**Signals Processing Systems**
Mekelweg 4,
2628 CD Delft
The Netherlands
https://sps.ewi.tudelft.nl/

**SPS-2023-00**

# M.Sc. Thesis

# Coded Cover for Acoustic Vector Sensors

**Chenyang Yan B.Sc.**

**Abstract**

This study delves into the application of coded covers in enhancing Acoustic Vector Sensor (AVS) performance for sound source localization. We initially explored the use of a coded mask inspired by ultrasound imaging. However, our analysis indicated that the coded mask primarily acts as a scaling factor without substantial benefits for direction of arrival (DOA) estimation.

In response, we investigated an alternative approach using a larger cover with spaced channels, positioned above the AVS sensor. In this setup, each channel operates independently, and their combined acoustic outputs are analyzed upon reaching the sensor. Importantly, the upper surface channels can be viewed as a virtual uniform linear array (ULA), allowing the application of compressed analysis methods.

Two distinct DOA estimation approaches were developed: the compressive sampling (CS) method and the compressive covariance sensing (CCS) method. Both methods were validated, showing improved accuracy in angle estimation. Notably, the CCS method exhibited the potential to expand the number of detectable sound sources using a single AVS.

In summary, while the coded mask did not offer significant advantages, the coded cover design, along with the CS and CCS methods, improved DOA accuracy and extended the detection capabilities of single AVS.

# Coded Cover for Acoustic Vector Sensors

Thesis

submitted in partial fulfillment of the
requirements for the degree of

Master of Science

in

Electrical Engineering

by

Chenyang Yan B.Sc.
born in Leshan, China

This work was performed in:

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

**Delft University of Technology**

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 **"Coded Cover for Acoustic Vector Sensors"** by **Chenyang Yan B.Sc.** in partial fulfillment of the requirements for the degree of **Master of Science**.

Dated: 25/10/2023

Chairman: _____

prof.dr.ir. Geert Leus

Advisor: _____

dr. Dani Fernandez Comesaña

Committee Members: _____

dr. Geethu Joseph

_____

dr. Nitin Meyers

# Abstract

This study delves into the application of coded covers in enhancing Acoustic Vector Sensor (AVS) performance for sound source localization. We initially explored the use of a coded mask inspired by ultrasound imaging. However, our analysis indicated that the coded mask primarily acts as a scaling factor without substantial benefits for direction of arrival (DOA) estimation.

In response, we investigated an alternative approach using a larger cover with spaced channels, positioned above the AVS sensor. In this setup, each channel operates independently, and their combined acoustic outputs are analyzed upon reaching the sensor. Importantly, the upper surface channels can be viewed as a virtual uniform linear array (ULA), allowing the application of compressed analysis methods.

Two distinct DOA estimation approaches were developed: the compressive sampling (CS) method and the compressive covariance sensing (CCS) method. Both methods were validated, showing improved accuracy in angle estimation. Notably, the CCS method exhibited the potential to expand the number of detectable sound sources using a single AVS.

In summary, while the coded mask did not offer significant advantages, the coded cover design, along with the CS and CCS methods, improved DOA accuracy and extended the detection capabilities of single AVS.

# Acknowledgments

Time flies, and in the blink of an eye, my master's journey of over two years at TU Delft is coming to an end. The past is irrevocable, but I look forward to a bright future.

Here, I would like to express my gratitude to my advisors, Geert Leus and Dani Fernandez. They have provided invaluable guidance in my academic pursuits. I am also thankful to my parents, who have been unwavering in their support of my dreams. And to all my friends, you have been a great source of strength during challenging times.

Chenyang Yan B.Sc.
Delft, The Netherlands
25/10/2023

# Contents

# List of Figures

# List of Tables

# 1 Introduction

Human ears possess the remarkable ability to hear sounds and determine their direction without the need for complex systems. In a similar fashion, the acoustic vector sensor (AVS) is capable of accomplishing the same task without requiring an excessive number of sensor elements. Unlike the acoustic pressure sensor (APS), commonly known as a microphone, which solely detects sound pressure, the AVS measures both the acoustic pressure and the particle velocity. This advanced sensor eliminates the necessity for multiple probes in traditional signal processing, significantly reducing the system's weight, size and cost. Even a single sensor is sufficient to determine the direction of arrival (DOA) [3].

In practice, the acoustic vector sensors are often surrounded by a layer of protective cover. These covers can distort and attenuate the acoustic waves and reduce the performance of the DOA measurements. However, could they also have a benefit? We already know that in ultrasound imaging, an aberrating layer, which is a specially-made mask, is used to encode the spatial measurement in order to provide more information about the location of the sources. Therefore, the question is whether we can apply this masking technique in ultrasound imaging to the acoustic vector sensor to improve the system.

## 1.1 Thesis Outline

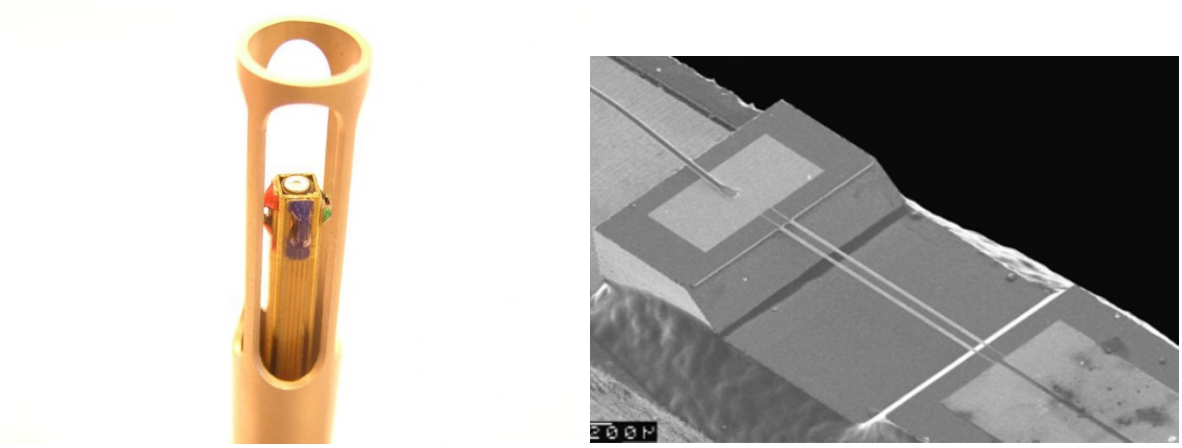
Chapter one provides a comprehensive introduction to the research, addressing its motivation and presenting an overview of the fundamental physics and signal principles relevant to the AVS. Additionally, we explore the distinctions between the application of a coded mask to AVS and its use in ultrasound imaging. Chapter two is dedicated to an in-depth investigation of the impact of implementing a coded mask on the AVS cover. Chapter three introduces an innovative cover structure designed to enhance the AVS DOA performance. We provide a detailed examination of this novel cover and discuss its potential for improvements in AVS DOA estimation. Chapter four shifts our focus to practical measurements, offering insights into experimental setups and presenting the results obtained from these experiments. Finally, in chapter five, we draw our conclusions and outline potential avenues for future research.

## 1.2 Basics of Acoustic Vector Sensor

Acoustic particle velocity sensors are transducers capable of capturing particle velocity in air. These transducers draw inspiration from hot-wire anemometers, where a wire is heated by an electrical current and cools when exposed to an acoustic flow. Due to temperature changes in the wire, its resistance varies, generating a variable electrical

signal proportional to the incident flow. By employing two closely spaced heated wires and measuring the temperature difference between them, it becomes possible to gauge the acoustic particle velocity. These sensors are compact in size, enabling the placement of three orthogonal sensors in close proximity to one another for the characterization of the sound field's acoustic particle velocity vector.

An acoustic vector sensor (AVS) is a device designed to measure both acoustic pressure and particle velocity. It is created by positioning multiple transducers closely together, comprising a combination of a microphone and three (or two) particle velocity transducers. These transducers are oriented orthogonally to each other and are typically located in three-dimensional ($\mathbb{R}^3$) or two-dimensional ($\mathbb{R}^2$) space. One such AVS, manufactured by Microflown Technologies, is depicted in Figure 1.1a. In this design, all the transducers are positioned on a pillar. At the top of the pillar, there is a microphone, while three particle velocity transducers are orthogonally placed on the sides of the pillar. Figure 1.1b provides a close-up view of the particle velocity transducer based on the principle of the hot-wire anemometer [1] [2].



(a) Acoustic vector sensor

(b) Acoustic particle velocity sensor

Figure 1.1: Acoustic vector sensor manufactured by Microflown Technologies. [1] [2]

### 1.2.1 Applications

In practice, the AVS and its arrays find applications in a wide range of scenarios. These applications can span both indoor and outdoor environments, involving near-field or far-field propagation of acoustic waves in air or underwater mediums. Depending on the specific nature of the environment, propagation characteristics, and the medium involved, AVS arrays have demonstrated their utility across a diverse array of fields. These fields include, but are not limited to, automotive technology, battlefield acoustics, underwater exploration, and many others. They are used for tasks such as sound visualization and imaging [5], acoustic intensity and impedance measurement [6], and sound source detection, localization, and tracking [7].

### 1.2.2 Sound Wave Propagation

In the scenario of a single point source operating under free-field conditions, the pressure at a point in space, denoted as $\mathbf{r}$, due to a harmonic point source with a volume velocity Q at the origin, can be described by the solution to the inhomogeneous Helmholtz equation [8]:

$$\nabla^2 p(\mathbf{r}) + k^2 p(\mathbf{r}) = -j\omega\rho Q\delta(\mathbf{r})e^{j\omega t} \tag{1.1}$$

where $p(\mathbf{r})$ is the complex acoustic pressure field at the point $\mathbf{r}$, $k$ is the wave number, defined as $k = \frac{\omega}{c}$, where $\omega$ is the angular frequency and $c$ is the speed of sound in the medium, and $\rho$ is the density of the medium.

The pressure at $\mathbf{r}$ is the solution of (1.1) and can be expressed as:

$$p(\mathbf{r}) = \frac{j\omega\rho Q e^{j(\omega t - kr)}}{4\pi r} \tag{1.2}$$

with

$$r = |\mathbf{r}| \tag{1.3}$$

The radial component of the particle velocity can be calculated from Euler's equation of motion:

$$v_r(r) = -\frac{1}{j\omega\rho}\frac{\partial p(\mathbf{r})}{\partial(r)} = \frac{jkQe^{j(\omega t - kr)}}{4\pi r}\left(1 + \frac{1}{jkr}\right) = \frac{p(\mathbf{r})}{\rho c}\left(1 + \frac{1}{jkr}\right) \tag{1.4}$$

In the context of far-field sources, the term $\left(1 + \frac{1}{jkr}\right)$ can be approximated as 1. Consequently, we can derive a linear relationship between pressure and particle velocity as follows:

$$v_r(r) = \frac{p(\mathbf{r})}{\rho c} \tag{1.5}$$

Furthermore, we can define the unit vector in the direction of the source as:

$$\mathbf{u} = \begin{bmatrix} u_x \\ u_y \end{bmatrix} = \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix} \tag{1.6}$$

Here, $\theta$ represents the incident angle of the source.

The relationship between the sound pressure, denoted as $p(\mathbf{r}, t)$, and the acoustic particle velocity, represented as $\mathbf{v}(\mathbf{r}, t)$, at time $t$, can be thus described as follows:

$$\mathbf{v}(\mathbf{r}, t) = -\frac{\mathbf{u}}{\rho c} p(\mathbf{r}, t) \tag{1.7}$$

where the term $\rho c$ represents the characteristic acoustic impedance of the medium.

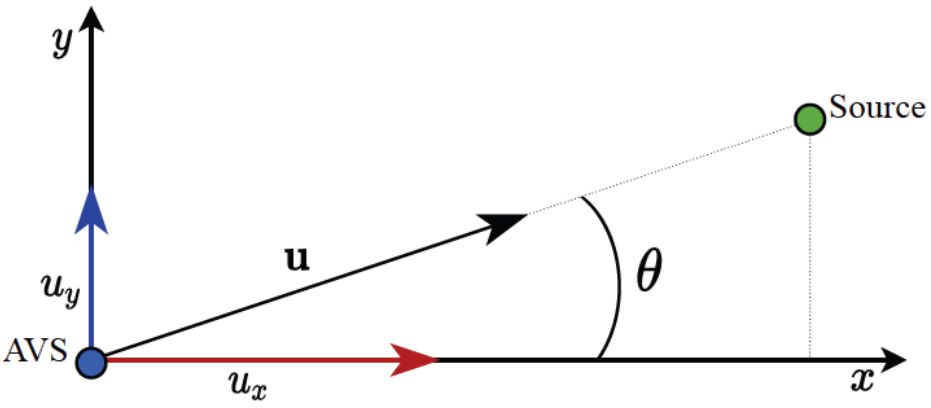



Figure 1.2: One 2D AVS and single source configuration in $\mathbb{R}^2$ [3]

### 1.2.3 Single AVS Measurement Model

In this section, we will elaborate on the measurement model involving a single AVS and a solitary source situated in the far field. For the purpose of our analysis, we consider a 2D AVS and a single-source configuration within the two-dimensional space ($\mathbb{R}^2$), as shown in Fig.1.2.

The direction of a plane wave impinging on an AVS shall be parameterized in terms of a two-dimensional ($\mathbb{R}^2$) spherical coordinate basis, with the azimuthal angle $\theta$ in the range $[0,\ 2\pi)$.

Given the linear relation between the pressure and particle velocity, a data model for the 2-D AVS can be obtained. The measurement data at time t can be written as:

$$\mathbf{y}(t) = e^{jk(\mathbf{r}^{\mathrm{T}}\mathbf{u})} \begin{bmatrix} 1 \\ \mathbf{u}/(\rho c) \end{bmatrix} s(t) + \underbrace{\begin{bmatrix} n_p(t) \\ \mathbf{n}_v(t) \end{bmatrix}}_{\mathbf{n}(t)} \tag{1.8}$$

In the provided expression, $\mathbf{y}(t)$ is a three-dimensional vector that encompasses observations from both the pressure sensor and the practical velocity sensors. The variables $s(t)$, $n_p(t)$, and $\mathbf{n}_v(t)$ correspond to the source signal and the additive noise components associated with the pressure sensor and particle velocity sensors, respectively. The exponential term represents the phase delay induced by far-field propagation. For the sake of analytical simplicity, we assume that the AVS is located at the origin point with $\mathbf{r} = 0$, and in this case, the exponential term simplifies to 1. We also scale the particle velocity measurements by $\rho c$. The data model can be then expressed as [9]:

$$\mathbf{y}(t) = \begin{bmatrix} 1 \\ \mathbf{u} \end{bmatrix} s(t) + \begin{bmatrix} n_p(t) \\ \mathbf{n_v}(t) \end{bmatrix} \tag{1.9}$$

The response vector of a single AVS is therefore:

$$\mathbf{a}(\theta) = \begin{bmatrix} 1 \\ \mathbf{u} \end{bmatrix} \tag{1.10}$$

### 1.2.4 Beam Pattern of a Single AVS

Given the measurement model in (1.9) for a single AVS with a single far-field source, we can analyze the beam pattern. The matched filter beam pattern for this model can be expressed as:

$$\mathrm{B}_{\mathrm{AVS}}(\theta) = |\mathbf{a}^{\mathrm{H}}(\theta)\mathbf{a}(\theta_0)| = 1 + \cos(\theta)\cos(\theta_0) + \sin(\theta)\sin(\theta_0), \tag{1.11}$$

where $\theta_0$ is the source direction.

