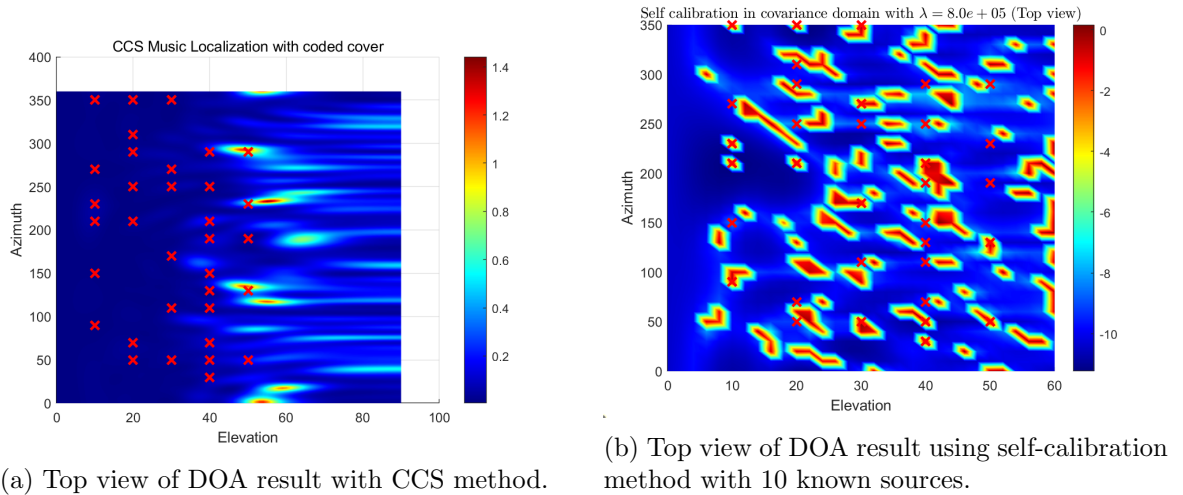


Figure 4.1: DOA estimation using geometrically perturbed data with 34 sources, including 10 known sources.

The geometric parameters can be categorized into three main types:

- l : The distance between the coded mask entry and the sensor array.
- ϵ : The positional deviation of the first sensor.

- ζ : The rotation angle of the planar array relative to the plane parallel to the coded mask.

To better illustrate, Figure 4.2 depicts these types of errors from an x-z plane perspective.

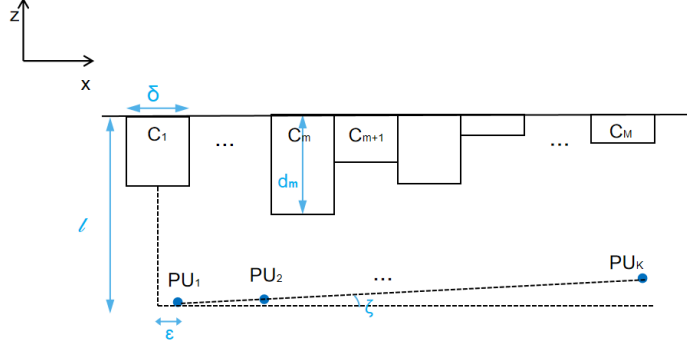


Figure 4.2: Visualization of geometric errors in the x-z plane.

Unlike receiver-based errors, which are typically row-wise and uniform per channel, geometric mismatches cause element-wise phase and gain distortions. As a result, the calibration matrix loses its diagonal structure, necessitating modifications to the optimization approach.

To address the problem caused by geometric mismatches, we first analyze the element-wise correction of the compression method and then propose a practical approach based on our compression model.

4.2 STLS-MMV Method with Perturbation in the Compression Matrix

When geometric mismatches occur, element-wise corrections in the compression matrix are required. To address this, we introduce a perturbation matrix $\Delta\Phi$ to model errors in the compression matrix. Since these perturbations affect individual elements, we consider two cases: (i) additive perturbation and (ii) multiplicative perturbation. The actual compression matrix can then be expressed as

$$\Phi + \Delta\Phi, \quad \text{or} \quad \Phi \circ \Delta\Phi.$$

By incorporating this perturbation model, we can not only correct geometric errors but also mitigate receiver errors and other unparameterized distortions. These unmodeled errors can be treated as noise within the perturbation matrix, compensating for the incomplete prior knowledge of the compression matrix in a real scenario.

Apart from the above, we observe that in most real-world cases, the perturbation matrix $\Delta\Phi$ exhibits a full-rank property. In addition, imposing a rank constraint on $\Delta\Phi$ would introduce non-convexity into the problem. Consequently, it is not a suitable constraint.

We avoid performing calibration directly in the covariance domain because the perturbation introduced becomes intertwined in the covariance computation. Instead, we

adopt the sparsity-cognizant total least-squares (STLS) method [30], which minimizes the norm of the error, providing a more practical and computationally efficient solution.

To jointly minimize the noise in the received data and the perturbation in the compression matrix while leveraging the sparsity of the signal matrix for DOA estimation, we formulate the following optimization problems:

Additive perturbation model

$$\underset{\bar{\mathbf{S}}, \Delta\mathbf{\Phi} \in \mathbb{C}^{Q \times M}}{\text{minimize}} \quad \frac{1}{2} \|\mathbf{Y} - (\mathbf{\Phi} + \Delta\mathbf{\Phi})\bar{\mathbf{A}}(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\psi}})\bar{\mathbf{S}}\|_F^2 + \gamma \|\Delta\mathbf{\Phi}\|_F^2 + \lambda \|\bar{\mathbf{S}}\|_{2,1}. \quad (4.1)$$

Multiplicative perturbation model

$$\underset{\bar{\mathbf{S}}, \Delta\mathbf{\Phi} \in \mathbb{C}^{Q \times M}}{\text{minimize}} \quad \|\mathbf{Y} - (\mathbf{\Phi} \circ \Delta\mathbf{\Phi})\bar{\mathbf{A}}(\bar{\boldsymbol{\theta}}, \bar{\boldsymbol{\psi}})\bar{\mathbf{S}}\|_F^2 + \gamma \|\Delta\mathbf{\Phi}\|_F^2 + \lambda \|\bar{\mathbf{S}}\|_{2,1}. \quad (4.2)$$

where the $\ell_{2,1}$ -norm serves as the sparse regularizer, promoting joint sparsity of the signals across different time instants. In the matrix $\bar{\mathbf{S}}$, sparsity is enforced solely in the gridded spatial domain, as the signal is not sparse in the time dimension. Given that the sources at different time instants are similar, we expect all columns to exhibit the same sparsity pattern. Therefore, we leverage the temporal correlation to achieve better spatial resolution. This approach is known as joint (or group) sparsity in the literature, and it can be enforced using the $\ell_{2,1}$ -norm. To begin, we compute the ℓ_2 -norm of each row in $\bar{\mathbf{S}}$. Then, we construct the $\bar{\mathbf{s}}^{(\ell_2)}$ vector, which stacks the ℓ_2 -norms of the rows of $\bar{\mathbf{S}}$. Finally, $\|\bar{\mathbf{S}}\|_{2,1} = \|\bar{\mathbf{s}}^{(\ell_2)}\|_1$ is computed.

Since the minimization problem involves a bilinear product of the optimization variables $\Delta\mathbf{\Phi}$ and $\bar{\mathbf{S}}$, it is inherently non-convex. To address this, we follow the approach in [30] and apply the alternating descent suboptimal algorithm.

The first step of the iterative block coordinate descent algorithm is to estimate $\bar{\mathbf{S}}_i$ while keeping $\Delta\mathbf{\Phi}_i$ fixed in the i th iteration. The solution is similar for both additive and multiplicative cases, since with a known perturbation matrix, the total transmission matrix is determined.

Specifically, we define

$$\bar{\mathbf{B}}_i = (\mathbf{\Phi} + \Delta\mathbf{\Phi}_i)\bar{\mathbf{A}}$$

for the additive case, and

$$\bar{\mathbf{B}}_i = (\mathbf{\Phi} \circ \Delta\mathbf{\Phi}_i)\bar{\mathbf{A}}$$

for the multiplicative case. Under this formulation, the minimization problem takes the same form:

$$\bar{\mathbf{S}}_i = \underset{\bar{\mathbf{S}}}{\text{argmin}} \quad \frac{1}{2} \|\mathbf{Y} - \bar{\mathbf{B}}_i \bar{\mathbf{S}}\|_F^2 + \lambda \|\bar{\mathbf{S}}\|_{2,1}. \quad (4.3)$$

Then with $\Delta\mathbf{\Phi}_i$ available, $\Delta\mathbf{\Phi}_{i+1}$ for the ensuing iteration is found by setting the first-order derivative of the cost wrt $\Delta\mathbf{\Phi}$ equal to zero.

4.2.1 First step to update signal matrix

The optimization problem in the first step can be solved using inverse-free sparse Bayesian learning (SBL), alternating direction method of multipliers (ADMM), and

the fast iterative shrinkage-thresholding algorithm (FISTA). These methods avoid the computation of a large matrix inverse, enabling a faster solution.

In this work, we adopt FISTA. Specifically, we use the multiple measurement vector (MMV) variant of FISTA. To reduce the computational complexity of the MMV model in (4.3), we project $\mathbf{Y}$ into an N -dimensional subspace with the fact that $N \ll L$.

To identify this subspace, we first compute the singular value decomposition (SVD) of $\mathbf{Y}$, yielding:

$$\mathbf{Y} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T.$$

Next, a reduced matrix that captures most of the signal power is constructed as:

$$\tilde{\mathbf{Y}} = \mathbf{U}\mathbf{\Sigma}\mathbf{D}_N = \mathbf{Y}\mathbf{V}\mathbf{D}_N,$$

where $\mathbf{D}_N = [\mathbf{I}_N \quad \mathbf{0}]^T$.

Similarly, the matrices $\mathbf{S}$ and $\mathbf{N}$ are transformed using these matrices, resulting in:

$$\tilde{\mathbf{S}} = \mathbf{\bar{S}}\mathbf{V}\mathbf{D}_N, \quad \tilde{\mathbf{N}} = \mathbf{N}\mathbf{V}\mathbf{D}_N.$$

With this, instead of solving the large problem described in (4.3), we now solve the following reduced problem:

$$\underset{\tilde{\mathbf{S}}}{\text{minimize}} \quad \frac{1}{2} \|\tilde{\mathbf{Y}} - \mathbf{\bar{B}}_i \tilde{\mathbf{S}}\|_F^2 + \lambda \|\tilde{\mathbf{S}}\|_{2,1}. \quad (4.4)$$

MMV-FISTA is a two-step method. The first step is solving the minimization problem associated with the proximity operator. This is similar to the iterative shrinkage-thresholding algorithm (ISTA) method. In ISTA, the proximal operator is directly applied at the v th step to the previous estimate $\tilde{\mathbf{S}}_i(v-1)$. However, in FISTA, the proximal operator is applied to a matrix $\mathbf{Z}(v)$, which is a specific linear combination of the two previous iterates, $\tilde{\mathbf{S}}_i(v-1)$ and $\tilde{\mathbf{S}}_i(v-2)$. The proximal problem of FISTA can be formulated as:

$$\tilde{\mathbf{S}}_i(v) = \text{prox}_{L_f \lambda \|\cdot\|_{2,1}} \left(\mathbf{Z}(v) + \frac{1}{L_f} \mathbf{\bar{B}}_i^T (\tilde{\mathbf{Y}} - \mathbf{\bar{B}}_i \mathbf{Z}(v)) \right), \quad (4.5)$$

where L_f is the Lipschitz constant, which is given by $\|\mathbf{\bar{B}}_i^H \mathbf{\bar{B}}_i\|_2$ (spectral norm which corresponds to the largest singular value). The solution of the proximity operator involves alternating between minimizing the Frobenius quadratic term using a small gradient step and the computation of the proximal operator associated with $l_{2,1}$ term [31]. The update steps for solving (4.4) are outlined in Algorithm 1. Note that $\mathbf{\bar{S}}$ can be easily computed from $\tilde{\mathbf{S}}$ through a simple matrix multiplication.

For the step of updating the perturbation matrix, we discuss each case separately as follows.

