# Study of MVDR Beamforming with Spatially Distributed Source: Theoretical Analysis and Efficient Microphone Array Geometry Optimization Method

Wenmeng Xiong, Changchun Bao, Maoshen Jia, José Picheral

# Study of MVDR beamforming with spatially distributed source: Theoretical analysis and efficient microphone array geometry optimization method

Wenmeng XIONG[1], Changchun BAO[1*], Maoshen JIA[1] and José PICHERAL[2]

[1]Speech and Audio Signal Processing Laboratory, Faculty of Information Technology,Beijing University of Technology, Beijing, 100124, China.

[2]Laboratoire des signaux et systèmes (Signal and System Laboratory), Université Paris-Saclay, CNRS, CentraleSupélec, 91190, Gif-sur-Yvette, France.

*Corresponding author(s). E-mail(s): baochch@bjut.edu.cn;
Contributing authors: wenmeng.xiong@bjut.edu.cn;
jiamaoshen@bjut.edu.cn; jose.picheral@centralesupelec.fr;

**Abstract**

In this paper, the minimum variance distortionless response (MVDR) beamforming technique is studied in the presence of a spatially coherently distributed (CD) source. In the first part, we propose the CD-MVDR beamforming in which the steering vector of the CD source model is used instead of the conventional point source model. We derive a theoretical expression of the white noise gain of CD-MVDR beamforming, which is inversely proportional to the square of the difference between the angular dispersion of the actual source and that of the CD-MVDR beamforming model. In the second part, based on the performance analyses, we propose an efficient optimization method for the microphone array geometry to reduce the impact of the CD source angular dispersion on the performance of the conventional MVDR beamforming. Simulation results validate our proposed theoretical expressions and show that with the proposed geometry, the conventional MVDR beamforming tends

to be significantly more robust to the angular dispersion of the CD source for both white noise gain (WNG) and directivity factor (DF).

**Keywords:** Microphone Array, Minimum Variance Distortionless Response, Beamforming, Coherent Distributed Source, Geometry Optimization, White Noise Gain

# 1 Introduction

s a famous signal enhancing technique, beamforming [4] has been widely used in speech enhancement, acoustic imaging, wireless communications, radar processing, etc. The beamforming techniques can be divided into two categories: fixed beamforming and adaptive beamforming. The fixed beamforming algorithms, including delay-and-sum beamforming, are independent of the received signals at the microphone array, while the adaptive beamforming algorithms are more flexible, the coefficients of the Beamformers can be changed according to the received signals, to preserve the signal of interest (SOI) while reducing the noise and the interference. Among the adaptive beamformings, the MVDR beamforming [2] is prevail for its convenient implementation. The MVDR beamforming has two different ways to make use of the signals received on the microphone array, leading to two objective functions: the one generally employed by the conventional MVDR filter minimizes the power of the output signals of the microphone array; the other one generally employed by the MVDR filter minimizes the power of the residual noise of the microphone array with the reconstructed interference plus noise covariance matrix [12]. Indeed, the latter one can achieve a better output signal to interference plus noise ratio (SINR) when the input signal to noise ratio (SNR) is high.

In addition, plenty of robust beamforming techniques have been proposed to improve the performance with point source model [18; 19; 25; 29]. However, in many practical applications, the source can no longer be considered a point source and angular dispersion should be taken into consideration. According to the statistical characteristics of source, the models of the spatially distributed source can be classified into two types: the incoherently distributed (ID) sources and the coherently distributed (CD) ones [23]. An ID source means that signals coming from different points of the same distributed source are assumed uncorrelated. The acoustic sources caused by the vibration of the front wheels and the rearview mirrors on a moving car can be considered examples of ID sources. A CD source means that the signal components from the same source are delayed and scaled replicas of one point in the source. The trumpet on a big loudspeaker can be considered an example of a CD source. The localization of the distributed source has been widely studied in the literature [6][27], but few works have been publicly reported on the enhancement techniques of a target distributed source. In [24] and [31] the proposed beamforming for a distributed source can achieve an optimal value of SINR due to

the Rayleigh quotient. However, the covariance matrix of the received signals emitted by the target source is required for the proposed method yet difficult to be obtained in practice. Thus, in our work, instead of utilizing the covariance matrix of the received signals, we assume that the steering vector of the distributed source is known a priori in the linear constraint in beamforming. We try to optimize the microphone array geometry to compensate for the performance degradation caused by the inaccuracy of the steering vector model.

The optimization of array geometry can be divided into two categories. The first category refers to choosing a subset of active microphones according to different criteria from a large scale of candidate microphones in a wireless acoustic sensor network (WASN), for the sake of improving the information transform efficiency and energy saving. In [28] the best subset of sensors is determined by minimizing the transmission cost while constraining the output noise power. In [8] the Camér-Rao Bound (CRB) is used as an optimization goal and the microphone subset selection problem is formulated as a sparsity based optimization problem. Similar work has been extended to correlated noise scenarios in [17]. In [5] and [22] the microphone subset selection criteria are the utility of each sensor and the SNR, respectively. The criteria are solved by greedy methods. The second category refers to the design of the geometry of the microphone array directly. For example, the V shape is optimum for ambiguity-free array [11]. The spiral array provides performance advantages over a wide frequency spectrum owing to the absence of inter-element spacing redundancy [1]. Beamforming with the famous differential microphone arrays (DMAs) can achieve a frequency invariant and super directive performance with a small inter-element spacing [3]. Recently, to make a compromise between the WNG and DF, a novel class of microphone arrays have been proposed [9] to decompose the steering vector of the source as a Kronecker product of two sub-steering vectors, one sub-steering vector corresponds to a microphone array with a smaller aperture and the other sub-steering vector corresponds to a microphone array with a larger inner space than the original one. The novel class of microphone array have been extended to circular [30], cuboid [26], and other geometries. Besides the Kronecker product method, one can also design optimal nonuniform linear DMAs with a specified target directivity pattern [7].

In our work, we first propose CD-MVDR beamforming by extending the conventional MVDR beamforming to the CD source scenario, and analyse theoretically the performance of CD-MVDR beamforming: with Taylor expansions, the WNG is found to be inversely proportional to the square of the difference between the angular dispersion of the actual source and that in the CD-MVDR beamforming model. As a second step, we propose a criterion to minimize the impact of the angular dispersion on the WNG across the whole angle sector of interest of the direction of arrival (DOA) of the SOI with a minimum inter-element space constraint. Although such a criterion is effective in theory, it is difficult to implement in practice due to the high computational burden. Therefore, an iterative algorithm is proposed to efficiently optimize

the array geometry, which locates the microphones one by one on the edge of the constraint region given by the located microphones. With the proposed locally-optimal geometry, both the WNG and the DF become more robust to the angular spread of the source.

This paper is organized as follows: the wide band CD source model and the conventional MVDR beamforming are briefly reviewed in section 2. The performance of CD-MVDR beamforming in the presence of the CD source is studied in section 3 and some simulation results are shown to validate the theoretical expressions. In section 4, we propose an array geometry optimization algorithm based on the results of the previous section. Finally, conclusions are given in section 5.

