Sound Localization Using an Acoustic Vector Sensor Array with a Coded Cover

Yuan Yuan*†, Dani Fernandez Comesaña†, Geert Leus*
*Delft University of Technology, The Netherlands
†Microflown Technologies, The Netherlands

***Abstract*—While the improvement of direction-of-arrival (DOA) estimation using coded covers with a single acoustic vector sensor (AVS) has been demonstrated, its extension to array-based systems remains relatively unexplored. To bridge this gap, we propose to extend the use of a coded cover from a single AVS setup to array-based acoustic measurement systems. The coded cover basically compresses the signals impinging on it from a large virtual array to a smaller physical array of sensors. To recover the uncompressed covariance matrix of the virtual array represented by the coded cover, we adopt compressed covariance sensing (CCS). Experimental results show that a 14 $\times$ 10 coded cover combined with an array of 12 probes that measure both pressure and particle velocity enables accurate localization of up to 100 sound sources in three dimensions, even under challenging conditions with a signal-to-noise ratio (SNR) as low as 10 dB. Furthermore, to address the impact of geometric mismatch—which distorts the compression matrix—we incorporate a grid-search-based calibration method to correct these perturbations. The proposed approach demonstrates both scalability and robustness, making it a promising solution for large-scale sound source localization.**

***Index Terms*—Acoustic particle velocity sensors, compressive covariance sensing, DOA estimation, geometric mismatch correction**

I. Introduction

Acoustic sound fields are characterized by both pressure variations and particle velocity. While acoustic pressure is a scalar quantity, particle velocity has both a magnitude and a direction, making it a vector quantity. Acoustic particle velocity sensors [1], which are transducers used to measure particle velocity in air, operate on the same principles as hot-wire anemometers. An acoustic vector sensor (AVS) has the ability to measure acoustic pressure and one, two or three particle velocity channels simultaneously. While individual sensors provide valuable data on acoustic pressure and particle velocity at a single point, scaling this approach to an array offers enhanced capabilities.

Current systems that employ a single AVS with a coded cover [2], [3] have shown improved detection accuracy through compressed sensing (CS) [4] and compressed covariance sensing (CCS) methods [5]. The model of the considered cover to improve the performance of an AVS is shown in Fig. 1. As the wave impinges on the cover, it passes through each channel independently. At the exit of each channel, we assume the presence of a point source. The sound then propagates from this point source to the position of the AVS in the near field. Notably, the CCS method applied to a single AVS with a coded cover has the ability to detect more sound sources than an AVS without cover. However, there remains a lack of evidence regarding the scalability and performance of these methods in more complex scenarios.

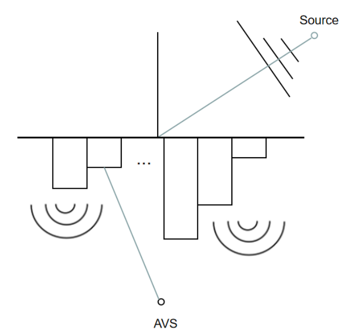


Fig. 1: The model of a single AVS with a coded cover [3]

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

Furthermore, this paper proposes calibration within the framework of CCS. Since CCS relies on an accurate system model, even a small mismatch in the geometric configuration can degrade performance. By integrating calibration techniques into the CCS process, we aim to correct this mismatch and enhance the accuracy of source localization.

II. Measurement Model

Assume K AVSs are placed behind a coded mask with M channels. Each sensor provides Q measurements, including one sound pressure measurement and $Q-1$ particle velocity measurements. To estimate both elevation and azimuth angles of N sources in 3D space, the mask must be configured in 2D or 3D. Specifically, we consider a 2D planar arrangement of the coded cover, as shown in Fig. 2. The mask is composed of J rows and I columns, leading to a total of $M = IJ$ channels. The AVS array positioned behind the planar mask also has a 2D planar arrangement. The localization includes the elevation, which is based on the deviation from the z-axis $\theta \in [0^\circ, 90^\circ)$ and azimuth from the x-axis $\psi \in [0^\circ, 360^\circ)$.