Figure 1.3 represents the beam pattern of a single AVS when the source direction is fixed at $\theta = 0$ degrees. In fact, the beam pattern of a single AVS is also referred to as velocity gain modulation (VGM). This modulation results in a doubled gain from the direction of arrival, thanks to the inclusion of two additional velocity channels.

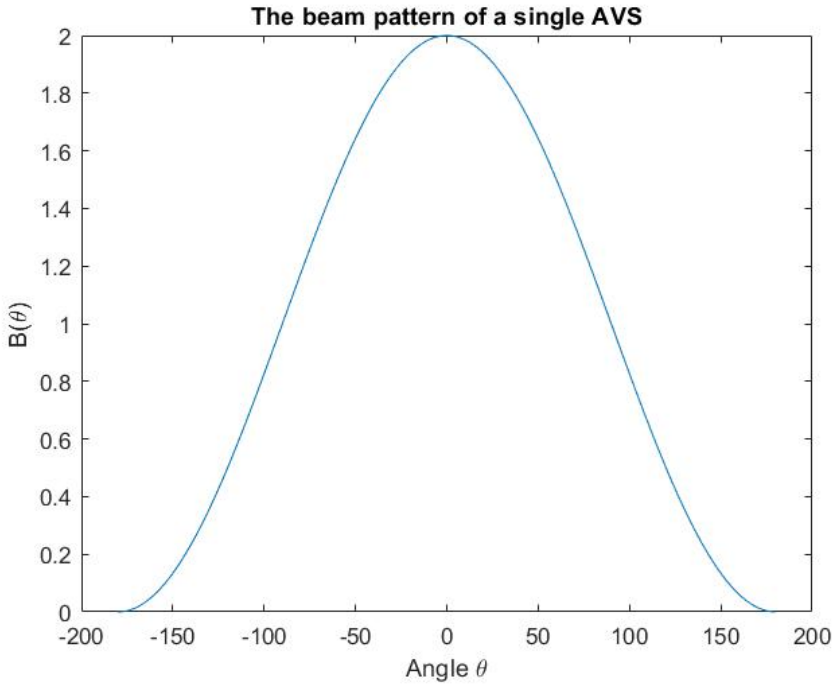



Figure 1.3: The beam pattern of a singel AVS

## 1.3 Comparing the Application of Coded Masks in AVS and Ultrasound Imaging

Ultrasound imaging primarily focuses on reconstructing an image by estimating the reflection coefficient from the available data. This involves solving an inverse problem, often using methods like least squares estimation [10]. On the other hand, AVS sensors are designed to estimate the direction of arrival of sound sources, which is a parameter estimation problem. These differing objectives result in variations in sensor configuration, signal processing techniques, and algorithms employed.

Furthermore, it is important to note that ultrasound sensors operate as active systems, emitting regular pulses of ultrasound waves. They function as both transmitters and receivers, allowing for the measurement of reflected signals [10] [4]. In contrast, AVS sensors operate as passive systems and do not transmit any signals. They solely act as receivers, detecting and measuring the acoustic pressure and particle velocity from the surrounding environment. This distinction in functionality means that in AVS sensor applications, it is not possible to rotate or translate the coded mask to obtain additional measurements with different spatial resolution. This is because the source signal changes over time and hence the source signal is not the same for two consecutive measurements.

Table 1.1: Comparison of Coded Mask Applications in Ultrasound Imaging and AVS

| **Name** | **System type** | **Channel number** | **Purpose** | **Problem Type** | **Movable Coded Mask** |
|---|---|---|---|---|---|
| Ultrasound Imaging | Active system | 1 | Image recovery | Inverse problem | Yes |
| DOA with AVS Sensor | Passive system | 3 (2-D case) | Angle estimation | Parameter estimation | No |

# Applying Coded Mask to AVS Sensor DOA Estimation

# 2

The coded mask employed in ultrasound imaging involves the placement of a thin plastic film in front of the ultrasound sensor. In this chapter, we explore the potential effectiveness of a similar coded mask when used with acoustic vector sensors (AVS). In ultrasound imaging, an ultrasonic wave is transmitted and received through a simple plastic coding mask. However, in the DOA estimation of an AVS, we only consider a reception through the code mask.

We begin by introducing the coded mask model adopted from ultrasound imaging. Subsequently, we conduct a comparative analysis of the performance of a single AVS with the mask and without the mask. In this analysis, we assume the presence of a single AVS and a single source, evaluating their respective capabilities in direction of arrival (DOA) estimation. Furthermore, we substantiate our findings through beam pattern analysis and an assessment of the Cramér-Rao bound for both methods.

## 2.1 Coded Mask Model

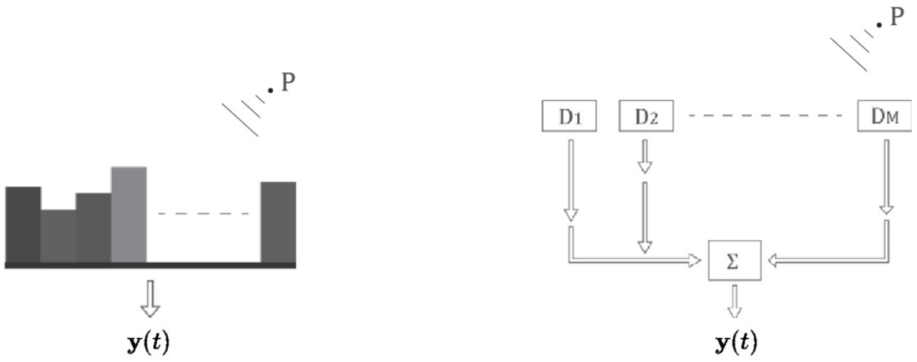
In ultrasound imaging, a coded mask is a plastic material placed over a sensor to modify its response characteristics. This modification is accomplished by varying the mask's thickness, which, in turn, impacts the time it takes for ultrasound waves to traverse it. This thickness variation introduces angle-dependent effects into the imaging process. In this section, we introduce the coded mask model, and in the following sections, we apply this model, borrowed from ultrasound imaging, to evaluate its impact on a single AVS.

In ultrasound imaging, we begin by dividing the mask's surface into small patches along the horizontal dimension. Each of these discrete points represents an individual sensor channel. For each channel, we introduce delays to the source signal, taking into account the thickness of the mask at that specific point and the speed of sound within the mask. Essentially, we assume that each channel receives the pulse-echo wave field at the interface between the mask and the medium as its input. The acoustic field then independently propagates through each channel before reaching the transducer. Each channel introduces a delay to the incident wavefield, and we approximate the spatial integration across the transducer surface by summing all delayed channel signals, as illustrated in Figure 2.1 [4].

Mathematically, we assume that there is a phase delay as the sound wave passes through each patch. From a mathematical perspective, the model for the i-th patch of the mask in the frequency domain can be expressed as follows:

$$d_m(\omega) = e^{-j2\omega\tau_m} \tag{2.1}$$

where $\tau_i$ represents the phase delay caused by the wavelength traveling through the




(a) The signal model of the coded mask.

(b) The physical model of the coded mask.

Figure 2.1: The physical model and signal model of the coded mask for ultrasound imaging [4].

$i$-th patch. It is calculated as the thickness of the mask divided by the velocity of the acoustic wave in the coded mask.

## 2.2 Measurement Model

In this section, we present a mathematical model in the time domain similar to ultrasound to depict the utilization of a coded mask with a single AVS in the presence of a single source. The coded mask is divided into $M$ patches. At each patch, we consider a unique phase delay $d_m$ determined by the sound velocity in the mask and the length of the patch. Additionally, we denote $\mathbf{r}_m$ as the position of the entry of the $m$-th patch, and the phase delay caused by the far-field propagation from the source to the entry of the $m$-th patch as $e^{jk\left(\mathbf{r}_m^{\mathrm{T}}\mathbf{u}\right)}$. The received signal following the $m$-th patch can be described using the following equations:

$$\mathbf{y}_m(t) = d_m(\omega)e^{jk(\mathbf{r}_m^{\mathrm{T}}\mathbf{u})}\begin{bmatrix} 1 \\ \mathbf{u} \end{bmatrix} s(t) \tag{2.2}$$

The acoustic vector sensor gathers signals from these $M$ patches and calculates their summation. To ensure a fair comparison with a single AVS, we compute the average across the $M$ channels:

$$\mathbf{y}(t) = \frac{1}{M}\sum_{i=1}^{M}\mathbf{y}_i(t) = \frac{1}{M}\left(\sum_{i=1}^{M} e^{-j\omega\tau_i}e^{jk(\mathbf{r}_i^{\mathrm{T}}\mathbf{u})}\right)\begin{bmatrix} 1 \\ \mathbf{u} \end{bmatrix} s(t) + \begin{bmatrix} n_p(t) \\ \mathbf{n_v}(t) \end{bmatrix} \tag{2.3}$$

In this equation, $\mathbf{y}(\mathrm{t})$ represents the overall received signal. $\mathrm{n_p(t)}$ and $\mathbf{n_v}(\mathrm{t})$ are assumed to be independent of the source's direction and unaffected by the presence of the coded mask.

Furthermore, we define the overall effect of the coded mask as $k(\theta)$, and the manifold after the application of the coded mask as $\mathbf{h}(\theta)$:

$$\mathbf{h}(\theta) = k(\theta)\mathbf{a}(\theta) \tag{2.4}$$

$$\mathrm{k}(\theta) = \frac{1}{M}\sum_{i=1}^{M} e^{-j\omega\tau_i} e^{jk\left(\mathbf{p}_i^{\mathrm{T}}\mathbf{u}\right)} \tag{2.5}$$

It is apparent that the effect of the coded mask is a scaling operation. This scaling effect is determined by the direction of the source signal. Notably, $0 \leq |k(\theta)| \leq 1$ due to the fact that the phase delays of each patch can either reinforce or attenuate each other when averaged. Consequently, unless all the phase delays can constructively add up, the overall norm of the combined signal cannot exceed 1.

In case of a scenario with $N$ sources located in the directions pointed by the unit vectors $\mathbf{u}_1\mathbf{u}_2 \ldots \mathbf{u}_{\mathrm{N}}$. The measurement model is:

$$\begin{aligned}\mathbf{y}(t) &= \begin{bmatrix} \mathbf{h}(\theta_1) & \mathbf{h}(\theta_2) & \ldots & \mathbf{h}(\theta_{\mathrm{N}}) \end{bmatrix} \begin{bmatrix} s_1(t) \\ s_2(t) \\ \cdot \\ \cdot \\ \cdot \\ s_N(t) \end{bmatrix} + \mathbf{n}(t) \\ &= \mathbf{H}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t)\end{aligned} \tag{2.6}$$

where $\boldsymbol{\theta} = \begin{bmatrix} \theta_1 & \theta_2 & \ldots & \theta_{\mathrm{N}} \end{bmatrix}$ collects the angle of the N sources.

## 2.3 Beam Pattern of the AVS Sensor with a Coded Mask

Using the measurement model in (2.3) for a single AVS with a coding mask and a single far-field source, we can examine the beam pattern. The matched filter beam pattern for this model can be expressed as:

$$\mathrm{B} = \left|\mathbf{h}^{\mathrm{H}}(\theta)\,\mathbf{h_0}(\theta)\right| = \left|k(\theta)^{\mathrm{H}}\,k(\theta_0)\right| \left|\mathbf{a}(\theta)^{\mathrm{H}}\,\mathbf{a}(\theta_0)\right| = \left|k(\theta)^{\mathrm{H}}\,k(\theta_0)\right| \mathrm{B}_{AVS} \tag{2.7}$$

In this equation, we've set $\theta_0$ as a constant, and $\theta$ ranges from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$. Here, $B_{\mathrm{AVS}}(\theta)$ represents the beam pattern of a single AVS in the absence of the coded mask.

Figure 2.2 illustrates the beam pattern of a single AVS with and without the coded mask, with $\theta_0$ fixed at 0 degrees. It's evident that the mask alters the shape of the beam pattern and reduces the received signal-to-noise ratio.

## 2.4 Performance Study Using MUSIC Algorithm

The MUSIC algorithm is a popular method used in array signal processing for DOA estimation. It's particularly effective when dealing with scenarios where there are multiple sources of signals and one needs to identify the angles from which these signals are arriving.

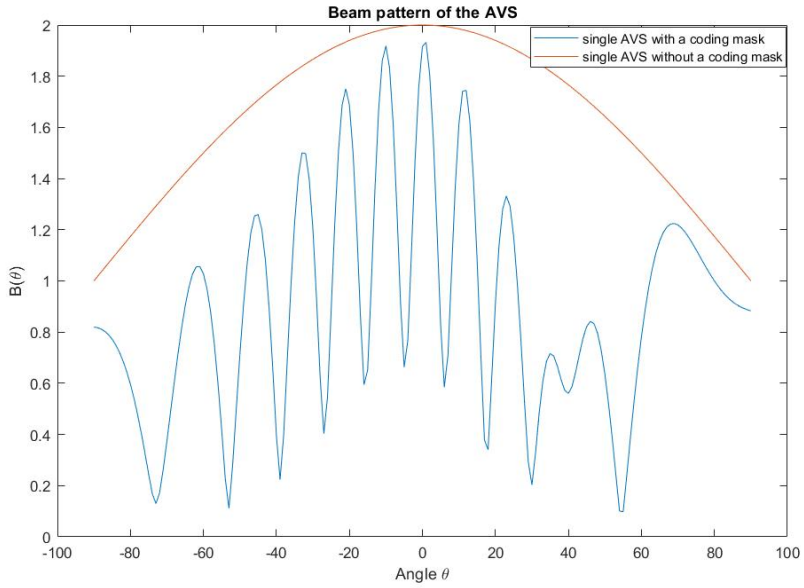



Figure 2.2: The beam pattern of a single AVS and a single AVS with the coded mask, with $\theta_0$ fixed at 0 degrees.

At its core, MUSIC relies on the idea of exploiting the orthogonality between the signal subspace (spanned by the signal vectors) and the noise subspace (spanned by the noise vectors). By doing so, it's capable of distinguishing signal sources from background noise, even in situations where the signals are weak compared to the noise.

The algorithm calculates a spatial spectrum, often called the MUSIC spectrum, that represents the likelihood of a signal originating from a particular direction. Peaks in this spectrum correspond to the angles at which signals are arriving, making it a powerful tool for DOA estimation [11].

As previously mentioned, the coded mask disrupts the shape of the beam pattern, rendering the traditional matched filter ineffective. However, the MUSIC algorithm relies on the orthogonality between the signal subspace and noise subspace, which remains valid in this scenario. Therefore, in this study, we employ the MUSIC algorithm to compare the performance of both the single AVS and the single AVS with the coded mask.

