Speech Communication 142 (2022) 61–76

Contents lists available at ScienceDirect

Speech Communication

journal homepage: www.elsevier.com/locate/specom

# Multistage approach for steerable differential beamforming with rectangular arrays

Gal Itzhak [a,*], Jacob Benesty [b], Israel Cohen [a]

[a] Andrew and Erna Viterby Faculty of Electrical and Computer Engineering, Technion – Israel Institute of Technology, Technion City, Haifa 3200003, Israel
[b] INRS-EMT, University of Quebec, 800 de la Gauchetiere Ouest, Montreal, QC H5A 1K6, Canada

ARTICLE INFO

*Keywords:*
Microphone arrays
Uniform rectangular arrays (URAs)
Differential beamforming
Two-dimensional (2-D) arrays

ABSTRACT

This paper presents a multistage rectangular approach for steerable differential beamforming. As a first step, we propose employing a two-dimensional (2-D) differentiation scheme that operates independently on the columns and rows of a uniform rectangular array (URA). This yields a differentials matrix controlled by two parameters, $P_c$ and $P_r$, which indicate the number of differential stages for the URA columns and rows. Then, as a second step, we design a rectangular differential beamformer and apply it to the vector form of the differentials matrix. We show that the proposed differentiation scheme may significantly improve the directivity of the resulted beamformer at the expense of white noise amplification. This improvement is heavily tied to selecting the $(P_c, P_r)$ configuration per the desired signal incident angle. Next, we propose four rectangular differential beamformers and analyze their performances in terms of the white noise gain (WNG) and directivity factor (DF) measures. In addition, we address reverberant scenarios with three distinct incident angles of the desired signal. We examine the performances of each beamformer in terms of four reduction factors calculated from the noisy and enhanced signals and investigate their quality and intelligibility. We demonstrate that the proposed rectangular differential beamformers outperform common linear and differential beamformers in these measures, mainly when the incident angle is far from the endfire direction. Finally, we compare the proposed approach to two existing rectangular differential beamforming approaches from the literature. We show the proposed method to be preferable, considering both the quality and intelligibility of the enhanced signals.

## 1. Introduction

Observed signals in communication systems are likely to be degraded by undesired noise and reverberations, which might heavily deteriorate the performance of such systems. These include, for example, speech and audio appliances, internet-of-things devices, and larger-scale systems such as radars and sonars. Sensor arrays are employed to simultaneously capture signals at different locations in space while aiming to attenuate undesirable signal components. The captured signals feed spatial filters, typically referred to as "beamformers". Beamformers, and their properties in time, frequency, and space, have been widely studied and optimized for different criteria and geometric configurations (Johnson and Dudgeon, 1992; Van Trees, 2004; Bai et al., 2014).

Differential microphone arrays (DMAs) have received particular attention within the sensor array processing framework due to their small physical size and frequency-invariant beampattern, two appealing characteristics for practical purposes DMA (Elko, 2004; Kolundzija et al., 2011; Sena et al., 2012; Benesty and Chen, 2013; Chen et al., 2014; Bernardini et al., 2018). DMAs were initially inspired by differentiating the acoustic pressure of successive microphones in the time domain (Weinberger et al., 1933; Olson, 1946), potentially in a multistage manner. These type of beamformers were later modified to overcome their inherent lack of robustness to microphone imperfections (Buck, 2002; Wu and Chen, 2016). These modifications typically considered the input observed signals in the short-time Fourier transform (STFT) domain, in which a DMA design was formulated as a single-stage linear equations system (Benesty et al., 2012, 2017).

Due to their simplicity and easy-to-analyze nature, differential uniform linear arrays (ULAs) have been most commonly addressed in the literature (Benesty et al., 2009; Borra et al., 2019; Benesty et al., 2019; Cohen et al., 2019). Unfortunately, they suffer from a few inherent drawbacks. For example, it is well known that to attain a high level of directivity, the desired signal is preferably in the endfire direction (Chen et al., 2014). In addition, ULAs suffer from a lower–upper

* Corresponding author.
*E-mail address:* galitz@campus.technion.ac.il (G. Itzhak).

https://doi.org/10.1016/j.specom.2022.06.006
Received 17 December 2021; Received in revised form 18 June 2022; Accepted 23 June 2022
Available online 27 June 2022
0167-6393/© 2022 Elsevier B.V. All rights reserved.

plane ambiguity: the beampattern of any ULA is always symmetric concerning the imaginary line connecting the sensors of the array. Therefore, more sophisticated geometric structures were explored, out of which differential uniform circular arrays (UCAs) have drawn the most attention (Huang et al., 2017; Buchris et al., 2018; Huang et al., 2020c). Other studies either assumed square-symmetric differential arrays (Zhao et al., 2021), whose design flexibility and hence their practicability is limited by construction, or exploited the Jacobi–Anger expansion approximation to refer to differential beamforming with arbitrarily-shaped planar arrays (Huang et al., 2018, 2020d). The latter approaches did not assume any regular array shape but merely required the positions of the array sensors to be either known in advance or measurable. While they are general, they are susceptible to selecting the expansion's reference point and may result in frequency-variant beampatterns as the array size increases. Therefore, they might not embody a proper beamforming design approach with symmetric array geometries, for which it may be possible to take advantage of the symmetry to circumvent these drawbacks.

Rectangular-shaped arrays are symmetric and valuable structures, which may be used to design differential beamformers with asymmetric beampatterns (Van Trees, 2004). On top of relaxing the aforementioned square-symmetry constraints, such arrays may also be designed flexibly. For example, a uniform rectangular beamformer may always be decomposed into two sub-beamformers by employing the Kronecker-product (KP) decomposition. This allows some flexibility: the KP decomposition is not unique, and each sub-beamformer may be independently designed for a different criterion (Huang et al., 2020a). An alternative parametric-configurable approach (Itzhak et al., 2021b) exploits the rectangular geometry to improve white noise robustness (i.e., to reduce white noise amplification) at the expense of array directivity. However, with this approach, the beam steering property of its corresponding beamformers (i.e., the ability to alter the direction to which the beampattern points) is not considered.

This paper presents a multistage rectangular approach for steerable differential beamforming. As a first step, we propose to employ a 2-D differentiation scheme that operates independently on the columns and rows of the observed signals of a URA. This yields a differentials matrix controlled by two parameters, $P_c$ and $P_r$, which indicate the number of differential stages for the URA columns and rows, respectively. Then, as a second step, we design a rectangular differential beamformer and apply it to the vector form of the differentials matrix. This approach may be seen as the URA generalization of the work in Huang et al. (2020b) and Itzhak et al. (2021a), which addresses ULAs. We show that the proposed differentiation scheme may significantly improve the directivity of the resulted beamformer at the expense of white noise amplification. This improvement is heavily tied to selecting the $(P_c, P_r)$ configuration per the desired signal incident angle. Next, we propose four rectangular differential beamformers and analyze their performance in the WNG and DF measures. In addition, we address reverberant scenarios with three distinct incident angles of the desired signal. We examine the versions of each presented beamformer in terms of four reduction factors calculated from the noisy and enhanced signals in the time domain and investigate their quality and intelligibility. We demonstrate that the proposed rectangular differential beamformers outperform common linear differential beamformers in these measures, mainly when the incident angle is far from the endfire direction. Finally, we compare the proposed approach to two existing rectangular differential beamforming approaches from the literature and show the proposed approach to be preferable — considering both the quality and intelligibility of the enhanced signals.

The rest of the paper is organized as follows. In Section 2, we discuss the rectangular array signal model. We define the signals of interest and formulate the signal model in vector and matrix forms. In Section 3, we present a multistage differential scheme that independently operates on both axes of the rectangular array. We formulate the resulting 2-D differentials matrix controlled by the $(P_c, P_r)$ configuration and express the corresponding signal-to-noise ratio (SNR) gains between the observed differentials and the noisy observations. Then, in Section 4, we discuss the application of a rectangular differential beamformer onto the vector form of the differentials matrix. We derive the standard performance measures, including the WNG, DF, and the power beampattern. Section 5 is dedicated to deriving four types of multistage rectangular differential beamformers. Section 6 consists of three parts. The first part analyzes the WNG and DF measures. The second part compares the proposed multistage rectangular differential beamformers to their existing linear counterparts through simulations with speech signals in reverberant environments and varying incident angles of the desired speech. The third part compares the proposed beamformers to existing rectangular differential beamformers from the literature. Finally, we summarize this study in Section 7.

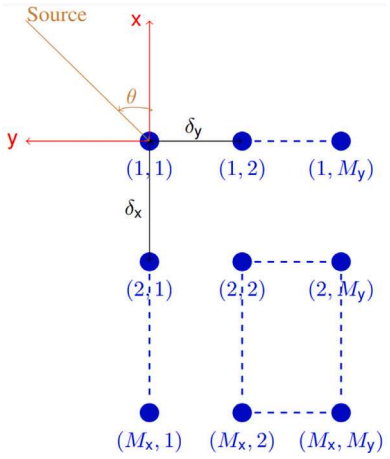


**Fig. 1.** Illustration of the studied rectangular microphone array. Dashed lines indicate an arbitrary number of microphones along an array axis.

## 2. Signal model

Consider a 2-D microphone URA. Given the Cartesian coordinate system with microphone $(1,1)$ as its origin, the URA is composed of $M_x$ omnidirectional sensors along the x (negative) axis with a uniform interelement spacing equal to $\delta_x$ and $M_y$ omnidirectional sensors along the y (negative) axis with a uniform interelement spacing equal to $\delta_y$. Thus, the total number of microphones is equal to $M_x M_y$, whose positions are denoted $(m_x, m_y)$, with $m_x = 1, 2, \ldots, M_x$ and $m_y = 1, 2, \ldots, M_y$. Notice that in the direction of the x axis, we have $M_y$ parallel ULAs composed of $M_x$ microphones each with a spacing $\delta_x$, while in the direction of the y axis, we have $M_x$ parallel ULAs composed of $M_y$ microphones each with a spacing $\delta_y$. An illustration of the 2-D URA studied in this paper is depicted in Fig. 1.