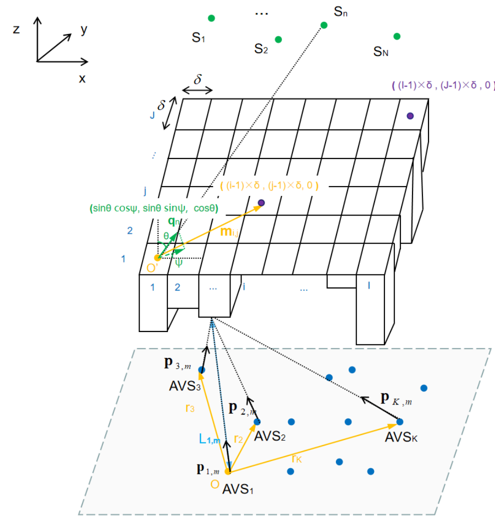


Fig. 2: 3D source localization model

The signal direction of the n-th source is represented by the unit vector $\boldsymbol{q}_n = [\sin\theta_n \cos\psi_n, \sin\theta_n \sin\psi_n, \cos\theta_n]^T$. The top of the coded cover can be viewed as a virtual array where every channel entry represents a virtual array element. The virtual element in the position (i, j) is characterized by the position vector $\boldsymbol{m}_{ij} = [(i-1)\times\delta, (j-1)\times\delta, 0]^T$. Here, δ is the side length of each square channel entry and also equals the spacing between the channels. In addition, the speed of sound is denoted as c and the source frequency as f. Hence, the spatial frequency or wavenumber is given by $k = 2\pi f/c$ and the channel spacing in wavelengths is $\triangle = \delta f/c$ (we will set it as $1/2$ to avoid aliasing problems).

The array response vector related to the n-th source that impinges on the virtual array can be expressed as

$$\boldsymbol{a}(\theta_n, \psi_n) = \begin{bmatrix} e^{jk\mathbf{q}_n^T\mathbf{m}_{1,1}} \\ e^{jk\mathbf{q}_n^T\mathbf{m}_{2,1}} \\ e^{jk\mathbf{q}_n^T\mathbf{m}_{3,1}} \\ \vdots \\ e^{jk\mathbf{q}_n^T\mathbf{m}_{I,J}} \end{bmatrix} \tag{1}$$

$$= \underbrace{\begin{bmatrix} 1 \\ e^{j2\pi\triangle\sin\theta_n\sin\psi_n} \\ \vdots \\ e^{j2\pi(J-1)\triangle\sin\theta_n\sin\psi_n} \end{bmatrix}}_{\boldsymbol{a}_y(\theta_n,\psi_n)} \otimes \underbrace{\begin{bmatrix} 1 \\ e^{j2\pi\triangle\sin\theta_n\cos\psi_n} \\ \vdots \\ e^{j2\pi(I-1)\triangle\sin\theta_n\cos\psi_n} \end{bmatrix}}_{\boldsymbol{a}_x(\theta_n,\psi_n)}, \tag{2}$$

where $\otimes$ denotes the Kronecker product. Based on this, the virtual array response matrices along the row and column directions can be denoted as:

$$\boldsymbol{A}_y(\boldsymbol{\theta}, \boldsymbol{\psi}) = [\boldsymbol{a}_y(\theta_1, \psi_1), \ldots, \boldsymbol{a}_y(\theta_N, \psi_N)] \in \mathbb{C}^{J\times N}, \tag{3}$$

$$\boldsymbol{A}_x(\boldsymbol{\theta}, \boldsymbol{\psi}) = [\boldsymbol{a}_x(\theta_1, \psi_1), \ldots, \boldsymbol{a}_x(\theta_N, \psi_N)] \in \mathbb{C}^{I\times N}. \tag{4}$$

Note that both these matrices have a Vandermonde structure. The virtual planar array matrix for all sources can finally be written as

$$\boldsymbol{A}(\boldsymbol{\theta}, \boldsymbol{\psi}) = [\boldsymbol{a}(\theta_1, \psi_1), \ldots, \boldsymbol{a}(\theta_N, \psi_N)] \tag{5}$$

$$= [\boldsymbol{a}_y(\theta_1, \psi_1) \otimes \boldsymbol{a}_x(\theta_1, \psi_1), \ldots, \boldsymbol{a}_y(\theta_N, \psi_N) \otimes \boldsymbol{a}_x(\theta_N, \psi_N)] \tag{6}$$

$$= \boldsymbol{A}_y(\boldsymbol{\theta}, \boldsymbol{\psi}) \odot \boldsymbol{A}_x(\boldsymbol{\theta}, \boldsymbol{\psi}) \in \mathbb{C}^{IJ\times N}, \tag{7}$$

where $\odot$ represents the column-wise Kronecker product or Khatri-Rao product.

Let us now discuss the compression matrix $\boldsymbol{\Phi}$ that describes the propagation from the virtual array to the sensor array. It can basically be constructed by calculating the distance from the entrance of each channel to each sensor. Assume the length of the m-th channel is d_m, and the distance between the exit of the m-th channel and the k-th AVS sensor is $L_{k,m}$. Since the distance from the coded mask to the AVS array is far shorter than 10 times the wavelength, a near-field model should be assumed. Specifically, the attenuation from the entrance of the m-th channel to the k-th AVS sensor can be represented by

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

Grouping the attenuation coefficients $\beta_{k,m}$ for $k = 1, 2, \ldots, K$ and a fixed channel m into a diagonal matrix, we obtain

$$\boldsymbol{E}_m = \mathrm{diag}(\beta_{1,m}, \ldots, \beta_{K,m}) \otimes \mathbf{I}_{Q\times Q}.$$