The covariance matrix $\mathbf{R}_{\mathrm{y}}$ of $\mathbf{y}(t)$ in (2.6) can be expressed as:

$$\mathbf{R}_{\mathrm{y}} = \mathbf{H}(\theta)\mathbf{R}_{\mathrm{s}}\mathbf{H}(\theta)^{\mathrm{H}} + \sigma_{\mathrm{n}}^2\mathbf{I}_3 = \mathbf{U}_{\mathrm{s}}\left(\Lambda_{\mathrm{s}} + \sigma_{\mathrm{s}}^2\mathbf{I}_{\mathrm{N}}\right)\mathbf{U}_{\mathrm{s}}^{\mathrm{H}} + \mathbf{U}_{\mathrm{n}}\left(\sigma_{\mathrm{n}}^2\mathbf{I}_{3-\mathrm{N}}\right)\mathbf{U}_{\mathrm{n}}^{\mathrm{H}} \tag{2.8}$$

In the context of the MUSIC algorithm, the relevant mathematical notations and variables are as follows: $N$ is the number of sources, $\sigma_s^2$ denotes the signal power, $\sigma_n^2$ denotes the noise power, $\mathbf{U}_s$ and $\mathbf{U}_n$ collect the orthogonal basis vectors associated with the signal and noise subspace, respectively.

The MUSIC algorithm evaluates the spatial spectrum at the angle $\theta$ using the following expression:

$$J_{\text{MUSIC}}(\theta) = \frac{|\mathbf{U}_n^{\text{H}}\mathbf{h}(\theta)|^2}{|\mathbf{h}(\theta)|^2} = \frac{\mathbf{h}(\theta)^{\text{H}}\mathbf{U}_n\mathbf{U}_n^{\text{H}}\mathbf{h}(\theta)}{\mathbf{h}(\theta)^{\text{H}}\mathbf{h}(\theta)} \tag{2.9}$$

In practice, the $N$ lowest local minima in this graph indicate the directions of the $N$ sound sources. When graphing this data, we often plot the inverse of $J_{\text{MUSIC}}(\theta)$ and search for the $N$ local maxima.

Figure 2.3 illustrates the performance comparison of the MUSIC algorithm between a single AVS with and without a coded mask. It can be observed from the figure that in the presence of only one source, a single AVS with and without a coded mask have nearly identical performance. This similarity arises because the coded mask is primarily considered as a scaling effect in this scenario.

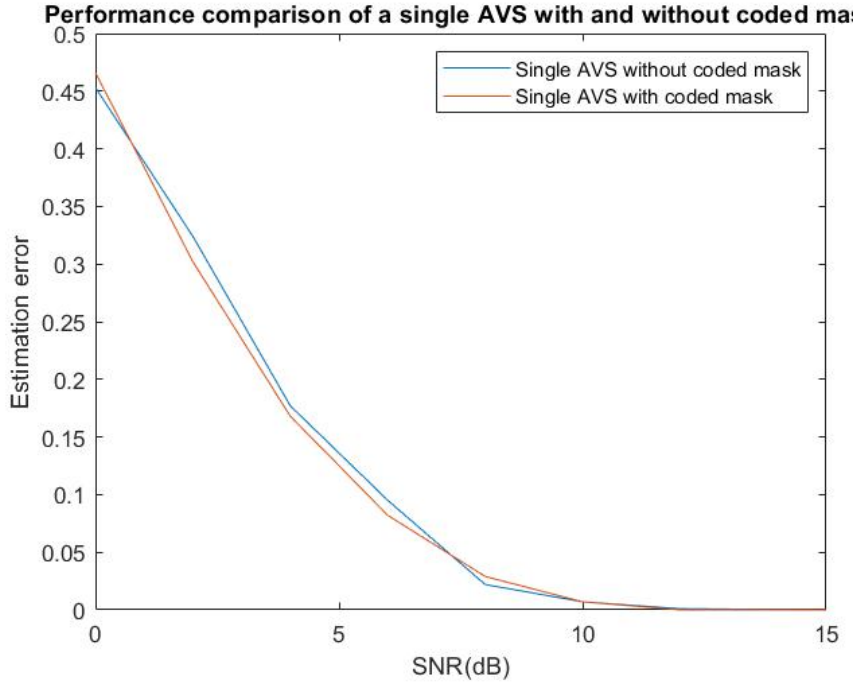



Figure 2.3: MUSIC performance comparison of a single AVS with and without a coded mask in the presence of a single source.

The above simulation results on DOA estimation indicates that the coded mask in ultrasound imaging does not offer a benefit in the application of AVS. This observation is further substantiated in the subsequent section through the verification using the Cramér-Rao Lower bound.

## 2.5 Cramér-Rao Lower Bound on DOA Estimation

The Cramér-Rao lower bound (CRLB) is a fundamental concept in statistical estimation theory. It serves as a lower limit on the variance of any unbiased estimator for a parameter. In the context of DOA estimation for acoustic sensors, the CRLB quanti-

fies the best achievable precision in estimating the DOA of a source signal given the available information and the inherent noise in the measurements.

The covariance matrix of a single source and a single AVS with coded mask in $\mathbb{R}^2$ with $\theta$ being the angle can be expressed as:

$$\mathbf{R}_{\mathrm{y}} = \mathbf{h}(\theta)\mathbf{h}^{\mathrm{H}}(\theta)\sigma_s + \sigma_n^2\mathbf{I}_3 \tag{2.10}$$

Based on [12] and Equation (2.10), with unknown parameters grouped as $\gamma = [\theta, \sigma_s, \sigma_n]^{\mathrm{T}}$, the CRLB on the DOA parameter of a single AVS with a coded mask and a single source can be expressed as:

$$\mathrm{CRB}_{\mathrm{mask}}(\theta) = \frac{\sigma^2}{2L}\frac{(|\mathbf{h}(\theta)|^2\sigma_s^2 + \sigma_n^2)}{|\mathbf{h}(\theta)|^2\sigma_s^4}\left\{\mathrm{Re}\left[\left(\mathbf{d}(\theta)^{\mathrm{H}}\Pi_c\mathbf{d}(\theta)\right)^{\mathrm{T}}\right]\right\}^{-1} \tag{2.11}$$

where $\Pi_c = \mathbf{I}_3 - \mathbf{h}(\theta)\mathbf{h}(\theta)^{\mathrm{H}}/|\mathbf{h}(\theta)|^2$ represents the projection matrix onto the noise subspace, and $\mathbf{d}(\theta) = \partial\mathbf{h}(\theta)/\partial\theta$ is the derivative of the array manifold vector $\mathbf{h}(\theta)$ with respect to $\theta$. The parameters $\sigma_n^2$ and $\sigma_s^2$ represent the noise power and the source power, respectively. The variable $L$ corresponds to the number of snapshots, which refers to the number of independent measurements used for DOA estimation.

Furthermore $\Pi_c = \mathbf{I}_3 - \mathbf{h}(\theta)\mathbf{h}(\theta)^{\mathrm{H}}/|\mathbf{h}(\theta)|^2$ and $\mathbf{d}(\theta)$ is given by:

$$\mathbf{d}(\theta) = \frac{\partial\mathbf{h}(\theta)}{\partial\theta} = \frac{\partial(\mathrm{k}(\theta)\mathbf{a}(\theta))}{\partial\theta} = \frac{\partial\mathrm{k}(\theta)}{\partial\theta}\mathbf{a}(\theta) + \frac{\partial\mathbf{a}(\theta)}{\partial\theta}\mathrm{k}(\theta), \tag{2.12}$$

and $\mathbf{d}(\theta)^{\mathrm{H}}\Pi_c\mathbf{d}(\theta)$ is calculated as:

$$\mathbf{d}(\theta)^{\mathrm{H}}\Pi_c\mathbf{d}(\theta) = \mathbf{d}(\theta)^{\mathrm{H}}\mathbf{d}(\theta) - \frac{\mathbf{d}(\theta)^{\mathrm{H}}\mathbf{h}(\theta)\mathbf{h}(\theta)^{\mathrm{H}}\mathbf{d}(\theta)}{|\mathbf{h}(\theta)|^2} = \mathbf{d}(\theta)^{\mathrm{H}}\mathbf{d}(\theta) - \frac{\mathbf{d}(\theta)^{\mathrm{H}}\mathbf{a}(\theta)\mathbf{a}(\theta)^{\mathrm{H}}\mathbf{d}(\theta)}{|\mathbf{a}(\theta)|^2} \tag{2.13}$$

Simplifying (2.13) further, we find that $\mathbf{d}(\theta)^{\mathrm{H}}\Pi_c\mathbf{d}(\theta)$ with a cover is $k(\theta)^2$ times greater than for a single AVS without a cover:

$$\mathbf{d}(\theta)^{\mathrm{H}}\Pi_c\mathbf{d}(\theta) = \frac{k(\theta)^2(\frac{\partial\mathbf{a}(\theta)}{\partial\theta})^{\mathrm{H}}\mathbf{a}(\theta)\mathbf{a}(\theta)^{\mathrm{H}}\frac{\partial\mathbf{a}(\theta)}{\partial\theta}}{|\mathbf{a}(\theta)|^2} - k(\theta)^2(\frac{\partial\mathbf{a}(\theta)}{\partial\theta})^{\mathrm{H}}\frac{\partial\mathbf{a}(\theta)}{\partial\theta} \tag{2.14}$$

Now we can derive the expression of the Cramér-Rao Bound (CRLB) on the DOA parameter of a single AVS with a coded mask and a single source:

$$\mathrm{CRB}_{\mathrm{mask}}(\theta) = \frac{\sigma^2}{2L\sigma_s^2}\left(1 + \frac{\sigma^2}{|\mathbf{a}(\theta)|^2\sigma_s^2}\right) = \mathrm{CRB}_{\mathrm{AVS}}(\theta) \tag{2.15}$$

where $\mathrm{CRB}_{\mathrm{AVS}}$ refers to the CRLB on the DOA parameter of a single AVS with a single source without a coded mask, which has been derived in [3]. It turns out that the CRLB on the DOA parameter for a single AVS and a single source remains unchanged when we apply the coded mask. The relationship between the CRLB and the SNR is illustrated in Figure 2.4, indicating that the CRLB decreases as the SNR increases, and a coded mask does not impact the CRLB of a single source.

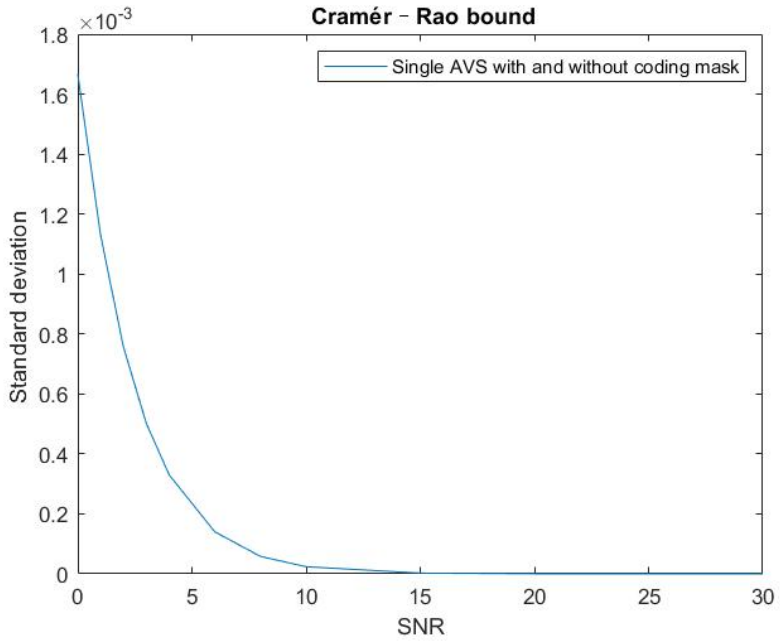



Figure 2.4: CRLB comparison of a single AVS and a single AVS with coded mask.

## 2.6 Conclusion

In this chapter, our primary objective is to explore the feasibility of adapting a coded mask originally designed for ultrasound imaging to enhance the performance of a single AVS for DOA estimation.

Our analysis of this new model includes an examination of the beam pattern, the results obtained through the MUSIC algorithm, and the CRLB. Surprisingly, our findings indicate that this new model with the coded mask does not offer any benefits for DOA estimation. This can be attributed to the fact that the mask functions primarily as a scalar in the measurement process.

To address this limitation, the next chapter will introduce an innovative design for the mask cover, which aims to overcome these challenges and potentially enhance the capabilities of the AVS system.

# Applying a Novel Coded Cover to AVS Sensor DOA Estimation

# 3

As discussed in the previous chapter, the application of the current coded mask model used in ultrasound imaging only results in a scaling effect when it comes to AVS sensor DOA estimation. It does not offer any significant benefits in terms of DOA estimation accuracy.

In this chapter, we present an innovative design for the AVS cover. This novel design is aimed at improving DOA estimation performance when using just one AVS sensor. Furthermore, by using the right processing, the upper limit of detectable signal sources by the AVS can be extended using the proposed cover for the AVS.

## 3.1 A Novel AVS Coded Cover

In this section, I will first explain the structure of this novel cover. Then based on the structure, I derive a physical propagation model of this cover. At last, the signal measurement model can be obtained.

### 3.1.1 Structure of the Novel Cover

Compared to the compact coded mask that is typically placed directly over the sensor, our coded cover has a relatively larger size, extending beyond the dimensions of the AVS. On this cover, we have integrated multiple channels of varying lengths to facilitate independent sound propagation. Our method entails the creation of these channels, each with a different length, ensuring that sound can pass through them independently while maintaining a consistent sound speed. This approach differs from the coded mask technique, where sound propagation is manipulated differently [10].

Another significant departure from the coded mask methodology lies in the orientation of our cover. We place the irregular side of the coded cover facing the AVS sensor, while the regular side faces the direction of the sound source. This arrangement allows us to utilize the channel entrances as a virtual ULA on top of the cover for compressive sensing purposes.

Additionally, we assume that at the outlet of each channel, the sound is re-emitted in the form of a point source to the location of the AVS sensor. This is different from the preview chapter, where we only consider the reception through a mask and not the transmission from the mask to the sensor. A physical propagation model of this cover can be seen in Figure. 3.1.

As depicted in Figure.3.1, there is a point source in the far-field, causing a plane wave to propagate from the source towards the upper side of the cover. As the wave reaches the cover, it passes through each channel independently. At the exit of each

channel, we assume the presence of another point source, and the sound propagates to the position of the AVS in the near field.

In this innovative cover scenario, the AVS sensor receives signals from each hypothetical point source, corresponding to the exit of each channel beneath the cover. Consequently, with this cover design, assuming the entrance above the cover forms a linear ULA, the cover essentially compresses the data from a virtual ULA into a three-channel 2D AVS, enabling a compressive analysis approach.

We assume there are a total of $M$ patches in the cover. We use $r_m$ to represent the distance from the source to the entry of the $m$-th patch. The length of the $m$-th patch is $d_m$, and the distance between the exit of the $m$-th patch and the position of the AVS is $L_m$. The angle of incidence at the $m$-th exit is $\phi_m$. The configuration of the coded mask is illustrated in Figure 3.2.