We assume that a farfield desired source signal (plane wave), on the same plane of the 2-D array, propagates from the azimuth angle, $\theta$, in an anechoic acoustic environment at the speed of sound, i.e., $c = 340$ m/s, and impinges on the above described array. Then, the corresponding steering matrix (of size $M_x \times M_y$) is (Van Trees, 2004):

$$\begin{aligned}\mathbf{D}_\theta(\omega) &= \left[\begin{array}{ccc} B_{\theta,1}(\omega)\,\mathbf{a}_\theta(\omega) & \cdots & B_{\theta,M_y}(\omega)\,\mathbf{a}_\theta(\omega)\end{array}\right] \\ &= \mathbf{b}_\theta^T(\omega) \otimes \mathbf{a}_\theta(\omega), \end{aligned} \tag{1}$$

where

$$\begin{aligned}\mathbf{a}_\theta(\omega) &= \left[\begin{array}{cccc} A_{\theta,1}(\omega) & A_{\theta,2}(\omega) & \cdots & A_{\theta,M_x}(\omega)\end{array}\right]^T \\ &= \left[\begin{array}{cccc} 1 & e^{-\jmath\varpi_{\theta,x}(\omega)} & \ldots & e^{-\jmath(M_x-1)\varpi_{\theta,x}(\omega)}\end{array}\right]^T \end{aligned} \tag{2}$$

is the steering vector associated with the x axis,

$$\mathbf{b}_\theta(\omega) = \begin{bmatrix} B_{\theta,1}(\omega) & B_{\theta,2}(\omega) & \cdots & B_{\theta,M_y}(\omega) \end{bmatrix}^T = \begin{bmatrix} 1 & e^{-\jmath\varpi_{\theta,y}(\omega)} & \cdots & e^{-\jmath(M_y-1)\varpi_{\theta,y}(\omega)} \end{bmatrix}^T \tag{3}$$

is the steering vector associated with the y axis,

$$\varpi_{\theta,x}(\omega) = \frac{\omega\delta_x\cos\theta}{c},$$
$$\varpi_{\theta,y}(\omega) = \frac{\omega\delta_y\sin\theta}{c},$$

the superscript $^T$ denotes the transpose operator, $\otimes$ is the KP operator, $\jmath = \sqrt{-1}$ is the imaginary unit, $\omega = 2\pi f$ is the angular frequency, and $f > 0$ is the temporal frequency.

Exploiting (1), the observed signal matrix of size $M_x \times M_y$ of the URA can be expressed in the frequency domain as (Benesty et al., 2017):

$$\begin{aligned}\mathbf{Y}(\omega) &= \mathbf{X}(\omega) + \mathbf{V}(\omega)\\ &= \mathbf{D}_\theta(\omega)X(\omega) + \mathbf{V}(\omega),\end{aligned} \tag{4}$$

where $X(\omega)$ is the zero-mean desired source signal and $\mathbf{V}(\omega)$ is the zero-mean additive noise signal matrix.

It is also convenient to express (4) in a vector form. Defining the steering vector $\mathbf{d}_\theta(\omega)$ of length $M_xM_y$, which is formed by concatenating the columns of $\mathbf{D}_\theta(\omega)$, by:

$$\mathbf{d}_\theta = \mathbf{b}_\theta \otimes \mathbf{a}_\theta, \tag{5}$$

we have

$$\begin{aligned}\mathbf{y}(\omega) &= \begin{bmatrix} \mathbf{y}_1^T(\omega) & \mathbf{y}_2^T(\omega) & \cdots & \mathbf{y}_{M_y}^T(\omega) \end{bmatrix}^T\\ &= \mathbf{x}(\omega) + \mathbf{v}(\omega)\\ &= \mathbf{d}_\theta(\omega)X(\omega) + \mathbf{v}(\omega),\end{aligned} \tag{6}$$

where

$$\begin{aligned}\mathbf{y}_{m_y}(\omega) &= \begin{bmatrix} Y_{m_y,1}(\omega) & \cdots & Y_{m_y,M_x}(\omega) \end{bmatrix}^T\\ &= \mathbf{x}_{m_y}(\omega) + \mathbf{v}_{m_y}(\omega)\\ &= B_{\theta,m_y}(\omega)\mathbf{a}_\theta(\omega)X(\omega) + \mathbf{v}_{m_y}(\omega),\end{aligned} \tag{7}$$

for $m_y = 1, 2, \ldots, M_y$, is the observed signal vector of length $M_x$ of the $m_y$th ULA parallel to the x axis, and $\mathbf{v}(\omega)$ and $\mathbf{v}_{m_y}(\omega)$ are defined in a similar manner. Dropping the dependence on $\omega$ to simplify the notation and assuming a distinct incident angle $\theta_s$, Eq. (6) becomes:

$$\mathbf{y} = \left(\mathbf{b}_{\theta_s} \otimes \mathbf{a}_{\theta_s}\right)X + \mathbf{v}, \tag{8}$$

where $\mathbf{b}_{\theta_s} \otimes \mathbf{a}_{\theta_s} = \mathbf{d}_{\theta_s}$ is the steering vector at the incident angle $\theta_s$, and the covariance matrix of $\mathbf{y}$ is:

$$\mathbf{\Phi}_\mathbf{y} = E\left(\mathbf{y}\mathbf{y}^H\right) = \phi_X\mathbf{d}_{\theta_s}\mathbf{d}_{\theta_s}^H + \mathbf{\Phi}_\mathbf{v}, \tag{9}$$

where $E(\cdot)$ denotes mathematical expectation, the superscript $^H$ is the conjugate-transpose operator, $\phi_X = E\left(|X|^2\right)$ is the variance of $X$, and $\mathbf{\Phi}_\mathbf{v} = E\left(\mathbf{v}\mathbf{v}^H\right)$ is the covariance matrix of $\mathbf{v}$. Assuming that the variance of the noise is approximately the same at all sensors, we can express (9) as:

$$\mathbf{\Phi}_\mathbf{y} = \phi_X\mathbf{d}_{\theta_s}\mathbf{d}_{\theta_s}^H + \phi_V\mathbf{\Gamma}_\mathbf{v}, \tag{10}$$

where $\phi_V$ is the variance of the noise at the reference microphone (i.e., the origin of the Cartesian coordinate system) and $\mathbf{\Gamma}_\mathbf{v} = \mathbf{\Phi}_\mathbf{v}/\phi_V$ is the pseudo-coherence matrix of the noise. From (10), we deduce that the input SNR is:

$$\text{iSNR} = \frac{\text{tr}\left(\phi_X\mathbf{d}_{\theta_s}\mathbf{d}_{\theta_s}^H\right)}{\text{tr}\left(\phi_V\mathbf{\Gamma}_\mathbf{v}\right)} = \frac{\phi_X}{\phi_V}, \tag{11}$$

where $\text{tr}(\cdot)$ denotes the trace of a square matrix. In the case of the spherically isotropic (diffuse) noise field, (10) becomes:

$$\mathbf{\Phi}_\mathbf{y} = \phi_X\mathbf{d}_{\theta_s}\mathbf{d}_{\theta_s}^H + \phi\mathbf{\Gamma}_\text{d}, \tag{12}$$

where $\phi$ is the variance of the diffuse noise and $\mathbf{\Gamma}_\text{d}$ is the pseudo-coherence matrix of the diffuse noise. We have

$$\mathbf{\Gamma}_\text{d} = \begin{bmatrix} \mathbf{\Gamma}_{\text{d},1} & \mathbf{\Gamma}_{\text{d},2} & \cdots & \mathbf{\Gamma}_{\text{d},M_y}\\ \mathbf{\Gamma}_{\text{d},2} & \mathbf{\Gamma}_{\text{d},1} & \cdots & \mathbf{\Gamma}_{\text{d},M_y-1}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{\Gamma}_{\text{d},M_y} & \mathbf{\Gamma}_{\text{d},M_y-1} & \cdots & \mathbf{\Gamma}_{\text{d},1} \end{bmatrix}, \tag{13}$$

which is a symmetric block Toeplitz matrix, and the elements of the $M_y$ symmetric Toeplitz matrices $\mathbf{\Gamma}_{\text{d},m_y}$, $m_y = 1, 2, \ldots, M_y$ (of size $M_x \times M_x$) are given by Johnson and Dudgeon (1992):

$$\left(\mathbf{\Gamma}_{\text{d},m_y}\right)_{ij} = \text{sinc}\left[\frac{\omega\sqrt{(i-j)^2\delta_x^2 + \left(m_y-1\right)^2\delta_y^2}}{c}\right], \tag{14}$$

with $i, j = 1, 2, \ldots, M_x$ and $\text{sinc}(x) = \sin x/x$.

Alternatively, in case of the spatially white sensor noise, the pseudo-coherence matrix of the noise reduces to $\mathbf{\Gamma}_\mathbf{v} = \mathbf{I}_{M_xM_y}$, where $\mathbf{I}_{M_xM_y}$ is the $M_xM_y \times M_xM_y$ identity matrix.

## 3. Multistage rectangular differentials

In this section, we propose a multistage differential scheme that operates on both axes of the rectangular array. It serves as the first step of the proposed differential beamforming approach. We point out that multistage differentials were previously presented in Huang et al. (2020b) and Itzhak et al. (2021a) in the context of linear beamforming, in which beam steering was not considered. In this work, we extend the concept of the multistage differential to rectangular arrays, which will be shown as critical in case the desired signal is not impinging on the array from the endfire direction.