# 2 Signal model and MVDR beamforming

## 2.1 The CD source model

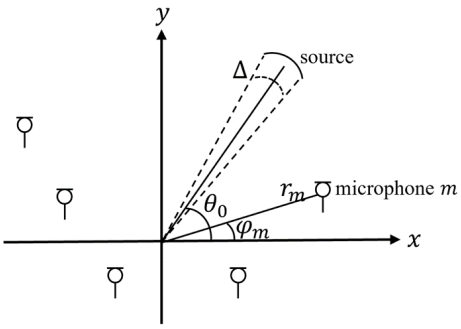



**Fig. 1**: Illustration of planar microphone array and spatially distributed source.

Let us consider a far-field wide band CD [23] source with angular spread $\Delta$ in an anechoic environment impinging on a microphone array from the DOA $\theta_0$ at the speed of sound, i.e., $c = 340m/s$. The microphone array is composed of $M$ sensors and assumed to be planar. With respect to the origin in the polar coordinate, the position of the $m^{th}$ microphone is given by the distance $r_m$ and the angular $\varphi_m$ (see Figure 1). The signal received at the $m^{th}$ microphone at moment $t$ is given by:

$$y_m(t) = \int_{-\pi}^{\pi} s(t + r_m \cos(\theta_0 + \phi - \varphi_m)/c) g(\phi) d\phi + n_m(t), \tag{1}$$

where $s(t)$ is the SOI, $n_m(t)$ is the noise observed at the $m^{th}$ microphone, $g(\phi)$ is introduced to describe the angular distribution of the CD source

$\int_{-\pi}^{\pi} g(\phi)d\phi = 1$. With the short time Fourier transform (STFT), assuming that the window length of the STFT is long enough, the wide band signal in frequency domain $\tilde{y}_m(\omega)$ can be given as:

$$\tilde{y}_m(\omega) = \int e^{j\omega r_m \cos(\theta_0+\phi-\varphi_m)/c} g(\phi)d\phi\tilde{s}(\omega) + \tilde{n}_m(\omega), \tag{2}$$

where $\tilde{y}_m(\omega)$, $\tilde{s}(\omega)$ and $\tilde{n}_m(\omega)$ are the STFT of $y_m(t)$, $s(t)$ and $n_m(t)$, respectively, $j$ is the imaginary unit with $j = \sqrt{-1}$. (2) can also be expressed in a vector form as:

$$\mathbf{y}(\omega) = \mathbf{c}(\theta_0, g, \omega)\tilde{s}(\omega) + \mathbf{n}(\omega), \tag{3}$$

where $\mathbf{y}(\omega) = [\tilde{y}_1(\omega), \ldots, \tilde{y}_M(\omega)]^T \in \mathbb{C}^{M\times 1}$, $\mathbf{n}(\omega) = [\tilde{n}_1(\omega), \ldots, \tilde{n}_M(\omega)]^T \in \mathbb{C}^{M\times 1}$, $\mathbf{c}(\theta_0, g, \omega) = [c_1(\theta_0, g, \omega), \ldots, c_M(\theta_0, g, \omega)]^T \in \mathbb{C}^{M\times 1}$, is the steering vector for a CD source which can be given by:

$$\mathbf{c}(\theta_0, g, \omega) = \int_{-\pi}^{\pi} \mathbf{a}(\theta_0 + \phi, \omega)g(\phi)d\phi, \tag{4}$$

where $\mathbf{a}(\theta_0, \omega)$ is the steering vector for a point source, which can be given by:

$$\mathbf{a}(\theta_0, \omega) = [e^{j\omega r_1 \cos(\theta_0-\varphi_1)/c}, ..., e^{j\omega r_M \cos(\theta_0-\varphi_M)/c}]^T, \tag{5}$$

where the superscript $^T$ is the transpose operator. The steering vector in (4) can be seen as a set of point sources impinging from $\theta_0 + \phi$ with the steering vector in (5) and weighted by the distribution function $g(\phi)$.

Assuming that the source and the additive noise are uncorrelated, with (3), the correlation matrix of the signals received at the microphone array is given by:

$$\begin{aligned}\mathbf{R}_y &\triangleq E[\mathbf{y}(\omega)\mathbf{y}^H(\omega)] \\ &= \sigma_s^2\mathbf{c}(\theta_0, g, \omega)\mathbf{c}^H(\theta_0, g, \omega) + \mathbf{R}_n(\omega),\end{aligned} \tag{6}$$

where $E[\cdot]$ denotes the mathematical expectation, the superscript $^H$ denotes the conjugate-transpose operator, $\sigma_s^2 \triangleq E[|S(\omega)|^2]$ is the power of the SOI, $\mathbf{R}_n(\omega) \triangleq E[\mathbf{n}(\omega)\mathbf{n}^H(\omega)]$ is the correlation matrix of the additive noise $\mathbf{n}(w)$ whose rank is $M$ in general. Notice that in the case of spatially white noise, $\mathbf{R}_n(\omega) = \sigma_n^2\mathbf{I}_M$, where $\sigma_n^2$ is the noise power, and $\mathbf{I}_M$ is a $M \times M$ diagonal matrix; while in the case of spherically isotropic noise, $\mathbf{R}_n(\omega) = \sigma_n^2\Gamma_n(\omega)$, where $[\Gamma_n(\omega)]_{ij} = \text{sinc}[\omega d_{ij}/c] = \sin(\omega d_{ij}/c)/(\omega d_{ij}/c)$, $[\Gamma_n(\omega)]_{ij}$ is the $(i, j)$th element of the matrix $\Gamma_n(\omega)$, and $d_{ij}$ is the distance between the $i^{th}$ and the $j^{th}$ microphones. Note that our work focuses on the CD source since no explicit expression exists for the steering vector in (4) for the ID source.

## 2.2 MVDR beamforming and its performance

The main idea of the conventional MVDR beamforming is given by minimizing the output power of the microphone array while preserving the target source. Assuming that a point source impinges from $\theta_0$, MVDR beamforming is formulated as:

$$\underset{\mathbf{h}(\omega)}{\operatorname{argmin}}\ \mathbf{h}^H(\omega)\mathbf{R}_y\mathbf{h}(\omega) \qquad \text{subject to} \quad \mathbf{h}^H(\omega)\mathbf{a}(\theta_0,\omega) = 1, \tag{7}$$

where $\mathbf{h}(\omega)$ can be considered as a finite impulse filter applied to each microphone output at angular frequency $w$.

In this paper, we investigate the performance of MVDR beamforming in the same sub-frequency band. Therefore, in the following, we omit the $\omega$ in the definitions for simplicity. The solution of (7) can be given as:

$$\mathbf{h} = \frac{\mathbf{R}_y^{-1}\mathbf{a}_0}{\mathbf{a}_0\mathbf{R}_y^{-1}\mathbf{a}_0}, \tag{8}$$

where $\mathbf{a}_0$ is a simplification of $\mathbf{a}(\theta_0,\omega)$.