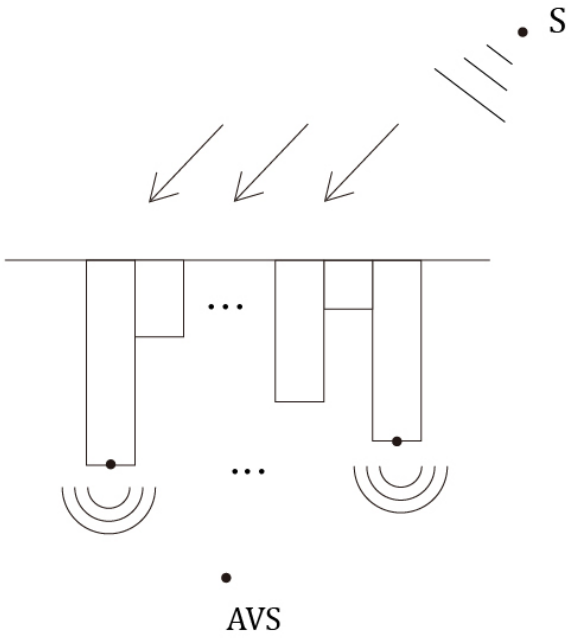



Figure 3.1: A physical model of the AVS: a single point source is placed in the far-field. Plane waves approach the patches and propagate independently. After passing through the channels, sound waves propagate in the near-field to the position of the AVS.

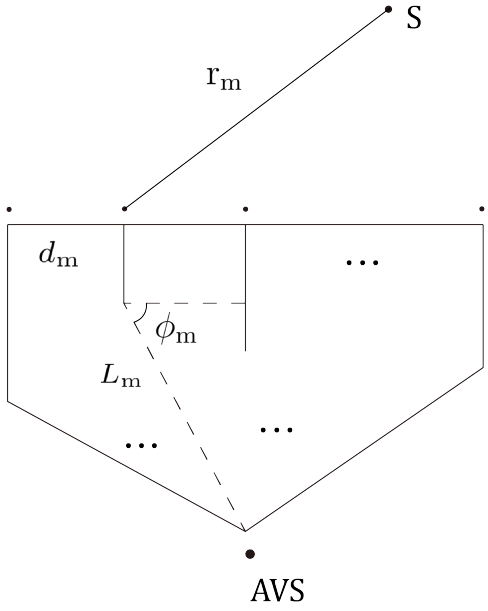



Figure 3.2: The AVS Configuration Model: outside the cover, $r_m$ represents the distance from the source to the entry of the $m$-th channel. Within the cover, there are $M$ patches, each characterized by its specific length $d_m$, its gap between the exit of the $m$-th patch and the AVS position $L_m$, and its incident angle at the $m$-th exit, denoted as $\theta_m$.

### 3.1.2 Physical propagation model

In Section 1.2.2, we covered the fundamental principles of sound propagation originating from a single point source. Now, we will proceed to formulate the propagation model while considering the presence of a coded cover.

During the initial phase of sound propagation, we will focus on far-field propagation and only consider the phase delay within each channel. Let's concentrate on a specific channel denoted by m. The pressure at the exit of the m-th patch can be expressed as:

$$\tilde{p}_m(t) = \frac{j\omega\rho Q}{4\pi r} e^{j[\omega t - k(r+d_m)]} \tag{3.1}$$

where $r$ is the distance from the source to the reference channel. As we assume a far-field source, the attenuation can be considered to be equal for all the channels.

Upon exiting each channel, we treat the exit of each channel as a distinct point source with a unique volume velocity $Q_m$, which is computed as follows:

$$Q_m = \frac{|\tilde{p}_m(t)|A_m}{\rho c} \tag{3.2}$$

$\mathrm{A_m}$ represents the cross-sectional area of the m-th patch.

By substituting the newly calculated $Q_m$ and the distance between the exist of the m-th patch and the AVS, $L_m$, into (1.2), we can derive the sound pressure at the AVS through the $m$-th channel as:

$$p_m(t) = \frac{kQ\omega\rho e^{j[\omega t - k(d_m + L_m)]}A_m}{(4\pi)^2 L_m r} \tag{3.3}$$

This equation reveals that the complex pressure received from the $m$-th channel encompasses both a geometric pressure distribution and a phase delay attributable to the total length of the channel.

Now, we can calculate the radial component of the particle velocity at the AVS through the $m$-th channel:

$$u_m(t) = \frac{p_m(t)}{\rho c} \tag{3.4}$$

Finally, we can decompose the particle velocity into its x and y components:

$$\begin{aligned} u_{x,m}(t) &= u_m(t)\cos(\theta_m) \\ u_{y,m}(t) &= u_m(t)\sin(\theta_m) \end{aligned} \tag{3.5}$$

### 3.1.3 Element Space Measurement Model

Based on the structure of the AVS cover and the physical model described in Sections 3.1.1 and 3.1.2, in this section, we derive the element space measurement model of a single AVS with this novel cover.

As shown in (3.3) and (3.5), the received pressure at the position of the AVS is composed of a geometric pressure distribution and a phase delay and the received particle velocity can be decompose into its x and y components. If a source signal $s(t)$ is transmitted, the overall received signal from the m-th patch at the position of the AVS will be denoted by:

$$\mathbf{y}_m(t) = \frac{\beta e^{-j2\pi f\left(\frac{L_m + d_m}{c}\right)} e^{jk\mathbf{r}_m\mathbf{u}(\theta)} \begin{bmatrix} 1 \\ \cos(\phi_m) \\ \sin(\phi_m) \end{bmatrix} s(t)}{L_m} \tag{3.6}$$

where $\mathbf{r}_m$ is the position of the entry of the $m$-th patch, $\beta$ is a normalization term related to the frequency and the volume velocity of the far-field source and can be considered as 1 if we implicitly adapt the noise variance accordingly. Besides, it is important to specifically emphasize the difference between $\theta$ and $\phi_m$. The first term

indicates the direction of the single source and the second term indicates the direction of the $m$-th virtual point source exiting the cover.

The received signal at the AVS is a summation of all the signals coming from M channels and can be expressed as:

$$\mathbf{y}(t) = \sum_{m=1}^{M} \mathbf{y}_m(t) = \sum_{m=1}^{M} \left( \frac{e^{-j2\pi f\left(\frac{L_m+d_m}{c}\right)} e^{jk\mathbf{r}_m\mathbf{u}(\theta)} \begin{bmatrix} 1 \\ \cos(\phi_m) \\ \sin(\phi_m) \end{bmatrix}}{L_m} \right) s(t) + \mathbf{n}(t) \tag{3.7}$$

The term $e^{jk\mathbf{r}_m\mathbf{u}(\theta)}$ is actually a phase delay caused by the spacing of the total M holes on the upper side of the cover. Since those five holes are equally spaced, we can assume them as a virtual ULA, and the array response of this virtual ULA can be expressed in a vector form:

$$\mathbf{h}(\theta) = \begin{bmatrix} e^{jk\mathbf{r}_1^\top\mathbf{u}(\theta)} \\ e^{jk\mathbf{r}_2^\top\mathbf{u}(\theta)} \\ \vdots \\ e^{jk\mathbf{r}_M^\top\mathbf{u}(\theta)} \end{bmatrix} = \begin{bmatrix} 1 \\ e^{j2\pi\Delta\sin(\theta)} \\ \vdots \\ e^{j2(M-1)\pi\Delta\sin(\theta)} \end{bmatrix} \tag{3.8}$$

where $\Delta$ is the spacing between the holes in wavelengths.

Besides, we can also define a compression matrix $\mathbf{\Phi}$, which contains the effect going from $M$ holes on the upper side of the cover to the three probes on the 2D AVS, which can be expressed as:

$$\mathbf{\Phi} = \begin{bmatrix} \frac{e^{-j2\pi f\frac{(L_1+d_1)}{c}}\mathbf{a}(\phi_1)}{L_1} & \frac{e^{-j2\pi f\frac{(L_2+d_2)}{c}}\mathbf{a}(\phi_2)}{L_2} & \cdots & \frac{e^{-j2\pi f\frac{(L_M+d_M)}{c}}\mathbf{a}(\phi_M)}{L_M} \end{bmatrix} \tag{3.9}$$

where $c$ is the sound velocity in the air and $\mathbf{a}(\phi_m) = \begin{bmatrix} 1 \\ \mathbf{u}(\phi_m) \end{bmatrix}$. Alternatively, (3.7) can be expressed in a more concise form:

$$\mathbf{y}(\mathrm{t}) = \mathbf{\Phi}\mathbf{h}(\theta)\mathrm{s}(\mathrm{t}) + \mathbf{n}(\mathrm{t}) \tag{3.10}$$

In this equation, $\mathbf{\Phi}$ is a $3 \times \mathrm{M}$ matrix that describes the compressive process of the cover. It is now evident that the cover functions as a compressive sampler, compressing the M receivers on the virtual ULA into three probe measurements on the AVS.

We can also define the response vector of the whole propagation as a $3 \times 1$ vector $\mathbf{b}$:

$$\mathbf{b}(\theta) = \mathbf{\Phi}\mathbf{h}(\theta) \tag{3.11}$$

Then the measurement at the AVS will become:

$$\mathbf{y}(t) = \mathbf{b}(\theta)s(t) + \mathbf{n}(t) \tag{3.12}$$

Now we extend the discussion to a scenario with $N$ sources located in the directions pointed by the unit vectors $\mathbf{u}_1\mathbf{u}_2 \dots \mathbf{u}_\mathrm{N}$. For this configuration, the measurement model of the AVS is:

$$\begin{aligned}\mathbf{y}(t) &= \begin{bmatrix} \mathbf{b}(\theta_1) & \mathbf{b}(\theta_2) & \dots & \mathbf{b}(\theta_N) \end{bmatrix} \begin{bmatrix} s_1(t) \\ s_2(t) \\ \cdot \\ \cdot \\ \cdot \\ s_N(t) \end{bmatrix} + \mathbf{n}(t) \\ &= \mathbf{B}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t)\end{aligned} \tag{3.13}$$

where $\boldsymbol{\theta} = \begin{bmatrix} \theta_1 & \theta_2 & \dots & \theta_\mathrm{N} \end{bmatrix}$ collects the angles of the N sources.

Until now, we have considered the element-space measurement model at a given time instance. Let's assume we have $L$ samples of data. Then, (3.13) can be modified as:

$$\begin{aligned}\mathbf{Y} &= \begin{bmatrix} \mathbf{b}(\theta_1) & \mathbf{b}(\theta_2) & \dots & \mathbf{b}(\theta_N) \end{bmatrix} \begin{bmatrix} s_1(1) & s_1(2) & \dots & s_1(L) \\ s_2(1) & s_2(2) & \dots & s_2(L) \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ s_N(1) & s_N(2) & \dots & s_N(L) \end{bmatrix} + \mathbf{N} \\ &= \mathbf{B}(\boldsymbol{\theta})\mathbf{S} + \mathbf{N}\end{aligned} \tag{3.14}$$

### 3.1.4 Covariance Measurement Model

In this section we extend the AVS array measurement model to the covariance domain. For this we assume that both $\mathbf{s}(t)$ and $\mathbf{n}(t)$ are independent and identically distributed (i.i.d.), zero-mean, circularly symmetric complex processes. Based on these assumptions, the covariance matrix, $\mathbf{R}_y$, of the measurement data described in (3.14) is evaluated as:

$$\mathbf{R}_y = E\left[\mathbf{y}(t)\mathbf{y}(t)^\mathrm{H}\right] = \mathbf{B}(\boldsymbol{\theta})\mathbf{R}_s\mathbf{B}(\boldsymbol{\theta})^\mathrm{H} + \mathbf{R}_n \tag{3.15}$$

where $\mathbf{R}_s$ and $\mathbf{R}_n$ are covariance matrices defined respectively by:

$$\mathbf{R}_s = E\left[\mathbf{s}(t)\mathbf{s}(t)^\mathrm{H}\right] = \sigma_s^2\mathbf{I}_M \tag{3.16}$$

$$\mathbf{R}_n = E\left[\mathbf{n}(t)\mathbf{n}(t)^\mathrm{H}\right] = \sigma_n^2\mathbf{I}_3 \tag{3.17}$$

Alternatively, we can also write $\mathbf{B}(\boldsymbol{\theta})$ as a multiplication of the compression matrix $\boldsymbol{\Phi}$ and the array response of the virtual ULA:

$$\mathbf{B}(\boldsymbol{\theta}) = \boldsymbol{\Phi}\mathbf{H}(\boldsymbol{\theta}) \tag{3.18}$$

with,

$$\mathbf{H}(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{h}(\theta_1) & \mathbf{h}(\theta_2) & \dots & \mathbf{h}(\theta_{\mathrm{N}}) \end{bmatrix} \tag{3.19}$$

In this case, the measurement covariance matrix in (3.15) can be written as:

$$\mathbf{R}_y = E\left[\mathbf{y}(t)\mathbf{y}(t)^{\mathrm{H}}\right] = \boldsymbol{\Phi}\mathbf{H}(\boldsymbol{\theta})\mathbf{R}_s\mathbf{H}(\boldsymbol{\theta})^{\mathrm{H}}\boldsymbol{\Phi}^{\mathrm{H}} + \mathbf{R}_n \tag{3.20}$$

Furthermore, we define $\mathbf{R}_{\mathrm{x}} = \mathbf{H}(\boldsymbol{\theta})\mathbf{R}_{\mathrm{s}}\mathbf{H}(\boldsymbol{\theta})^{\mathrm{H}}$ as the covariance measurement at the virtual ULA. The covariance measurement of the AVS can then also be expressed as:

$$\mathbf{R}_y = \boldsymbol{\Phi}\mathbf{R}_x\boldsymbol{\Phi}^{\mathrm{H}} \tag{3.21}$$

It is now evident that by introducing the AVS cover, we also compress the $M \times M$ covariance matrix of the virtual ULA into a $3 \times 3$ covariance matrix of the three probes of the AVS. Besides $\mathbf{R}_x$ has a Toeplitz structure which will be useful when discussing DOA estimation.

## 3.2 DOA Estimation Using the Compressive Sampling Method

In this research, we propose two DOA methods using the AVS cover. The first method is called the compressed-sampling-based method, which incorporates the compressed data at the AVS to perform DOA estimation. In this section, I will provide a brief introduction to the compressive sampling (CS) method and explain how to apply it to DOA estimation with a cover.

### 3.2.1 Compressive Sampling

Compressive sampling (CS) is a method for acquisition of sparse signals and reconstruction from compressed measurements [13]. CS theory asserts that one can recover certain signals and images from far fewer samples or measurements than traditional methods use. To make this possible, CS theory is founded on two fundamental principles: sparsity and incoherence, which enable the recovery of signals and images from significantly fewer samples or measurements compared to traditional methods [14] [15].

Our cover works as a compressive sampling in the spatial domain, which compresses or transforms the virtual ULA into an AVS of a much smaller number of sensors. However, instead of CS based reconstruction methods, we apply MUSIC estimation on the compressed data for the DOA task. This is because traditional CS reconstruction based methods usually involve iterative algorithms with high complexity [16] and related research has already shown that an angle-spectrum-based method such as MUSIC and MVDR perform reasonably well compared to a CS reconstruction approach [17] [18].

### 3.2.2 Compressed-sampling-based DOA Estimation

In this scenario, we utilize a covariance matrix derived from the compressed data. Because we have a perfect knowledge of the AVS cover's propagation model, we can apply various covariance-based DOA estimation methods, including the MUSIC algorithm

and the MVDR algorithm. For our purposes, we will focus on using the MUSIC algorithm for DOA estimation. It's worth noting that there are examples of applying the CS MVDR method in literature, as demonstrated in [19].