Let us consider the signal model given in (7). We define the first-order forward spatial difference of $\mathbf{y}_{m_y}$ $(m_y = 1, 2, \ldots, M_y)$ as:

$$\begin{aligned}\Delta Y_{m_y,i} &= Y_{m_y,i+1} - Y_{m_y,i}\\ &= Y_{m_y,(1),i},\ i = 1, 2, \ldots, M_x - 1,\end{aligned} \tag{15}$$

where $\Delta$ is the forward spatial difference operator. In a vector/matrix form, (15) is:

$$\mathbf{\Delta}_{(1)}\mathbf{y}_{m_y} = \mathbf{y}_{m_y,(1)}, \tag{16}$$

where

$$\mathbf{\Delta}_{(1)} = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0\\ 0 & -1 & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \ddots & \vdots\\ 0 & 0 & \cdots & -1 & 1 \end{bmatrix} \tag{17}$$

is a matrix of size $(M_x - 1) \times M_x$. In the same way, the second-order forward spatial difference of $\mathbf{y}_{m_y}$ is:

$$\begin{aligned}\Delta^2 Y_{m_y,i} &= \Delta\left(\Delta Y_{m_y,i}\right) = \Delta Y_{m_y,i+1} - \Delta Y_{m_y,i}\\ &= Y_{m_y,i+2} - 2Y_{m_y,i+1} + Y_{m_y,i} = Y_{m_y,(2),i}\end{aligned} \tag{18}$$

with $i = 1, 2, \ldots, M_x - 2$, which can be rewritten as:

$$\mathbf{\Delta}_{(2)}\mathbf{y}_{m_y} = \mathbf{y}_{m_y,(2)}, \tag{19}$$

where

$$\mathbf{\Delta}_{(2)} = \begin{bmatrix} 1 & -2 & 1 & 0 & \cdots & 0\\ 0 & 1 & -2 & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots\\ 0 & 0 & \cdots & 1 & -2 & 1 \end{bmatrix} \tag{20}$$

is a matrix of size $(M_x - 2) \times M_x$. More generally, let $p = 0, 1, \ldots, P_c$, with $1 \le P_c < M_x$. By definition, we write $\mathbf{\Delta}_{(0)} = \mathbf{I}_{M_x}$,where $\mathbf{I}_{M_x}$ is the $M_x \times M_x$ identity matrix. Therefore,

$$\mathbf{\Delta}_{(0)}\mathbf{y}_{m_y} = \mathbf{I}_{M_x}\mathbf{y}_{m_y} = \mathbf{y}_{m_y}. \tag{21}$$

We define the $p$th-order forward spatial difference of $\mathbf{y}_{m_y}$ as:

$$\begin{aligned}\Delta^p Y_{m_y,i} &= \Delta^{p-1}\left(\Delta Y_{m_y,i}\right)\\ &= \Delta^{p-1} Y_{m_y,i+1} - \Delta^{p-1} Y_{m_y,i}\\ &= \sum_{j=0}^{p} (-1)^{p-j} \binom{p}{j} Y_{m_y,i+j},\end{aligned} \tag{22}$$

where $i = 1, 2, \ldots, M_x - p$ and

$$\binom{p}{j} = \frac{p!}{j!\,(p-j)!}$$

is the binomial coefficient. In a vector/matrix form, (22) is:

$$\boldsymbol{\Delta}_{(P_c)} \mathbf{y}_{m_y} = \mathbf{y}_{m_y,(P_c)}, \tag{23}$$

where

$$\boldsymbol{\Delta}_{(P_c)} = \begin{bmatrix} \mathbf{c}^T_{(P_c)} & 0 & \cdots & 0 \\ 0 & \mathbf{c}^T_{(P_c)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{c}^T_{(P_c)} \end{bmatrix} \tag{24}$$

is a matrix of size $(M_x - P_c) \times M_x$, with

$$\mathbf{c}_{(P_c)} = \begin{bmatrix} (-1)^{P_c} \binom{P_c}{0} & (-1)^{P_c-1} \binom{P_c}{1} & \cdots & -\binom{P_c}{P_c-1} & 1 \end{bmatrix}^T \tag{25}$$

being a vector of length $P_c + 1$.

Now, substituting (7) (with $\theta = \theta_s$) into (22), it can be shown that:

$$\begin{aligned}\boldsymbol{\Delta}_{(P_c)} \mathbf{y}_{m_y} &= B_{\theta_s,m_y} \boldsymbol{\Delta}_{(P_c)} \mathbf{a}_{\theta_s} X + \boldsymbol{\Delta}_{(P_c)} \mathbf{v}_{m_y}\\ &= B_{\theta_s,m_y} \tau^{P_c}_{\theta_s,x} \mathbf{a}_{\theta_s,(P_c)} X + \mathbf{v}_{m_y,(P_c)}\\ &= \mathbf{y}_{m_y,(P_c)},\end{aligned} \tag{26}$$

where

$$\tau_{\theta_s,x} = e^{-\jmath \varpi_{\theta_s,x}} - 1, \tag{27}$$

$$\mathbf{a}_{\theta_s,(P_c)} = \begin{bmatrix} 1 & e^{-\jmath \varpi_{\theta_s,x}} & \cdots & e^{-\jmath (M_x - P_c - 1) \varpi_{\theta_s,x}} \end{bmatrix}^T \tag{28}$$

is the steering vector of length $M_x - P_c$ of the ULA at $\theta = \theta_s$, and $\mathbf{v}_{m_y,(P_c)} = \boldsymbol{\Delta}_{(P_c)} \mathbf{v}_{m_y}$. In an analogous manner, Eqs. (15)–(28) can be rewritten with the roles of $x$ and $y$ axes interchanged. That is, we may differentiate over the rows of $\mathbf{Y}$ instead of over its columns. Define:

$$\tau_{\theta_s,y} = e^{-\jmath \varpi_{\theta_s,y}} - 1, \tag{29}$$

$$\mathbf{b}_{\theta_s,(P_r)} = \begin{bmatrix} 1 & e^{-\jmath \varpi_{\theta_s,y}} & \cdots & e^{-\jmath (M_y - P_r - 1) \varpi_{\theta_s,y}} \end{bmatrix}^T, \tag{30}$$

with $1 \le P_r < M_y$. Recalling the matrix form in (4), we may define the 2-D differentials matrix of $\mathbf{Y}$ with $P_c$ column differentials and $P_r$ row differentials, $\mathbf{Y}_{(P_c,P_r)}$, by:

$$\begin{aligned}\mathbf{Y}_{(P_c,P_r)} &= \boldsymbol{\Delta}_{(P_c)} \mathbf{Y} \boldsymbol{\Delta}^T_{(P_r)}\\ &= \boldsymbol{\Delta}_{(P_c)} [\mathbf{X} + \mathbf{V}] \boldsymbol{\Delta}^T_{(P_r)}\\ &= \tau^{P_c}_{\theta_s,x} \tau^{P_r}_{\theta_s,y} \left(\mathbf{b}^T_{\theta_s,(P_r)} \otimes \mathbf{a}_{\theta_s,(P_c)}\right) X\\ &\quad + \boldsymbol{\Delta}_{(P_c)} \mathbf{V} \boldsymbol{\Delta}^T_{(P_r)}.\end{aligned} \tag{31}$$

Applying the (column-wise) vectorization operator, $\mathrm{vec}[\cdot]$, to $\mathbf{Y}_{(P_c,P_r)}$, we obtain:

$$\begin{aligned}\mathbf{y}_{(P_c,P_r)} &= \mathrm{vec}\left[\mathbf{Y}_{(P_c,P_r)}\right]\\ &= \tau^{P_c}_{\theta_s,x} \tau^{P_r}_{\theta_s,y} \left(\mathbf{b}_{\theta_s,(P_r)} \otimes \mathbf{a}_{\theta_s,(P_c)}\right) X\\ &\quad + \mathrm{vec}\left[\boldsymbol{\Delta}_{(P_c)} \mathbf{V} \boldsymbol{\Delta}^T_{(P_r)}\right]\\ &= \tau^{P_c}_{\theta_s,x} \tau^{P_r}_{\theta_s,y} \mathbf{d}_{\theta_s,(P_c,P_r)} X\\ &\quad + \left(\boldsymbol{\Delta}_{(P_r)} \otimes \boldsymbol{\Delta}_{(P_c)}\right) \mathbf{v},\end{aligned} \tag{32}$$

where $\mathbf{d}_{\theta_s,(P_c,P_r)} = \mathbf{b}_{\theta_s,(P_r)} \otimes \mathbf{a}_{\theta_s,(P_c)}$ is the 2-D differential steering vector of length $(M_x - P_c)(M_y - P_r)$. We deduce that the $(M_x - P_c)(M_y - P_r) \times (M_x - P_c)(M_y - P_r)$ covariance matrix of $\mathbf{y}_{(P_c,P_r)}$ is:

$$\begin{aligned}\boldsymbol{\Phi}_{\mathbf{y}_{(P_c,P_r)}} &= \phi_X \left|\tau_{\theta_s,x}\right|^{2P_c} \left|\tau_{\theta_s,y}\right|^{2P_r}\\ &\quad \times \mathbf{d}_{\theta_s,(P_c,P_r)} \mathbf{d}^H_{\theta_s,(P_c,P_r)}\\ &\quad + \phi_V \left(\boldsymbol{\Delta}_{(P_r)} \otimes \boldsymbol{\Delta}_{(P_c)}\right) \boldsymbol{\Gamma}_{\mathbf{v}}\\ &\quad \times \left(\boldsymbol{\Delta}_{(P_r)} \otimes \boldsymbol{\Delta}_{(P_c)}\right)^T.\end{aligned} \tag{33}$$

We immediately obtain the SNR corresponding to $\mathbf{y}_{(P_c,P_r)}$ by:

$$\begin{aligned}\mathrm{SNR}_{\mathbf{y}_{(P_c,P_r)}} &= \mathrm{tr}\Big(\phi_X \left|\tau_{\theta_s,x}\right|^{2P_c} \left|\tau_{\theta_s,y}\right|^{2P_r}\\ &\quad \times \mathbf{d}_{\theta_s,(P_c,P_r)} \mathbf{d}^H_{\theta_s,(P_c,P_r)}\Big)\\ &\quad \times \Big[\mathrm{tr}\Big(\phi_V \left(\boldsymbol{\Delta}_{(P_r)} \otimes \boldsymbol{\Delta}_{(P_c)}\right) \boldsymbol{\Gamma}_{\mathbf{v}}\\ &\quad \times \left(\boldsymbol{\Delta}_{(P_r)} \otimes \boldsymbol{\Delta}_{(P_c)}\right)^T\Big)\Big]^{-1}\\ &= \phi_X \left|\tau_{\theta_s,x}\right|^{2P_c} \left|\tau_{\theta_s,y}\right|^{2P_r}\\ &\quad \times (M_x - P_c)(M_y - P_r)\\ &\quad \times \Big[\mathrm{tr}\Big(\phi_V \left(\boldsymbol{\Delta}_{(P_r)} \otimes \boldsymbol{\Delta}_{(P_c)}\right) \boldsymbol{\Gamma}_{\mathbf{v}}\\ &\quad \times \left(\boldsymbol{\Delta}_{(P_r)} \otimes \boldsymbol{\Delta}_{(P_c)}\right)^T\Big)\Big]^{-1},\end{aligned} \tag{34}$$

and the SNR gain between $\mathbf{y}_{(P_c,P_r)}$ and $\mathbf{y}$:

$$\mathcal{G}_{(P_c,P_r)} = \frac{\left|\tau_{\theta_s,x}\right|^{2P_c} \left|\tau_{\theta_s,y}\right|^{2P_r} (M_x - P_c)(M_y - P_r)}{\mathrm{tr}\left(\left(\boldsymbol{\Delta}_{(P_r)} \otimes \boldsymbol{\Delta}_{(P_c)}\right) \boldsymbol{\Gamma}_{\mathbf{v}} \left(\boldsymbol{\Delta}_{(P_r)} \otimes \boldsymbol{\Delta}_{(P_c)}\right)^T\right)}. \tag{35}$$