Note that we use the Kronecker product here to indicate the linear relation between the pressure and the particle velocity measurements used in the AVS. The compression matrix can now be written as the following $QK \times M$ matrix:

$$\boldsymbol{\Phi} = [\boldsymbol{E}_1\mathbf{u}_1, \boldsymbol{E}_2\mathbf{u}_2, \ldots, \boldsymbol{E}_M\mathbf{u}_M]. \tag{8}$$

where $\mathbf{u}_m = [1, \mathbf{p}_{1,m}^T, \ldots, 1, \mathbf{p}_{K,m}^T]^T$, with $\mathbf{p}_{k,m}$ the directional unit vector pointing from the k-th sensor to the exit of the m-th channel. Note that the length of $\mathbf{p}_{k,m}$ is $Q-1$.

After obtaining $\boldsymbol{A}(\boldsymbol{\theta}, \boldsymbol{\psi})$ and $\boldsymbol{\Phi}$, the total array response matrix is:

$$\begin{aligned} \mathbf{B}(\boldsymbol{\theta}, \boldsymbol{\psi}) &= \boldsymbol{\Phi}\mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\psi}) \\ &= [\boldsymbol{\Phi}\mathbf{a}(\theta_1, \psi_1), \boldsymbol{\Phi}\mathbf{a}(\theta_2, \psi_2), \ldots, \boldsymbol{\Phi}\mathbf{a}(\theta_N, \psi_N)] \\ &= [\mathbf{b}(\theta_1, \psi_1), \mathbf{b}(\theta_2, \psi_2), \ldots, \mathbf{b}(\theta_N, \psi_N)]. \end{aligned}$$

The received signals at the AVS array can finally be expressed as:

$$\mathbf{y}(t) = \boldsymbol{\Phi}\mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\psi})\mathbf{s}(t) + \mathbf{n}(t) = \mathbf{B}(\boldsymbol{\theta}, \boldsymbol{\psi})\mathbf{s}(t) + \mathbf{n}(t) \tag{9}$$

$$= \boldsymbol{\Phi}\mathbf{x}(t) + \mathbf{n}(t) \in \mathbb{C}^{QK\times 1}, \tag{10}$$

where $\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\psi})\mathbf{s}(t)$ is the noiseless measurement vector at the virtual array. The measurement covariance matrix $\mathbf{R_y}$ can thus be calculated as

$$\mathbf{R_y} = \boldsymbol{\Phi}\mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\psi})\mathbf{R_s}\mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\psi})^H\boldsymbol{\Phi}^H + \mathbf{R_n} \tag{11}$$

$$= \mathbf{B}(\boldsymbol{\theta}, \boldsymbol{\psi})\mathbf{R_s}\mathbf{B}(\boldsymbol{\theta}, \boldsymbol{\psi})^H + \mathbf{R_n} \tag{12}$$

$$= \boldsymbol{\Phi}\mathbf{R_x}\boldsymbol{\Phi}^H + \mathbf{R_n} \in \mathbb{C}^{QK\times QK}, \tag{13}$$

where $\mathbf{R_s} = \mathbb{E}\{\mathbf{s}(t)\mathbf{s}(t)^H\}$, $\mathbf{R_n} = \mathbb{E}\{\mathbf{n}(t)\mathbf{n}(t)^H\}$ and $\mathbf{R_x} = \mathbb{E}\{\mathbf{x}(t)\mathbf{x}(t)^H\}$ are the covariance matrices of the corresponding signals.

III. CCS Principle

Suppose now that $\mathbf{R_x}$ lives in a subspace that is characterized by a set of basis matrices $\mathcal{S} = \{\mathbf{\Sigma}_0, \mathbf{\Sigma}_1, \ldots, \mathbf{\Sigma}_{|\mathcal{S}|-1}\}$. In other words, we can write $\mathbf{R_x}$ as

$$\mathbf{R_x} = \sum_{m=0}^{|\mathcal{S}|-1} \alpha_m \mathbf{\Sigma}_m. \tag{14}$$

The structure of this subspace will be discussed in the next section, but let us for now assume that such a subspace exists and that its dimension $|\mathcal{S}|$ is much smaller than $M^2 = I^2J^2$.