The two important optimization goals of beamforming are given as follows:

**WNG:** The white noise gain is defined as the ratio between the output SNR and the input SNR of the microphone array in the scenario of spatially white noise, which also describes the robustness of the Beamformer:

$$\mathcal{G}_W(\mathbf{h}) = \frac{\sigma_n^2\mathbf{h}^H\mathbf{R}_s\mathbf{h}}{\sigma_s^2\mathbf{h}^H\mathbf{R}_n\mathbf{h}}. \tag{9}$$

In the case of one target source, the WNG can be given as $1/\mathbf{h}^H\mathbf{h}$.

**DF:** The directivity factor is defined as the ratio between the output SNR and the input SNR of the microphone array in the scenario of a diffuse sound field, which also indicates the beampattern gain at the look direction to the average beampattern gain at other directions:

$$\mathcal{G}_D(\mathbf{h}) = \frac{1}{\mathbf{h}^H\Gamma_n(\omega)\mathbf{h}}. \tag{10}$$

# 3 CD-MVDR beamforming

In this section, we first extend the conventional MVDR beamforming for the CD source scenario as CD-MVDR beamforming, and then study its performance.

## 3.1 CD-MVDR beamforming

The CD-MVDR beamforming is designed to enhance the signal of a CD source located at $\theta_B$ with a generalized steering vector $\mathbf{c}(\theta_B, g, \omega)$. In this case the

beamforming $\mathbf{h}(\omega)$ is the solution of:

$$\min_{\mathbf{h}(\omega)} \mathbf{h}^H(\omega)\mathbf{R}_y\mathbf{h}(\omega) \qquad \text{subject to} \quad \mathbf{h}^H(\omega)\mathbf{c}(\theta_B, \omega) = 1. \tag{11}$$

In the following, the steering vector model $\mathbf{c}(\theta_B, g, \omega)$ in the CD-MVDR beamforming is shortened as $\mathbf{c}_B$. As a consequence, the CD-MVDR beamforming filter is given as:

$$\mathbf{h} = \frac{\mathbf{R}_y^{-1}\mathbf{c}_B}{\mathbf{c}_B\mathbf{R}_y^{-1}\mathbf{c}_B}. \tag{12}$$

## 3.2 CD-MVDR beamforming performance

To further investigate the performance of CD-MVDR beamforming, we first assume that the angular spread of the CD source is not too large ($\leq 10^\circ$) in the far-field, and the distribution shape is symmetrical. Thus, $\int_{-\pi}^{\pi} g(\phi)d\phi = 1$, $\int_{-\pi}^{\pi} \phi g(\phi)d\phi = 0$. Considering the second order *Taylor approximation* in the angular spread, the steering vector for the CD source in (4) can be approximated by:

$$\begin{aligned} \mathbf{c}_0 &\approx \int_{-\pi}^{\pi} (\mathbf{a}_0 + \dot{\mathbf{a}}_0\phi + \ddot{\mathbf{a}}_0\phi^2)g(\phi)d\phi \\ &= \mathbf{a}_0 + \frac{1}{2}\ddot{\mathbf{a}}_0\delta_0^2, \end{aligned} \tag{13}$$

where the steering vector $\mathbf{c}(\theta_0, g, \omega)$ of the actual source is simplified as $\mathbf{c}_0$, $\mathbf{a}_0$ is the steering vectors of the point source with the same DOA of $\mathbf{c}_0$, $\dot{\mathbf{a}}_0 = \frac{\partial \mathbf{a}(\theta)}{\partial \theta}\,|_{\theta_0}$, $\ddot{\mathbf{a}}_0 = \frac{\partial^2 \mathbf{a}(\theta)}{\partial \theta^2}\,|_{\theta_0}$, $\delta_0^2 = \int \phi^2 g(\phi)d\phi$ is the angular dispersion of the distributed source. The first order term in $\phi$ is canceled due to the spatial symmetry of the distribution function of the source. Note that if the distribution function of one source is simple such as Uniform, Gaussian, or Raised Cosine, its angular dispersion can be expressed explicitly as:

$$\begin{aligned} \text{Uniform}: &\quad \delta_0^2 = \Delta_0^2/12, \\ \text{Gaussian}: &\quad \delta_0^2 = \Delta_0^2, \\ \text{Raised cosine}: &\quad \delta_0^2 = \frac{\pi^2 - 6}{12\pi^2}\Delta_0^2, \end{aligned} \tag{14}$$

where $\Delta_0$ is a parameter of the angular distribution function of the actual distributed source. For Uniform and Raised cosine distribution, $\Delta_0$ is the length of the support of the source; for Gaussian distribution, $\Delta_0$ is the standard deviation of the Gaussian function.

Similarly, the second order approximation of the steering vector in CD-MVDR beamforming (11) can be given as:

$$\mathbf{c}_B \approx \mathbf{a}_B + \frac{1}{2}\ddot{\mathbf{a}}_B \delta_B^2. \tag{15}$$

In the case that the *a-priori* information of the central DOA is well estimated, that is, $\theta_0 = \theta_B$ and thus $\mathbf{a}_0 = \mathbf{a}_B$, the relationship between (13) and (15) can be given as:

$$\mathbf{c}_0 = \mathbf{c}_B + \frac{1}{2}\ddot{\mathbf{a}}_0(\delta_0^2 - \delta_B^2). \tag{16}$$

We focus on the study of the CD-MVDR beamforming performance when the angular spread is roughly known. We aim to express the performance with respect to the angular dispersion error $\delta_0^2 - \delta_B^2$ where $\delta_0^2$ is the dispersion of the actual source and $\delta_B^2$ is the dispersion used by the CD-MVDR beamforming. One can notice that the case $\delta_B^2 = 0$ addresses the performance study of the standard MVDR beamforming in presence of a CD source.

In addition, with the *Woodbury formula* [14], the inverse of the correlation matrix $\mathbf{R}_y$ in (8) can be given by:

$$\mathbf{R}_y^{-1} = \mathbf{R}_n^{-1} - \frac{\mathbf{R}_n^{-1}\mathbf{c}_0\mathbf{c}_0^H\mathbf{R}_n^{-1}}{1 + \mathbf{c}_0^H\mathbf{R}_n^{-1}\mathbf{c}_0}. \tag{17}$$

Introducing (17) and (16) into (8), and keeping the second order *Taylor approximation* in $\delta$, the MVDR Beamformer can be given by:

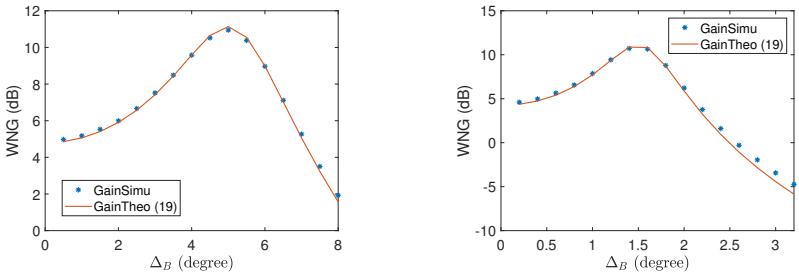
$$\mathbf{h} \approx \mathbf{h}_0 + \Delta\mathbf{h}, \tag{18}$$

where $\mathbf{h}_0 = \mathbf{R}_n^{-1}\mathbf{c}_B / \mathbf{c}_B^H\mathbf{R}_n^{-1}\mathbf{c}_B$ is the term depending on the central DOA of the source, the array geometry and the noise on the microphone array; $\Delta\mathbf{h} = (\delta_0^2 - \delta_B^2) \cdot \left(\mathbf{R}_n^{-1}\mathbf{c}_B \mathcal{R}eal\{\ddot{\mathbf{a}}_B\mathbf{R}_n^{-1}\mathbf{c}_B\} - \frac{1}{2}\mathbf{R}_n^{-1}\ddot{\mathbf{a}}_B\mathbf{c}_B^H\mathbf{R}_n^{-1}\mathbf{c}_B - \frac{1}{2}\mathbf{R}_n^{-1}\mathbf{c}_B\ddot{\mathbf{a}}_B^H\mathbf{R}_n^{-1}\mathbf{c}_B\right) / \mathbf{c}_B^H\mathbf{R}_n^{-1}\mathbf{c}_B$ is the term due to $(\delta_0^2 - \delta_B^2)$. Note that when the conventional MVDR beamforming is used, that is, the angular spread in the steering vector $\mathbf{c}_B$ in CD-MVDR beamforming equals to 0, $\mathbf{c}_B$ shrinks to $\mathbf{a}_B$, and $(\delta_0^2 - \delta_B^2)$ shrinks to $\delta_0^2$. In this case, $\Delta\mathbf{h}$ depends explicitly on the angular dispersion of the actual source.

Similarly, introducing (18) into (9), the approximation of WNG as an explicit function of the angular spread can be deduced as:

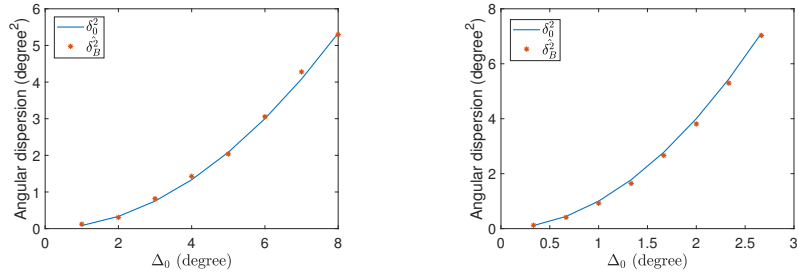
$$\tilde{\mathcal{G}}_W(\mathbf{h}, \delta_0^2, \delta_B^2) \approx \frac{M^2}{M\sigma_n^2 + (\delta_0^2 - \delta_B^2)^2 \cdot \left(0.25M^2\|\ddot{\mathbf{a}}_B\|^2 + 0.25M\|\mathbf{c}_B^H\ddot{\mathbf{a}}_B\|^2 - M\mathcal{R}eal\{\ddot{\mathbf{a}}_B^H\mathbf{c}_B\}^2 + 0.5M\mathcal{R}eal\{(\ddot{\mathbf{a}}_B^H\mathbf{c}_B)^2\}\right)}. \tag{19}$$

From (19) we can see that the WNG is inversely proportional to the square of the angular dispersion difference between the actual source and CD-MVDR




(a) Uniform CD source (b) Gaussian CD source

**Fig. 2**: WNG vs. angular spread model in CD-MVDR beamforming (Uniform CD source with $\Delta_0 = 5°$ and Gaussian CD source with $\Delta_0 = 1.5°$.)




(a) Uniform CD source (b) Gaussian CD source

**Fig. 3**: Angular dispersion vs. angular spread of actual CD source

beamforming model. Besides, (19) proves that the WNG can be used to correct the angular dispersion in the CD-MVDR beamforming model such that:

$$\hat{\delta_B^2} = \underset{\delta_B^2}{\operatorname{argmax}}\, \mathcal{G}_W(\mathbf{h}, \delta_0^2, \delta_B^2), \tag{20}$$

where $\mathcal{G}_W(\mathbf{h}, \delta_0^2, \delta_B^2)$ here is obtained in (9).

## 3.3 Numerical simulation results

In this subsection, we validate our theoretical results on CD-MVDR beamforming in (19) and (20). A female speaker reading a book as the source [20] impinges on the array from $\theta_0 = 45°$, a uniformly distributed source propagation model is generated according to (3). The total snapshots number is $2^{17}$, the sampling rate of the source is 16kHz, and the window length of the STFT is 256. A uniform linear array with $M = 13$ microphones is used. We work in the sub-frequency band of 1000Hz, $d = 0.2\lambda = 0.068m$, where $\lambda$ is the wavelength of the source. In all the simulations, the signal to noise ratio (SNR) is set to 10 dB. 100 Monte Carlo experiments are executed.

In Figure 2, we plot the WNG of CD-MVDR beamforming as a function of the angular spread $\Delta_B$ in the CD-MVDR beamforming steering vector model for a uniform CD source with $\Delta_0 = 5°$ (Figure 2(a)) and a Gaussian CD source with $\Delta_0 = 1.5°$ (Figure 2(b)), respectively. The angular distribution shapes in the CD-MVDR beamforming models are identical to the actual sources. We can see that our theoretical results obtained with (19) are generally consistent with the simulation results. The WNG reaches a maximum value when the angular dispersion of the actual source and that in the CD-MVDR beamforming model approaches, and then decreases when the gap between the actual angular dispersion and that in the model increases.

In Figure 3, we plot the actual and estimated angular dispersion as a function of the angular spread, for both uniform CD source and Gaussian CD source. The blue lines represent the actual angular dispersions obtained directly with (14), while the red stars represent the estimated angular dispersions obtained with (20). The distribution function in the CD-MVDR beamforming model is set to be uniform. We can see that by maximizing (20) the angular dispersion can be well estimated, even if the distribution function of the actual source and that of the model in CD-MVDR beamforming are mismatched (see Figure 3(b)).

# 4 Array geometry optimization algorithm

In the previous section, we can see that CD-MVDR beamforming requires *a-priori* information on the angular dispersion of the CD source. The performance of CD-MVDR beamforming degrades if the angular dispersion in its model is mismatched with that of the actual source. In this section, we optimize the microphone array geometry to improve the performance of the conventional MVDR beamforming in the CD source scenario, where the a priori information of the angular dispersion of the CD source is dispensable. For simplification, the microphone array geometry optimizations are given with the WNG. We can see in the simulations in section 4.1 that the proposed geometries can also improve the DF.

In the following, we seek to maximize the WNG by minimizing the influence of the angular spread on the WNG. Firstly, we define:

$$\alpha_m = \omega r_m \cos(\theta_0 - \varphi_m)/c,$$
$$\beta_m = \omega r_m \sin(\theta_0 - \varphi_m)/c,$$

where $m = 1, ..., M$.