Based on the compressive measurement of (3.15), we perform an eigenvalue decomposition:

$$\mathbf{R}_y = \mathbf{B}(\boldsymbol{\theta})\mathbf{R}_s\mathbf{B}(\boldsymbol{\theta})^{\mathrm{H}} + \mathbf{R}_n = \mathbf{U}_s\left(\boldsymbol{\Lambda}_s + \sigma^2\mathbf{I}_N\right)\mathbf{U}_s^{\mathrm{H}} + \mathbf{U}_n\left(\sigma_n^2\mathbf{I}_{3-N}\right)\mathbf{U}_n^{\mathrm{H}} \tag{3.22}$$

Here, we assume that the source number is $N < 3$. The CS MUSIC spectrum can be expressed as follows:

$$J_{\mathrm{MUSIC}}(\theta) = \frac{\left|\mathbf{U}_n^{\mathrm{H}}\mathbf{b}(\theta)\right|^2}{|\mathbf{b}(\theta)|^2} = \frac{\mathbf{b}(\theta)^{\mathrm{H}}\mathbf{U}_n\mathbf{U}_n^{\mathrm{H}}\mathbf{b}(\theta)}{\mathbf{b}(\theta)^{\mathrm{H}}\mathbf{b}(\theta)} \tag{3.23}$$

Simulations were conducted to evaluate the performance of this method. Initially, a single source was placed at a 30-degree angle and additional white noise was introduced at the AVS. The virtual ULA was configured with 5 elements. A total of 1000 independent simulations were executed, and the mean square errors were compared. The results, showcasing the performance of the MUSIC algorithm with and without a cover for a single AVS, can be observed in Figure 3.3.

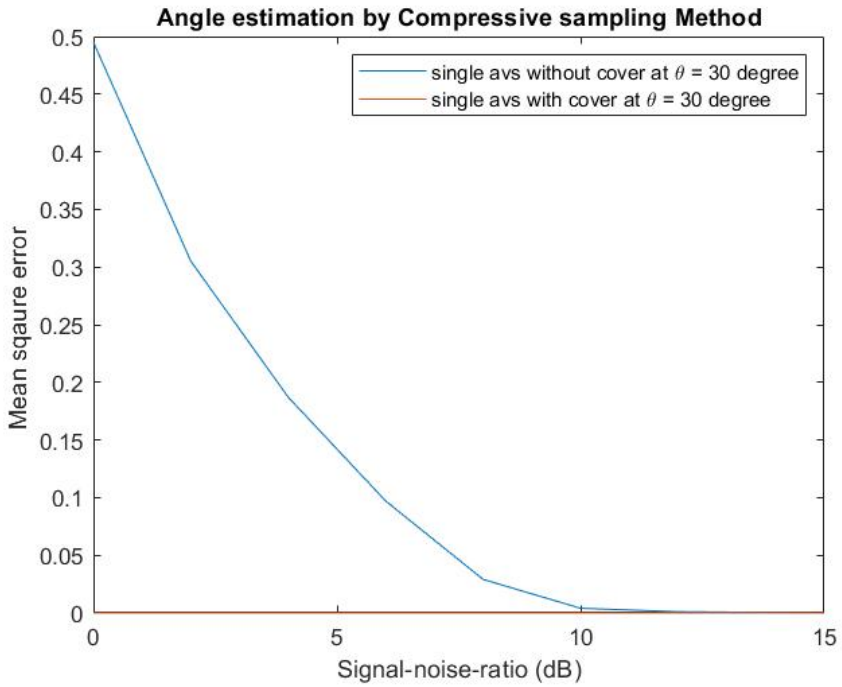



Figure 3.3: Simulation results for angle estimation with and without cover: a single source is placed at a 30-degree angle. The performance of the MUSIC algorithm is compared at different SNRs. The AVS with a cover demonstrates a significant improvement in angle estimation accuracy.

Figure 3.3 illustrates that a single AVS with the cover consistently outperforms the uncovered AVS across varying SNRs, with the mean square error consistently reaching

zero. This suggests that with the cover, we can achieve highly accurate DOA estimation even in very low SNR conditions. It's important to note that the SNR used here is a received SNR and can be defined as:

$$\text{SNR} = 10\log_{10}\left(\frac{E[|\mathbf{y}(t)|^2]}{E[|\mathbf{n}(t)|^2]}\right) \tag{3.24}$$

However, we still had concerns about whether this cover design could provide benefits across the entire range of angles of interest. To investigate this further, we conducted another simulation where the angle of the source was scanned through the entire range of angles, and the SNR was fixed at -5 dB. The results of this simulation are depicted in Figure 3.4. The simulation results shows that the AVS with cover has a better performance over the whole range of angles.

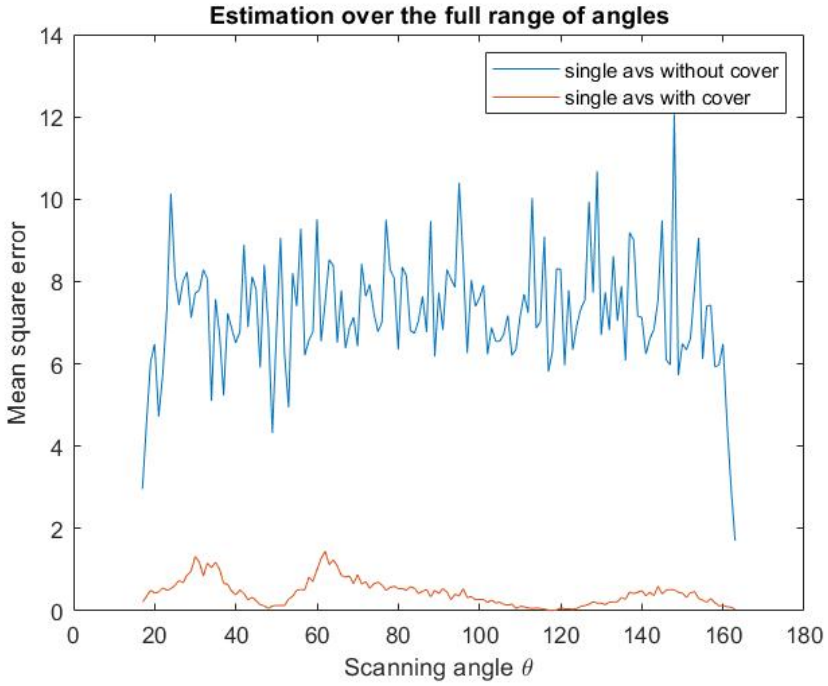
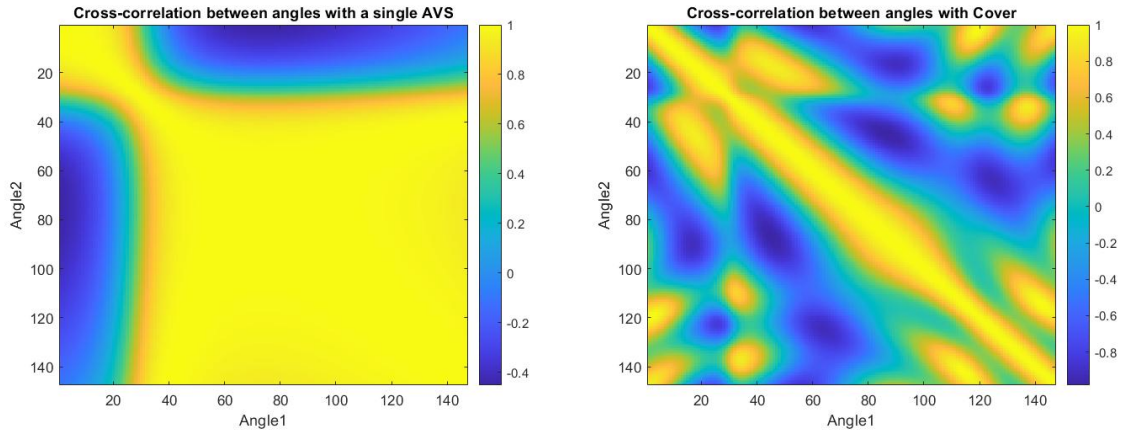



Figure 3.4: Simulation results for angle estimation across the entire range of angles: a single source is scanned across angles, and the SNR is set to -5 dB. The performance of the AVS with the cover outperforms a single AVS across the entire range of angles.

To understand why the AVS with a cover can achieve better performance, let's examine the cross-correlation between angles for the case of a single AVS with and without a cover. Figure 3.5 shows the cross-correlation for a single AVS with and without a cover.




(a) Cross-correlation between angles in the case of a single AVS.

(b) Cross-correlation between angles in the case of a single AVS with cover.

Figure 3.5: Comparison of cross-correlation between angles in the presence and absence of the cover: it is observed that the presence of a cover significantly reduce the cross-correlation between angles. In other words, the cover helps improve the spatial resolution.

As depicted in Figure 3.5, the cross-correlation between angles with the cover (Figure 3.5b) is notably lower than the cross-correlation between angles without the cover (Figure 3.5a). This signifies that the cover provides a significantly more distinct spatial response from different angles, resulting in a substantial improvement in spatial resolution and more accurate DOA estimation. This enhancement in spatial resolution is also evident in the MUSIC spectrum. Figure 3.6 illustrates the MUSIC spectrum at $\text{SNR} = -5$ with and without the cover. It is evident that the presence of an AVS cover leads to a much sharper peak in the MUSIC spectrum, indicating an improvement in spatial resolution.

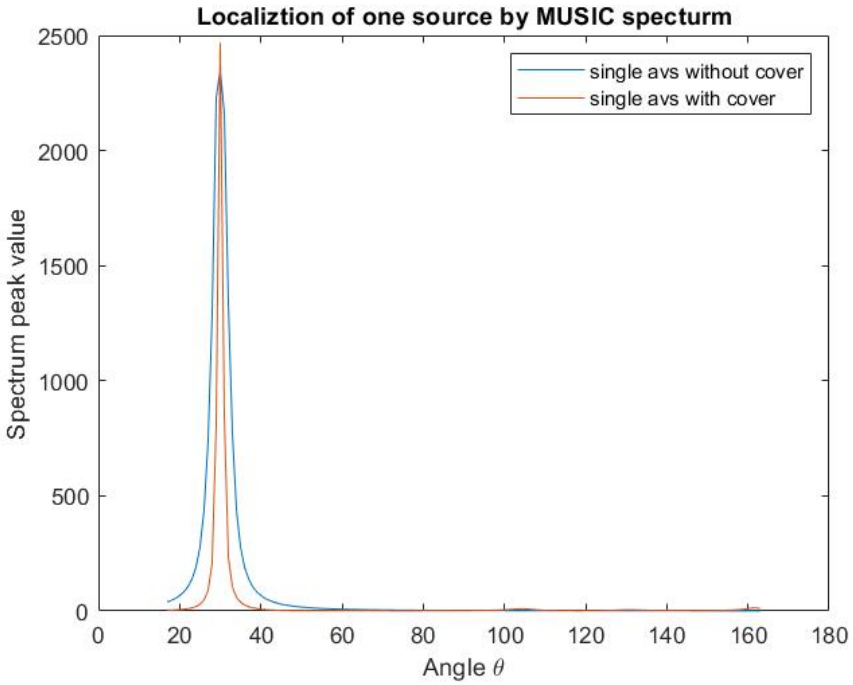



Figure 3.6: MUSIC Spectrum with a single source at 30-degree: the presence of an AVS cover results in a sharper peak, signifying an improvement in spatial resolution.

## 3.3 DOA Estimation Using the Compressive Covariance Sensing Method

In addition to directly utilizing the compressed data, second-order statistics such as covariance and power spectrum can also be effectively compressed and used for reconstruction. This approach is commonly referred to as the compressive covariance sensing (CCS) Method [20]. In the following section, we will provide an in-depth exploration of the CCS Method and demonstrate its application in the context of DOA estimation with an AVS cover

### 3.3.1 Compressive Covariance Sensing

The CCS method utilizes the compressed data and its corresponding steering vector for DOA estimation. In contrast, the CCS method is primarily concerned with the reconstruction of second-order statistics, including covariance and power spectrum. Unlike the CS Method, which necessitates a sparse original signal, CCS aims to recover the second-order statistics of a signal from the compressed signal without any sparsity constraints. However, CCS does not recover the original signal itself [21].

CCS operates in two key phases: 1. Compression Phase: in this stage, it subsamples an original signal, resulting in a set of compressed samples. 2. Recovery Phase: following compression, the CCS method employs a recovery algorithm to reconstruct

the covariance of the original signal from the compressed data [22].

The compression process can be mathematically modeled as an under-determined system. This under-determined problem is subsequently transformed into an over-determined problem by assuming specific structures for the original covariance matrix, such as Toeplitz or circulant matrices [23]. With this transformed over-determined model, the covariance matrix can be accurately reconstructed using a least squares (LS) method. Because many signal processing techniques are based on second-order statistics, CCS has been used in many applications such as power spectrum estimation, wideband cognitive radio, and DOA estimation [22].

### 3.3.2 Virtual ULA Covariance Estimation and Compression Limits Study

As discussed in the last section, the specific structure of the covariance matrix will decide how much compression we can achieve. Therefore, in the case of (3.21), we need to first determine the structure of the covariance matrix $\mathbf{R}_x$ of the virtual array.

Since $\mathbf{R}_x$ is a positive semi-definite HT matrix, it's reasonable to assume that it belongs to a specific region defined by positive semi-definite matrices and the subspace spanned by a group of HT matrices. We refer to one of these subspaces as the standard Toeplitz basis, represented as:

$$\mathcal{S}_T = \{\mathbf{\Sigma}_0, \mathbf{\Sigma}_1, \ldots, \mathbf{\Sigma}_{2M-2}\} = \{\mathbf{I}_M\} \cup \{\mathbf{T}_1, \ldots, \mathbf{T}_{M-1}\} \cup \left\{\tilde{\mathbf{T}}_1, \ldots, \tilde{\mathbf{T}}_{M-1}\right\} \tag{3.25}$$

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 $-J$. This subspace essentially captures our prior knowledge, and the smaller its size, the greater the potential for achieving compression [24].

With this covariance subspace, we can decompose the covariance matrix $\mathbf{R}_x$ into a linear combination:

$$\mathbf{R}_x = \sum_{m=0}^{2M-2} \alpha_m \mathbf{\Sigma}_m = \alpha_0 \mathbf{I}_M + \sum_{m=1}^{M-1} (\alpha_m \mathbf{T}_m + \alpha_{m+M-1} \tilde{\mathbf{T}}_{m+M-1}) \tag{3.26}$$

Since $\mathcal{S}_T$ is a linearly independent set of matrices, then the following conduction should hold:

$$\sum_{m=0}^{2M-2} \alpha_m \mathbf{\Sigma}_m = \sum_{m=0}^{2M-2} \beta_m \mathbf{\Sigma}_m \Rightarrow \alpha_m = \beta_m \quad \forall m \tag{3.27}$$

In this case, $\mathcal{S}_T$ should be a basis for the covariance subspace and a covariance matrix can be uniquely decomposed by the basis of the subspace. Consequently, knowing $\alpha_\mathrm{s}$ is equal to knowing the covariance matrix $\mathbf{R}_\mathrm{x}$.