Plugging (14) into (13), we then get

$$\mathbf{R_y} - \mathbf{R_n} = \sum_{m=0}^{|\mathcal{S}|-1} \alpha_m \mathbf{\Phi}\mathbf{\Sigma}_m\mathbf{\Phi}^H \tag{15}$$

$$= \sum_{m=0}^{|\mathcal{S}|-1} \alpha_m \overline{\mathbf{\Sigma}}_m, \tag{16}$$

where the set of compressed basis matrices is given by

$$\overline{\mathcal{S}} = \{\overline{\mathbf{\Sigma}}_0, \overline{\mathbf{\Sigma}}_1, \ldots \overline{\mathbf{\Sigma}}_{|\mathcal{S}|-1}\} \quad \text{with} \quad \overline{\mathbf{\Sigma}}_m = \mathbf{\Phi}\mathbf{\Sigma}_m\mathbf{\Phi}^H.$$

If the compression matrix $\mathbf{\Phi}$ is able to preserve all the second-order statistical information of $\mathbf{x}$, it is called an $\mathcal{S}$-covariance sampler. Such a sampler is able to preserve the linear independence among the matrices in the set $\overline{\mathcal{S}}$, i.e., if there is a unique set of weights $\{\alpha_m\}$ to describe $\mathbf{R_x}$, then this is also the unique set of weights to describe $\mathbf{R_y} - \mathbf{R_n}$. The idea of compressed covariance sensing (CCS) [6] therefore is to estimate the set of weights $\{\alpha_m\}$ based on (16) and to use them to reconstruct $\mathbf{R_x}$ based on (14). Note that the smaller $|\mathcal{S}|$, the higher the compression ratio achieved by CCS.

To estimate the α_m values, we can use either a maximum-likelihood (ML) approach [7], [8] or a linear least squares (LLS) method. ML involves a high computational cost and requires an accurate statistical characterization of the observations. For this reason, we consider a simple LLS cost based on (16):

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \|\mathbf{R_y} - \mathbf{R_n} - \sum_{m=0}^{|\mathcal{S}|-1} \alpha_m \overline{\mathbf{\Sigma}}_m\|_F^2.$$

The solution is described in [5]. First, we vectorize the matrices $\mathbf{R_y}$, $\mathbf{R_n}$, $\mathbf{R_x}$ and $\mathbf{\Sigma}_m$ into the respective vectors $\boldsymbol{\sigma}_\mathbf{y}$, $\boldsymbol{\sigma}_\mathbf{n}$, $\boldsymbol{\sigma}_\mathbf{x}$ and $\boldsymbol{\sigma}_m$ and we define $\boldsymbol{\alpha} = [\alpha_0, \alpha_1, \ldots, \alpha_{|\mathcal{S}|-1}]^T$. From (14), we can then obtain a linear relation between $\boldsymbol{\sigma}_x$ and $\boldsymbol{\alpha}$:

$$\boldsymbol{\sigma}_\mathbf{x} = \sum_{m=0}^{|\mathcal{S}|-1} \alpha_m \boldsymbol{\sigma}_m = \mathbf{S}\boldsymbol{\alpha}, \tag{17}$$

where $\mathbf{S} = [\boldsymbol{\sigma}_0, \boldsymbol{\sigma}_1, \ldots, \boldsymbol{\sigma}_{|\mathcal{S}|-1}]$. And from (13), we can derive also a linear relation between $\boldsymbol{\sigma}_\mathbf{y} - \boldsymbol{\sigma}_\mathbf{n}$ and $\boldsymbol{\alpha}$:

$$\boldsymbol{\sigma}_\mathbf{y} - \boldsymbol{\sigma}_\mathbf{n} = (\mathbf{\Phi}^* \otimes \mathbf{\Phi})\boldsymbol{\sigma}_x = (\mathbf{\Phi}^* \otimes \mathbf{\Phi})\mathbf{S}\boldsymbol{\alpha}. \tag{18}$$

Hence, $\boldsymbol{\alpha}$ can be estimated by solving the above equation with LLS: $\hat{\boldsymbol{\alpha}} = [(\mathbf{\Phi}^* \otimes \mathbf{\Phi})\mathbf{S}]^+(\boldsymbol{\sigma}_y - \boldsymbol{\sigma}_n)$, where $^+$ denotes the pseudo-inverse operation. Finally, $\mathbf{R_x}$ can be obtained as

$$\hat{\mathbf{R}}_\mathbf{x} = \text{vec}^{-1}\{\mathbf{S}[(\mathbf{\Phi}^* \otimes \mathbf{\Phi})\mathbf{S}]^+(\boldsymbol{\sigma}_y - \boldsymbol{\sigma}_n)\}.$$

Any covariance-based DOA estimation algorithm can be applied to localize the sources using $\hat{\mathbf{R}}_\mathbf{x}$. In this paper, we will focus on the multiple signal classification (MUSIC) algorithm for this [9].

IV. Covariance Subspace Characterization

In this section, we will discuss the specific structure of $\mathbf{R_x}$, which in turn specifies the set of basis matrices $\mathcal{S}$.