STLS-MMV with element-wise additive perturbation For the additive case, we can derive the update formula directly from the original form of the minimization problem:

$$\Delta \Phi_{i+1} = \underset{\Delta \Phi}{\text{argmin}} \quad \frac{1}{2} \|\mathbf{Y} - \Phi \mathbf{\bar{A}} \mathbf{\bar{S}}_i - \Delta \Phi \mathbf{\bar{A}} \mathbf{\bar{S}}_i\|_F^2 + \gamma \|\Delta \Phi\|_F^2. \quad (4.6)$$

The optimal solution to this quadratic problem is obtained in closed form as:

$$\Delta \Phi_{i+1} = [\mathbf{Y} - \Phi \mathbf{\bar{A}} \mathbf{\bar{S}}_i] \mathbf{\bar{S}}_i^H \mathbf{\bar{A}}^H \left(2\gamma \mathbf{I}_{M \times M} + \mathbf{\bar{A}} \mathbf{\bar{S}}_i \mathbf{\bar{S}}_i^H \mathbf{\bar{A}}^H \right)^+. \quad (4.7)$$

Algorithm 1 MMV FISTA method

```

1: Input:  $\tilde{\mathbf{Y}}, N, \bar{\mathbf{B}}_i, \bar{\mathbf{A}}, \mathbf{D}_N, \mathbf{V}$ , coefficient:  $\lambda$  and the number of iterations:  $V$ 
2: Output:  $\bar{\mathbf{S}}_i$ 
3: Initialize  $\lambda > 0$ ,  $L_f = \|\bar{\mathbf{B}}_i^H \bar{\mathbf{B}}_i\|_2$ ,  $t(1) = 1$ ,  $\tilde{\mathbf{S}}_i(0) = \mathbf{Z}(1) = \mathbf{0}_{M \times N}$ 
4: for  $v = 1, \dots, V$  do
5:   calculate  $\mathbf{Q}(v) = \mathbf{Z}(v) - \frac{1}{L_f}(\bar{\mathbf{B}}_i^H \bar{\mathbf{B}}_i \mathbf{Z}(v) - \bar{\mathbf{B}}_i^H \tilde{\mathbf{Y}})$ 
6:   compute  $\tilde{S}_i^{(g,n)}(v) = Q^{(g,n)}(v)(1 - \frac{\lambda}{\|\mathbf{q}^{(g)}(v)\|_2})^+$  where  $Q^{(g,n)}(v)$  is the  $(g, n)$ th element
      of the matrix  $\mathbf{Q}(v)$  and  $\mathbf{q}^{(g)}(v)$  is the  $g$ th row of matrix  $\mathbf{Q}(v)$ . The function  $()^+$  is defined
      as  $(a)^+ = \max(a, 0)$  and we use the convention  $0/0 = 0$ .
7:   update  $t(v+1) = 0.5(1 + \sqrt{1 + 4t(v)^2})$ 
8:   update  $\mathbf{Z}(v+1) = \tilde{\mathbf{S}}_i(v) + \frac{t(v)-1}{t(v+1)}(\tilde{\mathbf{S}}_i(v) - \tilde{\mathbf{S}}_i(v-1))$ 
9: end for
10:  $\bar{\mathbf{S}}_i = \tilde{\mathbf{S}}_i(V) \mathbf{D}_N^T \mathbf{V}^T$ 

```

Since γ can be zero in the unperturbed case, we use the pseudoinverse instead of the direct inverse.

STLS-MMV with element-wise productive perturbation Since $\Delta\Phi$ is difficult to extract directly from the problem, we first vectorize all the terms:

$$\min_{\phi \in \mathbb{C}^{QM \times 1}} \left\| \text{vec}(\mathbf{Y}) - (\bar{\mathbf{A}} \bar{\mathbf{S}}_i)^H \otimes \mathbf{I}_Q \text{diag}(\text{vec}(\Phi)) \phi \right\|_2^2 + \gamma \|\phi\|_2^2. \quad (4.8)$$

Then it can be solved by equalizing the first derivative to 0, i.e.

$$-2(\mathbf{F}_i^H (\text{vec}(\mathbf{Y}) - \mathbf{F}_i \phi) + 2\gamma \phi) = \mathbf{0}.$$

Here, for simplicity, we introduce a matrix $\mathbf{F}_i = ((\bar{\mathbf{A}} \bar{\mathbf{S}}_i)^T \otimes \mathbf{I}_Q) \text{diag}(\text{vec}(\Phi)) \in \mathbb{C}^{QL \times QM}$.

From this, we get:

$$\phi_{i+1} = (2\gamma \mathbf{I}_{QM \times QM} + \mathbf{F}_i^H \mathbf{F}_i)^+ \mathbf{F}_i^H \text{vec}(\mathbf{Y}). \quad (4.9)$$

After the unvectorization process, we obtain:

$$\Delta\Phi_{i+1} = \text{vec}^{-1}(\phi_{i+1}).$$

The whole process iterates between these two steps until convergence.

The whole process of the MMV FISTA based S-TLS solver is tabulated as Algorithm 2.

Algorithm 2 STLS with MMV FISTA method

```

1: Input:  $\mathbf{Y}, \Phi, \bar{\mathbf{A}}$ , coefficients:  $\lambda, \gamma$  and the number of iterations:  $iter$ 
2: Output:  $\Delta\Phi_i$  and  $\bar{\mathbf{S}}_i$  upon convergence
3: Initialize  $\Delta\Phi_1 = \mathbf{0}_{2K \times M}$  for additive case and  $\Delta\Phi_1 = \mathbf{1}_{2K \times M}$  for multiplicative case
4: Perform SVD:  $\mathbf{Y} = \mathbf{U}\Sigma\mathbf{V}^T$ , and compute  $\tilde{\mathbf{Y}} = \mathbf{U}\Sigma\mathbf{D}_N = \mathbf{Y}\mathbf{V}\mathbf{D}_N$  with  $\mathbf{D}_N = [\mathbf{I}_N \quad \mathbf{0}]^T$ 
5: for  $i = 1, \dots, iter$  do
6:   Solve minimization problem of (4.3) using Algorithm 1
7:   Update  $\Delta\Phi_{(i+1)}$  as in (4.7) or (4.9)
8: end for

```

Comparing the computational load of the two methods, we notice that for the multi-

plicative case, the pseudoinversion is involved with a larger matrix due to vectorization. Especially in the 3D case, this significantly slows down the speed of computation compared to the additive case with the same number of iterations. Therefore, we first test the algorithm in the linear array case and then move on to the case with a planar array.

4.3 Simulation Results for STLS with the MMV-FISTA Method

(1) ULA with a Linear Coded Cover in the 2D Case We consider the setting where a ULA with four PU probes is spaced at 7.5 cm. A half-wavelength spaced linear coded cover with 14 channels is applied, and the SNR is set to 20 dB.

One source: We first test the case where a single source is present at an angle of 72° . The data consists of 1000 time snapshots, with regularization parameters set to $\lambda = 0.2$ and $\gamma = 0.5$. Each MMV-FISTA algorithm runs for 300 iterations, while the compression matrix is updated over 50 iterations.

The introduced errors are as follows: the actual distance l is 14.5 cm instead of 15 cm, the rotation ζ is 2 degrees closer to the mask, and the first sensor is shifted by $0.2 \times \delta$ from the origin. The scanning range for the STLS method with additive perturbation optimization spans from 30° to 150° , with a resolution of 1° .

When only geometric error is present, Figure 4.3 shows that STLS calibration enables the source to be localized with an error of 1° . Applying CCS afterward, as illustrated in Figure 4.4, further improves the accuracy, reducing the error to 0.9° . In contrast, without calibration, the DOA estimated using CCS deviates more from the true direction compared to the calibrated result.

When uncertainties in the compression matrix are introduced, we first use an additive complex Gaussian noise with a variance of 0.1. Figure 4.5 presents the localization results after applying the STLS method with additive perturbation optimization. The target is accurately recovered at the correct source direction. In contrast, Figure 4.6 illustrates the result without calibration, where the peak deviates slightly (by approximately 0.2°) from the true target direction. This deviation is expected, as the optimization is performed over QM entries. The MUSIC algorithm requires precise knowledge of the compression matrix, and recovering the perturbation accurately is challenging due to the large number of unknowns.

Further, we analyze the effect of increased additive noise in the compression matrix. The geometric mismatching is same as before. Suppose the variance of the perturbation increases to 1. The new target is now located at 130° .

Due to the increased perturbation in the compression matrix, a stronger regularization is required to constrain its norm. Therefore, we adjust the regularization parameter from 0.5 to 5. The localization results, shown in Figure 4.7, demonstrate that our proposed method remains effective despite the increased noise.

For comparison, Figure 4.8 presents the CCS localization result with and without calibration. As expected, the uncalibrated data leads to a slight deviation from the true target direction. This further highlights the necessity of the proposed optimization

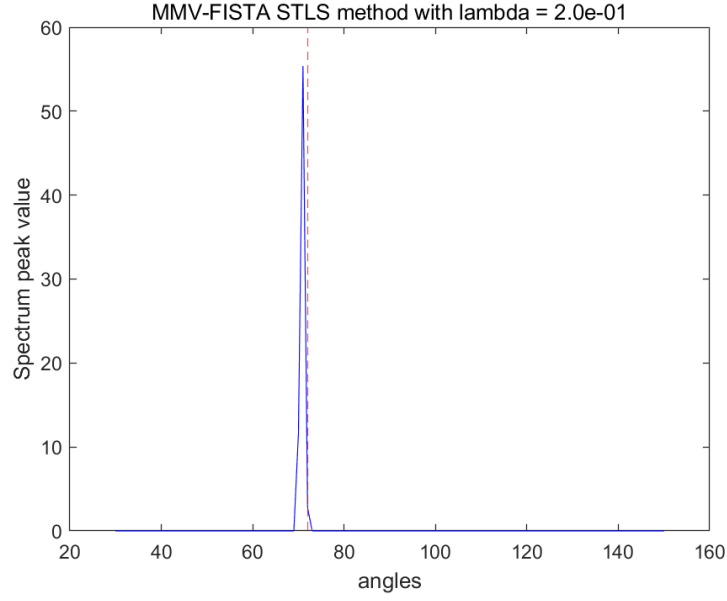


Figure 4.3: Localization result using STLS with additive perturbation optimization under geometric mismatch in compression matrix, showing accurate source recovery at 72° .

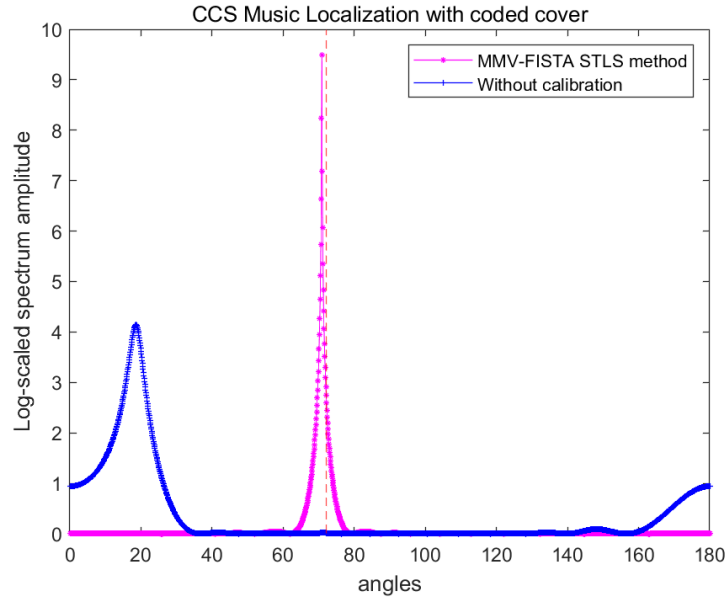


Figure 4.4: CCS localization results with and without STLS calibration, showing improved accuracy with STLS calibration.

approach for handling perturbations in the compression matrix.

Then we also tested the recovery performance of geometric mismatch correction with uncalibrated sensors. Here, the geometric mismatch is the same as before. The random receiving errors in gain are drawn from a uniform distribution within the range $[-3, 3]$

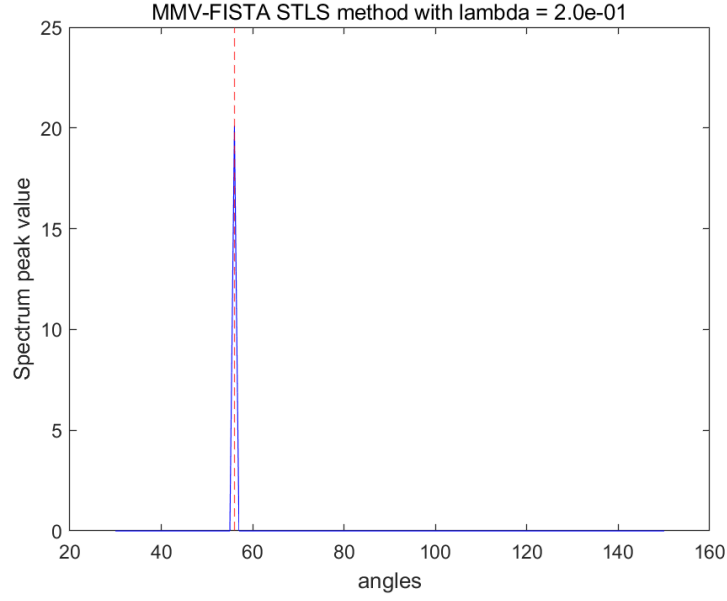


Figure 4.5: Localization result using STLS with additive perturbation optimization under geometric mismatch and small perturbation in compression matrix, showing accurate source recovery at 56° .