Further, when the conventional MVDR beamforming is used, expression (19) remains valid with $\delta_B^2 = 0$ and $\mathbf{c}_B$ becomes equal to $\mathbf{a}_B$. We note the terms concerning the angular dispersion in $\tilde{\mathcal{G}}_W$ in (19) as $\mathcal{F}$ and they are given as:

$$\mathcal{F} = 0.25M^2\ddot{\mathbf{a}}_B^H\ddot{\mathbf{a}}_B + 0.25M\|\mathbf{a}_B^H\ddot{\mathbf{a}}_B\|^2 - M\mathcal{R}eal\{\ddot{\mathbf{a}}_B^H\mathbf{a}_B\}^2 + 0.5M\mathcal{R}eal\{(\ddot{\mathbf{a}}_B^H\mathbf{a}_B)^2\}. \tag{21}$$

To explore the intrinsic structure of $\mathcal{F}$, we introduce $\alpha_m$ and $\beta_m$ in (21). With *Euler's formula*, $\mathcal{F}$ can be given as:

$$\mathcal{F} = (\frac{\delta\omega}{c})^4 \left[ -\frac{1}{4}M(\sum \beta_m^2)^2 + \frac{1}{4}M^2 \sum \beta_m^4 + \frac{1}{4}M^2 \sum \alpha_m^2 - \frac{1}{4}M(\sum \alpha_m)^2 \right]. \tag{22}$$

According to *Cauchy inequality*, it can be easily seen that $\mathcal{F} \geq 0$, and the global minimum of $\mathcal{F}$ can be obtained when $\alpha_1 = ... = \alpha_m = ... = \alpha_M$, and $\beta_1 = ... = \beta_m = ... = \beta_M$. However, it is not possible to put all the microphones in the same place in practice. It is generally necessary to impose a constraint on the minimum distance between the microphones, for the sake of eliminating a coherence too high [21], alleviating poor noise sensitivity at low frequencies [10], as well as reducing the mutual coupling [15][16], etc. Therefore, in this paper, we also give a minimum distance between two neighbor microphones as a constraint for the optimization of the microphone geometry. Intuitively, a more compact array with a smaller aperture is more robust to the angular spread for the MVDR beamforming.

In addition, we can see that $\mathcal{F}$ is a function of $\theta_0$, since $\alpha_m$ and $\beta_m$ are both function of $\theta_0$. Therefore, minimizing $\mathcal{F}$ will lead to an optimal geometry for a fixed $\theta_0$. If the DOA of the source is changed, the performance of the MVDR beamforming will degrade as well. In this paper, we propose a criterion which optimizes the geometry globally across the whole angle sector of the possible DOA:

$$\mathcal{F}_g = \int_0^{2\pi} \mathcal{F} d\theta. \tag{23}$$

After tedious calculations, the global criterion $\mathcal{F}_g$ can be given by:

$$\begin{aligned} \mathcal{F}_g =& \frac{1}{4}\pi^3 M^2 (\|\mathbf{x}\|_2^2 + \|\mathbf{y}\|_2^2) - \frac{1}{4}\pi^3 M \left(\overline{\mathbf{x}}^2 + \overline{\mathbf{y}}^2\right) \\ &+ \frac{1}{4}\pi^5 M^2 \left(\frac{3}{4}\|\mathbf{x}\|_4^4 + \frac{3}{4}\|\mathbf{y}\|_4^4 + \frac{3}{2}\|\mathbf{x} \odot \mathbf{y}\|_2^2\right) \\ &- \frac{1}{4}\pi^5 M \left(\frac{3}{4}\|\mathbf{x}\|_2^4 + \frac{3}{4}\|\mathbf{y}\|_2^4 + \overline{\mathbf{x} \odot \mathbf{y}}^2 + \frac{1}{2}\|\mathbf{x}\|_2^2\|\mathbf{y}\|_2^2\right), \end{aligned} \tag{24}$$

where $\mathbf{x} = [x_1, ...x_m, ...x_M]$, $\mathbf{y} = [y_1, ...y_m, ...y_M]$, $x_m$ and $y_m$ are the Cartesian coordinates of the $m^{th}$ microphone, respectively, $\|\cdot\|_p$ is the operator of p-norm, $\bar{\cdot}$ is the mean value operator, $\odot$ is the element wise product operator. Therefore, the optimization of the microphone array geometry means finding the $\mathbf{x}$ and $\mathbf{y}$ satisfying:

$$\begin{aligned} &\{\mathbf{x}_{opt}, \mathbf{y}_{opt}\} = \arg\min_{\mathbf{x},\mathbf{y}} \|\mathcal{F}_g\|_2^2 \\ &\text{s.t.} \qquad \|(x_m, y_m) - (x_n, y_n)\|_2 \geq d, \quad \text{where} \qquad m, n \in [1, ..., M], m \neq n, \end{aligned} \tag{25}$$

**Table 1:** Iterative method for array optimization

**Parameters:** M,d
**Initialization:** $x_1 = y_1 = 0, x_2 = d, y_2 = 0$
$x_3 = ... = x_M = y_3 = ... = y_M = 0$
**For** $m = 3 : M$
$\{x_m, y_m\} = \arg\min_{x_m, y_m} \|\mathcal{F}_g(x_1, ..., x_m, y_1, ..., y_m)\|_2^2$
s.t. $\{x_m, y_m\} \in \text{edge}\{\mathcal{C}(1) \cup ... \cup \mathcal{C}(m-1)\}$
**end**

where $d$ is the minimum inter element space constraint that we impose a-priori. The criterion above is not convex and can be solved by greedy search across the whole region of interest for all the microphones with an extremely high computational burden. To overcome this problem, here we propose to solve (25) with an iterative method which can find at least one local solution of the optimization problem: Intuitively, from (22) we can see that all the microphones have the trend to be located as close as possible. Without loss of generality, we assume that the first microphone is placed at the origin and the second on the $x$ axis at $(d, 0)$. Defining that $\mathcal{C}(m)$ is the circle centered at the $m^{th}$ microphone with a radius $d$, the $m^{th}$ microphone will be located on the edge of the union of $\mathcal{C}(1)$,$\mathcal{C}(2)$,...,$\mathcal{C}(m-1)$, at the same time the criterion of minimizing $\mathcal{F}_g$ is satisfied for the first $m$ microphones. The iterative optimization algorithm is formulated as in table 1.

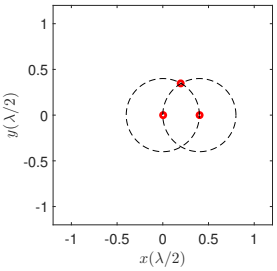
## 4.1 Numerical simulation results

In this subsection, we present the locally-optimal geometry obtained with the algorithm in (25) with various numbers of microphones and show the advantages of the proposed geometry by comparing the performance of the conventional MVDR beamforming using various array geometries.

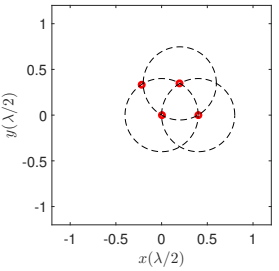
In Figure 4, the locally-optimal geometries are plotted for various values of the microphone number $M$. The red dots represent the microphones and the black dashed lines represent the circles centered at the microphones with radius $d = 0.2\lambda$; where $d$ is the minimum inter-element space constraint. We can see that the locally-optimal geometry is fan-shaped, and Figure 4(e) and 4(h) show that the locally-optimal geometry has the trend to be concentric.