Assuming all the N sources are independent, i.e., the source covariance matrix $\mathbf{R_s}$ is diagonal, the covariance matrix $\mathbf{R_x}$ can be written as

$$\mathbf{R_x} = \mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\psi})\boldsymbol{R}_\mathbf{s}\mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\psi})^H$$
$$= \sum_{n=1}^{N} \mathbf{a}(\theta_n, \psi_n)\mathbf{a}(\theta_n, \psi_n)^H \sigma_n^2$$
$$= \sum_{n=1}^{N} [\mathbf{a}_y(\theta_n, \psi_n)\mathbf{a}_y(\theta_n, \psi_n)^H \otimes \mathbf{a}_x(\theta_n, \psi_n)\mathbf{a}_x(\theta_n, \psi_n)^H]\sigma_n^2,$$

where σ_n^2 is the signal power of the n-th source. This matrix turns out to be a positive semidefinite (PSD) (I, J), 2-level Toeplitz (2LT) matrix, as defined in [10]. The rank of this matrix is N. When $N < \min(I, J)$, all the sources can be recovered surely based on Theorem 1 in [10]. According to Theorem 2, if such a decomposition of the covariance matrix with $N \geq \min(I, J)$ exists, it can be uniquely found under the weak condition:

$$N \leq IJ - \max(I, J) = M - \max(I, J). \tag{19}$$

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

Let us now characterize this set of (I, J) 2LT matrices based on the subspace of Hermitian Toeplitz (HT) matrices. A basis for the subspace of HT matrices of size $J \times J$ is:

$$\mathcal{S}' = \{\mathbf{\Sigma}'_0, \mathbf{\Sigma}'_1, \ldots \mathbf{\Sigma}'_{2J-2}\} \tag{20}$$

$$= \{\mathbf{I}_J\} \cup \{\mathbf{T}'_1, \ldots, \mathbf{T}'_{J-1}\} \cup \{\tilde{\mathbf{T}}'_1, \ldots, \tilde{\mathbf{T}}'_{J-1}\}, \tag{21}$$

where $\mathbf{T}'_j$ denotes the HT matrix with all zeros except for the entries on the diagonals $+j$ and $-j$, which have ones, and $\tilde{\mathbf{T}}'_j$ represents the HT matrix with all zeros except for the entries on the diagonal $+j$, which have the imaginary unit $\jmath$, and those on the diagonal $-j$, which have $-\jmath$. As a result, for the n-th source, we can write

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

Next, a basis for the subspace of HT matrices of size $I \times I$ is:

$$\mathcal{S}'' = \{\mathbf{\Sigma}''_0, \mathbf{\Sigma}''_1, \ldots \mathbf{\Sigma}''_{2I-2}\} \tag{22}$$

$$= \{\mathbf{I}_I\} \cup \{\mathbf{T}''_1, \dots, \mathbf{T}''_{I-1}\} \cup \{\tilde{\mathbf{T}}''_1, \dots, \tilde{\mathbf{T}}''_{I-1}\}, \qquad (23)$$

where $\mathbf{T}''_i$ and $\tilde{\mathbf{T}}''_i$ are similarly defined as before. And as before, for the n-th source we can then write

$$\boldsymbol{a}_x(\theta_n, \psi_n)\boldsymbol{a}_x(\theta_n, \psi_n)^H = \sum_{i=0}^{2I-2} \alpha''_i \boldsymbol{\Sigma}''_i.$$

From the derivation above, the full covariance matrix $\mathbf{R}_\mathbf{x}$ for a single source n has the form

$$\boldsymbol{a}(\theta_n, \psi_n)\boldsymbol{a}(\theta_n, \psi_n)^H \sigma_n^2 = \sum_{i=0}^{2I-2} \sum_{j=0}^{2J-2} \alpha''_i \alpha'_j \sigma_n^2 \boldsymbol{\Sigma}''_i \otimes \boldsymbol{\Sigma}'_j \qquad (24)$$

$$= \sum_{m=1}^{(2I-1)(2J-1)} \alpha_{m,n} \boldsymbol{\Sigma}_m. \qquad (25)$$

Here, $\alpha_{m,n}$ contains all unknown scalar information in the two HT subspaces as well as the signal power, defined as $\alpha''_i \alpha'_j \sigma_n^2$. The basis matrices $\boldsymbol{\Sigma}_m$ that yield the subspace of $\mathbf{R}_\mathbf{x}$ for such a single source n can thus be defined as:

$$\mathcal{S} = \{\boldsymbol{\Sigma}_m \in \mathbb{C}^{M \times M} | \boldsymbol{\Sigma}_m = \boldsymbol{\Sigma}''_i \otimes \boldsymbol{\Sigma}'_j, \boldsymbol{\Sigma}''_i \in \mathcal{S}'', \boldsymbol{\Sigma}'_j \in \mathcal{S}'\}. \qquad (26)$$