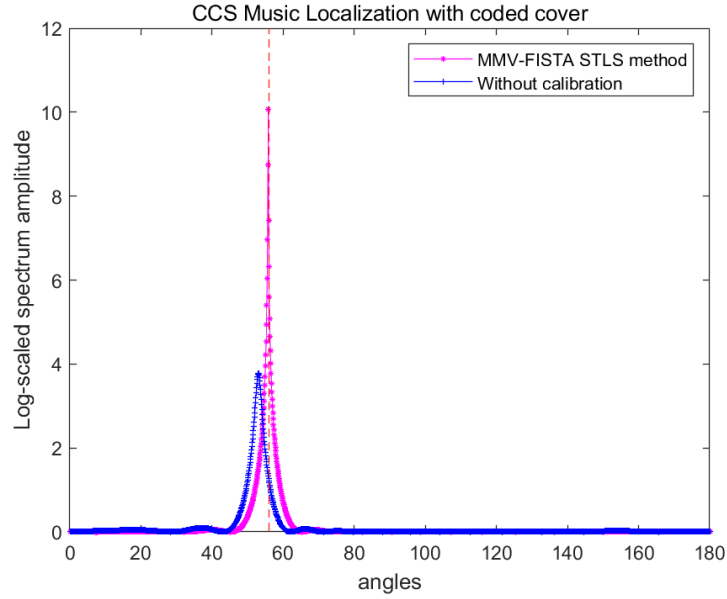


Figure 4.6: CCS localization results with and without STLS calibration under small perturbations, showing improved accuracy with STLS calibration.

dB, while the phase errors are selected from a uniform distribution over $[-20^\circ, 20^\circ]$.

We observed that, with the same phase and amplitude receiving errors introduced earlier, the STLS MMV-FISTA method was able to recover the target with some de-

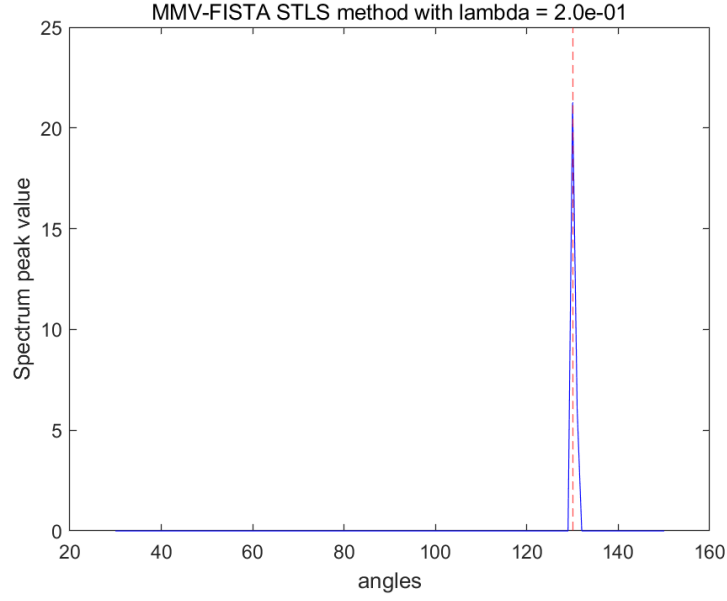


Figure 4.7: Localization result using STLS with additive perturbation optimization under geometric mismatch and larger perturbation in compression, showing accurate source recovery at 130°

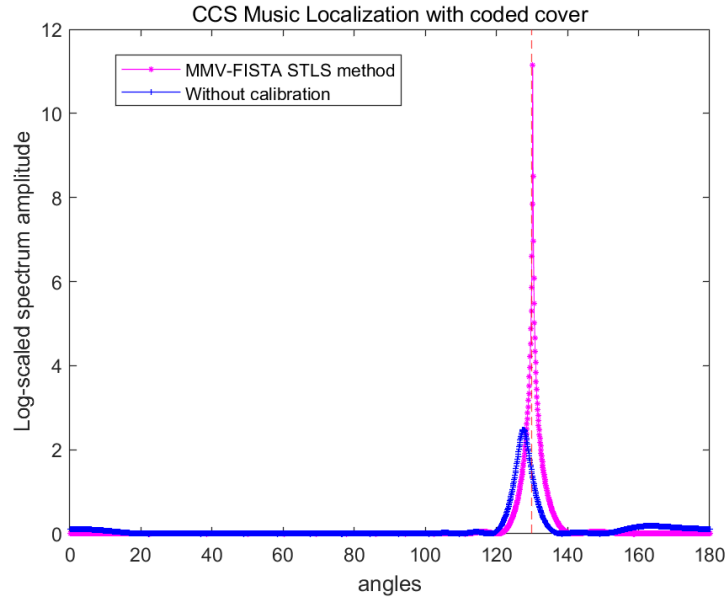


Figure 4.8: CCS localization results with and without STLS calibration under larger perturbations, showing improved accuracy with STLS calibration.

viation, as shown in Figure 4.9. In contrast, the uncalibrated case failed to locate any source, demonstrating the effectiveness of the STLS correction method. The improved localization after applying STLS calibration can be seen in Figure 4.10.

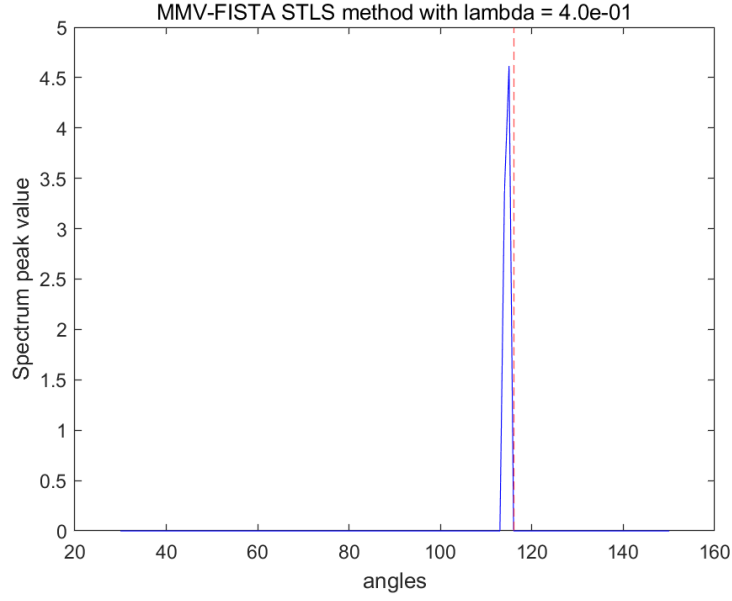


Figure 4.9: Localization result using STLS MMV-FISTA under geometric mismatch with uncalibrated sensors

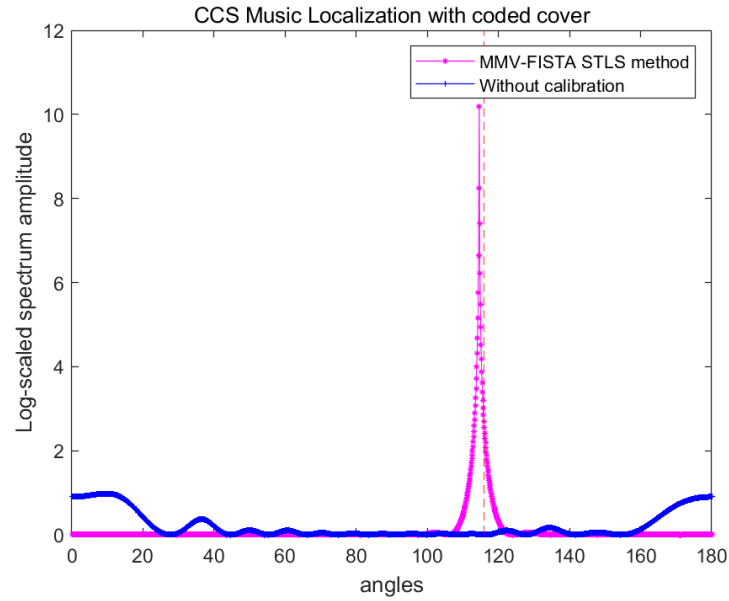


Figure 4.10: Localization result with and without STLS correction under geometric mismatch with uncalibrated sensors, showing improved target recovery.

Two sources: We set one source at 66° and the other at 125° . Initially, we only account for geometric mismatches, which remain the same as previously described. The regularization parameters are set to $\lambda = 0.075$ and $\gamma = 0.5$. The number of iterations for updating the compression matrix is 40, with 100 updates for the signal matrix in

each iteration.

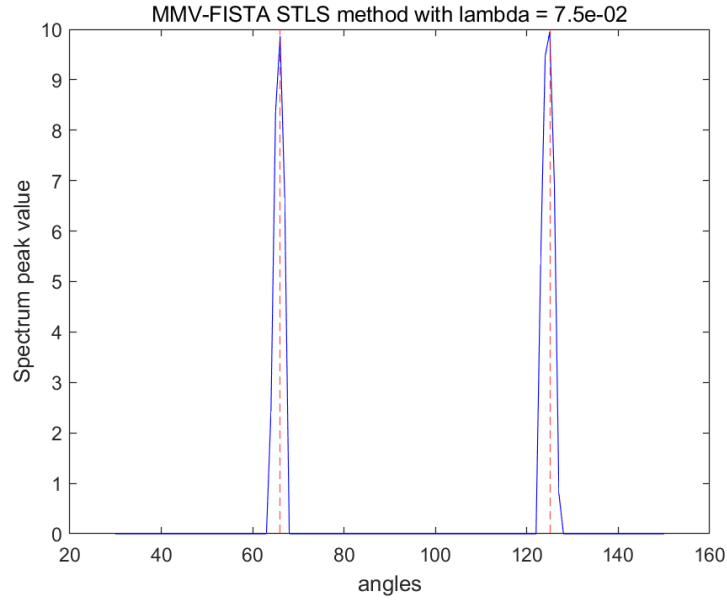


Figure 4.11: Localization result using STLS, successfully recovering both targets.

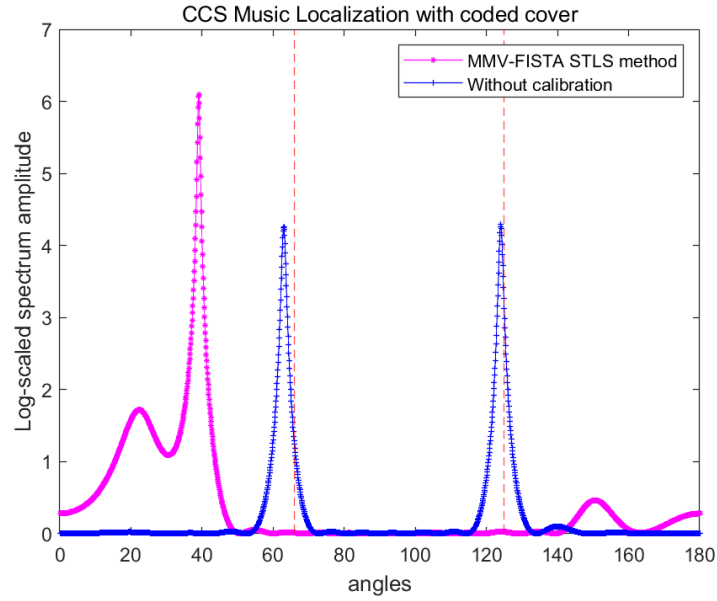


Figure 4.12: CCS localization result after STLS calibration. Due to an inaccurate perturbation matrix, performance degrades, but STLS already provides accurate localization.

As shown in Figure 4.11, after all iterations, the STLS method successfully recovers both targets. However, in Figure 4.12, the CCS result after calibration is significantly degraded due to the use of an inaccurate perturbation matrix. Despite this, as previously mentioned, the STLS result already provides accurate localization of both targets

and outperforms the uncalibrated scenario. Therefore, with the STLS method, applying further correction via CCS is unnecessary, as it may lead to degraded results.

When additional perturbation noise is introduced, we analyze the impact on localization accuracy. Figure 4.13 presents the results after adding noise with a variance of 0.5 to the perturbation matrix, in addition to the geometric mismatches.