Addressing the spatially white noise case, we obtain:

$$\begin{aligned}\mathcal{W}_{(P_c,P_r)} &= \frac{\left|\tau_{\theta_s,x}\right|^{2P_c} \left|\tau_{\theta_s,y}\right|^{2P_r} (M_x - P_c)(M_y - P_r)}{\mathrm{tr}\left(\left(\boldsymbol{\Delta}_{(P_r)} \otimes \boldsymbol{\Delta}_{(P_c)}\right)\left(\boldsymbol{\Delta}_{(P_r)} \otimes \boldsymbol{\Delta}_{(P_c)}\right)^T\right)}\\ &= \frac{\left|\tau_{\theta_s,x}\right|^{2P_c} \left|\tau_{\theta_s,y}\right|^{2P_r} (M_x - P_c)}{\mathrm{tr}\left((\mathbf{c}^T_{(P_r)} \otimes \boldsymbol{\Delta}_{(P_c)})(\mathbf{c}_{(P_r)} \otimes \boldsymbol{\Delta}^T_{(P_c)})\right)}\\ &= \frac{\left|\tau_{\theta_s,x}\right|^{2P_c} \left|\tau_{\theta_s,y}\right|^{2P_r} (M_x - P_c)}{\binom{2P_r}{P_r} \boldsymbol{\Delta}_{(P_c)} \boldsymbol{\Delta}^T_{(P_c)}}\\ &= \frac{\left|\tau_{\theta_s,x}\right|^{2P_c}}{\binom{2P_c}{P_c}} \times \frac{\left|\tau_{\theta_s,y}\right|^{2P_r}}{\binom{2P_r}{P_r}}\\ &= \mathcal{W}_{(P_c)} \times \mathcal{W}_{(P_r)},\end{aligned} \tag{36}$$

where $\mathcal{W}_{(P_c)}$ and $\mathcal{W}_{(P_r)}$ are the linear white noise gains with respect to the columns and the rows of the URA, respectively, and we used the two following identities in the third row:

$$(\mathbf{Q} \otimes \mathbf{R})(\mathbf{S} \otimes \mathbf{T}) = (\mathbf{QS}) \otimes (\mathbf{RT})$$

for any four matrices $\mathbf{Q}, \mathbf{R}, \mathbf{S}$ and $\mathbf{T}$ whose matrix products are well defined, and

$$\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}.$$

Substituting (13) into (35), the diffuse noise gain is given by:

$$\mathcal{D}_{(P_c,P_r)} = \frac{\left|\tau_{\theta_s,\mathrm{x}}\right|^{2P_c}\left|\tau_{\theta_s,\mathrm{y}}\right|^{2P_r}(M_\mathrm{x}-P_c)(M_\mathrm{y}-P_r)}{\mathrm{tr}\left(\left(\boldsymbol{\Delta}_{(P_r)}\otimes\boldsymbol{\Delta}_{(P_c)}\right)\boldsymbol{\Gamma}_\mathrm{d}\left(\boldsymbol{\Delta}_{(P_r)}\otimes\boldsymbol{\Delta}_{(P_c)}\right)^T\right)}. \tag{37}$$

## 4. Multistage rectangular differential beamforming

In this section, we present a rectangular differential beamforming approach that operates on the multistage rectangular differentials computed in the previous part.

Next, we would like to apply a differential beamformer $\mathbf{w}_{(P_c,P_r)}$ of length $(M_\mathrm{x}-P_c)(M_\mathrm{y}-P_r)$ to the vector $\mathbf{y}_{(P_c,P_r)}$. Then, the beamformer output signal is:

$$\begin{aligned} Z_{(P_c,P_r)} &= \mathbf{w}^H_{(P_c,P_r)}\mathbf{y}_{(P_c,P_r)} \\ &= X_{\mathrm{fd},(P_c,P_r)} + V_{\mathrm{rn},(P_c,P_r)}, \end{aligned} \tag{38}$$

where $Z_{(P_c,P_r)}$ is the estimate of $X$,

$$X_{\mathrm{fd},(P_c,P_r)} = \tau^{P_c}_{\theta_s,\mathrm{x}}\tau^{P_r}_{\theta_s,\mathrm{y}}\left(\mathbf{w}^H\mathbf{d}_{\theta_s,(P_c,P_r)}\right)X \tag{39}$$

is the filtered desired signal, and:

$$V_{\mathrm{rn},(P_c,P_r)} = \mathbf{w}^H_{(P_c,P_r)}\mathbf{v}_{(P_c,P_r)} \tag{40}$$

is the residual noise, where $\mathbf{v}_{(P_c,P_r)} = \left(\boldsymbol{\Delta}_{(P_r)}\otimes\boldsymbol{\Delta}_{(P_c)}\right)\mathbf{v}$. Consequently, the variance of $Z_{(P_c,P_r)}$ is:

$$\begin{aligned} \phi_{Z_{(P_c,P_r)}} &= \mathbf{w}^H_{(P_c,P_r)}\boldsymbol{\Phi}_{\mathbf{y}_{(P_c,P_r)}}\mathbf{w}_{(P_c,P_r)} \\ &= \phi_{X_{\mathrm{fd},(P_c,P_r)}} + \phi_{V_{\mathrm{rn},(P_c,P_r)}}, \end{aligned} \tag{41}$$

where

$$\begin{aligned} \phi_{X_{\mathrm{fd},(P_c,P_r)}} = \phi_X &\left|\tau_{\theta_s,\mathrm{x}}\right|^{2P_c}\left|\tau_{\theta_s,\mathrm{y}}\right|^{2P_r} \\ &\times\left|\mathbf{w}^H_{(P_c,P_r)}\mathbf{d}_{\theta_s,(P_c,P_r)}\right|^2, \end{aligned} \tag{42}$$

$$\phi_{V_{\mathrm{rn},(P_c,P_r)}} = \mathbf{w}^H_{(P_c,P_r)}\boldsymbol{\Phi}_{\mathbf{v}_{(P_c,P_r)}}\mathbf{w}_{(P_c,P_r)}, \tag{43}$$

and $\boldsymbol{\Phi}_{\mathbf{v}_{(P_c,P_r)}}$ is the correlation matrix of $\mathbf{v}_{(P_c,P_r)}$ which is given by:

$$\begin{aligned} \boldsymbol{\Phi}_{\mathbf{v}_{(P_c,P_r)}} &= \phi_V\left(\boldsymbol{\Delta}_{(P_r)}\otimes\boldsymbol{\Delta}_{(P_c)}\right)\boldsymbol{\Gamma}_\mathbf{v} \\ &\quad\times\left(\boldsymbol{\Delta}_{(P_r)}\otimes\boldsymbol{\Delta}_{(P_c)}\right)^T \\ &= \phi_V\boldsymbol{\Gamma}_{\mathbf{v}_{(P_c,P_r)}}. \end{aligned} \tag{44}$$

Finally, it can be derived from (39) that the distortionless constraint is given by:

$$\mathbf{w}^H_{(P_c,P_r)}\mathbf{d}_{\theta_s,(P_c,P_r)} = \tau^{-P_c}_{\theta_s,\mathrm{x}}\tau^{-P_r}_{\theta_s,\mathrm{y}}. \tag{45}$$

Now, let us relate the most commonly used performance measures corresponding to $\mathbf{w}_{(P_c,P_r)}$. The output SNR and SNR gain are, respectively,

$$\begin{aligned} \mathrm{oSNR}\left(\mathbf{w}_{(P_c,P_r)}\right) = \left|\tau_{\theta_s,\mathrm{x}}\right|^{2P_c}&\left|\tau_{\theta_s,\mathrm{y}}\right|^{2P_r}\times\frac{\phi_X}{\phi_V} \\ &\times\frac{\left|\mathbf{w}^H_{(P_c,P_r)}\mathbf{d}_{\theta_s,(P_c,P_r)}\right|^2}{\mathbf{w}^H_{(P_c,P_r)}\boldsymbol{\Gamma}_{\mathbf{v}_{(P_c,P_r)}}\mathbf{w}_{(P_c,P_r)}}, \end{aligned} \tag{46}$$

and

$$\begin{aligned} \mathcal{G}\left(\mathbf{w}_{(P_c,P_r)}\right) = \left|\tau_{\theta_s,\mathrm{x}}\right|^{2P_c}&\left|\tau_{\theta_s,\mathrm{y}}\right|^{2P_r} \\ &\times\frac{\left|\mathbf{w}^H_{(P_c,P_r)}\mathbf{d}_{\theta_s,(P_c,P_r)}\right|^2}{\mathbf{w}^H_{(P_c,P_r)}\boldsymbol{\Gamma}_{\mathbf{v}_{(P_c,P_r)}}\mathbf{w}_{(P_c,P_r)}}. \end{aligned} \tag{47}$$