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

For multiple sources, since each independent source shares the same subspace, the weight of a basis matrix will be given by the sum of the respective weights for the N sources. In other words, we obtain

$$\begin{aligned}\mathbf{R}_\mathbf{x} &= \sum_{n=1}^{N} \boldsymbol{a}(\theta_n, \psi_n)\boldsymbol{a}(\theta_n, \psi_n)^H \sigma_n^2 \\ &= \sum_{m=1}^{(2I-1)(2J-1)} \sum_{n=1}^{N} \alpha_{m,n} \boldsymbol{\Sigma}_m \\ &= \sum_{m=1}^{(2I-1)(2J-1)} \alpha_m \boldsymbol{\Sigma}_m,\end{aligned}$$

where α_m is the unknown scalar that contains all source information. To keep the system of equations in (18) overdetermined, the following constraints on I and J must be satisfied:

$$(2I-1)(2J-1) \leq (QK)^2.$$

This will allow for the design of $M = IJ$ coded cover channels, and the recovery of $M - \max(I, J)$ sources will be guaranteed.

V. Calibration of Geometric Mismatch

Accurate DOA estimation depends on the precise geometric configuration of the sensing system. Small deviations in the geometry, such as the sensor positions in the array or array rotations, can significantly distort the compression model and degrade performance. In this section, we examine the impact of geometric mismatch and propose a method for compensating these distortions.

A. Categories of Geometric Mismatch

The geometric system parameters can be categorized into three main types:

- l: The distance between the coded mask entry and the sensor array.
- ϵ: The positional deviation of the first sensor. We restrict this to deviations in the x-direction, although this can be extended.
- ζ: The rotation angle of the planar array relative to the plane parallel to the coded mask. For simplicity, we limit ourselves to rotations around the y-axis, although this can be extended.

Fig. 3 depicts these system parameters from an x-z plane perspective.

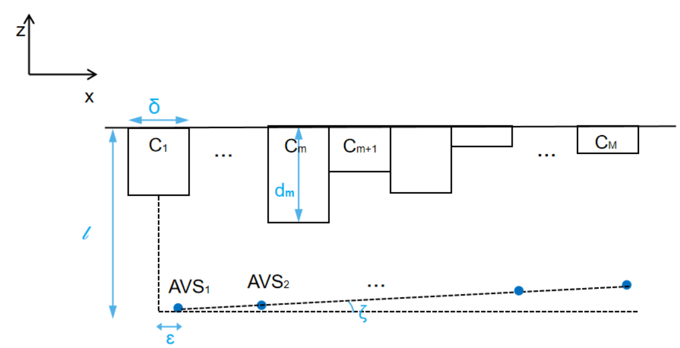


Fig. 3: Visualization of the geometric parameters in the x-z plane.

B. Grid-Search Method

Unlike receiver-based errors [11], which typically act row-wise on the compression matrix and are uniform across channels, a geometric mismatch introduces element-wise phase and gain distortions on the compression matrix. Consequently, applying the calibration matrix directly to the compression matrix requires element-wise corrections [12]. However, in our case, the large size of the compression matrix renders this approach computationally impractical. Therefore, we propose a practical alternative based on the structure of our compression model.

When considering perturbations solely due to geometric mismatch, we propose using a grid-search method to optimize the geometric parameters accordingly. The main geometric parameters that need to be optimized have been introduced in Section V-A.

For each parameter configuration, the corresponding compression matrix $\boldsymbol{\Phi}^i$ is computed. The CCS method is then applied to recover the covariance matrix by constructing:

$$\mathbf{R}_\mathbf{x}^i = \mathrm{vec}^{-1}\left\{\mathbf{S}\left[\left(\boldsymbol{\Phi}^{i*} \otimes \boldsymbol{\Phi}^i\right)\mathbf{S}\right]^+ (\boldsymbol{\sigma}_\mathbf{y} - \boldsymbol{\sigma}_\mathbf{n})\right\}.$$

And the corresponding reconstruction error is then computed as:

$$e_i = \left\|\boldsymbol{\sigma}_\mathbf{y} - \boldsymbol{\sigma}_\mathbf{n} - \left(\boldsymbol{\Phi}^{i*} \otimes \boldsymbol{\Phi}^i\right)\boldsymbol{\sigma}_\mathbf{x}^i\right\|_2.$$

Fig. 4: RECT-7.5

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

VI. Simulation Tests

In this section, we first perform simulations to validate the effectiveness of the CCS method in 3D sound source localization. Then, we evaluate the performance of the grid search method for correcting geometric mismatch.

The sensors we used are called PU probes. It is a sound intensity probe that integrates two sensing technologies: a traditional microphone (P) and a Microflown sensor (U) [13], [14]. It can basically be viewed as a 1D AVS. The Microflown PU probe provides a distinctive capability to capture acoustic information from one direction within a sound field. One of the planar array product series from Microflown is the Hand-Held Array. It includes the RECT-7.5, RECT-1.8, and BiHex models. The first two arrays both feature a 3×4 configuration with 12 PU probes.