In Figure 5, the WNG and DF as a two-dimensional function of the angular spread and the DOA of the actual source are illustrated. We can see that in Figure 5(a) the WNG is robust to the DOA of the source, as the proposed locally-optimal geometry is obtained by minimizing the impact of the angular spread of the CD source on the WNG of MVDR beamforming. On the contrary, in Figure 5(b) the DF is less robust to the DOA than the WNG. Similarly, Figure 5(c) and Figure 5(d) show the WNG and DF with spiral geometry, where the aperture of the spiral geometry is approximately equivalent to the proposed locally-optimal geometry. We can see the advantage of the proposed geometry for both the WNG and DF.

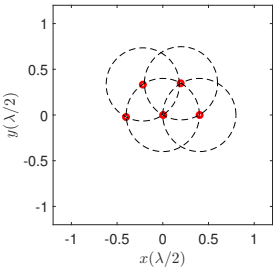
In Figure 6, we compare the WNG and the DF of the MVDR beamforming with a Fermat spiral array (Figure 6(a)), an arbitrary array (Figure 6(b)), and



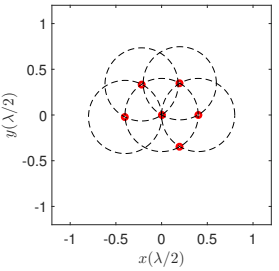

(a) M3



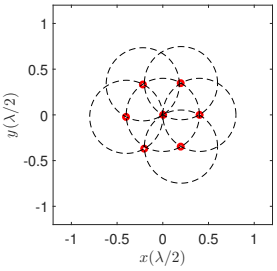

(b) M4



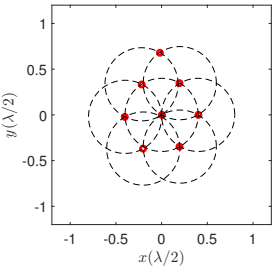

(c) M5



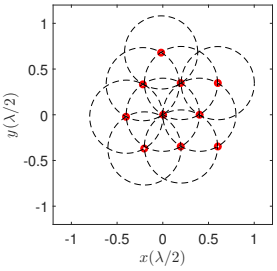

(d) M6



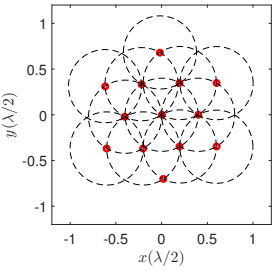

(e) M7




(f) M8

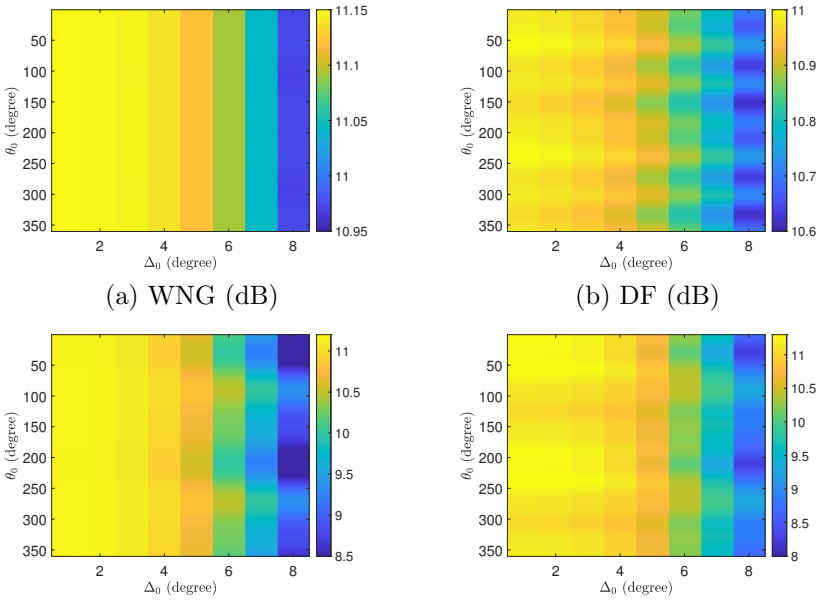



(g) M10




(h) M13

**Fig. 4**: Locally-optimal geometry with different number of microphones (inter-element space constraint $d = 0.2\lambda$)




(a) WNG (dB) (b) DF (dB)

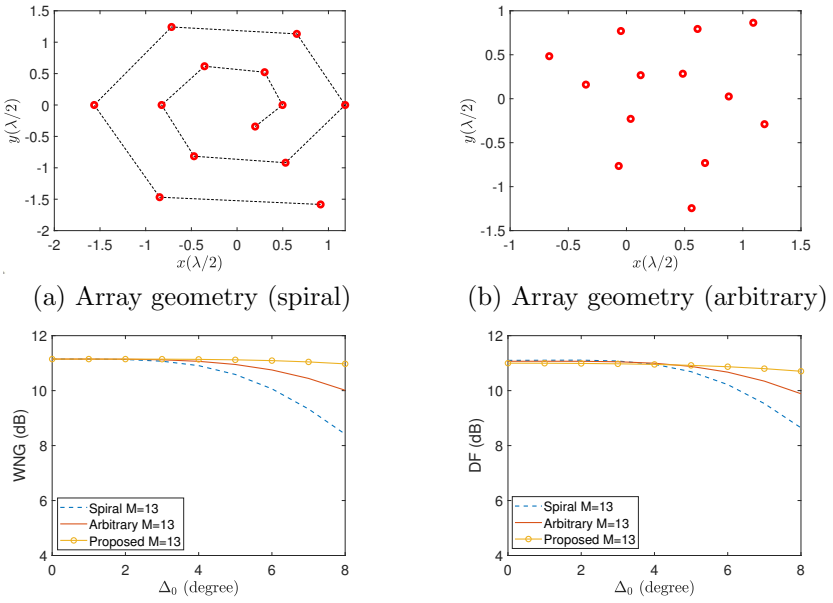
(c) WNG (dB) (d) DF (dB)

**Fig. 5**: WNG and DF vs. angular spread and DOA (uniformly distributed source, (a) and (b): the locally-optimal geometry with $M = 13$, (c) and (d): the spiral geometry with $M = 13$)

the proposed array. All the microphone arrays compose of $M = 13$ microphones and have the same minimum inter-element space constraint. The spherically isotropic noise is generated according to the method in [13]. In Figure 6(c) we can see that the performance degradation with the proposed geometry is significantly less than the others, while the performance degradation with the spiral geometry is the largest, which can be explained by the fact that under the same minimum inter-element space constraint, the proposed geometry has the minimum array aperture as well as the minimum value of $\mathcal{F}$ in (22). Moreover, in Figure 6(d) we can see that, even though our proposed criterion for the optimization of the array is based on the WNG, with the locally-optimal geometry the DF degrades significantly less than the others.