The covariance matrix of the AVS measurements is related with the covariance matrix of the virtual ULA by (3.21). As a result, we can also decompose $\mathbf{R}_\mathrm{y}$ as:

$$\mathbf{R}_\mathrm{y} = \mathbf{\Phi} \mathbf{R}_\mathrm{x} \mathbf{\Phi}^\mathrm{H} = \sum_{m=0}^{2M-2} \alpha_m \bar{\mathbf{\Sigma}}_m, \quad \text{where } \bar{\mathbf{\Sigma}}_m = \mathbf{\Phi} \mathbf{\Sigma}_m \mathbf{\Phi}^\mathrm{H} \tag{3.28}$$

$\bar{\mathbf{\Sigma}}_m$ belongs to a set of matrices:

$$\overline{\mathcal{S}} = \left\{ \bar{\mathbf{\Sigma}}_0, \bar{\mathbf{\Sigma}}_1, \ldots, \bar{\mathbf{\Sigma}}_{2M-2} \right\} \subset \mathbb{C}^{3\times 3} \tag{3.29}$$

The set $\overline{\mathcal{S}}$ is an extension of the covariance basis $\mathcal{S}$. While the matrices in $\overline{\mathcal{S}}$ are Hermitian, they do not have a general Toeplitz structure. If the compression operation retains all relevant information, then $\overline{\mathcal{S}}$ remains linearly independent. Knowing $\overline{\mathcal{S}}$ is essentially equivalent to knowing the $\alpha_s$ values, which, in turn, equates to knowing $\alpha_s$. On the contrary, if the compression is too severe and causes a loss of linear independence, then certain second-order information regarding $\mathbf{R}_x$ cannot be retrieved.

Therefore, in practice, we would like to use the sampler, or the compression matrix $\mathbf{\Phi}$ that is able to preserve all the second-order statistical information of $\mathbf{x}$ and this kind of samplers are called a $\mathcal{S}$-covariance sampler.

The connection between the covariance basis $\mathcal{S}$ and the expanded basis $\overline{\mathcal{S}}$ is described in (3.30), while the link between the covariances is demonstrated in (3.31). The function $\phi$ plays a crucial role in this process, as it transforms the covariance measurements from the virtual ULA into the corresponding covariance measurements obtained from the AVS probes.

$$\text{Span}(\mathcal{S}_T) \xrightarrow{\phi} \text{Span}(\overline{\mathcal{S}}) \tag{3.30}$$

$$\mathbf{R}_y \xrightarrow{\phi} \phi(\mathbf{R}_x) = \mathbf{\Phi}\mathbf{\Sigma}\mathbf{\Phi}^{\text{H}} \tag{3.31}$$

As outlined in [24], a matrix $\mathbf{\Phi}$ defines an $\mathcal{S}$-covariance sampler if the associated function $\phi$ is invertible. Considering that the covariance matrix $\mathbf{R}_x$ is an $M \times M$ Hermitian Toeplitz (HT) matrix, when we refer to (3.21), there will be a total of $2M-1$ unknowns on the right side. Conversely, $\mathbf{R}_y$ is a Hermitian matrix, which implies there are 9 known values on the right side. To maintain the invertibility of the function $\phi$, the maximum value for $M$ should be constrained to 5, which means that we can at most estimate 4 sources with the cocver. Under this constraint, the compression ratio, denoted as $\rho$, can be defined as follows:

$$\rho = \frac{3}{M} \tag{3.32}$$

Once known that $\mathbf{\Phi}$ is a $\mathcal{S}$-covariance sampler, we can estimate the covariance matrix $\mathbf{R}_x$ by estimating the coefficients $\alpha_s$ first. The following cost function will be used to estimate the $\alpha_s$'s:

$$\hat{\boldsymbol{\alpha}} = \min_{\boldsymbol{\alpha}} \left\| \mathbf{R}_y - \sum_{m=0}^{M-1} \mathbf{\Phi}\alpha_m \mathbf{\Sigma}_m \mathbf{\Phi}^{\text{H}} \right\|_2^2 \tag{3.33}$$

$\boldsymbol{\alpha} = [\alpha_1 \quad \alpha_2 \ldots \alpha_M]^{\text{T}}$. To minimize this cost function, several algorithms can be employed, including methods such as least squares (LS) estimation, convex optimization, and maximum likelihood (ML) estimation, as described in [23] and [25]. After estimating $\alpha_s$'s, $\hat{\mathbf{R}}_x$ can be obtained using (3.26).

### 3.3.3 Compressive-Covariance-based DOA Estimation

In CCS-based DOA Estimation, our first task is to estimate the semi-definite Hermitian Toeplitz (HT) covariance matrix $\mathbf{R}_x$. To gauge the accuracy of our estimated covariance matrix $\hat{\mathbf{R}}_x$ in comparison to the true covariance matrix $\mathbf{R}_x$ of the virtual ULA, we employ the normalized correlation coefficient:

$$\rho_c = \mathrm{vec}(\mathbf{R}_x)^{\mathrm{H}} \mathrm{vec}(\hat{\mathbf{R}}_x) / \left( \|\mathbf{R}_x\|_{\mathrm{F}} \|\hat{\mathbf{R}}_x\|_{\mathrm{F}} \right) \tag{3.34}$$

The normalized correlation coefficient $\rho_c$ quantifies the proximity of the estimated covariance matrix to the true covariance. A value closer to 1 indicates a more accurate estimation.

Once the covariance matrix $\mathbf{R}_x$ is estimated, we can apply various covariance-based DOA methods for angle estimation. One such method is the MUSIC algorithm. This involves performing an eigenvalue decomposition on the estimated covariance matrix:

$$\mathbf{R}_x = \mathbf{H}(\boldsymbol{\theta}) \mathbf{R}_s \mathbf{H}(\boldsymbol{\theta})^{\mathrm{H}} + \mathbf{R}_n = \mathbf{U}_s (\boldsymbol{\Lambda}_s + \sigma^2 \mathbf{I}_N) \mathbf{U}_s^{\mathrm{H}} + \mathbf{U}_n (\sigma_n^2 \mathbf{I}_{M-N}) \mathbf{U}_n^{\mathrm{H}} \tag{3.35}$$

Here, $N$ represents the number of detectable sources ($N < M$). The MUSIC spectrum is given by:

$$J_{\mathrm{MUSIC}}(\theta) = \frac{\left| \mathbf{U}_n^{\mathrm{H}} \mathbf{h}(\theta) \right|^2}{|\mathbf{h}(\theta)|^2} = \frac{\mathbf{h}(\theta)^{\mathrm{H}} \mathbf{U}_n \mathbf{U}_n^{\mathrm{H}} \mathbf{h}(\theta)}{\mathbf{h}(\theta)^{\mathrm{H}} \mathbf{h}(\theta)} \tag{3.36}$$

In our simulation, our initial goal is to assess the influence of the CCS-based method in a scenario with a single source. We introduce a single Gaussian noise source positioned at a 30-degree angle. The virtual ULA comprises 5 elements, and we perform 1000 independent experiments. In the phase of estimating $\mathbf{R}_x$, we analyze the connection between the mean normalized correlation coefficient and the SNR, as depicted in Figure 3.7.

In the figure, it's evident that as the SNR improves, the mean normalized correlation coefficient approaches 1. This trend signifies that with higher SNR, the estimation of $\mathbf{R}_x$ becomes more accurate.

Furthermore, we want to compare the results of the CCS-based method and a single AVS. Figure 3.8 presents the results. When using the CCS-based method, the performance is consistently better at all SNRs compared to a single AVS. However, it's important to note that the exceptional performance observed in the compressive sampling method, as shown in Figure 3.3, is not replicated here. This discrepancy can be attributed to the fact that a substantial number of samples are required for the covariance matrix $\mathbf{R}_x$ to exhibit Toeplitz-like characteristics.

The comparison of the MUSIC spectrum between a single AVS and the CCS-based method is shown in Figure 3.9. Notably, the spectrum peak with the CCS-based method is significantly sharper than with a single AVS without cover. This enhancement in spatial resolution is achieved because the method is equivalent to utilizing an array with $M$ elements for DOA estimation, thereby improving the spatial resolution.

Figure 3.10 illustrates the cross-correlation between angles for a single AVS and a $M$-element ULA. This comparison further demonstrates how the CCS-based method

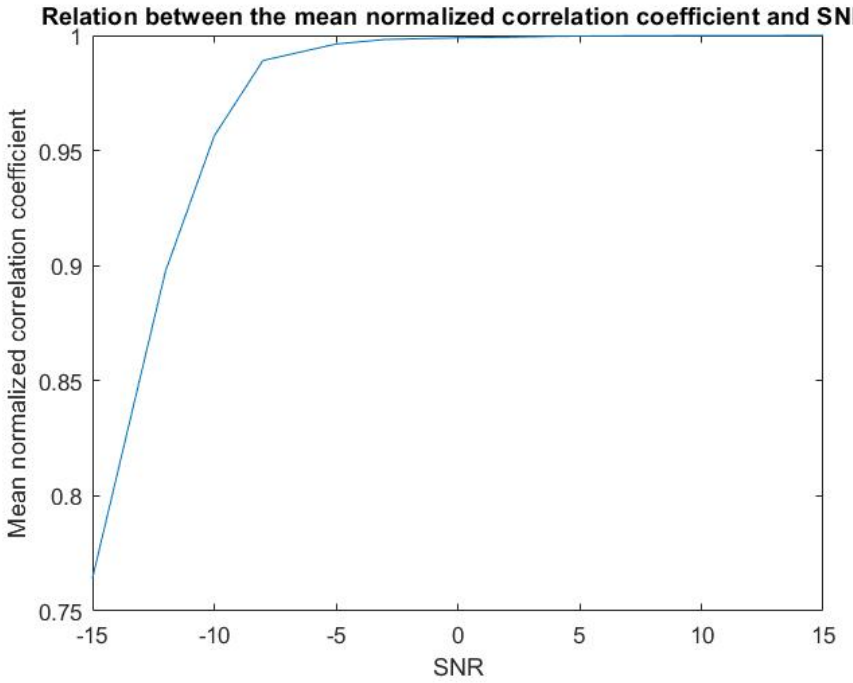



Figure 3.7: Relationship between the mean normalized correlation coefficient and SNR: as SNR increases, the mean normalized correlation coefficient approaches 1, indicating more accurate $\mathbf{R}_x$ estimation.

emulates the capabilities of a larger ULA, resulting in a more distinct spatial response and improved spatial resolution.

In addition to the improved spatial resolution, the CCS-based method represents a significant breakthrough. With the recovery of a larger dimension covariance matrix $\mathbf{R}_x$, it becomes possible to estimate more sources than with a single AVS. The covariance matrix of a single AVS is limited to dimensions of $3 \times 3$, allowing the detection of at most 2 sources. However, with a ULA consisting of $M = 5$ elements, we can now estimate up to 4 sources, which aligns with the compression ratio of $1/2$ in this situation. Figures 3.11 and 3.12 illustrate a scenario with 4 sources, where the CCS-based method successfully estimates all of them, while a single AVS without cover would fail in this situation.

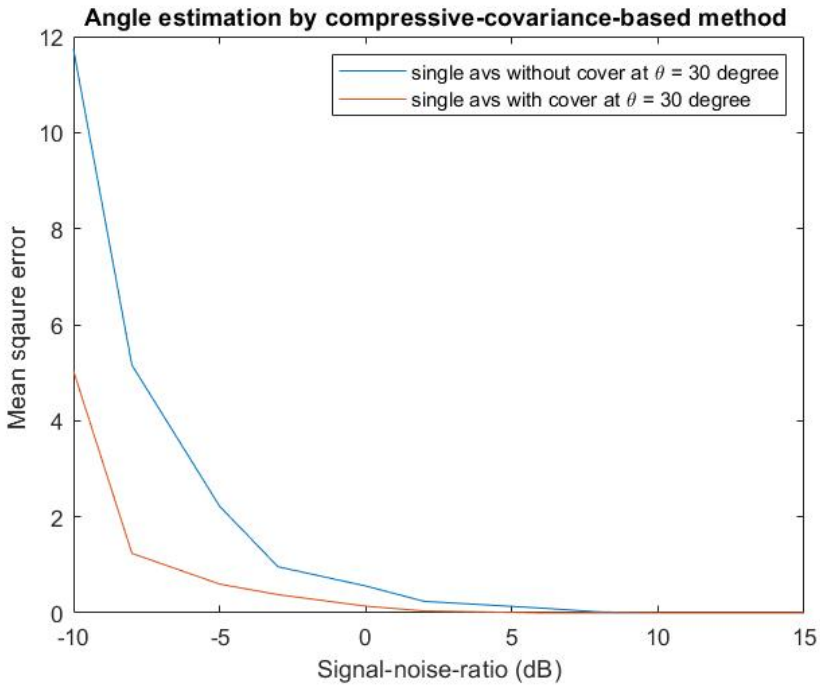



Figure 3.8: Simulation results for angle estimation with and without cover: a single source is placed at a 30-degree angle. The performance of the MUSIC algorithm is compared at different SNRs. The CCS-based MUSIC consistently outperforms the non-compressive method across all SNRs, although it doesn't reach the same level of performance as the compressive sampling method shown in Figure 3.3.

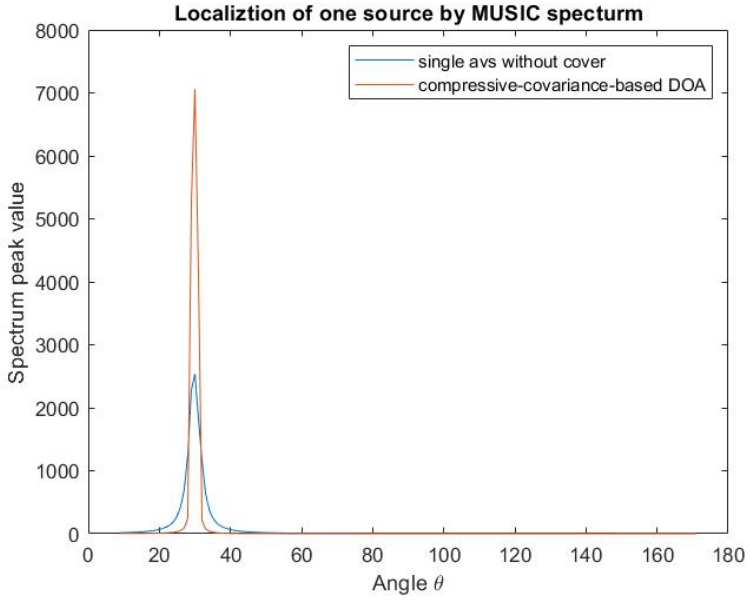
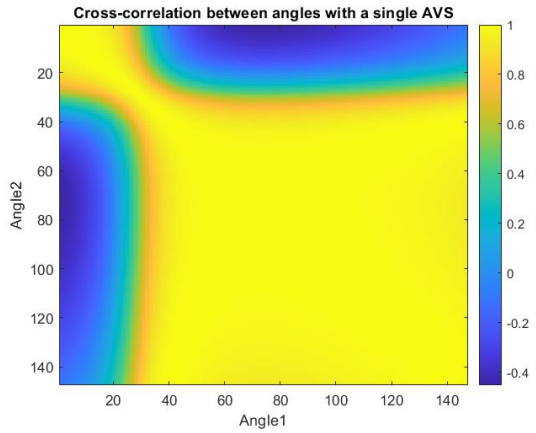
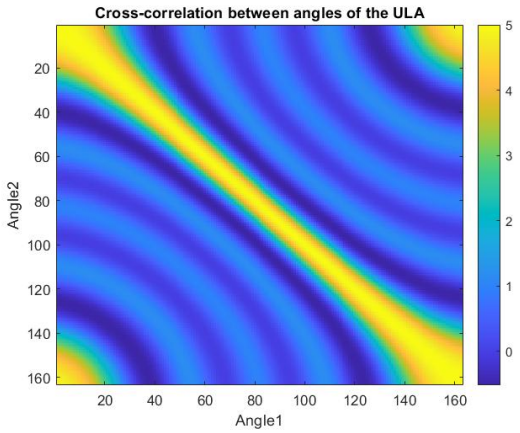



Figure 3.9: MUSIC Spectrum with a Single Source at 30 Degrees: the spectrum peak obtained with the compressive-covariance-based method is significantly sharper than that of a single AVS, leading to improved spatial resolution.