In the simulations, we focused only on the RECT-7.5, where the orientation of each probe is perpendicular to the array plane, which is towards the z axis in the model described before. The sensor spacing is kept as 75 mm. Fig. 4 shows a picture of the RECT-7.5. Thus, we have $K = 12$ and $Q = 2$.

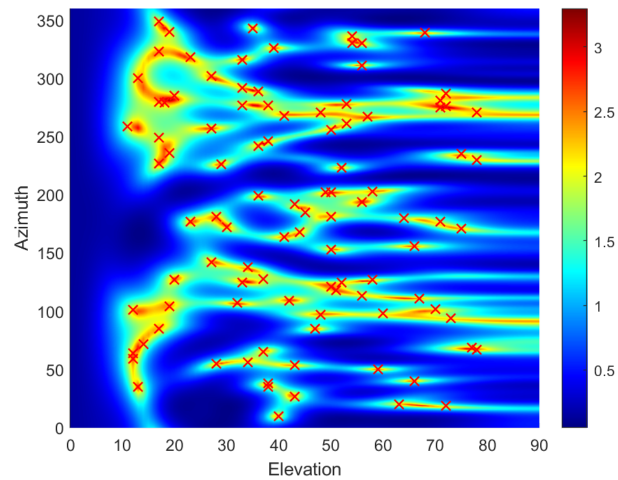
A. Simulation Results of CCS for 3D Source Localization

We test the case when a 10×14 planar coded cover is used. The maximum number of identifiable sources for the considered set-up is $140 - 14 = 126$. The RECT-7.5 is positioned 15 cm behind the entry plane of the coded cover. The length of the coded cover channels ranges from 1 cm to 10 cm. The signal frequency is 10 kHz, so the channel spacing is chosen to be 1.7 cm to ensure it is half the wavelength.

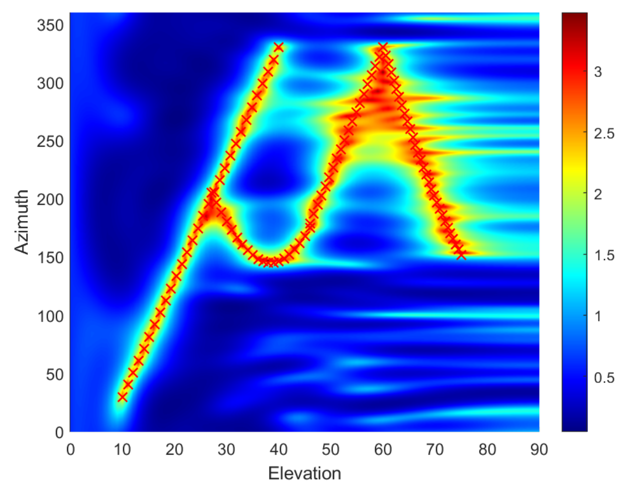
Fig. 5 shows that when 100 sources are present, the CCS method combined with the MUSIC algorithm is able to localize them all effectively.

B. Simulation Results after Grid Search

The proposed algorithm is next evaluated using 30 randomly placed sources in 3D space, with an SNR of 20 dB. The introduced geometric mismatch is defined as follows: the actual distance between the coded cover and the array is 15.2



(a) 100 random sources



(b) 100 sources in a specific pattern

Fig. 5: MUSIC localization of 100 sources with CCS when SNR = 10dB

cm instead of the nominal 15 cm. Additionally, the array experiences a rotational deviation around the y axis of $2°$, resulting in a non-parallel alignment between the two planes.

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

A comparison between Fig. 6 and Fig. 7 demonstrates the effectiveness of grid-search calibration in the 3D scenario. Once the optimal parameters are identified, the algorithm accurately localizes the sources. The distance l that best fits the model is found to be 15.2 cm, and the rotation adjustment corresponds precisely to the introduced deviation of $2°$.

VII. Conclusion

In this paper, we proposed a planar coded cover for an array of acoustic vector sensors and developed a CCS algorithm for 3D sound source localization. With the proposed setup, we

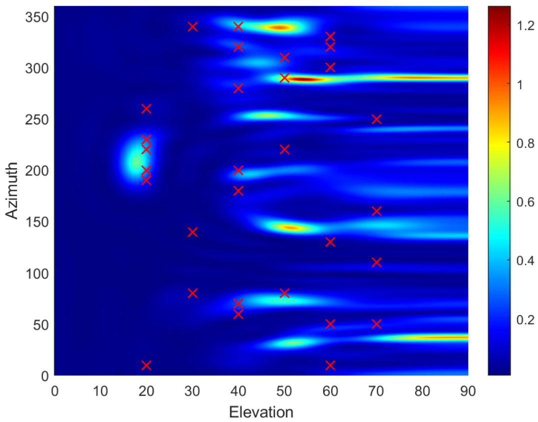


Fig. 6: CCS localization result using the nominal (incorrect) compression matrix in the 3D case, where the detected peaks do not align with the true source positions

successfully estimated 100 sources under an SNR of 10 dB, demonstrating the extension of the CCS localization capability from 2D to 3D. To address geometric mismatch, we further proposed a grid-search-based correction method, which was validated through simulations.