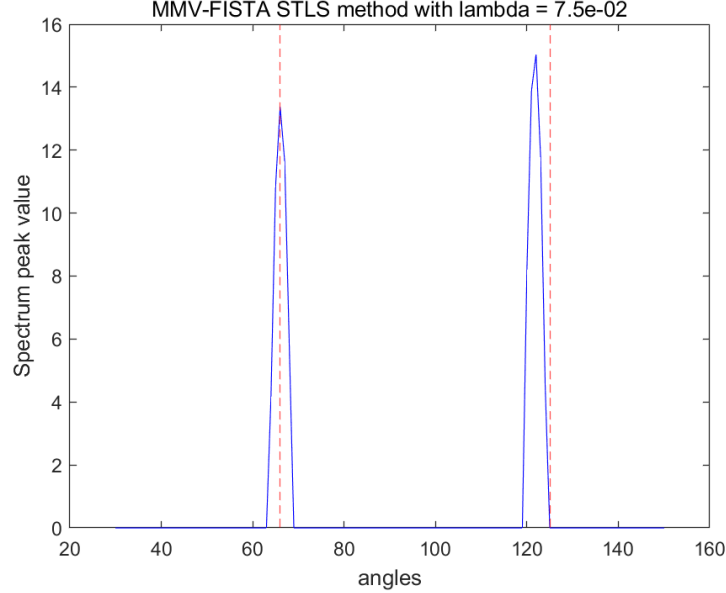


Figure 4.13: Localization result using STLS with added perturbation noise (variance = 0.5). The first target is exactly recovered, while the second peak has a mismatch of 3 degrees.

As shown in Figure 4.13, the first target is correctly localized, but the second peak exhibits a mismatch of 3 degrees. However, compared to the CCS localization results in Figure 4.14, where only the first peak is roughly matched and the second target is missing, the STLS method still demonstrates a superior ability to recover multiple sources despite increased perturbation noise.

Three sources: Figure 4.15 illustrates the localization results when geometric mismatches occur. The STLS method can still roughly recover three sources. The true source locations are at 66° , 83° , and 125° , while the recovered sources are at 66° , 81° , and 122° .

As shown in Figure 4.15, the STLS method successfully estimates the target locations with small deviations. In contrast, Figure 4.16 illustrates that without proper calibration, only the first source is roughly estimated, while the other sources are poorly matched.

(2) RECT-7.5 with a Planar Coded Cover in the 3D Case Now, we move on to the 3D case. The introduced geometric mismatching is described as follows: l is 14.5 cm instead of 15 cm, the rotation is 2 degrees, and the deviation of the first sensor is

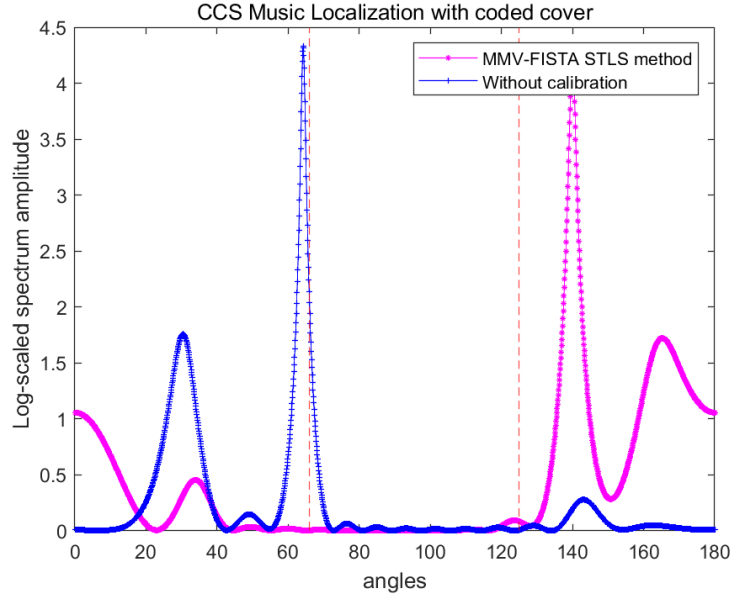


Figure 4.14: CCS localization result after STLS with added perturbation noise. Only the first peak is roughly matched, and the second target is missing.

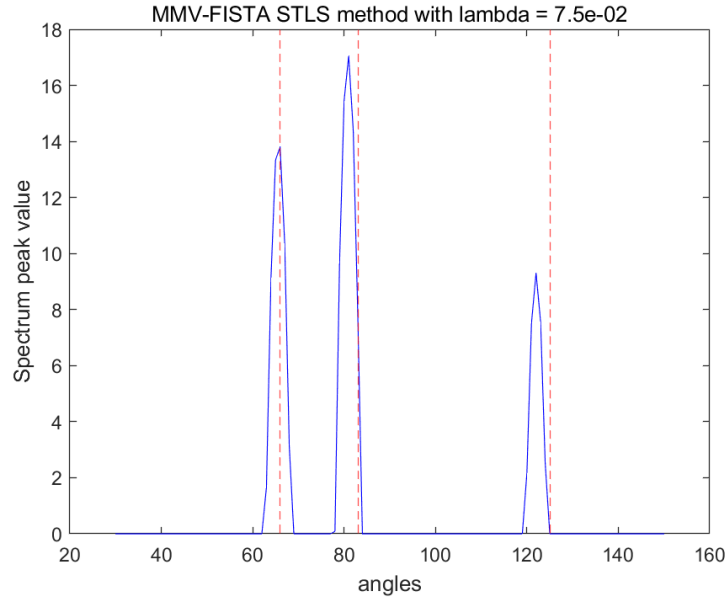


Figure 4.15: Localization result using STLS under geometric mismatches. Three sources are roughly recovered.

0.2 times the channel length away from the origin. The SNR is set to 20 dB, and the number of time instances is 10000.

Due to the high computational load in 3D space, we reduce the resolution of the scanning grid. Specifically, the elevation angle is scanned from 0° to 50° with a 2°

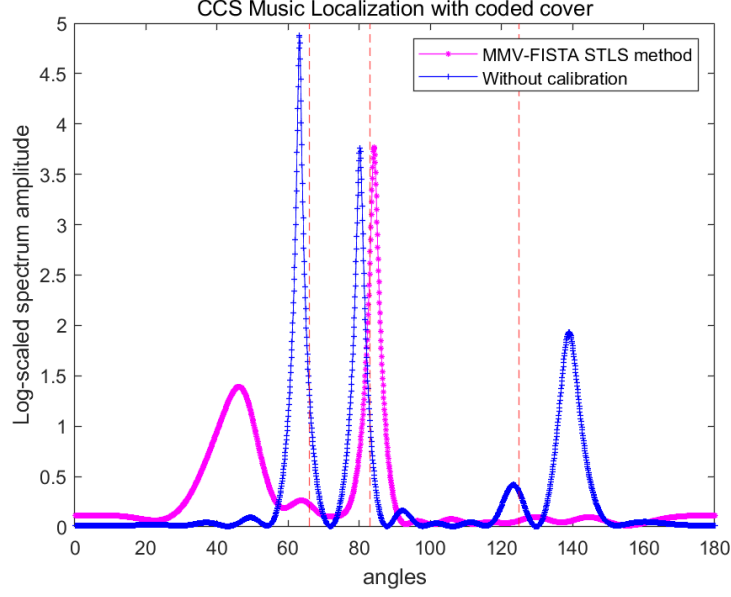


Figure 4.16: CCS localization result after STLS under geometric mismatches. Only the first source is roughly estimated, while the others are poorly matched.

resolution, while the azimuth angle is scanned from 0° to 350° with a 10° resolution.

When a single source is located at $(20^\circ, 30^\circ)$, Figure 4.17 shows the recovery result using $\lambda = 0.04$ and $\gamma = 0.5$ after 20 updates of the compression matrix and 30 iterations of signal matrix updates per compression matrix update. Compared to Figure 4.18, we observe that with perturbations in the compression matrix, the CCS method fails to recover the single target, whereas the STLS method successfully localizes it. There is some spread of the target point due to the lower resolution of the scanning grid.

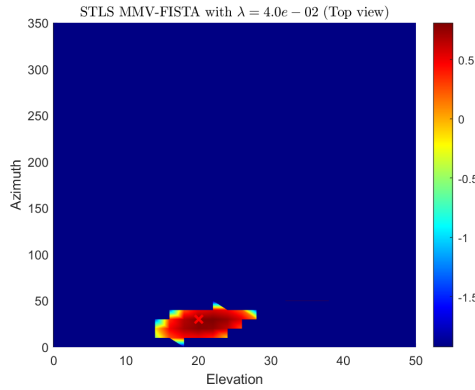


Figure 4.17: STLS recovery result for a single source at $(20, 30)$.

When three sources are randomly placed in the 3D case, Figure 4.19 shows the result of the STLS MMV-FISTA method. The located area is divided into four regions, each corresponding to one target source. Some deviation is observed at the center of the regions, which is expected, as similar phenomena were observed in the linear array

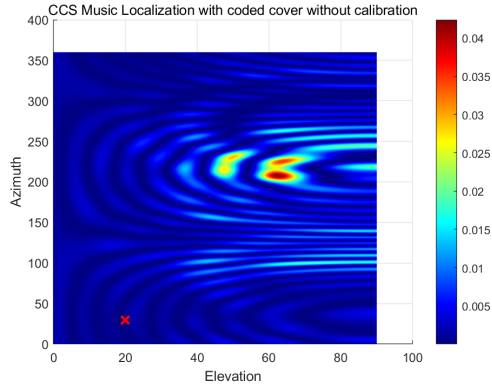


Figure 4.18: CCS recovery result without calibration, failing to recover the single target.

case. In contrast to the complete failure of the CCS method shown in Figure 4.20, our algorithm is able to at least identify the regions where the targets are located.

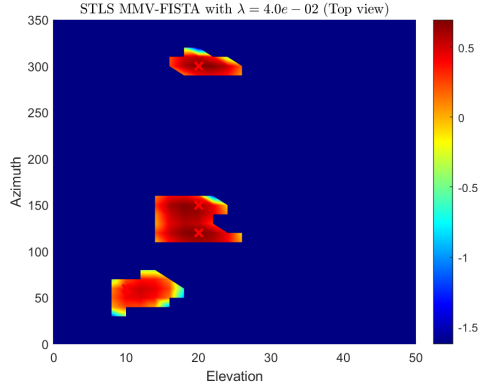


Figure 4.19: STLS MMV-FISTA recovery result for four sources in the 3D case.

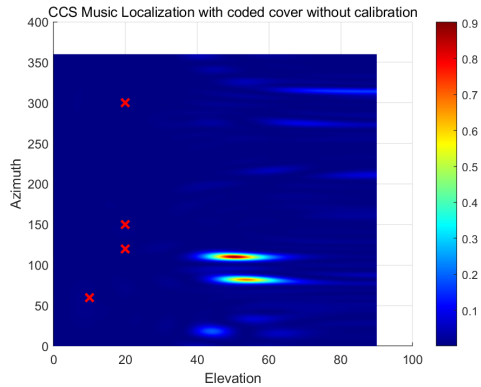


Figure 4.20: CCS recovery result without calibration, showing complete failure to recover the sources.

Next, we introduce some unexpected noise into the compression matrix. In this case,

four sources are randomly placed in 3D space. In addition to the geometric mismatching introduced previously, we also add noise to the compression matrix with a variance of 0.1. Figure 4.21 shows the recovered result using the STLS-MMV-FISTA method, where the four sources are located near the center of the targeted areas, demonstrating successful recovery. In contrast, Figure 4.22 shows the result without any calibration, where no sources are properly located.

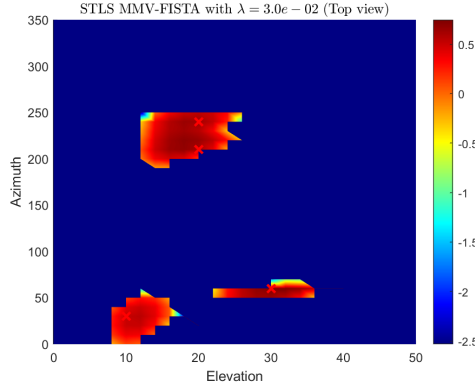


Figure 4.21: Recovered result with STLS-MMV-FISTA method after adding noise to the compression matrix (variance 0.1).

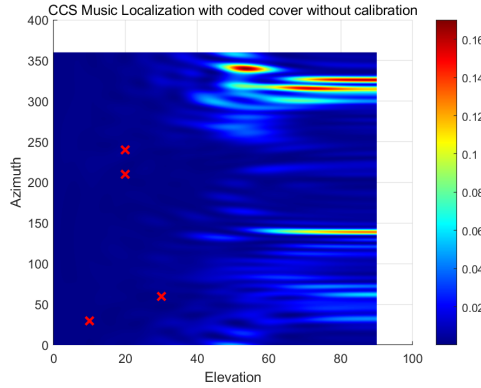


Figure 4.22: Recovery result without calibration, showing poor performance due to the added noise in the compression matrix.

Remark 3: In the simulation, we ran both algorithms with either additive or multiplicative perturbations. The results show no significant difference between the two. However, since the multiplicative case requires more computational effort, we only present the results from the additive case. In addition, compared to the self-calibration method in the element space domain, our proposed method is more suitable for cases involving compression matrix perturbation. This is because our method does not require any known source for DOA estimation. Additionally, it does not rely on prior knowledge of the noise power, making it more flexible and applicable in scenarios where such information is unavailable. However, due to the inherent limitations of the original STLS algorithm in recovering sources without compression [30], calibrating and

estimating multiple sources simultaneously becomes more challenging after applying compressed sensing. That's why we considered the case with no more than five sources.