In Figure 7 and Figure 8, we compare the proposed method with the model-driven microphone subset selection (MSS) method proposed in [28]. As one of the state-of-the-art methods, the MSS method selecting the most informative subset of microphones from a WASN for MVDR beamforming. Its main idea is as follows:

$$\underset{\mathbf{h},\mathbf{p}\in\{0,1\}^M}{\operatorname{argmin}} \|\operatorname{diag}(\mathbf{p})\mathbf{c}\|_1$$



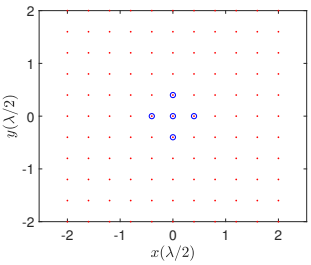

(a) Array geometry (spiral) (b) Array geometry (arbitrary)

(c) WNG (d) DF

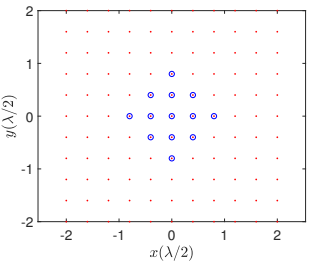
**Fig. 6**: WNG and DF vs. angular spread ($\theta_0 = 45°$, uniformly distributed source, $M = 13$)

$$\text{s.t.}\mathbf{h}^H\mathbf{R}_n\mathbf{h} \leq \frac{\beta}{\alpha},$$
$$\mathbf{h}^H\mathbf{a} = 1, \tag{26}$$

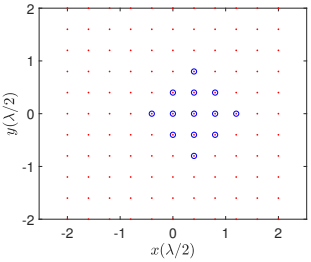
where $\mathbf{p} = [p_1, p_2, ..., p_M] \in \{0, 1\}^M$, and $p_i = 1$ means that the $i^{th}$ microphone is selected, $\mathbf{c} = [c_1, c_2, ..., c_M]^T$ denote the pairwise transmission cost between each microphone and the fusion center (FC) of the WASN, $\beta$ denotes the output noise power after beamforming when all microphones are used (i.e., $\mathbf{p} = \mathbf{1}_M$), $\alpha \in (0, 1]$ is an adaptive factor to control the output noise power compared to $\beta$. Figure 7 illustrates the array geometry examples obtained by the MSS method in different noise scenarios and with different values of $\alpha$. The FC is placed at the origin or at $(d, 0)$. The red dots represent the candidate microphones and the blue circles represent the selected microphones. $\alpha = 0.15$ in spatially white noise scenario (Figure 7(b) and 7(c)) and $\alpha = 0.25$ in spherically isotropic noise scenario (Figure 7(e) and 7(f)) are considered in Figure 8 to compare the SNR gain performance with our proposed geometry. With this configuration, the number of microphones is always equal to 13. We can see from Figure 7 and Figure 8 that the advantages of the proposed method lie in the following aspects: 1) The MSS method concerns choosing informative microphones from a large number of candidate microphones, while the proposed method only needs a small number of microphones. 2) For the



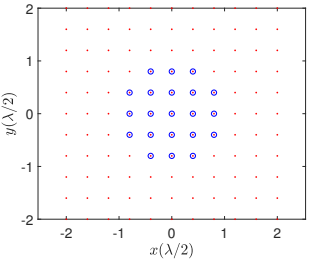

(a) Spatially white noise, $\alpha = 0.07$, $\theta_0 = 45°$, FC $(0, 0)$



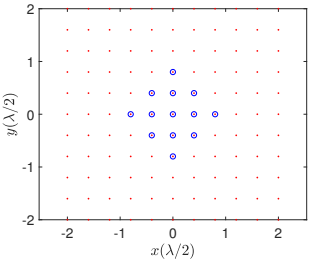

(b) Spatially white noise, $\alpha = 0.15$, $\theta_0 = 45°$, FC $(0, 0)$



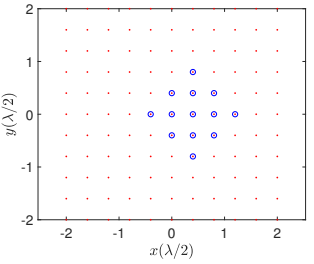

(c) Spatially white noise, $\alpha = 0.15$, $\theta_0 = 45°$, FC $(d, 0)$



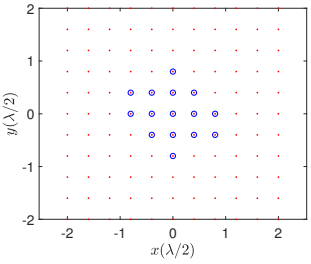

(d) Spatially white noise, $\alpha = 0.2$, $\theta_0 = 45°$, FC $(0, 0)$



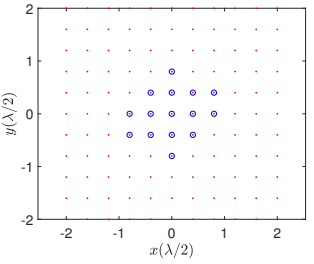

(e) Spherically isotropic noise, $\alpha = 0.25$, $\theta_0 = 45°$, FC $(0, 0)$




(f) Spherically isotropic noise, $\alpha = 0.25$, $\theta_0 = 45°$, FC $(d, 0)$

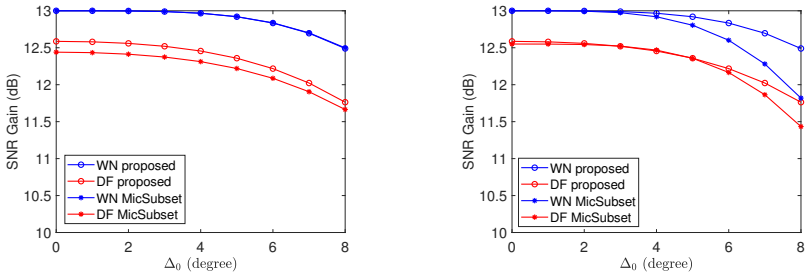



(g) Spherically isotropic noise, $\alpha = 0.3$, $\theta_0 = 45°$, FC $(0, 0)$




(h) Spherically isotropic noise, $\alpha = 0.3$, $\theta_0 = 135°$, FC $(0, 0)$

**Fig. 7**: Array geometries obtained by the model-driven MSS method in [28] in different noise scenarios (inter element space $d = 0.2\lambda$)




(a) FC $(0, 0)$ (b) FC $(d, 0)$

**Fig. 8**: WNG and DF vs. angular spread ($\theta_0 = 45°$, uniformly distributed source, spatially white noise, $M = 13$)

MSS method, the parameter $\alpha$ is introduced to control the output power of the network and can thus impact the array geometry (see the subfigures in Figure 7); and the position of the FC can also impact the array geometry (see Figure 7(e) and Figure 7(f)). The proposed method does not have such parameters. 3) The array geometry obtained by the MSS method is sensitive to the DOA of the SOI in some scenarios, see Figure 7(g) and Figure 7(h) for example, while the proposed criterion in our paper is robust to the DOA of the SOI. 4) Figure 8 illustrates that if the FC is placed at the origin, the proposed method performs approximately as well as the MSS method in the spatially white noise scenario and outperforms slightly the MSS method in the spherically isotropic noise scenario. However, if the FC is placed with a bias with respect to the origin, for example at $(d, 0)$, with the MSS method the robustness of beamforming to the angular distribution of the SOI decreases evidently.