Consequently, we deduce that the WNG is given by:

$$\begin{aligned} \mathcal{W}\left(\mathbf{w}_{(P_c,P_r)}\right) = \left|\tau_{\theta_s,\mathrm{x}}\right|^{2P_c}&\left|\tau_{\theta_s,\mathrm{y}}\right|^{2P_r} \\ &\times\frac{\left|\mathbf{w}^H_{(P_c,P_r)}\mathbf{d}_{\theta_s,(P_c,P_r)}\right|^2}{\mathbf{w}^H_{(P_c,P_r)}\boldsymbol{\Xi}_{(P_c,P_r)}\mathbf{w}_{(P_c,P_r)}}, \end{aligned} \tag{48}$$

and the DF:

$$\begin{aligned} \mathcal{D}\left(\mathbf{w}_{(P_c,P_r)}\right) = \left|\tau_{\theta_s,\mathrm{x}}\right|^{2P_c}&\left|\tau_{\theta_s,\mathrm{y}}\right|^{2P_r} \\ &\times\frac{\left|\mathbf{w}^H_{(P_c,P_r)}\mathbf{d}_{\theta_s,(P_c,P_r)}\right|^2}{\mathbf{w}^H_{(P_c,P_r)}\boldsymbol{\Gamma}_{\mathrm{d},(P_c,P_r)}\mathbf{w}_{(P_c,P_r)}}, \end{aligned} \tag{49}$$

where

$$\boldsymbol{\Xi}_{(P_c,P_r)} = \left(\boldsymbol{\Delta}_{(P_r)}\otimes\boldsymbol{\Delta}_{(P_c)}\right)\left(\boldsymbol{\Delta}_{(P_r)}\otimes\boldsymbol{\Delta}_{(P_c)}\right)^T, \tag{50}$$

$$\boldsymbol{\Gamma}_{\mathrm{d},(P_c,P_r)} = \left(\boldsymbol{\Delta}_{(P_r)}\otimes\boldsymbol{\Delta}_{(P_c)}\right)\boldsymbol{\Gamma}_\mathrm{d}\left(\boldsymbol{\Delta}_{(P_r)}\otimes\boldsymbol{\Delta}_{(P_c)}\right)^T. \tag{51}$$

Finally, we may define the power beampattern by:

$$\begin{aligned} \left|\mathcal{B}_\theta\left(\mathbf{w}_{(P_c,P_r)}\right)\right|^2 = \left|\tau_{\theta,\mathrm{x}}\right|^{2P_c}&\left|\tau_{\theta,\mathrm{y}}\right|^{2P_r} \\ &\times\left|\mathbf{w}^H_{(P_c,P_r)}\mathbf{d}_{\theta,(P_c,P_r)}\right|^2, \end{aligned} \tag{52}$$

where $\tau_{\theta,\mathrm{x}} = e^{-\jmath\varpi_{\theta,\mathrm{x}}} - 1$ and $\tau_{\theta,\mathrm{y}} = e^{-\jmath\varpi_{\theta,\mathrm{y}}} - 1$.

We end this part by referring to the work in Itzhak et al. (2021b), which differs significantly from this work. The main idea (Itzhak et al., 2021b) is to employ a 2-D mean operator to improve white noise robustness at the expense of the array directivity. In addition, with this approach, the beam steering property of its corresponding beamformers is not considered. In contrast, in our proposed method, we exploit the rectangular array geometry to improve the array directivity while allowing beam steering, that is, while roughly maintaining the same level of performance regardless of the desired signal incident angle.

## 5. Optimal rectangular differential beamformers

This section proposes and derives four types of multistage rectangular differential beamformers: two maximum SNR gain beamformers and two null-constrained beamformers.

### 5.1. Maximum SNR gain rectangular differential beamformers

Let us start with the maximization of the WNG measure. Recalling (48), we deduce that the maximum WNG (MWNG) beamformer may be derived from:

$$\begin{aligned} &\min_{\mathbf{w}_{(P_c,P_r)}}\ \mathbf{w}^H_{(P_c,P_r)}\boldsymbol{\Xi}_{(P_c,P_r)}\mathbf{w}_{(P_c,P_r)} \\ &\text{s. t.}\quad \mathbf{w}^H_{(P_c,P_r)}\mathbf{d}_{\theta_s,(P_c,P_r)} = \tau^{-P_c}_{\theta_s,\mathrm{x}}\tau^{-P_r}_{\theta_s,\mathrm{y}}, \end{aligned} \tag{53}$$

with the distortionless constraint of (45) taken into account. Then, it is straightforward to derive the solution:

$$\begin{aligned} \mathbf{w}_{\mathrm{MWNG}(P_c,P_r)} = &\frac{1}{\left(\tau^{P_c}_{\theta_s,\mathrm{x}}\tau^{P_r}_{\theta_s,\mathrm{y}}\right)^*} \\ &\times\frac{\boldsymbol{\Xi}^{-1}_{(P_c,P_r)}\mathbf{d}_{\theta_s,(P_c,P_r)}}{\mathbf{d}^H_{\theta_s,(P_c,P_r)}\boldsymbol{\Xi}^{-1}_{(P_c,P_r)}\mathbf{d}_{\theta_s,(P_c,P_r)}}. \end{aligned} \tag{54}$$

We proceed by considering Eq. (49). The maximum DF (MDF) beamformer is derived from:

$$\begin{aligned} &\min_{\mathbf{w}_{(P_c,P_r)}}\ \mathbf{w}^H_{(P_c,P_r)}\boldsymbol{\Gamma}_{\mathrm{d},(P_c,P_r)}\mathbf{w}_{(P_c,P_r)} \\ &\text{s. t.}\quad \mathbf{w}^H_{(P_c,P_r)}\mathbf{d}_{\theta_s,(P_c,P_r)} = \tau^{-P_c}_{\theta_s,\mathrm{x}}\tau^{-P_r}_{\theta_s,\mathrm{y}}. \end{aligned} \tag{55}$$

**Table 1**
The DF in dB units between $\mathbf{y}_{(P_c,P_r)}$ and $\mathbf{y}$ for varying values of $(P_c, P_r)$ and three distinct values of the incident angle $\theta_s$. Gray background color indicates appropriate configuration selections. Simulation parameters: $f = 4$ kHz, $M_x = 5$, $M_y = 4$, $\delta_x = 1$ cm and $\delta_y = 1.2$ cm.

| $\theta_s = 15^o$ | | | | | | $\theta_s = 45^o$ | | | | | | $\theta_s = 75^o$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $P_r$ | | | | | | $P_r$ | | | | | | $P_r$ | | | |
| | | 0 | 1 | 2 | 3 | | | 0 | 1 | 2 | 3 | | | 0 | 1 | 2 | 3 |
| $P_c$ | 0 | 0.0 | -6.8 | -16.1 | -26.2 | $P_c$ | 0 | 0.0 | 1.8 | 1.1 | -0.3 | $P_c$ | 0 | 0.0 | 4.4 | 6.3 | 7.4 |
| | 1 | 4.4 | -0.3 | -8.2 | -17.2 | | 1 | 1.8 | 5.7 | 6.4 | 6.0 | | 1 | -6.9 | -0.4 | 2.9 | 5.1 |
| | 2 | 6.3 | 3.0 | -3.8 | -12.0 | | 2 | 1.1 | 6.4 | 8.1 | 8.6 | | 2 | -16.2 | -8.3 | -4.0 | -1.0 |
| | 3 | 7.5 | 5.2 | -0.8 | -8.3 | | 3 | -0.4 | 6.0 | 8.5 | 9.7 | | 3 | -26.3 | -17.4 | -12.2 | -8.5 |
| | 4 | 8.3 | 6.9 | 1.5 | -5.4 | | 4 | -2.2 | 5.0 | 8.2 | 10.0 | | 4 | -36.8 | -27.0 | -21.6 | -16.9 |

**Table 2**
The WNG in dB units between $\mathbf{y}_{(P_c,P_r)}$ and $\mathbf{y}$ for varying values of $(P_c, P_r)$ and three distinct values of the incident angle $\theta_s$. Gray background color indicates appropriate configuration selections. Simulation parameters: $f = 4$ kHz, $M_x = 5$, $M_y = 4$, $\delta_x = 1$ cm and $\delta_y = 1.2$ cm.

| $\theta_s = 15^o$ | | | | | | $\theta_s = 45^o$ | | | | | | $\theta_s = 75^o$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $P_r$ | | | | | | $P_r$ | | | | | | $P_r$ | | | |
| | | 0 | 1 | 2 | 3 | | | 0 | 1 | 2 | 3 | | | 0 | 1 | 2 | 3 |
| $P_c$ | 0 | 0.0 | -15.8 | -33.4 | -51.4 | $P_c$ | 0 | 0.0 | -7.2 | -16.2 | -25.6 | $P_c$ | 0 | 0.0 | -4.6 | -11.0 | -17.8 |
| | 1 | -6.1 | -21.9 | -39.5 | -57.5 | | 1 | -8.7 | -15.9 | -24.9 | -34.3 | | 1 | -17.4 | -22.0 | -28.4 | -35.2 |
| | 2 | -14.0 | -29.8 | -47.4 | -65.4 | | 2 | -19.2 | -26.4 | -35.4 | -44.8 | | 2 | -36.5 | -41.2 | -47.5 | -54.4 |
| | 3 | -22.3 | -38.2 | -55.7 | -73.8 | | 3 | -30.2 | -37.4 | -46.4 | -55.8 | | 3 | -56.1 | -60.8 | -67.1 | -74.0 |
| | 4 | -30.9 | -46.7 | -64.3 | -82.3 | | 4 | -41.4 | -48.6 | -57.6 | -67.0 | | 4 | -76.0 | -80.6 | -87.0 | -93.8 |

The solution is therefore given by:

$$\mathbf{w}_{\mathrm{MDF}(P_c,P_r)} = \frac{1}{\left(\tau_{\theta_s,\mathrm{x}}^{P_c}\tau_{\theta_s,\mathrm{y}}^{P_r}\right)^*} \times \frac{\boldsymbol{\Gamma}_{\mathrm{d},(P_c,P_r)}^{-1}\mathbf{d}_{\theta_s,(P_c,P_r)}}{\mathbf{d}_{\theta_s,(P_c,P_r)}^H\boldsymbol{\Gamma}_{\mathrm{d},(P_c,P_r)}^{-1}\mathbf{d}_{\theta_s,(P_c,P_r)}}. \tag{56}$$

We end this part by noting that in case the desired signal incident angle $\theta_s$ equals either 0°, 180°, 90°, or 270° the solution is undefined. In the first two cases, to achieve a valid solution, we will never differentiate with respect to the $y$-axis. Therefore, in these cases, we always have $P_r = 0$. The same issue and solution apply to the two other cases concerning $x$-axis differentiation, in which we always have $P_c = 0$. In addition, for the sake of mathematical completeness, we define:

$$\tau_{0,\mathrm{y}}^0 = \tau_{\pi,\mathrm{y}}^0 = \tau_{\pi/2,\mathrm{x}}^0 = \tau_{3\pi/2,\mathrm{x}}^0 = 1. \tag{57}$$