4.4 Grid Search Method for Geometric Mismatching with calibrated sensors

From the simulation results with the STLS method, we observed the limited performance of the STLS method in compressed sensing. This limitation arises because the method requires the optimization of many parameters. Therefore, when considering perturbations solely due to geometric mismatching, we propose using a grid search method to optimize the parameters accordingly. The main geometric parameters that need to be optimized have been introduced in the beginning. They are l , ϵ and ζ .

For each parameter configuration, the corresponding compression matrix Φ^i is computed. The CCS method is then applied to recover the covariance matrix by constructing:

$$\mathbf{R}_x^i = \text{vec}^{-1} \left\{ \mathbf{S} [(\Phi^{i*} \otimes \Phi^i) \mathbf{S}]^+ (\mathbf{r}_0 - \mathbf{r}_n) \right\}.$$

The corresponding reconstruction error is computed as:

$$e_i = \|\mathbf{r}_0 - \mathbf{r}_n - (\Phi^{i*} \otimes \Phi^i) \sigma_x^i\|_2.$$

Then the best set of geometric parameters is selected by finding the minimum error.

4.5 Simulation Results of Grid Search Method for Geometric Mismatching with Calibrated Sensors

In this section, we apply the Grid Search Method for geometric mismatching with calibrated sensors. We consider two cases: one with a ULA and the other with a planar array. Geometric errors are introduced in both cases to demonstrate the effectiveness of our proposed method.

(1) ULA with a Linear Coded Cover in the 2D Case Consider a scenario where there are 4 PU probes, with a half-wavelength spaced linear coded cover containing 14 channels, and an SNR of 20 dB. 13 sources are randomly distributed between 30° and 150° . The simulation introduces human-induced mismatching, described as follows: the actual distance between the coded cover and the PU array is 14.6 cm instead of the intended 15 cm. Additionally, the first channel is not exactly aligned with the first sensor as initially set. Instead, it is offset by 0.5 times the channel length δ from the original position.

The search range for the distance is set between 14.5 cm and 15.5 cm, with a resolution of 0.1 cm. The deviation for the origin is searched from $-\delta$ to δ , also with a resolution of 0.1δ .

Through comparing Figure 4.23 and Figure 4.24, we can clearly see that after the best-fit parameters are identified through grid search, the targets are exactly matched by the algorithm. The search results show that the best-fitting value for the distance

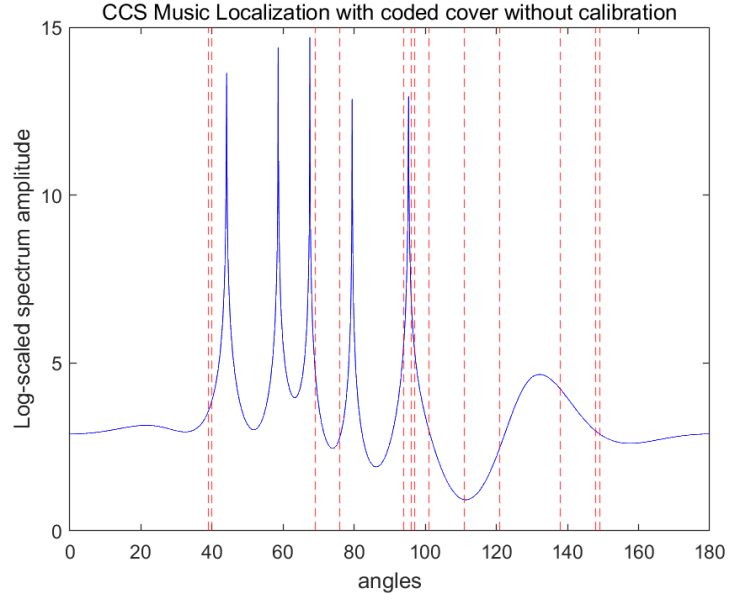


Figure 4.23: CCS localization result using ideal compression matrix, where all peaks do not correspond to the true target positions.

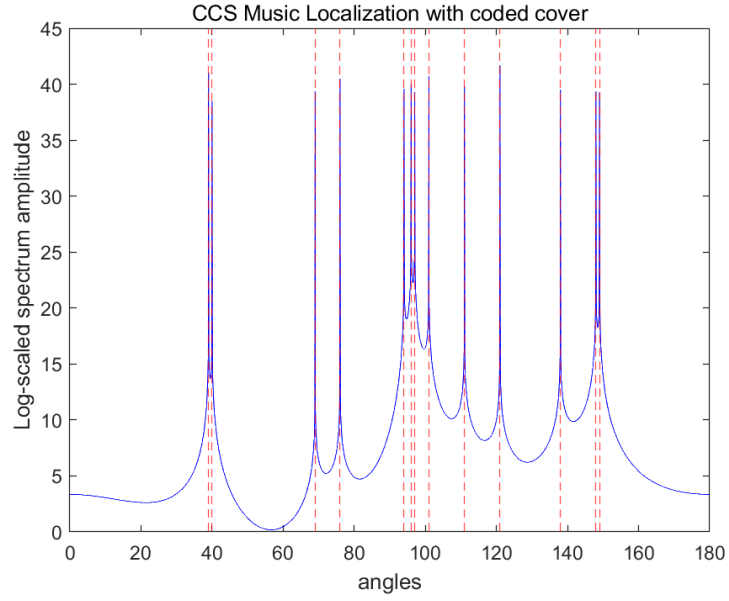


Figure 4.24: CCS localization result after applying grid-search calibration, where the targets are exactly matched by the algorithm.

parameter l is 14.6 cm, and the origin mismatch is 0.5δ , which corresponds exactly to the introduced error.

(2) RECT-7.5 with a Planar Coded Cover in the 3D Case The proposed algorithm is evaluated using 40 randomly placed sources in 3D space, with an SNR of 20 dB. The introduced geometric mismatching is defined as follows: the actual distance between the coded cover and the array is 15.3 cm instead of the nominal 15 cm. Additionally, the array experiences a rotational deviation around the y -axis of 2° , resulting in non-parallel alignment between the two planes.

The search range for l spans from 14.5 cm to 15.5 cm, with a resolution of 0.1 cm. For the rotational deviation, the search range extends from -5° to 5° with a resolution of 1° .

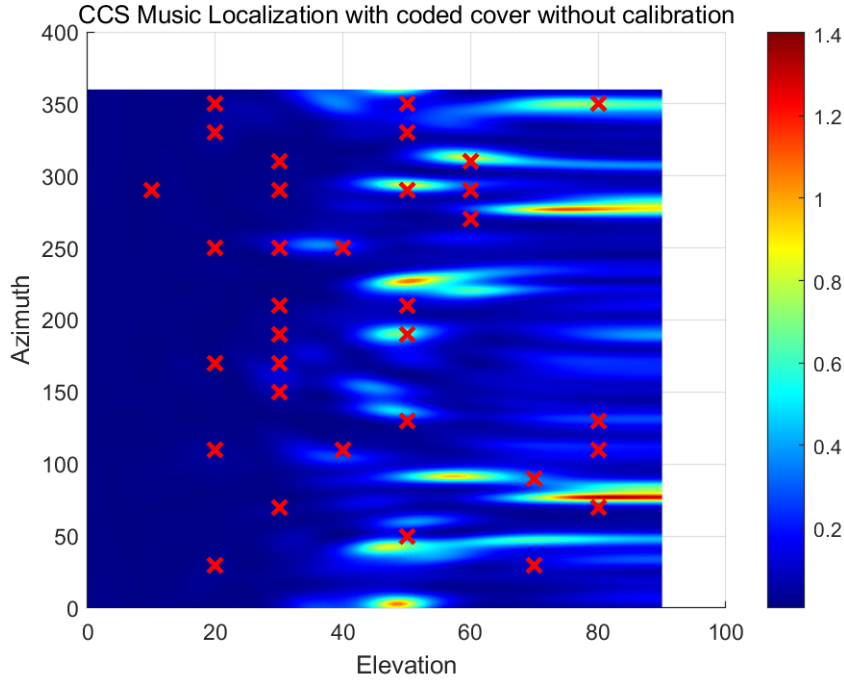


Figure 4.25: CCS localization result using the ideal compression matrix in the 3D case, where the detected peaks do not align with the true source positions.

A comparison between Figure 4.25 and Figure 4.26 demonstrates the effectiveness of grid-search calibration in the 3D scenario. Once the optimal parameters are identified, the algorithm accurately localizes the sources. The best-fitting distance l is found to be 15.3 cm, and the rotation adjustment corresponds precisely to the introduced deviation of 2° .

4.6 Self-Calibration integrated Grid Search Method for Geometric Mismatching with uncalibrated sensors

The previous geometric calibration assumes fully calibrated sensors without channel-specific reception errors. When sensor calibration is incomplete, a joint optimization approach can be applied to simultaneously calibrate the data and search for the optimal

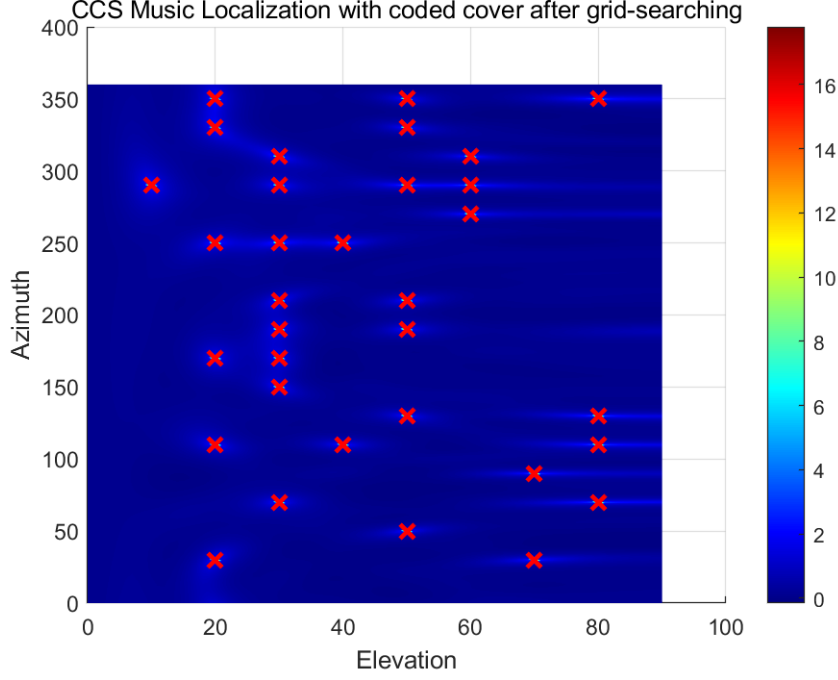


Figure 4.26: CCS localization result after applying grid-search calibration in 3D, showing precise alignment between the detected peaks and the true source positions.

compression matrix.

For the i -th set of geometric parameters, the self-calibration problem with the corresponding compression matrix Φ^i is formulated as:

$$\underbrace{\begin{bmatrix} \text{diag}(\mathbf{r}_0 - \mathbf{r}_n) & -\mathbf{B}_{\text{co}}^i(\boldsymbol{\theta}, \boldsymbol{\psi}) \end{bmatrix}}_{\mathbf{G}_{\text{co}}^i} \underbrace{\begin{bmatrix} \text{vec}(\mathbf{C}) \\ \boldsymbol{\sigma}_s^i \end{bmatrix}}_{\boldsymbol{\gamma}_{\text{co}}^i} = \mathbf{0}.$$

The optimization problem is defined as:

$$\min_{\mathbf{C}, \boldsymbol{\sigma}_s^i} \quad \|\mathbf{G}_{\text{co}}^i \boldsymbol{\gamma}_{\text{co}}^i\|_2^2 + \lambda \|\boldsymbol{\sigma}_s^i\|_0, \quad (4.10)$$

$$\text{s.t.} \quad \mathbf{C} = \mathbf{c}\mathbf{c}^H, \quad (4.11)$$

$$\boldsymbol{\sigma}_s^i \succeq \mathbf{0}, \quad (4.12)$$

$$\mathbf{c}(1) = 1, \quad (4.13)$$

$$\boldsymbol{\sigma}_s^i(k) = 1, \quad k \in \mathcal{K}. \quad (4.14)$$

The corresponding reconstruction error is then computed as:

$$e_i = \|\mathbf{G}_{\text{co}}^i \boldsymbol{\gamma}_{\text{co}}^i\|_2.$$

The optimal geometric parameter and correction matrix $\mathbf{C}$ are determined by finding the smallest error. Using the appropriate compression and correction matrices, the CCS method can then enhance the accuracy of DOA estimation.