(a) Cross-correlation between angles of a single AVS.




(b) Cross-correlation between angles of a virtual ULA.

Figure 3.10: Comparison of cross-correlation between angles: this figure compares the cross-correlation between angles for a single AVS and an $M$-element ULA. It illustrates how the CCS-based method approximates the capabilities of a larger ULA, resulting in a more distinctive spatial response and improved spatial resolution.

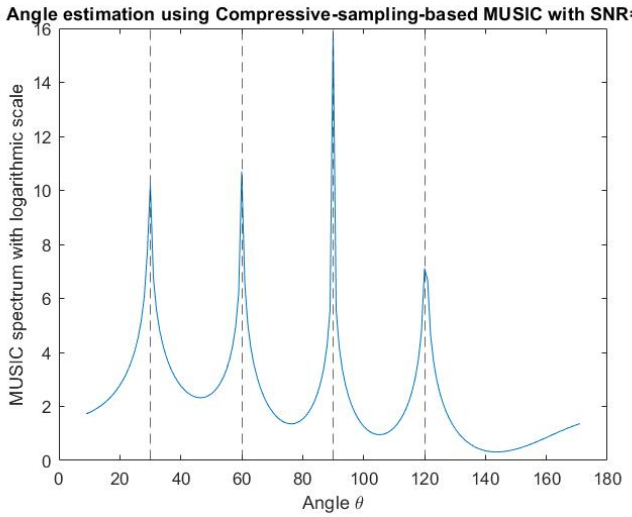



Figure 3.11: Estimation of 4 sources using CCS: the four sources are located at 30-degree, 60-degree, 90-degree, and 120-degree. The CCS method successfully estimates all 4 sources, effectively doubling the capacity compared to a single AVS.

## 3.4 Conclusion

In this chapter, we have introduced a novel structure for the AVS cover and have analyzed it from the perspective of sound propagation physics and the measurement model. Unlike the coded masks used in ultrasound imaging, this innovative cover not

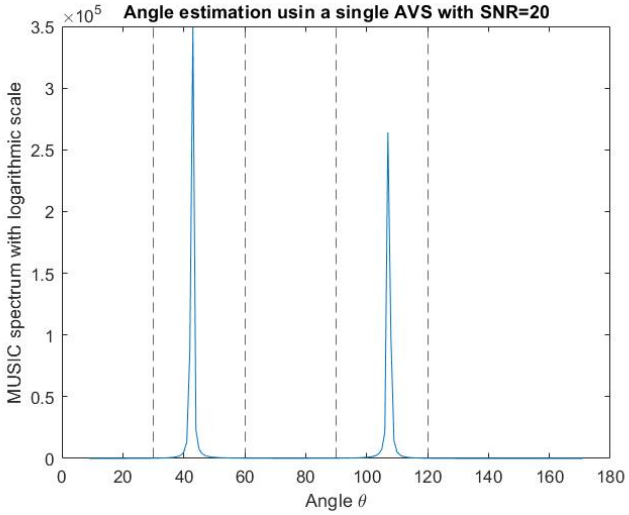



Figure 3.12: A single AVS is not able to estimate 4 sources.

only considers the reception through the mask but also accounts for the propagation from the cover to the AVS. Consequently, a transfer function within the cover needs to be modeled. This transfer function can be viewed as a compression matrix, enabling a compressive analysis.

Another key distinction is the orientation of the cover. We position the irregular side of the cover towards the AVS and leave the flat side facing outward. This arrangement allows us to assume that all entries of the patches form a virtual ULA because they are equally spaced. This configuration transforms the signal reception into a WSS process at the virtual array. We can approximate the covariance matrix of the virtual array as an HT matrix. This requires fewer entries to be known to recover the entire covariance matrix. Therefore, when the covariance at the AVS is perfectly known, we can recover the complete covariance matrix of the virtual ULA under compressive limits.

Two distinct DOA methods based on this structure are proposed. The first is a CS-based DOA method, where we utilize the entire steering vector of the system for DOA estimation. Throughout the chapter, we employ the MUSIC algorithm. Notably, this method exhibits significantly improved performance due to reduced cross-correlation between different angles, facilitated by the novel cover. However at most two sources can be recovered.

The second method is CCS-based DOA, where we initially estimate the covariance matrix of the virtual ULA and then apply it to covariance-based DOA estimation. With this approach, we not only achieve enhanced spatial resolution but also expand the range of sources that can be estimated using a single sensor. This is because we recover a covariance matrix with a larger dimension. In this case, at most 5 sources can be recovered.

# 4 Prototype Development

In this research, we didn't just propose a theory and conduct simulations. We actually manufactured a coded cover based on the theory presented in Chapter 3, and we conducted various experiments to validate if the theory and simulations align with real-world results.

In this chapter, we will first introduce the process of creating a physical coded cover. To prevent sound from bypassing the coded cover and reaching the sensor, we also designed a protective enclosure, colloquially referred to as a "box", to house the sensor. Using this coded cover and the protective enclosure, we conducted experiments to assess the ability of a single AVS with a coded mask to localize both single and multiple sources.

## 4.1 Crafting the Coded Cover and Protective Enclosure

In reality, we use a material called the plastic glass to create the coded cover and the box. The use of plastic glass (often referred to as acrylic or Plexiglas) to create coded covers and boxes with specific sound propagation characteristics is a practical application of acoustic engineering. This approach takes advantage of the properties of plastic glass to control and manipulate the behavior of sound waves.

Plastic glass is favored for its relatively high density, which lends itself the capacity to effectively obstruct the passage of sound waves. This property becomes especially advantageous when the objective is to construct an enclosure or cover that isolates sound from the surrounding environment.

Within the plastic glass, channels can be intricately designed to serve as conduits for guiding and directing sound waves along specific trajectories. These channels can assume diverse shapes and configurations, allowing us to attain the desired acoustic outcomes. By meticulously controlling the geometry and dimensions of these channels, we exercise influence over the manner in which sound propagates within the material.

In practical terms, we crafted a coded cover using a 10mm thick sheet of plastic glass. This coded cover was intentionally designed with 5 evenly spaced holes. Over these holes, we securely attached channels constructed from plastic glass, each possessing unique lengths.

To safeguard against sound bypassing the coded cover and reaching the AVS, we engineered a box in which the coded cover served as one side, providing a designated space for the placement of the AVS. Consequently, sound waves could only propagate to the AVS through the five channels situated on the side of the coded cover. However, there remained the possibility of sound reflecting within the box once it entered. To mitigate this, we incorporated sound-absorbing materials on the internal surfaces of the box, excluding the side with the coded cover. This measure was adopted to prevent

sound from undergoing unwanted reflections within the box.

A model of the experimental box is presented in Figure 4.1. One side of this box corresponds to the AVS cover, with the irregular side of the AVS cover facing inward, as illustrated in Figure 3.1. At the bottom of the box, there are two openings. The larger opening serves as the entry point for the AVS sensor, where the AVS is securely affixed to a tripod. The smaller opening is designed for attaching the box to a tripod for experimental convenience.

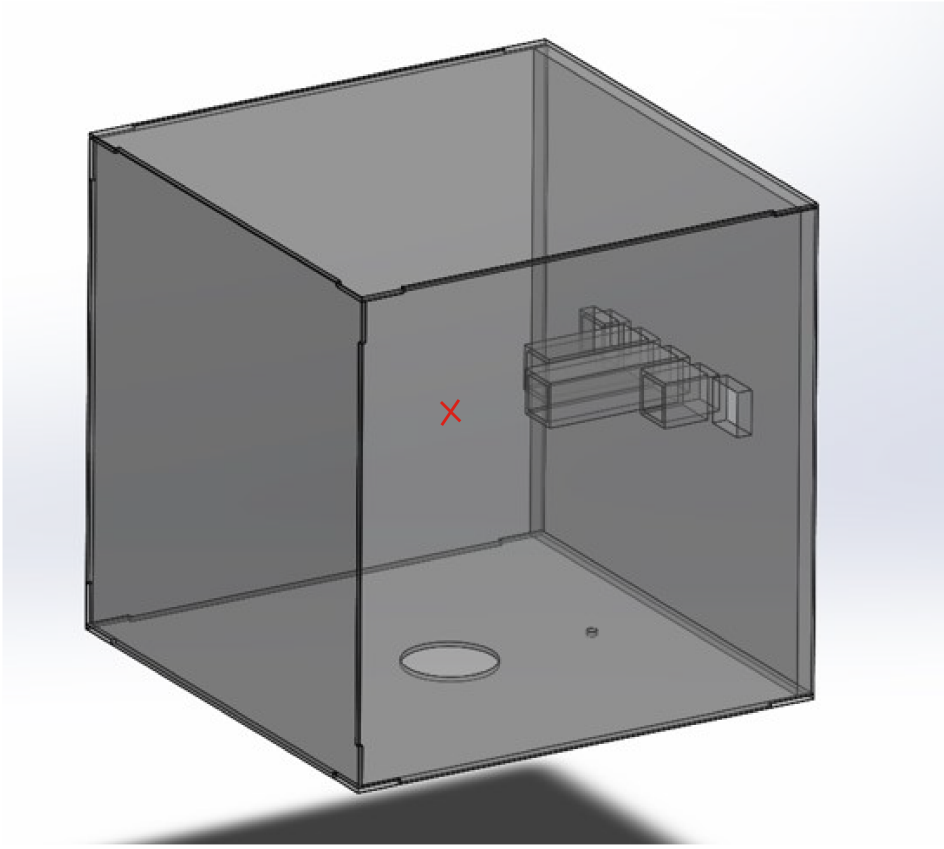


Figure 4.1: The experimental box model illustrates the fundamental concept of our experimental setup. One side of the box, featuring 5 channels, is designated as the coded cover. The AVS is inserted through a large opening at the bottom of the box, as indicated by the red cross. This placement ensures that the AVS is in the same plane as the patches, enabling the collection of sound signals, as described in our experiments.

## 4.2 Experimental Setup

All the experiments are conducted in an acoustic room, where all the walls are covered with sound absorption materials to create a highly non-reflective environment. The AVS sensor and the box are supported by tripods, and the tripod's height is meticulously adjusted to align the AVS at the same height as the five channels in the coded cover. Figures 4.2 and 4.3 shows the experimental setup from a distant view and a close-up view, respectively.

Figure 4.2: Experimental setup from a distant view

The compression matrix, as denoted in (3.9), can be derived through measurements of the transfer function from each patch's entry point to the AVS sensor's location. Furthermore, the virtual ULA's array response can be computed by considering the spacing between the holes on the cover. Consequently, we did not measure the transfer function for every individual angle. Our measurements were primarily focused on obtaining the compression matrix. This matrix was subsequently used in conjunction with the virtual ULA's array response to calculate the estimated manifold for the entire propagation process, as explained in (3.12).

To conduct our experiments, we positioned speakers at various angles on the same plane as the AVS and coded cover, with their positions pre-determined. Uncorrelated white noise was played through each speaker. Subsequently, we collected signals at the location of the AVS sensor, enabling us to apply the two methods outlined in Chapter 3 for DOA estimation.

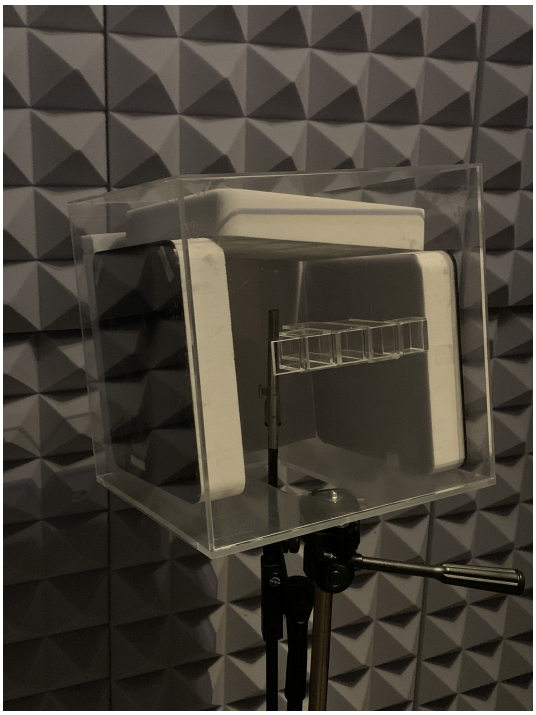


Figure 4.3: Experimental setup from a close-up view

## 4.3 Measurement Results

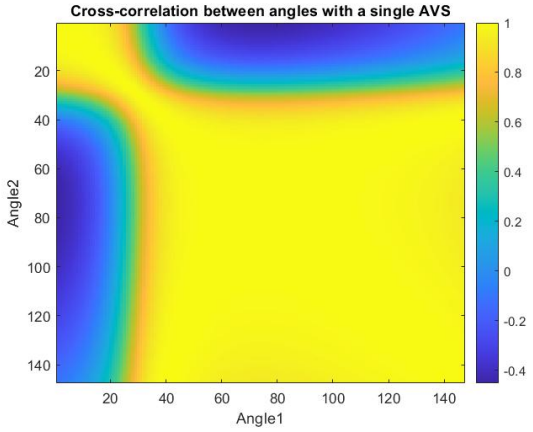
In this section, we delve into the measurement results of our experiments. Specifically, we compare the CS-based method and CCS-based method with the outcomes obtained using a single AVS. To facilitate our analysis, all signals received by the AVS sensor are subjected to a short time Fourier transform (STFT) to transform them into the frequency domain. The primary objective of this chapter is to validate whether our theoretical studies align with real-world measurements. Consequently, we restrict our DOA estimation to a single frequency bin. It's worth noting that a more robust performance could be attained through the utilization of coherent signal-subspace processing methods [26].

### 4.3.1 Single Source DOA

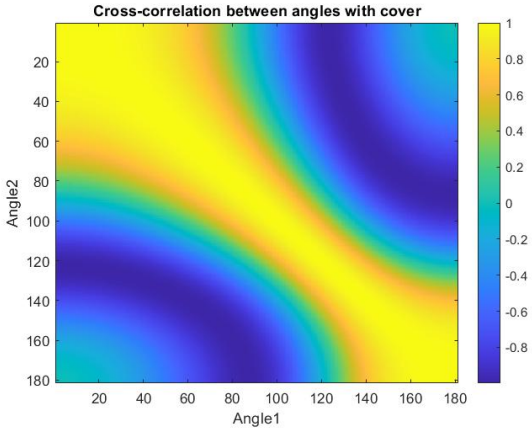
In this section, we present the results of an experiment involving a single source placed at a far-field distance. We will compare the performance of a single AVS without a cover with the results obtained from the CS-based method and CCS-based method.