REFERENCES

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

[2] C. Yan, "Coded cover for acoustic vector sensors," M.S. thesis, Faculty of Electrical Engineering, Mathematics and Computer Science, Delft Univ. of Technology, Delft, The Netherlands, 2023. [Online]. Available: http://resolver.tudelft.nl/uuid:7c6480c0-c8fb-4a45-b2a3-ba24fb32d1ea

[3] C. Yan, C. Yan, D. F. Comesaña and G. Leus, "Coded Cover for Acoustic Vector Sensors," in *Proceedings of the 2024 58th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, USA, 2024, pp. 1210–1214, doi: 10.1109/IEEECONF60004.2024.10942950.

[4] E. J. Candes and M. B. Wakin, "An Introduction To Compressive Sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, March 2008, doi: 10.1109/MSP.2007.914731.

[5] D. Romero, D. D. Ariananda, Z. Tian, and G. Leus, "Compressive covariance sensing: Structure-based compressive sensing beyond sparsity," *IEEE Signal Process. Mag.*, vol. 33, no. 1, pp. 78–93, 2016.

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

[7] J. P. Burg, D. G. Luenberger, and D. L. Wenger, "Estimation of structured covariance matrices," *Proc. IEEE*, vol. 70, no. 9, pp. 963–974, Sept. 1982.

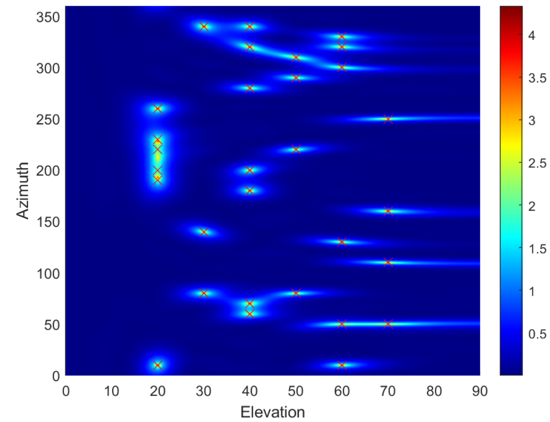
[8] D. Romero and G. Leus, "Wideband spectrum sensing from compressed measurements using spectral prior information," *IEEE Trans. Signal Process.*, vol. 61, no. 24, pp. 6232–6246, 2013.

[9] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

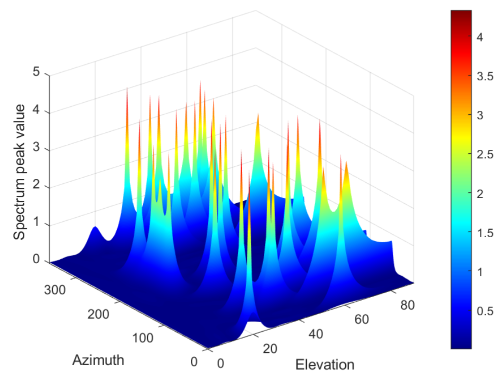
[10] Z. Yang, L. Xie, and P. Stoica, "Vandermonde decomposition of multilevel Toeplitz matrices with application to multidimensional super-resolution," *IEEE Trans. Inf. Theory*, vol. 62, no. 6, pp. 3685–3701, Jun. 2016.

[11] K. N. Ramamohan, S. P. Chepuri, D. F. Comesaña, and G. Leus, "Self-calibration of acoustic scalar and vector sensor arrays," *IEEE Trans. Signal Process.*, vol. 71, pp. 61–75, 2023, doi: 10.1109/TSP.2022.3214383.

[12] H. Zhu, G. Leus, and G. B. Giannakis, "Sparsity-cognizant total least-squares for perturbed compressive sampling," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2002–2016, May 2011, doi: 10.1109/TSP.2011.2109956.



(a) Grid-search calibration result (top view)



(b) Grid-search calibration result (3D view)

Fig. 7: CCS localization results after applying grid-search calibration in 3D, showing a precise alignment between the detected peaks and true source positions

[13] H.-E. de Bree, *The Microflown E-book*, 2009. [Online]. Available: https://www.microflown.com

[14] H.-E. De Bree, P. Leussink, T. Korthorst, H. Jansen, T. Lammerink, and M. Elwenspoek, "The μ-flown - a novel device measuring acoustical flows," in *Proc. Int. Solid-State Sensors Actuators Conf. (TRANSDUCERS'95)*, vol. 1, IEEE, 1995, pp. 536–539.