4.7 Simulation Results of Self-Calibration integrated Grid Search Method with uncalibrated sensors

(1) ULA with a Linear Coded Cover in the 2D Case In the ULA case, perturbations are introduced by distance mismatching and rotation of the linear array along the y -axis. Although the linear array and the linear coded mask lie in the same plane, they are not exactly parallel. The distance l is set to 15.2 cm instead of 15 cm, and the rotation is 2° instead of being exactly parallel (0° rotation). The search range for l spans from 14.5 cm to 15.5 cm with a resolution of 0.1 cm. The rotation angle is searched from -5° to 5° with a 1° resolution.

The receiving error per channel follows the same uniform distribution introduced earlier. Thirteen unknown sources are randomly selected from a one-dimensional scanning range of 30° to 150° with 5° spacing, considering only one reference channel for self-calibration.

Figure 4.27 presents the localization outcome after applying the proposed method. By integrating self-calibration with the grid search approach, the method effectively mitigates geometric mismatches and improves source localization accuracy in 2D space. The optimization process determines the best distance parameter as $l = 15.2$ cm and the optimal rotation correction as 2° , demonstrating successful geometric mismatch compensation. The high localization precision underscores the robustness and effectiveness of the proposed method in the linear mask case.

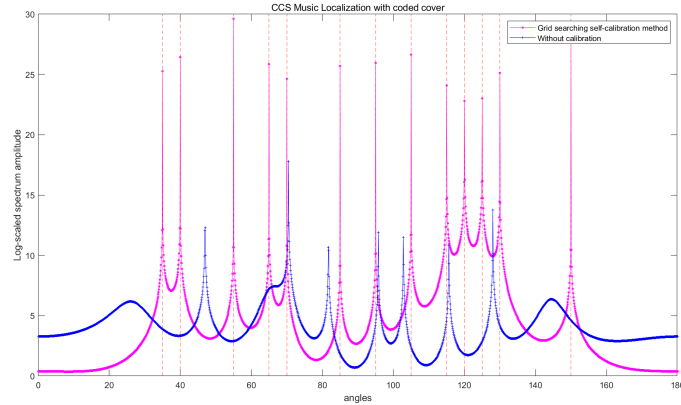
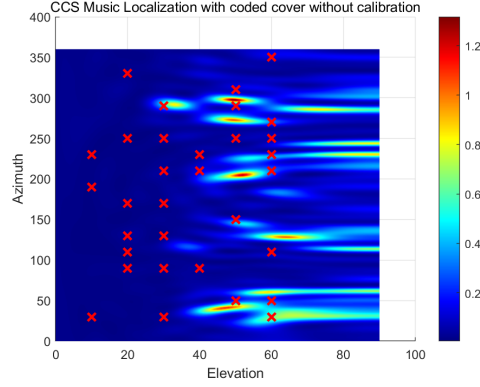


Figure 4.27: Localization result in 2D space using grid searching and self-calibration under $\text{SNR} = 20$ dB.

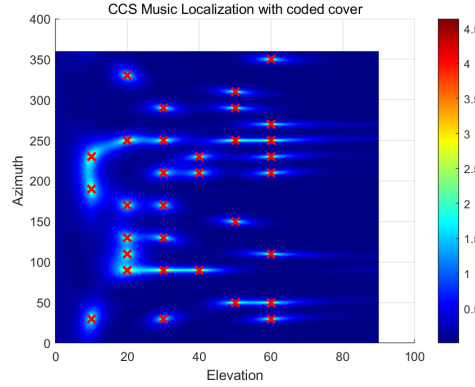
(2) RECT-7.5 with a Planar Coded Cover in the 3D Case With the consideration of mismatches from the distance l , we introduce a mismatch where l is set to 14.8 cm instead of the previous 15 cm.

Due to the high computational demands of 3D space optimization, the scanning range has been reduced. Specifically, l is scanned from 14.7 cm to 15.3 cm with a resolution of 0.1 cm. The SNR is set to 20 dB, and the first receiving channel is used as the reference channel. A total of 33 sources are randomly distributed in 3D space.

Figure 4.28 illustrates the recovery result under these conditions. The grid searching method combined with self-calibration effectively compensates for geometric mismatches, achieving accurate source localization in 3D space. Despite the increased spatial complexity, the approach maintains high localization accuracy, highlighting its robustness and effectiveness in three-dimensional environments. After grid searching, the best obtained distance parameter for l is 14.8 cm, demonstrating the feasibility of recovering geometric mismatching.



(a) Localization result in 3D space without calibration.



(b) Localization result in 3D space using grid searching and self-calibration.

Figure 4.28: Comparison of 3D localization results with and without grid-searching integrated self-calibration under $\text{SNR} = 20$ dB.

4.8 Conclusion

In this section, we evaluated the feasibility of the STLS-MMV-FISTA method in the context of compressed sensing. Simulation results revealed that its localization capability is significantly limited, with deviations in source estimation that are sometimes unavoidable. To address this issue, we introduced a grid search method to correct geometric errors.

To address this, we adopted different grid search approaches for calibrated and uncalibrated sensors: CCS for the former and self-calibration for the latter, aiming to minimize estimation errors. The proposed methodology was first validated in 2D simulations before being extended to the 3D case. The results demonstrate the effectiveness of our approach in correcting reception errors and mitigating geometric mismatches.

Experimental Results

In this chapter, we describe the process of 3D printing the coded cover and assembling the protective enclosure. The experimental setup consists of the RECT-7.5 array, a planar coded cover, and a protective enclosure designed to fully cover the planar array. Using this configuration, we can conduct experiments to evaluate the localization capability of the RECT-7.5 array with the designed planar coded cover.

Due to the limitations of the available facilities, the experiments were conducted in a constrained near-field environment, resulting in spherical waves impinging on the channels of the coded mask. To make the results applicable for comparison with far-field theoretical models, necessary corrections were applied, which will be discussed in the following sections.

The CCS method could not be applied in this case, as the Toeplitz structure of the covariance matrix is lost in the near-field scenario. However, through the experiments, we were still able to validate the localization results based on the STLS with MMV-FISTA method and the self-calibration integrated grid-searching method.

5.1 3D Printing of the Coded Cover and Assembly of the Protective Enclosure

For this project, we ideally aimed to use plastic glass (also known as acrylic or Plexiglas) to construct both the coded cover and the protective enclosure. Plastic glass is highly favored in acoustic engineering due to its relatively high density, which effectively blocks sound waves. This material is particularly useful for creating structures that isolate sound and control its propagation.

Given the precision required for the 1.7 cm spacing between each hole, handcrafting is not ideal due to the potential for errors and the labor-intensive nature of the process. To address this challenge, we opted for 3D printing technology. Specifically, we selected Polylactic Acid (PLA) material, which is known for its rigidity. PLA enhances the structural integrity of the model, ensuring a secure fit while minimizing noise interference, making it an excellent choice for applications that demand precision and stability. The channel thickness was set to 2.5 mm. The model was first designed in SolidWorks, as shown in Figure 5.1, and the final 3D-printed product is displayed in Figure 5.2.

To prevent sound from bypassing the coded cover and reaching the planar array, we designed a protective enclosure, with the coded cover serving as one side. This design ensured that sound could only reach the array through the 140 channels on the coded cover. However, there were still multiple reflections within the box, which could cause incorrect source detection due to the reflected directions. To mitigate this, we applied sound-absorbing materials to the four sides of the box, excluding the side

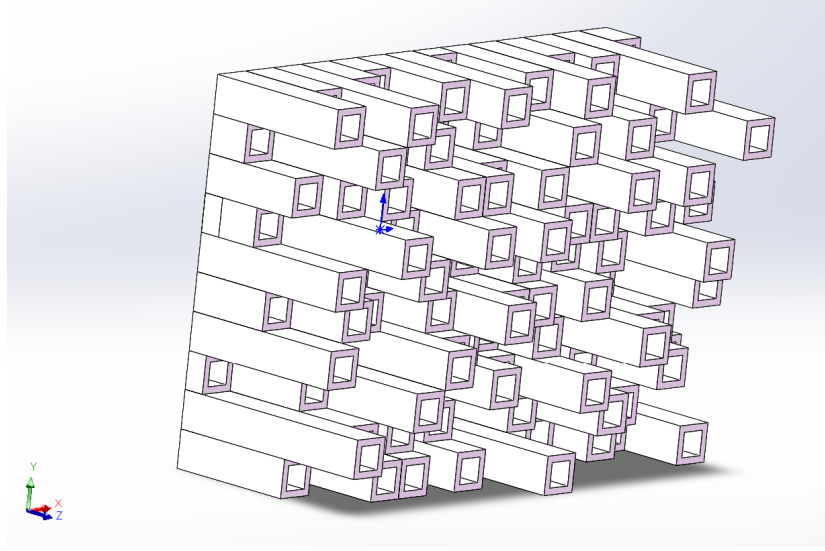
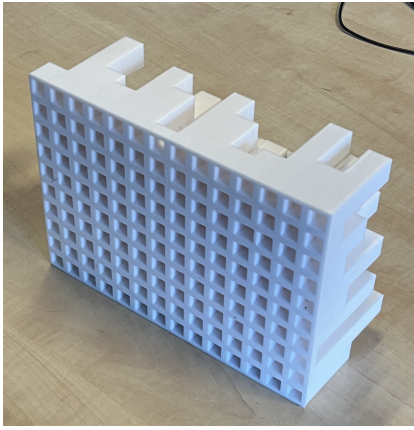
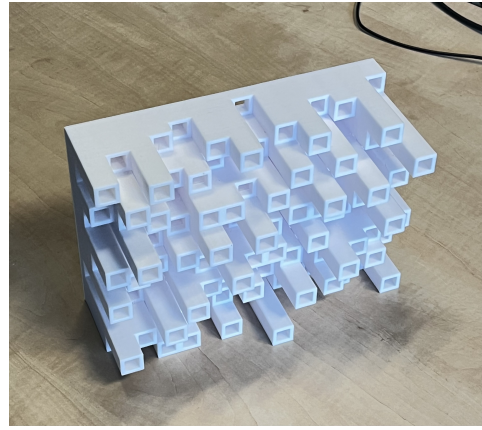


Figure 5.1: The model of coded cover in SolidWorks



(a) Front view of 3D printed coded cover



(b) Back view of 3D printed coded cover

Figure 5.2: 3D printed coded cover in two different views

with the coded cover. The design of the enclosure box, including all required length specifications, is illustrated in Figure 5.3.

The box itself was constructed from wood, with a thickness of 1.8 cm. We chose wood as the material for its excellent balance of durability and sound-damping properties. The relatively thick wood walls help isolate sound effectively while maintaining structural integrity. The physical object is shown in Figure 5.4. The inside view of the enclosure with the RECT-7.5 inside is shown in Figure 5.5.

Due to the unavailability of an anechoic chamber, the experiment was conducted in a corner where two adjacent walls were covered with sound-absorbing materials. However, the ceiling, floor, and the remaining two walls were left uncovered, resulting in an environment that was not fully anechoic. Despite these limitations, the setup still provided sufficiently controlled conditions to validate the proposed methods.

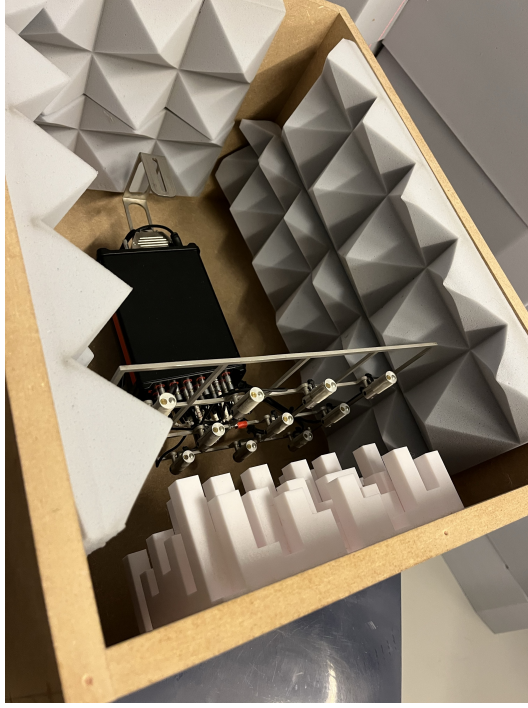


Figure 5.5: Inside view of the enclosure with the RECT-7.5 inside.

coded cover. Since the positions of all channels are known, the distance between the source and each channel can be computed directly. The corresponding scanning vector is then given by:

$$\mathbf{a}(\theta) = \begin{bmatrix} \frac{1}{r_1} e^{jkr_1} \\ \frac{1}{r_2} e^{jkr_2} \\ \vdots \\ \frac{1}{r_M} e^{jkr_M} \end{bmatrix},$$

where k is the spatial frequency of the source signal, also called the wave number.

With this, we can construct a dictionary matrix $\mathbf{A}(\theta)$ to further implement STLS with MMV-FISTA method and the self-calibration integrated grid-searching method.