In Figure 3.5, we observed a decrease in the cross-correlation between angles when the coded mask is applied. In practical terms, we can obtain the compression matrix $\mathbf{\Phi}$ by measuring the transfer function of each patch and calculating the response vector of the entire propagation using (3.12). Simultaneously, we can compare the cross-correlation of the response vector of a single AVS with and without a coded mask. The results are shown in Figure 4.4. Similarly, the AVS with a cover exhibits lower cross-correlation between angles, indicating that the cover contributes to enhanced spatial

resolution.




(a) Cross-correlation between angles of a single AVS without cover.




(b) Cross-correlation between angles of a single AVS with cover

Figure 4.4: Comparison of cross-correlation between angles in the presence and absence of the cover: it is observed that the presence of a cover significantly reduces the cross-correlation between angles. In other words, the cover helps improve the spatial resolution.

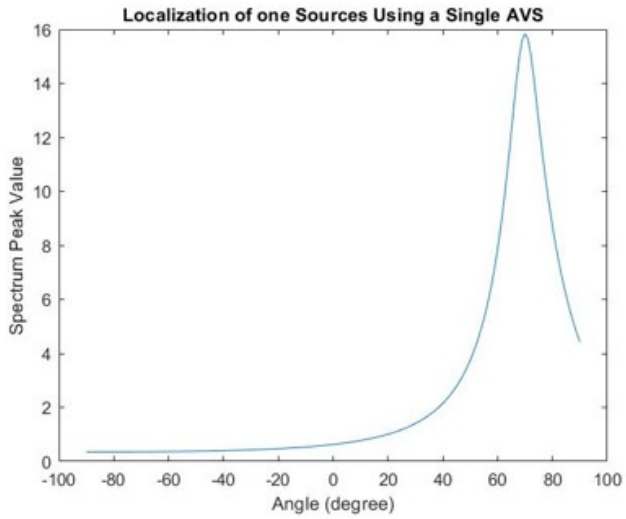



Figure 4.5: Real measurement result with a single AVS without cover

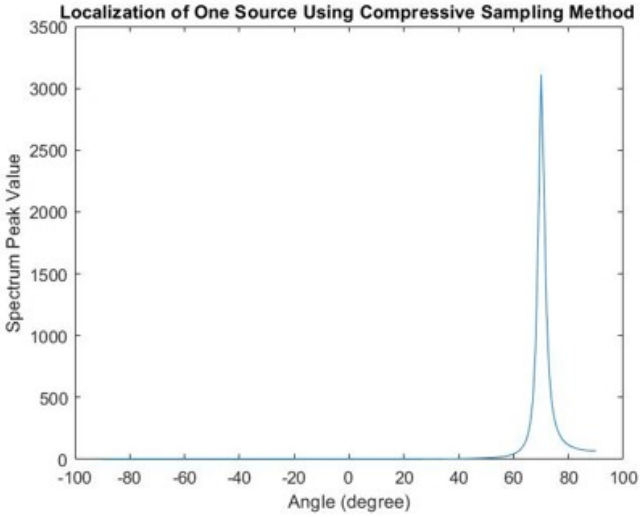



Figure 4.6: Real measurement result with a single AVS with cover by the CS-based method

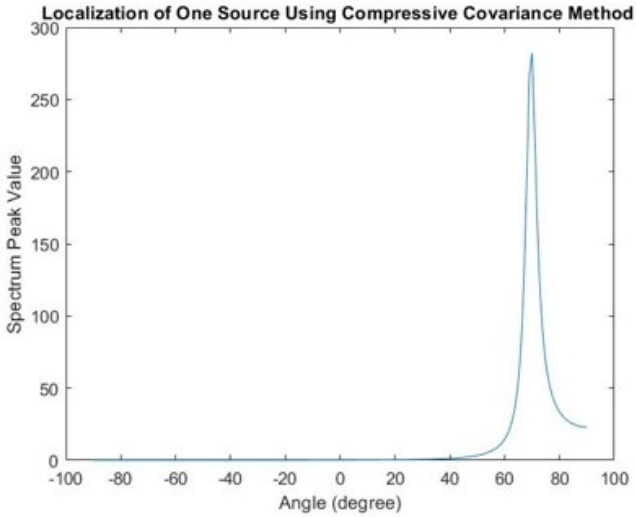


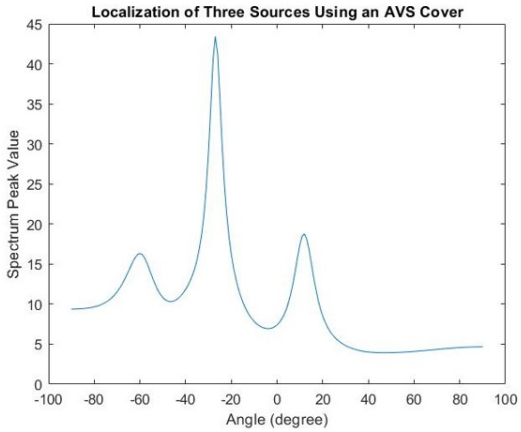

Figure 4.7: Real measurement result with a single AVS with cover by the CCS-based method

Figures 4.5, 4.6, 4.7 depict the results of DOA estimation for a single source using the three methods: a single AVS without a cover, the CS-based method, and the CCS-based method. The latter two methods involve measurements with the cover. The single source is located at an angle of 70 degrees. It is evident that both the CS-based method and the CCS-based method using the coded cover exhibit a sharper peak in comparison to a single AVS without a cover.

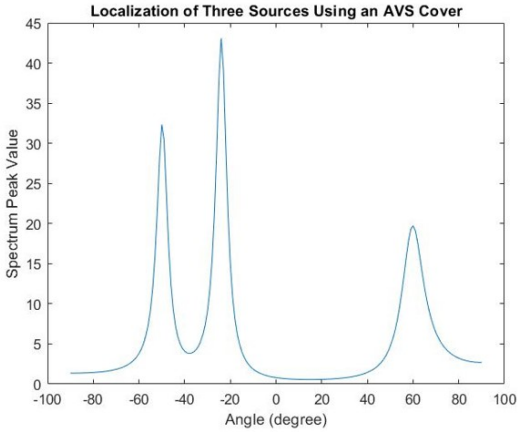
### 4.3.2 Multiple Sources DOA Estimation

In section 3.3.3, we introduced another advantage of the CCS-based method: its ability to detect more sources compared to what a single AVS can achieve. In this section,

we provide real measurement results to validate this concept. We conducted two experiments, with each experiment involving three sources placed at distinct angles. In the first experiment, the three sources were positioned at -60 degrees, -30 degrees, and 10 degrees. In the second experiment, the three sources were situated at -50 degrees, -15 degrees, and 60 degrees. It turns out that using the CCS-based method, we can successfully estimate the angles in these two experiments. The measurement results are shown in Figure 4.8.




(a) DOA estimation of 3 sources using the CCS-based method, experiment 1.




(b) DOA estimation of 3 sources using the CCS-based method, experiment 2.

Figure 4.8: Angle estimation of 3 sources using the CCS-based method in real measurements.

Therefore, we can conclude that by employing an AVS cover and the CCS-based method, we can detect the directions of at least three signal sources in practice. This surpasses the number of sources, which is limited to two, that a single AVS without cover can detect.

## 4.4 Conclusion

In this section, we validated the theories behind the coded cover through real experiments. We fabricated a physical AVS cover and designed an AVS box to enhance the accuracy of our experiments. Our findings confirm that the AVS cover significantly enhances spatial resolution. The precision of both the CS-based and CCS-based methods surpasses that of a single AVS, as indicated by the observation of a sharper peak in the results obtained through the MUSIC algorithm. Notably, we achieve successful estimation of three sources using the CCS-based method, providing evidence that the coded cover can expand the number of detectable sources when compared to a single AVS.

# 5 Conclusions and Future Study

## 5.1 Conclusions

The AVS is a sensor capable of measuring both sound pressure in the air and particle velocity. This unique capability enables it to perform DOA estimation tasks using a single sensor. In this study, we draw inspiration from the concept of imaging with a coded mask, as applied in ultrasound imaging, and investigate its potential applicability to improve the DOA estimation performance of a single AVS. We explore two different structures to achieve this.

The first structure involves a coded mask, a concept already utilized in the field of ultrasound imaging. It consists of placing a thin plastic film in front of the AVS sensor, and in this configuration, we focus on the reception of sound signals through this coded mask. We have conducted a performance study of this method using beam pattern analysis, MUSIC performance analysis, and the CRLB on DOA estimation.

Our findings reveal that, in this case, the coded mask primarily acts as a scaling effect. It results in nearly identical performance when using a MUSIC algorithm with the same SNR. Additionally, the same conclusion is reached when calculating the CRLB on DOA estimation, with or without the coded mask. This suggests that the coded mask, as currently implemented, does not significantly enhance DOA estimation performance for the AVS.

The second structure, introduced in this study, is referred to as a coded cover. In the coded cover configuration, we consider that sound independently propagates through different patches and is subsequently received by the AVS sensor. In this configuration, our goal is to not only improve the accuracy of DOA estimation with a single source but also expand the number of sources that can be detected with a single AVS. This is achieved by assuming the entries of the patches as a virtual ULA, and employing compressive methods to estimate the covariance matrix of this virtual ULA.

Two compressive-based DOA estimation methods have been proposed, utilizing the AVS cover to enhance the performance. The first method is referred to as the CS method, while the second method is known as the CCS method.

The CS method directly employs the compressed data collected by the single AVS along with the entire propagation manifold for DOA estimation. In contrast, the CCS method initially estimates the covariance matrix for the virtual ULA formed by the AVS cover and then leverages it for DOA estimation.

Both methods have been extensively validated through simulations and real-world measurements. They exhibit significant improvements in spatial resolution, leading to enhanced DOA estimation accuracy. Furthermore, the CCS method demonstrates the remarkable capability to estimate the directions of up to 4 sources with a single AVS, showcasing its potential for expanding the number of detectable sources.

## 5.2 Future Study

In this study, we have focused on analyzing the performance of 2D AVS with a coded cover for sound source localization. Firstly, we can extend our research to a 3D environment. Currently, we have primarily addressed sound source localization within a 2D plane. However, real-world acoustic environments often involve more dimensions. In future research, we can adapt our methods to 3D acoustic spaces, enhancing our understanding of sound propagation and reception in three-dimensional space.

Furthermore, our current study has primarily focused on the application of a single AVS with a coded cover. In future investigations, we can explore the potential benefits of configuring multiple single AVS units with covers as an array or placing multiple AVS sensors beneath a shared coded cover. Such an approach would require cooperative processing and data fusion techniques, ultimately improving the accuracy and robustness of sound source localization.

# Bibliography


[1] H.-E. de Bree, *The Microflown E-Book*. Microflown Technologies, Arnhem, 2007.

[2] H.-E. de Bree, "An overview of microflown technologies," *Acta acustica united with Acustica*, vol. 89, no. 1, pp. 163–172, 2003.

[3] K. Nambur Ramamohan, *Direction of arrival estimation and self-calibration techniques using an array of acoustic vector sensors: Theory, algorithms and applications*. PhD thesis, Delft University of Technology, 2022.

[4] P. van der Meulen, P. Kruizinga, J. G. Bosch, and G. Leus, "Coding mask design for single sensor ultrasound imaging," *IEEE Transactions on Computational Imaging*, vol. 6, pp. 358–373, 2019.

[5] D. Fernandez Comesaña, S. Steltenpool, G. Pousa, H.-E. Bree, and K. Holland, "Scan and paint: Theory and practice of a sound field visualization method," *ISRN Mechanical Engineering*, vol. 2013, 01 2013.

[6] E. Tijs and E. Druyvesteyn, "Spatial variation of impedance and intensity measured close to a sample in the near field of a spherical sound source," in *Symposium on the Acoustics of Poro-Elastic Materials*, 2011.

[7] M. Penhale and A. Barnard, "Direction of arrival estimation in practical scenarios using moving standard deviation processing for localization and tracking with acoustic vector sensors," *Applied Acoustics*, vol. 168, p. 107421, 2020.

[8] F. Jacobsen and P. M. Juhl, *Fundamentals of general linear acoustics*. John Wiley & Sons, 2013.

[9] A. Nehorai and E. Paldi, "Acoustic vector-sensor array processing," *IEEE Transactions on Signal Processing*, vol. 42, no. 9, pp. 2481–2491, 1994.

[10] P. Kruizinga, P. van der Meulen, A. Fedjajevs, F. Mastik, G. Springeling, N. de Jong, J. G. Bosch, and G. Leus, "Compressive 3d ultrasound imaging using a single sensor," *Science advances*, vol. 3, no. 12, p. e1701423, 2017.

[11] H. Tang, "Doa estimation based on music algorithm," 2014.

[12] M. Hawkes and A. Nehorai, "Effects of sensor placement on acoustic vector-sensor array performance," *IEEE Journal of Oceanic Engineering*, vol. 24, no. 1, pp. 33–40, 1999.

[13] Y. Wang, G. Leus, and A. Pandharipande, "Direction estimation using compressive sampling array processing," in *2009 IEEE/SP 15th Workshop on Statistical Signal Processing*, pp. 626–629, 2009.

[14] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[15] E. Candès, "Compressive sampling," *Proceedings oh the International Congress of Mathematicians, Vol. 3, 2006-01-01, ISBN 978-3-03719-022-7, pags. 1433-1452*, vol. 3, 01 2006.

[16] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Transactions on signal processing*, vol. 53, no. 7, pp. 2477–2488, 2005.

[17] Y. Wang, A. Pandharipande, and G. Leus, "Compressive sampling based mvdr spectrum sensing," in *2010 2nd International Workshop on Cognitive Information Processing*, pp. 333–337, IEEE, 2010.

[18] Y. Wang, G. Leus, and A. Pandharipande, "Direction estimation using compressive sampling array processing," in *2009 IEEE/SP 15th Workshop on Statistical Signal Processing*, pp. 626–629, IEEE, 2009.

[19] Y. Wang, A. Pandharipande, and G. Leus, "Compressive sampling based mvdr spectrum sensing," in *2010 2nd International Workshop on Cognitive Information Processing*, pp. 333–337, 2010.

[20] D. Romero and G. Leus, "Compressive covariance sampling," in *2013 Information Theory and Applications Workshop (ITA)*, pp. 1–8, 2013.

[21] D. Romero, D. D. Ariananda, Z. Tian, and G. Leus, "Compressive covariance sensing: Structure-based compressive sensing beyond sparsity," *IEEE Signal Processing Magazine*, vol. 33, pp. 78–93, 01 2016.

[22] C. Park and B. Lee, "Online compressive covariance sensing," *Signal Processing*, vol. 162, pp. 1–9, 2019.

[23] D. Romero, D. D. Ariananda, Z. Tian, and G. Leus, "Compressive covariance sensing: Structure-based compressive sensing beyond sparsity," *IEEE signal processing magazine*, vol. 33, no. 1, pp. 78–93, 2015.

[24] D. Romero, R. López-Valcarce, and G. Leus, "Compression limits for random vectors with linearly parameterized second-order statistics," *IEEE Transactions on Information Theory*, vol. 61, no. 3, pp. 1410–1425, 2015.

[25] D. Romero and G. Leus, "Wideband spectrum sensing from compressed measurements using spectral prior information," *IEEE Transactions on Signal Processing*, vol. 61, no. 24, pp. 6232–6246, 2013.

[26] H. Hung and M. Kaveh, "Focussing matrices for coherent signal-subspace processing," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 8, pp. 1272–1281, 1988.