### 5.2. Null-constrained rectangular differential beamformers

We now turn to null-constrained beamformers. In practice, in order to give a desired shape to a beampattern or attenuate directional interferences, spatial null constraints may be required. Therefore, with $N$ distinct null constraints (53) is transformed into:

$$\min_{\mathbf{w}} \mathbf{w}_{(P_c,P_r)}^H\boldsymbol{\Xi}_{(P_c,P_r)}\mathbf{w}_{(P_c,P_r)}$$
$$\text{s. t.} \quad \mathbf{C}^H\left(\boldsymbol{\Delta}_{(P_r)}\otimes\boldsymbol{\Delta}_{(P_c)}\right)^T\mathbf{w}_{(P_c,P_r)} = \boldsymbol{\beta}, \tag{58}$$

where $\boldsymbol{\beta}$ is the first column of the identity matrix of size $(N+1)\times(N+1)$, and $\mathbf{C}$ is a constraint matrix of size $M_xM_y\times(N+1)$:

$$\mathbf{C} = \begin{bmatrix} \mathbf{d}_{\theta_s} & \mathbf{d}_{\theta_1} & \cdots & \mathbf{d}_{\theta_N} \end{bmatrix}, \tag{59}$$

whose first column is the steering vector in the direction of the desired signal, and the remaining independent columns are the steering vectors in the directions of the desired nulls. The resulting null-constrained maximum WNG (NCMWNG) beamformer is given by:

$$\mathbf{w}_{\mathrm{NCMWNG}(P_c,P_r)} = \boldsymbol{\Xi}_{(P_c,P_r)}^{-1}\left(\boldsymbol{\Delta}_{(P_r)}\otimes\boldsymbol{\Delta}_{(P_c)}\right)\mathbf{C} \times\left[\mathbf{C}^H\left(\boldsymbol{\Delta}_{(P_r)}\otimes\boldsymbol{\Delta}_{(P_c)}\right)^T\boldsymbol{\Xi}_{(P_c,P_r)}^{-1} \times\left(\boldsymbol{\Delta}_{(P_r)}\otimes\boldsymbol{\Delta}_{(P_c)}\right)\mathbf{C}\right]^{-1}\boldsymbol{\beta}. \tag{60}$$

In a similar manner, we may optimize with respect to the DF criterion:

$$\min_{\mathbf{w}} \mathbf{w}_{(P_c,P_r)}^H\boldsymbol{\Gamma}_{\mathrm{d},(P_c,P_r)}\mathbf{w}_{(P_c,P_r)}$$
$$\text{s. t.} \quad \mathbf{C}^H\left(\boldsymbol{\Delta}_{(P_r)}\otimes\boldsymbol{\Delta}_{(P_c)}\right)^T\mathbf{w}_{(P_c,P_r)} = \boldsymbol{\beta}. \tag{61}$$

Then, the null-constrained maximum DF (NCMDF) beamformer is obtained as:

$$\mathbf{w}_{\mathrm{NCMDF}(P_c,P_r)} = \boldsymbol{\Gamma}_{\mathrm{d},(P_c,P_r)}^{-1}\left(\boldsymbol{\Delta}_{(P_r)}\otimes\boldsymbol{\Delta}_{(P_c)}\right)\mathbf{C} \times\left[\mathbf{C}^H\left(\boldsymbol{\Delta}_{(P_r)}\otimes\boldsymbol{\Delta}_{(P_c)}\right)^T\boldsymbol{\Gamma}_{\mathrm{d},(P_c,P_r)}^{-1} \times\left(\boldsymbol{\Delta}_{(P_r)}\otimes\boldsymbol{\Delta}_{(P_c)}\right)\mathbf{C}\right]^{-1}\boldsymbol{\beta}. \tag{62}$$

## 6. Simulations

### 6.1. Performance study

In this part, we investigate the performance of each of the four rectangular differential beamformers presented in the former section in terms of the WNG and DF measures.

Let us assume the existence of the following URA: $M_x = 5$, $M_y = 4$, $\delta_x = 1$ cm and $\delta_y = 1.2$ cm. We will exploit this URA example through the entire section. To begin with, it is valuable to evaluate the WNG and DF between the observed differentials vector and the observed signal vector, that is, $\mathcal{W}_{(P_c,P_r)}$ and $\mathcal{D}_{(P_c,P_r)}$. Recalling Eqs. (36) and (37), we realize that the $(P_c, P_r)$ configuration should be set with respect to the rectangular array structure and the desired signal incident angle $\theta_s$. Tables 1 and 2, respectively, show all possible values $\mathcal{D}_{(P_c,P_r)}$ and $\mathcal{W}_{(P_c,P_r)}$ for three distinct values of $\theta_s$: 15°, 45° and 75° with $f = 4$ kHz. We note that to maximize both measures, appropriate values of the $(P_c, P_r)$ configuration (marked in gray color) should be chosen. In essence, the $(P_c, P_r)$ configuration is set to maximize the DF while deteriorating the WNG as little as possible. Since these measures contradict with each other, very high values of DF imply extreme white noise amplification (e.g., when the WNG is lower than −35 dB) and are therefore impractical. We observe that the appropriate values are obtained in accordance with $\theta_s$: when $\theta_s = 15°$, $P_r$ should be set to zero whereas $P_c$ sets the WNG-DF trade-off; when $\theta_s = 75°$, the roles of $P_r$ and $P_c$ interchange; and when $\theta_s = 45°$ it is recommended that $P_c$

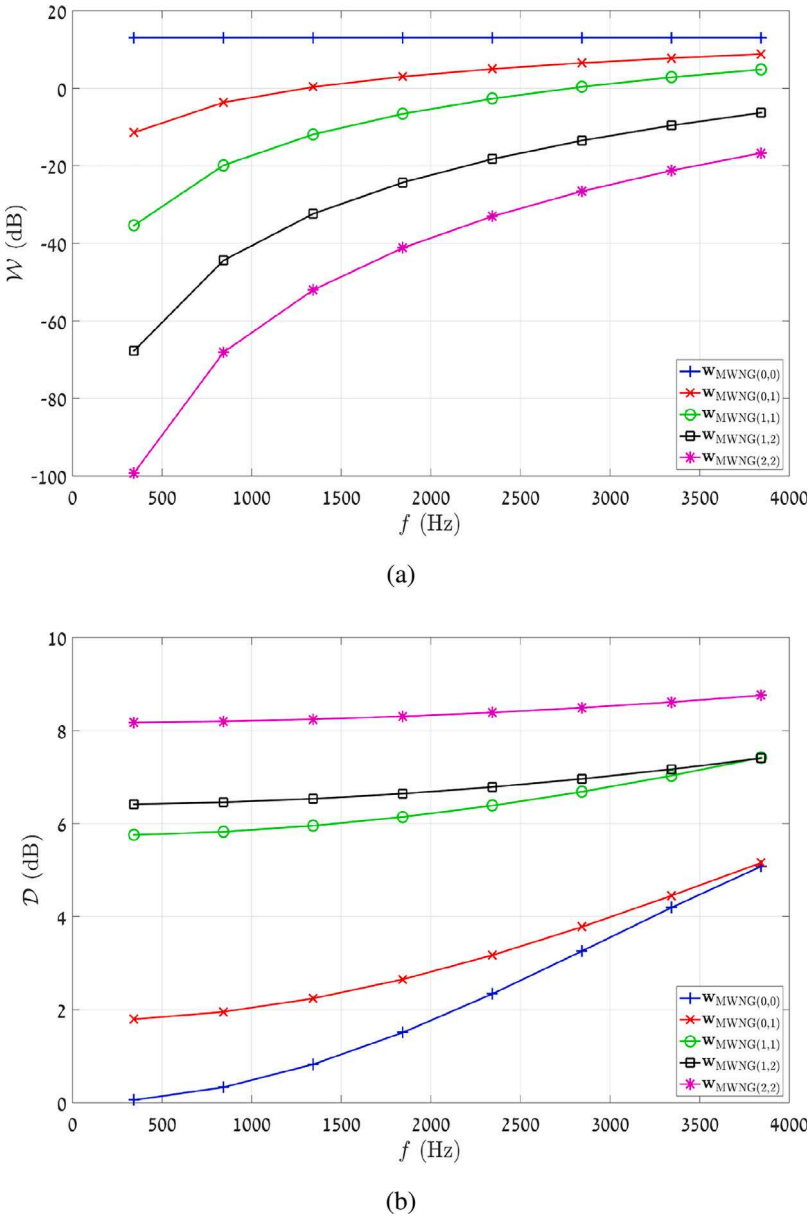


**Fig. 2.** WNG and DF measures with the MWNG rectangular differential beamformer, $\mathbf{w}_{\mathrm{MWNG}(P_c,P_r)}$, with varying values of $(P_c, P_r)$. (a) WNG and (b) DF. Simulation parameters: $\theta_s = 45°$, $M_x = 5$, $M_y = 4$, $\delta_x = 1$ cm and $\delta_y = 1.2$ cm.

and $P_r$ satisfy $|P_r - P_c| \leq 1$. We deduce that the multistage rectangular differential approach enables controlled directivity gains and beam steering flexibility.