5.3 Measurement Results

In this section, we implement the self-calibration integrated grid-searching and STLS-MMV-FISTA methods to estimate the source location. For comparison, we also apply the CS method to evaluate whether leveraging the sparsity of the signal leads to any improvement. This is particularly relevant since residual high side lobes may still be present due to reflections within the enclosure. We aim to mitigate these effects using sparsity-based methods. To obtain a preliminary comparison between calibrated and uncalibrated data, we performed a rough calibration of the sensors in an office environment.

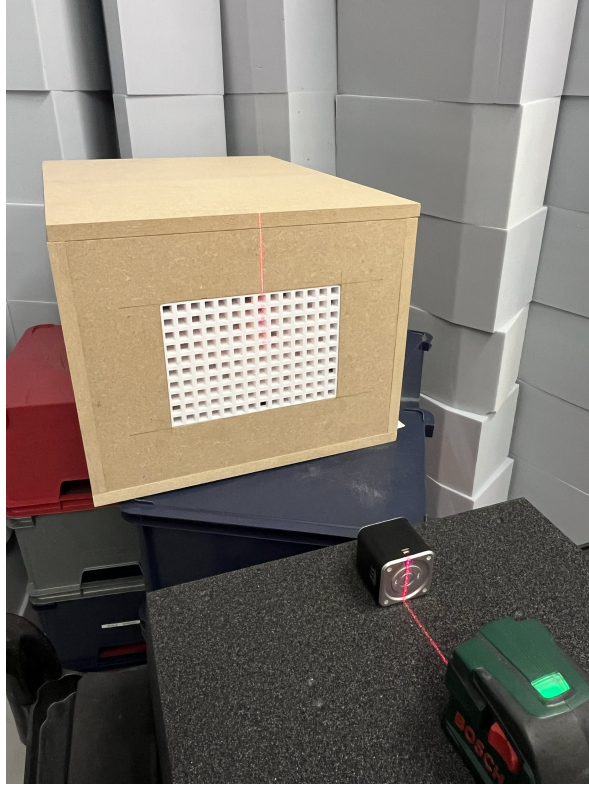


Figure 5.6: Experimental setup with the coded cover

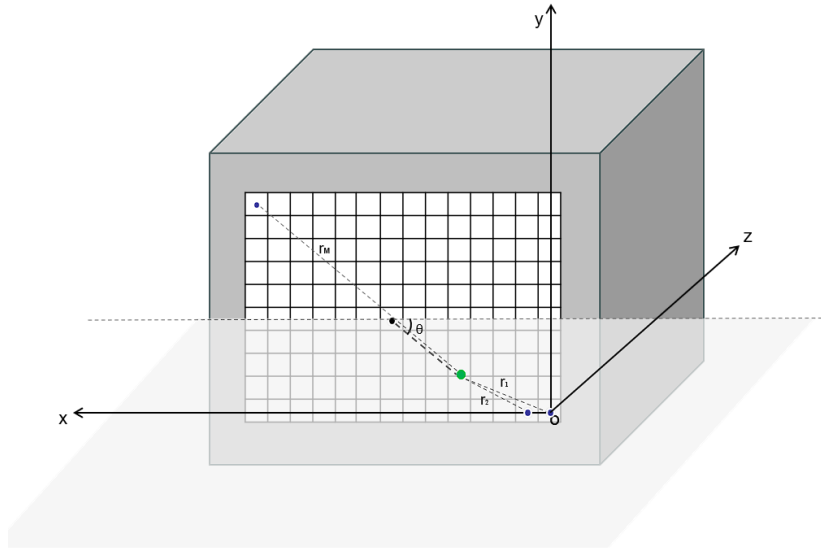


Figure 5.7: Near-field propagation setup modulation

Before processing the received data, a short time Fourier transform (STFT) preprocessing is necessary to extract the target frequency component. The sampling frequency

is set to 25 kHz, and the applied time-frequency analysis window is a Hanning window of length 4096, with a 3072-point overlap.

The scanning resolution is set to 1° , covering an angular range from 1° to 179° . The grid-searching method is applied to optimize the correction parameters: the search range for l is from 14.8 cm to 15.2 cm with a resolution of 0.1 cm, while the search range for the rotation angle ζ spans from -4° to 0° , with a resolution of 1° . This refinement is motivated by the observation that the array may be slightly twisted in one direction due to uneven weight distribution.

In each iteration of grid-searching, self-calibration for phase and gain errors is performed alongside DOA estimation, and the recovered signal is selected based on the minimum error criterion.

5.3.1 Calibrated Sensors

We first examine the scenario where all sensors are calibrated. The source is positioned 30 cm away from the planar coded cover and aligned with the 9th channel of the 5th row. The true source direction is 93.24° .

After applying the CS with MUSIC method, a peak is observed at 92° , deviating by 1.24° from the actual source direction, as shown in Figure 5.8. Despite the sensors are calibrated, the target cannot be precisely localized, and multiple high side lobes introduce ambiguity, potentially leading to the false detection of additional sources. For instance, there are high side lobes peaking at 70° and 124° in the MUSIC spectrum.

The STLS-MMV-FISTA method produces a better result, with a single sharp peak at 92° , as shown in Figure 5.9. This slight deviation is expected, as observed in the simulation results from Section 4.3. STLS in compressed sensing may introduce minor localization errors due to the optimization of a large number of unknowns. However, compared to the CS method, it significantly improves localization sharpness and effectively mitigates side lobes.

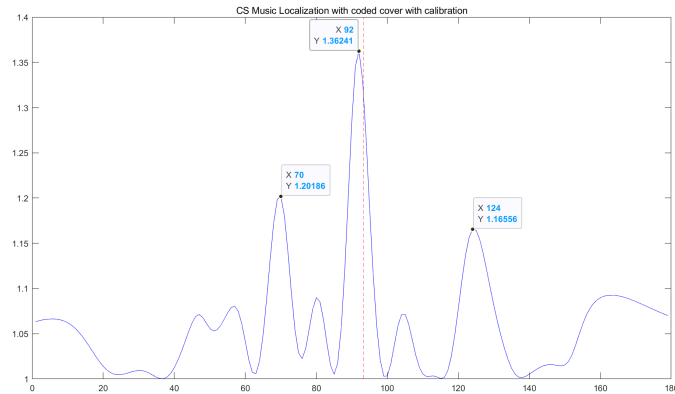


Figure 5.8: Localization result using the CS method with calibrated sensors.

Figure 5.10 presents the DOA estimation result using the self-calibration integrated grid-searching method. After self-calibration, the peak shifts closer to the actual target,

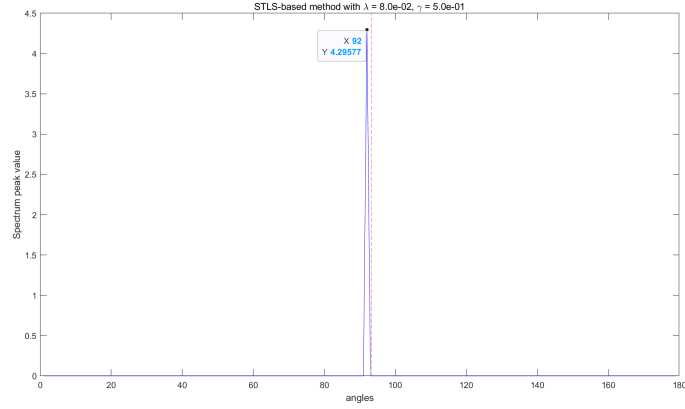


Figure 5.9: Localization result using the STLS-MMV-FISTA method with calibrated sensors.

with an error of only 0.24° . The optimal geometric parameter set obtained is $l = 15.0$ cm and $\zeta = -2.0^\circ$.

To further refine the estimation, CS with the MUSIC method can be applied using the corrected received data and the updated compression matrix incorporating the best geometric parameter set. The enhanced estimation result is shown in Figure 5.11. It is evident that the peak is significantly sharper compared to the previous cases.

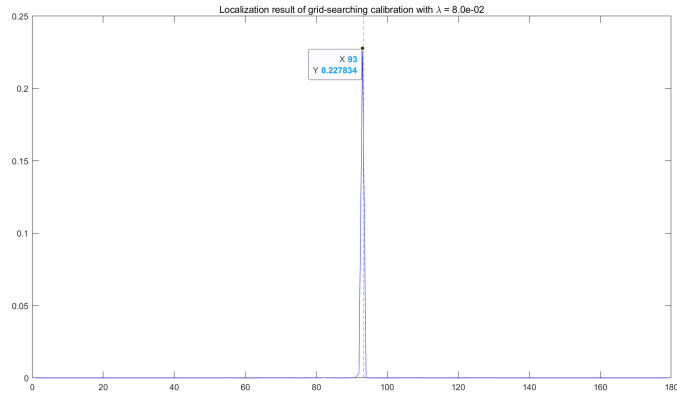


Figure 5.10: DOA estimation result using self-calibration with grid-searching.

5.3.2 Uncalibrated Sensors

Building upon the previous case, we now evaluate the performance of the proposed calibration with the DOA estimation algorithm using uncalibrated sensors. The source signal was positioned 25 cm away from the coded cover, aligned with the 6th channel of the fifth row. The true source direction is 82.25° .

Figure 5.12 presents the DOA estimation result obtained using CS with the MUSIC localization method. The estimated direction is 84° , with noticeably higher side lobes

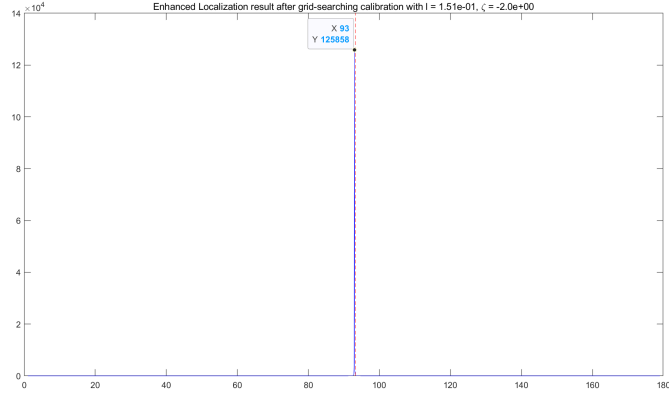


Figure 5.11: Enhanced DOA estimation result using CS with MUSIC after self-calibration.

compared to the calibrated case. In contrast, the STLS method, as shown in Figure 5.13, also estimates the target at 84° , but with a significantly sharper peak and no side lobes.

When applying self-calibration, the estimated peak is located at 81° , bringing it closer to the true source direction, as illustrated in Figure 5.14. Further improvement is achieved by incorporating the optimal calibration and geometric parameter set. The final enhanced result, obtained after applying the CS method, is shown in Figure 5.15.

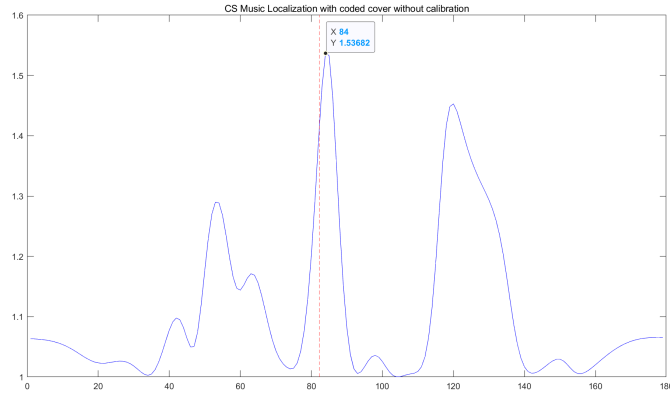


Figure 5.12: DOA estimation result using CS with MUSIC for uncalibrated sensors.

5.4 Conclusion

In this study, we conducted real-world experiments using the CS method, the STLS-MMV-FISTA method, and the self-calibration method. Despite certain experimental constraints, the results demonstrate the effectiveness of the proposed approach. In particular, under non-ideal conditions, methods leveraging signal sparsity effectively mitigate side lobes, leading to improved localization accuracy.

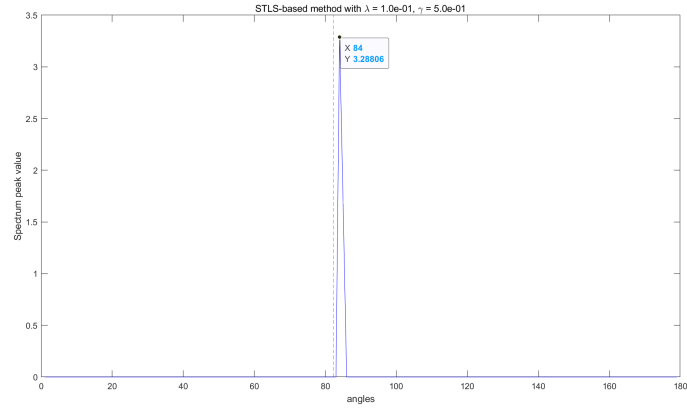


Figure 5.13: DOA estimation result using STLS for uncalibrated sensors.