## 5 Conclusions

In this paper, we have first extended the conventional MVDR beamforming into the CD source scenario as the CD-MVDR beamforming; and we have found that the WNG of CD-MVDR beamforming is inversely proportional to the square of the angular dispersion difference between the actual source and the CD-MVDR beamforming model. In the second part, we have proposed an efficient optimization method of microphone array geometry based on the minimization of the WNG of the conventional MVDR beamforming. We have found that the locally-optimal geometry is fan-shaped and has the trend to be concentric with one microphone at the center. Numerical results have shown that the proposed locally-optimal geometry is more robust than the MVDR beamforming in the CD source scenario for both the WNG and the DF.

## Data availability statement

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

## Declaration

The authors declare that they have no conflict of interest.

## Fundings

This work was supported by the National Natural Science Foundation of China (Grant No.61831019, No.61971015, and No.62101013).

## References

[1] M. R. Aldeman. A hybrid spiral microphone array design for performance and portability. *Applied Acoustics*, 170(4–5), 2020.

[2] J. Benesty, J. Chen, and Y. Huang. Microphone array signal processing. *Springer-Verlag Berlin Heidelberg*, 2008.

[3] Jacob Benesty and Jingdong Chen. Study and design of differential microphone arrays. *Springer Topics in Signal Processing*, 6, 2013.

[4] Jacob Benesty, Jingdong Chen, Yiteng Huang, and Boaz Rafaely. Microphone array signal processing. *Journal of the Acoustical Society of America*, 125(6):4097–4098, 2009.

[5] Alexander Bertrand and Marc Moonen. Efficient sensor subset selection and link failure response for linear mmse signal estimation in wireless sensor networks. In *2010 18th European Signal Processing Conference*, pages 1092–1096, 2010.

[6] T. F. Brooks and W. M. Humphreys. A deconvolution approach for the mapping of acoustic sources (damas) determined from phased microphone arrays. *Journal of Sound and Vibration*, 294(4):856–879, 2006.

[7] Jingdong Chen, Jacob Benesty, and Chao Pan. On the design and implementation of linear differential microphone arrays. *Journal of the Acoustical Society of America*, 136(6):3097, 2014.

[8] Sundeep Prabhakar Chepuri and Geert Leus. Sparsity-promoting sensor selection for non-linear measurement models. *IEEE Transactions on Signal Processing*, 63(3):684–698, 2015. doi: 10.1109/TSP.2014.2379662.

[9] Israel Cohen, Jacob Benesty, and Jingdong Chen. Differential kronecker product beamforming. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, pages 892–902, 2019.

[10] R. M. M. Derkx and K. Janse. Theoretical analysis of a first-order azimuth-steerable superdirective microphone array. *IEEE Transactions on Audio Speech and Language Processing*, 17(1):150–162, 2009.

[11] Houcem Gazzah and Karim Abed-Meraim. Optimum ambiguity-free directional and omnidirectional planar antenna arrays for doa estimation. *IEEE Transactions on Signal Processing*, 57(10):3942–3953, 2009. doi: 10.1109/TSP.2009.2023943.

[12] Yujie Gu, Nathan A. Goodman, Shaohua Hong, and Yu Li. Robust adaptive beamforming based on interference covariance matrix sparse reconstruction. *Signal Processing*, 96:375 – 381, 2014.

[13] E.A.P. Habets and S.Gannot. Generating sensor signals in isotropic noise fields. *J. Acoust. Soc. Amer.*, 122:3464–3470, 2007.

[14] Roger Horn and Charles Johnson. Matrix analysis. *Cambridge University Press*, 2012.

[15] C. Liu and P. P. Vaidyanathan. Super nested arrays: Linear sparse arrays with reduced mutual coupling—part i: Fundamentals. *IEEE Transactions on Signal Processing*, 64(15):3997–4012, 2016.

[16] C. Liu and P. P. Vaidyanathan. Super nested arrays: Linear sparse arrays with reduced mutual coupling—part ii: High-order extensions. *IEEE Transactions on Signal Processing*, 64(16):4203–4217, 2016.

[17] S. Liu, S. P. Chepuri, M. Fardad, E. Maşazade, G. Leus, and P. K. Varshney. Sensor selection for estimation with correlated measurement noise. *IEEE Transactions on Signal Processing*, 64(13):3509–3522, 2016. doi: 10.1109/TSP.2016.2550005.

[18] Y. Liu, C. Liu, Y. Zhao, and J. Zhu. Robust adaptive wideband beamforming against direction and sensor location errors. *Circuits, Systems, and Signal Processing*, 38:664–681, 2019.

[19] S. Mohammadzadeh and O. Kukrer. Robust adaptive beamforming with improved interferences suppression and a new steering vector estimation based on spatial power spectrum. *Circuits Systems and Signal Processing*, 2019.

[20] Vassil Panayotov, Guoguo Chen, Daniel Povey, and Sanjeev Khudanpur. Librispeech: An asr corpus based on public domain audio books. In *ICASSP 2015 - 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2015.

[21] K. U. Simmer and A. Wasiljeff. Adaptive microphone arrays for noise suppression in the frequency domain. *Second Cost Workshop on Adaptive Algorithm in Communication*, 1992.

[22] Joseph Szurley, Alexander Bertrand, Marc Moonen, Peter Ruckebusch, and Ingrid Moerman. Energy aware greedy subset selection for speech enhancement in wireless acoustic sensor networks. In *2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*, pages 789–793, 2012.

[23] S. Valaee, B. Champagne, and P. Kabal. Parametric localization of distributed sources. *IEEE Transactions on Signal Processing*, 43(9): 2144–2153, 1995.

[24] X. Wang, I. Cohen, J. Chen, and J. Benesty. On robust and high directive beamforming with small-spacing microphone arrays for scattered sources. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 27 (4):842–852, 2019.

[25] X. Wang, W. Liu, M. Jin, and S. Ding. Parameter-free landweber iteration method for robust adaptive beamforming. *Circuits Systems and Signal Processing*, 39(4), 2020.

[26] Xuehan Wang, Jacob Benesty, Gongping Huang, Jingdong Chen, and Israel Cohen. Design of kronecker product beamformers with cuboid

microphone arrays. 09 2019.

[27] W. Xiong, J. Picheral, G. Chardon, and S. Marcos. Sparsity-based localization of spatially coherent distributed sources. In *IEEE International Conference on Acoustics*, pages 3241–3245, 2016.

[28] J. Zhang, S. P. Chepuri, R. C. Hendriks, and R. Heusdens. Microphone subset selection for mvdr beamformer based noise reduction. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 26(3):550–563, 2018. doi: 10.1109/TASLP.2017.2786544.

[29] P. Zhang, Z. Yang, G. Jing, and T. Ma. Adaptive beamforming via desired signal robust removal for interference-plus-noise covariance matrix reconstruction. *Circuits, Systems, and Signal Processing*, 40(1):401–417, 2021.

[30] Jiahong Zhao and Christian H. Ritz. Co-prime circular microphone arrays and their application to direction of arrival estimation of speech sources. In *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019.

[31] Zhi Zheng, Yueping Fu, and Wen-Qin Wang. Sparse array beamforming design for coherently distributed sources. *IEEE Transactions on Antennas and Propagation*, 69(5):2628–2636, 2021.