We move on to analyzing the WNG and DF performance measures of each of the beamformers presented in the former section with $\theta_s = 45°$. We begin with the MWNG rectangular beamformer, $\mathbf{w}_{\mathrm{MWNG}(P_c,P_r)}$, whose WNG and DF measures are depicted in Fig. 2. We observe the following. For $(P_c, P_r) = (0, 0)$, $\mathbf{w}_{\mathrm{MWNG}(0,0)}$ is the well-known Delay-and-Sum (DS) rectangular beamformer, whose frequency-independent WNG equals to the number of array sensors. As the configurations of $(P_c, P_r)$ change to increase $\mathcal{D}_{(P_c,P_r)}$, according to the selected configurations of Tables 1 and 2, the DF performance improves, but the WNG performance degrades. We note that the DF improvement of $\mathbf{w}_{\mathrm{MWNG}(P_c,P_r)}$ complies with $\mathcal{D}_{(P_c,P_r)}$: the DFs of $\mathbf{w}_{\mathrm{MWNG}(0,0)}$ and $\mathbf{w}_{\mathrm{MWNG}(0,1)}$ are rather close and low; $\mathbf{w}_{\mathrm{MWNG}(1,1)}$ and $\mathbf{w}_{\mathrm{MWNG}(1,2)}$ exhibit a significant DF improvement with respect to the former configurations; with $\mathbf{w}_{\mathrm{MWNG}(2,2)}$ the DF is further improved at the expense of a significantly worse WNG performance. We deduce that, indeed, the configuration of the parameters $(P_c, P_r)$ highly affects the WNG-DF performance of the rectangular differential beamformer, with a strong correlation to the gain between the observed differentials vector and the raw observed signal vector.

Next, we focus on $\mathbf{w}_{\mathrm{MDF}(P_c,P_r)}$, whose WNG and DF performances are depicted in Fig. 3. We observe a similar WNG-DF trade-off as the previous beamformer upon changing the configuration of $(P_c, P_r)$. In contrast to the former case, and as one would expect, the DF performance of $\mathbf{w}_{\mathrm{MDF}(P_c,P_r)}$ is significantly better than $\mathbf{w}_{\mathrm{MWNG}(P_c,P_r)}$'s at the expense of the WNG performance. This is true regardless of the selection of $(P_c, P_r)$. In addition, we note that with the higher values of $(P_c, P_r)$, the WNG performance is poor, particularly in low frequencies. This unappealing behavior implies that, in practice, $\mathbf{w}_{\mathrm{MDF}(P_c,P_r)}$ should only be designed with a modest $\mathcal{D}_{(P_c,P_r)}$ performance, that is, with small values of $P_c$ and $P_r$.

We turn to the NCMWNG and NCMDF rectangular differential beamformers, whose performances are shown in Figs. 4 and 5. We note that both beamformers are designed with $N = 2$ distinct nulls located

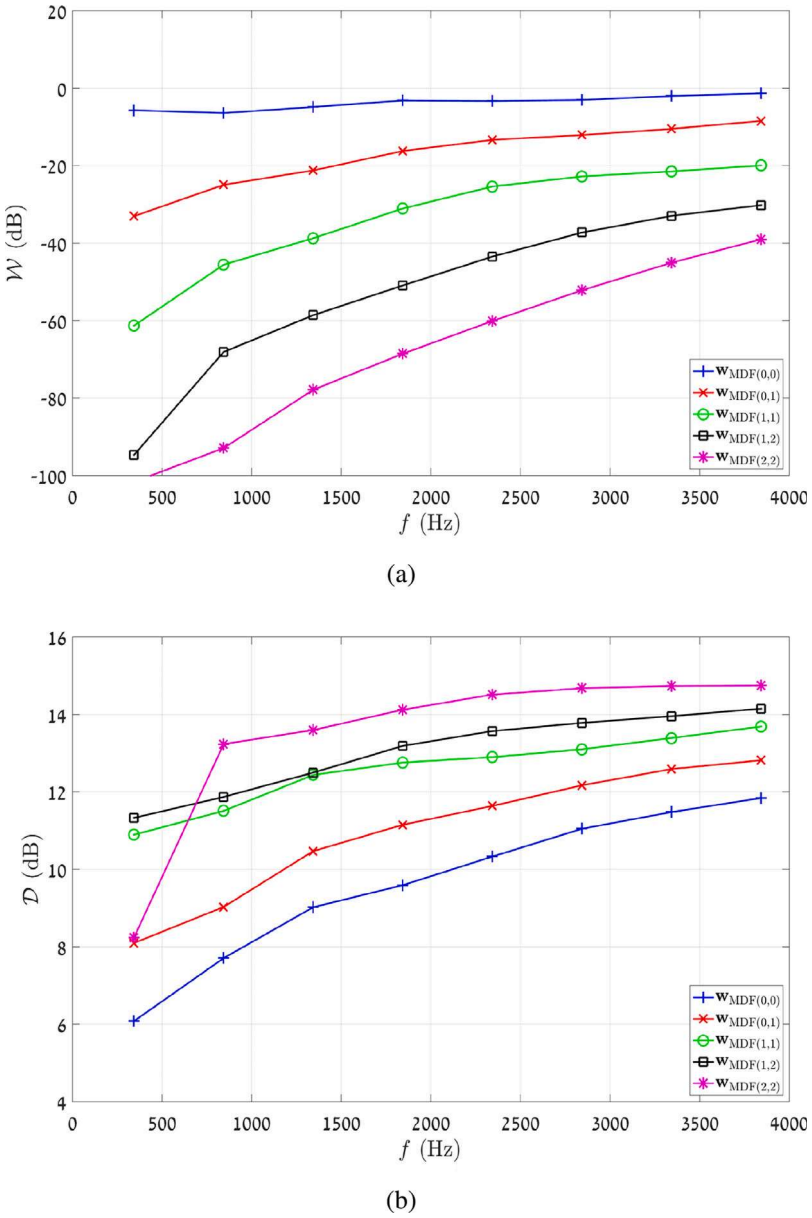


**Fig. 3.** WNG and DF measures with the MDF rectangular differential beamformer, $\mathbf{w}_{\mathrm{MDF}(P_c,P_r)}$, with varying values of $(P_c, P_r)$. (a) WNG and (b) DF. Simulation parameters: $\theta_s = 45°$, $M_x = 5$, $M_y = 4$, $\delta_x = 1$ cm and $\delta_y = 1.2$ cm.

at 160° and −90°. Considering $\mathbf{w}_{\mathrm{NCMWNG}(P_c,P_r)}$, we clearly observe a DF performance improvement with respect to $\mathbf{w}_{\mathrm{MWNG}(P_c,P_r)}$. This improvement is at the expense of worse WNG performance. Turning to $\mathbf{w}_{\mathrm{NCMDF}(P_c,P_r)}$, we note that its directivity is better than $\mathbf{w}_{\mathrm{NCMWNG}(P_c,P_r)}$'s but its WNG performance is indeed worse. Nevertheless, the DF performance of $\mathbf{w}_{\mathrm{NCMDF}(P_c,P_r)}$ is worse than $\mathbf{w}_{\mathrm{MDF}(P_c,P_r)}$'s, but its WNG performance is preferable. We deduce that the null-constrained versions of the MWNG and MDF rectangular differential beamformers enable additional WNG-DF performance tuning flexibility.

We end this part by accentuating the relation between the appropriate selection of the $(P_c, P_r)$ configuration and the desired signal incident angle $\theta_s$. Specifically, we focus on the MWNG rectangular differential beamformer $\mathbf{w}_{\mathrm{MWNG}(P_c,P_r)}$. However, this behavior is similar to the three other proposed beamformers. The corresponding WNG and DF measures as a function of the frequency $f$ and the desired signal incident angle $\theta_s$ are depicted in Fig. 6 with $(P_c, P_r) = (2, 0)$, $(P_c, P_r) = (1, 1)$ and $(P_c, P_r) = (0, 2)$. We observe that the selected configuration greatly influences both performance measures. For example, when $\theta_s$ is near 0° the $(P_c, P_r) = (2, 0)$ configuration is the appropriate selection between the three (and in general $P_r$ should be set to 0 whereas $P_c$ sets the WNG-DF trade-off). Similarly, when $\theta_s$ is near 90° it is best to set $P_c$ to 0 and tune the performance trade-off with $P_r$. Finally, when $\theta_s$ is near 45°, $(P_c, P_r) = (1, 1)$ is the most appropriate configuration among the three, implying that in this case both $P_c$ and $P_r$ may be greater than 0. We note that this behavior complies with Tables 1 and 2 and is valid with each of the proposed beamformers. In addition, when $\theta_s$ is not in the vicinity of 0° or 90°, the exact $(P_c, P_r)$ configuration should be carefully set according to the array geometry (that is, the interelement spacing and the number of microphones along each axis), and the desired level of array directivity. This can easily be done in practice by assessing the SNR gains, as in Tables 1 and 2, for any given scenario.

### 6.2. Speech signals simulations in reverberant environments

In this part, we demonstrate the performance of the proposed rectangular differential beamformers on speech signals in practical