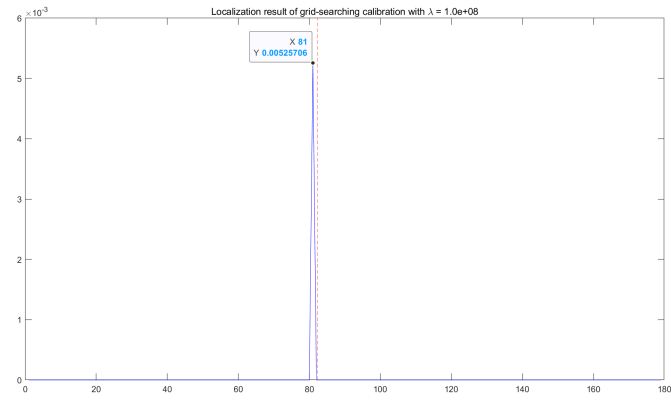


Figure 5.14: DOA estimation result after self-calibration.

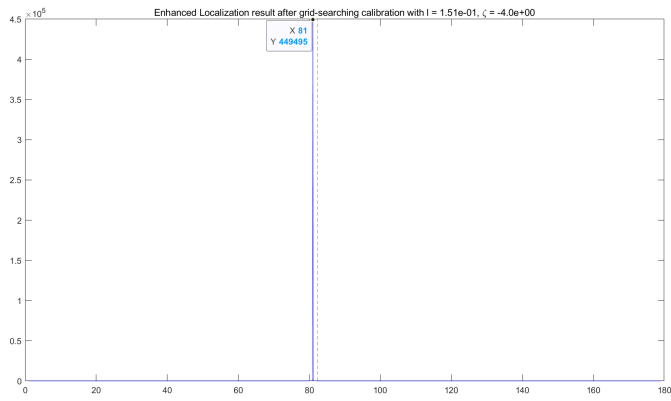


Figure 5.15: Enhanced DOA estimation result using CS with optimized calibration parameters.

Conclusions and Future Directions

6

6.1 Conclusion

This thesis investigates the application of advanced acoustic sensing technologies for DOA estimation and sound source localization, utilizing novel techniques such as CS, CCS, self-calibration, and STLS-MMV-FISTA methods.

In Chapter 2, we explored the extension of the CCS method from a single AVS system to an array-based system, demonstrating that a 14×10 coded cover can effectively localize 100 sound sources in three-dimensional space, even with an SNR as low as 10 dB. This chapter emphasizes the scalability and robustness of the CCS method in large-scale localization tasks, offering a foundation for further advancements in DOA estimation techniques.

Chapter 3 introduced a self-calibration approach within the framework of compressed sensing. We discussed the challenges of phase and gain errors in the receiving channels and presented a solution by leveraging the sparsity of the signals for self-calibration. A theoretical analysis of the ambiguity and identifiability was also investigated. The integration of self-calibration with CCS was shown to significantly improve localization resolution and reduce side lobes, demonstrating its potential for enhancing DOA estimation accuracy.

In Chapter 4, we investigated the use of the STLS-MMV-FISTA method to address geometric mismatches and estimation errors. We explored the effectiveness of grid-searching methods for calibrating and correcting geometric errors. The grid-searching approach, when applied to both calibrated and uncalibrated sensors, helped improve localization accuracy and robustness, especially in 3D scenarios.

Finally, Chapter 5 focused on real-world experiments where we tested the CS method, STLS-MMV-FISTA, and self-calibration techniques under practical conditions. Despite experimental constraints, the results demonstrated that leveraging the sparsity of the signals effectively mitigated side lobes and improved localization accuracy. These experimental results confirmed the applicability of the proposed methods in real-world settings, showing promising performance under non-ideal conditions.

6.2 Future Directions for Research

In the preceding sections, we have demonstrated the benefits of extending the CS and CCS methods with a single AVS to an array of acoustic measurement sensors. Moreover, we have shown that with a specially designed coded cover, source localization can be extended from 2D to 3D. In our simulation, with a 14×10 coded cover, 100 sources in 3D space can be localized effectively using the CCS method when SNR is only 10dB. Moving forward, it is necessary to conduct real-world experiments to validate

the theoretical simulations.

Additionally, several aspects warrant further investigation. For instance, exploring the use of irregularly shaped coded covers could potentially enable the localization of more sources. However, such modifications may alter the subspace structure of the covariance matrix, necessitating further analysis.

Another consideration is the impact of sensor placement. What if the sensors were arranged in a 3D configuration, such as a spherical or other geometric shape? For example, Microflow's Scan&Paint 3D system [32] utilizes a spherical array with 12 sensors, as shown in Figure 6.1. Possibly we could design a cover which could localize sources across the entire 3D space rather than just in the forward direction. Addressing these questions could further enhance the applicability of the proposed methods.



Figure 6.1: Scan&Paint 3D system from Microflow

Regarding calibration, further research is needed to identify methods that are well-suited for compressed sensing frameworks, particularly for cases where multiple types of errors need to be calibrated. Another important direction is investigating potential extensions of the STLS method to enhance its localization capabilities. Based on our research findings, both MMV-based and covariance-based models show promise for future exploration.

Bibliography

- [1] J. P. Kitchens. “Acoustic vector-sensor array performance”. Master’s thesis. Massachusetts Institute of Technology, 2008.
- [2] H.-E. de Bree. “An overview of Microflow technologies”. In: *Acta Acustica united with Acustica* 89.1 (2003), pp. 163–172.
- [3] Hans-Elias. *The Microflow E-book*. Available online at <https://www.microflow.com>. 2009.
- [4] F. Jacobsen and P. M. Juhl. *Fundamentals of General Linear Acoustics*. John Wiley & Sons, 2013.
- [5] A. Nehorai and E. Paldi. “Acoustic vector-sensor array processing”. In: *IEEE Transactions on Signal Processing* 42.9 (1994), pp. 2481–2491.
- [6] Chenyang Yan. “Coded Cover for Acoustic Vector Sensors”. Supervisors: G. J. T. Leus and Daniel Fernandez Comesana; Committee members: G. Joseph, N. J. Jonathan Myers. Master’s thesis. Faculty of Electrical Engineering, Mathematics and Computer Science: Delft University of Technology, 2023. URL: <http://resolver.tudelft.nl/uuid:7c6480c0-c8fb-4a45-b2a3-ba24fb32d1ea>.
- [7] D. Romero et al. “Compressive Covariance Sensing: Structure-based compressive sensing beyond sparsity”. In: *IEEE Signal Processing Magazine* 33.1 (2016), pp. 78–93. DOI: [10.1109/MSP.2015.2486805](https://doi.org/10.1109/MSP.2015.2486805).
- [8] Y. Wang, G. Leus, and A. Pandharipande. “Direction estimation using compressive sampling array processing”. In: *2009 IEEE/SP 15th Workshop on Statistical Signal Processing*. 2009, pp. 626–629.
- [9] E. Candes. “Compressive sampling”. In: *Proceedings of the International Congress of Mathematicians*. Vol. 3. 2006, pp. 1433–1452. ISBN: 978-3-03719-022-7.
- [10] E. J. Candes and M. B. Wakin. “An introduction to compressive sampling”. In: *IEEE Signal Processing Magazine* 25.2 (2008), pp. 21–30.
- [11] Y. Wang, G. Leus, and A. Pandharipande. “Direction estimation using compressive sampling array processing”. In: *2009 IEEE/SP 15th Workshop on Statistical Signal Processing*. IEEE, 2009, pp. 626–629.
- [12] Y. Wang, A. Pandharipande, and G. Leus. “Compressive sampling based MVDR spectrum sensing”. In: *2010 2nd International Workshop on Cognitive Information Processing*. IEEE. 2010, pp. 333–337.
- [13] S. F. Cotter et al. “Sparse solutions to linear inverse problems with multiple measurement vectors”. In: *IEEE Transactions on Signal Processing* 53.7 (2005), pp. 2477–2488.
- [14] D. Romero and G. Leus. “Compressive covariance sampling”. In: *2013 Information Theory and Applications Workshop (ITA)*. IEEE, 2013, pp. 1–8.
- [15] Zai Yang, Lei Xie, and Petre Stoica. “Vandermonde Decomposition of Multilevel Toeplitz Matrices With Application to Multidimensional Super-Resolution”. In: *IEEE Transactions on Information Theory* 62.6 (June 2016), pp. 3685–3701. DOI: [10.1109/TIT.2016.2553041](https://doi.org/10.1109/TIT.2016.2553041).
- [16] N.R. Krishnaprasad. “Acoustic Vector Sensor Based Source Localization”. G.J.T. Leus (mentor). Master’s thesis. Delft, The Netherlands: Delft University of Tech-

- nology, 2016. URL: <http://resolver.tudelft.nl/uuid:cfd81ca7-e44c-4d0d-8680-a9150f9d1162>.
- [17] Li-Juan Sun et al. “Cooperative Calibration Scheme for Mobile Wireless Sensor Network”. In: *2019 15th International Conference on Mobile Ad-Hoc and Sensor Networks (MSN)*. 2019, pp. 143–150.
 - [18] Florentin Delaine, Bérengère Lebental, and Hervé Rivano. “In Situ Calibration Algorithms for Environmental Sensor Networks: A Review”. In: *IEEE Sensors Journal* 19.15 (2019), pp. 5968–5978.
 - [19] Jielong Yang, Xionghu Zhong, and Wee Peng Tay. “A Dynamic Bayesian Non-parametric Model for Blind Calibration of Sensor Networks”. In: *IEEE Internet of Things Journal* 5.5 (2018), pp. 3942–3953.
 - [20] Clément Dorffer et al. “Informed Nonnegative Matrix Factorization Methods for Mobile Sensor Network Calibration”. In: *IEEE Transactions on Signal and Information Processing over Networks* 4.4 (2018), pp. 667–682.
 - [21] Clément Dorffer and Matthieu Puigt. “Outlier-robust calibration method for sensor networks”. In: *2017 IEEE International Workshop of Electronics, Control, Measurement, Signals and their Application to Mechatronics (ECMSM)*. 2017, pp. 1–6.
 - [22] Yuzhi Wang et al. “A Deep Learning Approach for Blind Drift Calibration of Sensor Networks”. In: *IEEE Sensors Journal* 17.13 (2017), pp. 4158–4171.
 - [23] Jielong Yang, Wee Peng Tay, and Xionghu Zhong. “A dynamic Bayesian non-parametric model for blind calibration of sensor networks”. In: *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2017, pp. 4207–4211.
 - [24] Rosaria Rossini et al. “WSNs Self-Calibration Approach for Smart City Applications Leveraging Incremental Machine Learning Techniques”. In: *2016 8th IFIP International Conference on New Technologies, Mobility and Security (NTMS)*. 2016, pp. 1–7.
 - [25] Clément Dorffer et al. “Nonlinear mobile sensor calibration using informed semi-nonnegative matrix factorization with a Vandermonde factor”. In: *2016 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*. 2016, pp. 1–5.
 - [26] K. N. Ramamohan et al. “Self-Calibration of Acoustic Scalar and Vector Sensor Arrays”. In: *IEEE Transactions on Signal Processing* 71 (2023), pp. 61–75. DOI: [10.1109/tsp.2022.3214383](https://doi.org/10.1109/tsp.2022.3214383).
 - [27] J. R. Fienup. “Reconstruction of an object from the modulus of its Fourier transform”. In: *Optics Letters* 3 (1978), pp. 27–29.
 - [28] A. J. Weiss and B. Friedlander. “Eigenstructure methods for direction finding with sensor gain and phase uncertainties”. In: *Circuits, Systems, and Signal Processing* 9.3 (1990). Revised, pp. 271–300.
 - [29] P. Pal and P. P. Vaidyanathan. “On application of LASSO for sparse support recovery with imperfect correlation awareness”. In: *Proc. IEEE Conf. Rec. 46th Asilomar Conf. Signals, Syst. Comput.* 2012, pp. 958–962.

- [30] H. Zhu, G. Leus, and G. B. Giannakis. “Sparsity-Cognizant Total Least-Squares for Perturbed Compressive Sampling”. In: *IEEE Transactions on Signal Processing* 59.5 (May 2011), pp. 2002–2016. DOI: [10.1109/TSP.2011.2109956](https://doi.org/10.1109/TSP.2011.2109956).
- [31] Alexandre Gramfort, Matthieu Kowalski, and Matti Hämäläinen. “Mixed-norm estimates for the m/EEG inverse problem using accelerated gradient methods”. In: *Physics in Medicine and Biology* 57.7 (Mar. 2012), pp. 1937–1961. DOI: [10.1088/0031-9155/57/7/1937](https://doi.org/10.1088/0031-9155/57/7/1937).
- [32] D. Fernandez Comesana. “Scan-based sound visualisation methods using sound pressure and particle velocity”. Ph.D. thesis. University of Southampton, 2014.