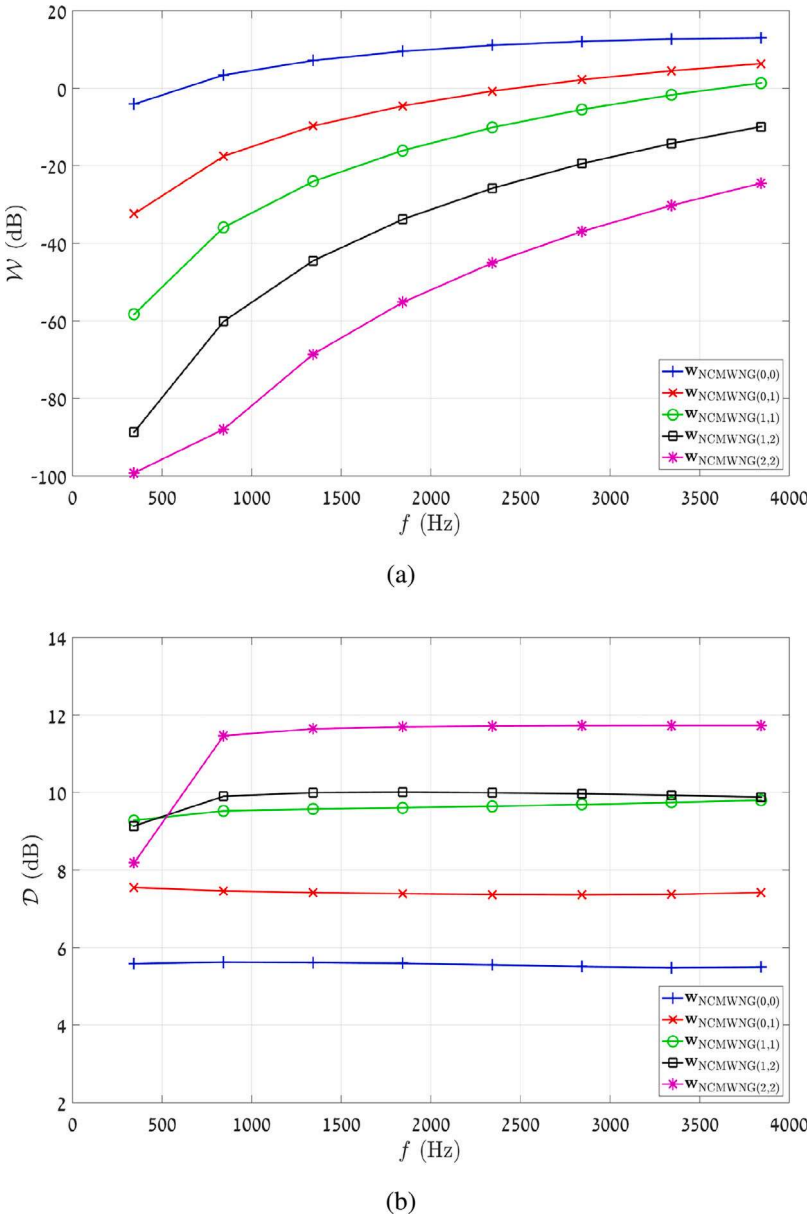


**Fig. 4.** WNG and DF measures with the null-constrained MWNG rectangular differential beamformer, $\mathbf{w}_{\mathrm{NCMWNG}(P_c,P_r)}$, with varying values of $(P_c, P_r)$. (a) WNG and (b) DF. Simulation parameters: $\theta_s = 45°$, $M_x = 5$, $M_y = 4$, $\delta_x = 1$ cm and $\delta_y = 1.2$ cm.

simulated scenarios in reverberant environments. The reverberant simulations are performed as follows. We use a room impulse response (RIR) generator (Habets, 2008) to simulate the reverberant noise-free signal received in a URA consisting of $M_x \times M_y = 5 \times 4$ microphones. The URA is located on the $z = 1.5$ m plane, where the first microphone is located at the $(x, y) = (2\,\mathrm{m}, 2\,\mathrm{m})$ coordinate, with $\delta_x = 1$ cm and $\delta_y = 1.2$ cm. The RIR generator is based on the image method of Allen and Berkley (1979). We simulate three distinct scenarios in a $6 \times 6 \times 3$ m room, which differ by the value of the desired speech signal incident angle $\theta_s$: 15°, 45° or 75°. In all scenarios, the desired speech signal source is located on the same plane as the URA, that is, the $z = 1.5$ m plane, and we set $T_{60} = 600$ ms, where $T_{60}$ is defined by Sabin–Franklin's formula (Pierce, 2019). In addition, two uncorrelated directional interferences are located on the $z = 1.5$ m plane in the same null directions of $\mathbf{w}_{\mathrm{NCMDF}(P_c,P_r)}$ and $\mathbf{w}_{\mathrm{NCMDF}(P_c,P_r)}$ from the former part, that is, $160^0$ and $-90°$. The former interference is 10 dB weaker than the latter, which for itself is approximately 4.9 dB weaker than the desired signal. On top of the directional interferences, two uncorrelated noise fields are present: a white thermal Gaussian noise, which is 30 dB weaker than the more powerful interference, and a spherically-isotropic diffuse noise, which is 3 dB stronger than the more powerful interference. The desired speech signal, $x(t)$, is a concatenation of 24 speech signals (12 speech signals per gender) with varying dialects taken from the TIMIT database (Anon, 1993). It is sampled at a sampling rate of $f_s = 1/T_s = 16$ kHz within the signal duration $T$.

The noisy observed signal is transformed into the STFT domain using 75% overlapping time frames and a Hamming analysis window of length 256 (16 ms). The discrete Fourier-transform length is set to 256 as well. Next, rectangular differential beamformers with different configurations are independently applied to the noisy signal to yield clean estimates in the STFT domain, followed by an inverse STFT procedure to obtain time-domain enhanced signals.

Next, we are interested in objectively quantifying the performance of each of the four beamformers. We shall do that by individually examining the power ratio between the noise and reverberation components of the first microphone and their respective components in the

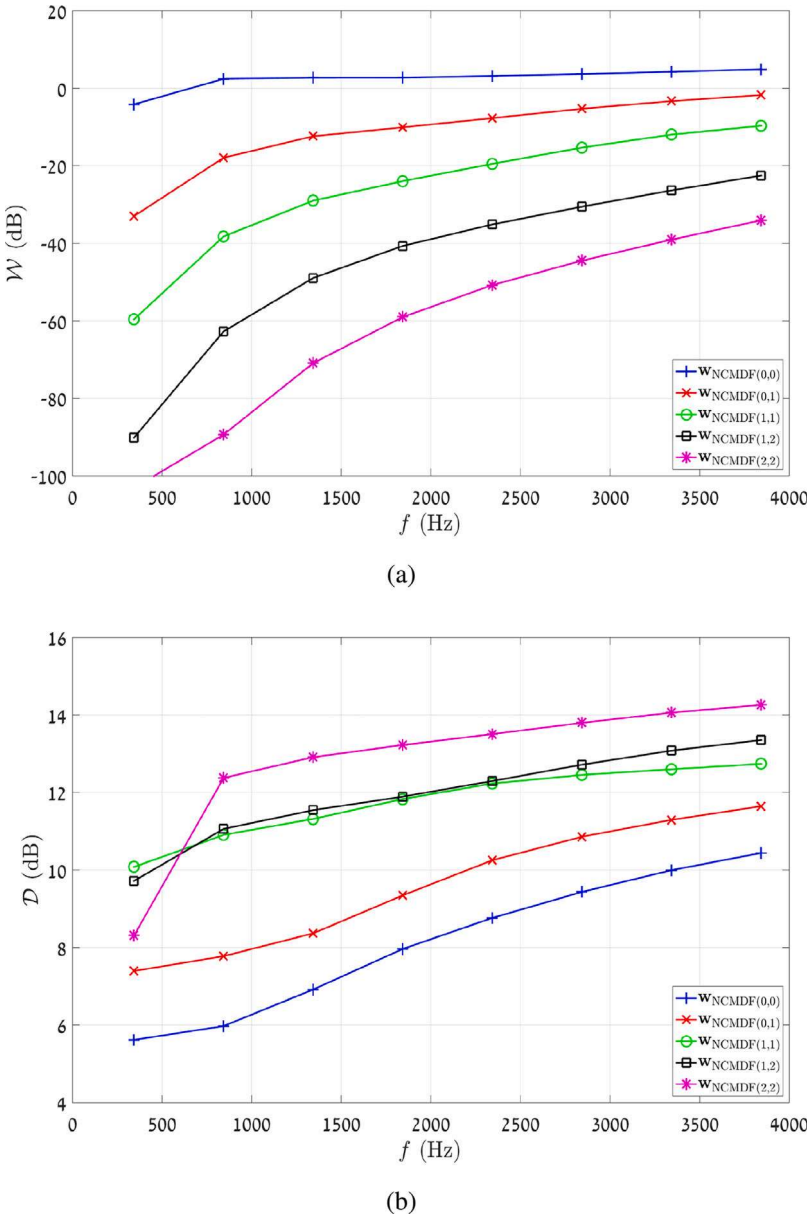


**Fig. 5.** WNG and DF measures with the null-constrained MDF rectangular differential beamformer, $\mathbf{w}_{\mathrm{NCMDF}(P_c,P_r)}$, with varying values of $(P_c, P_r)$. (a) WNG and (b) DF. Simulation parameters: $\theta_s = 45°$, $M_x = 5$, $M_y = 4$, $\delta_x = 1$ cm and $\delta_y = 1.2$ cm.

time-domain enhanced signals. This includes the white thermal Gaussian noise, the diffuse noise, the reverberant directional interferences, and the desired speech signal reverberations. Formulating the noisy observed signal in the time domain in microphone $m$ we have

$$\begin{aligned} y_m &= x_{\mathrm{d}} * g_{\mathrm{d},m} + v_{\mathrm{i},1} * g_{\mathrm{i},1,m} + v_{\mathrm{i},2} * g_{\mathrm{i},2,m} \\ &\quad + v_{\mathrm{d},m} + v_{\mathrm{w},m} \\ &= x_m + x_{\mathrm{r},m} + v_{\mathrm{r},1,m} + v_{\mathrm{r},2,m} + v_{\mathrm{d},m} + v_{\mathrm{w},m}, \end{aligned} \tag{63}$$

where $*$ is the linear convolution operator, $x_{\mathrm{d}}$ is the desired speech signal, $v_{\mathrm{i},1}$ and $v_{\mathrm{i},2}$ are, respectively, the two directional interferences, $v_{\mathrm{d},m}$ is the additive diffuse noise and $v_{\mathrm{w},m}$ is white noise in microphone $m$. Additionally, $g_{\mathrm{d},m}$ is the RIR from the desired signal source to microphone $m$, $g_{\mathrm{i},1,m}$ and $g_{\mathrm{i},2,m}$ are, respectively, the RIR from the directional interference sources to microphone $m$, whereas $x_m$, $x_{\mathrm{r},m}$, $v_{\mathrm{r},1,m}$ and $v_{\mathrm{r},2,m}$ are the direct path desired signal, its reverberations and the two reverberant interferences as received in microphone $m$, respectively. Lastly, we define the same components of the second row of (63) with respect to the time-domain enhanced speech signals by using the subscript $\mathrm{f}$. For example, $x_{\mathrm{f}}$ is the enhanced direct path desired signal and $v_{\mathrm{w,f}}$ is the white noise component in the enhanced speech signal.

We now address the white thermal noise and the diffuse noise. Using the notations of (63), we define the Diffuse Noise Reduction (DNR) factor by:

$$\mathrm{DNR} = \frac{E\left[v_{\mathrm{d},1}^2\right]}{E\left[v_{\mathrm{d},f}^2\right]}, \tag{64}$$

and the White Noise Reduction (WNR) factor by:

$$\mathrm{WNR} = \frac{E\left[v_{\mathrm{w},1}^2\right]}{E\left[v_{\mathrm{w},f}^2\right]}. \tag{65}$$

To demonstrate the advantages of the proposed multistage rectangular differential approach compared to the linear approach of Huang et al. (2020b), in the following, we compare the performance of the

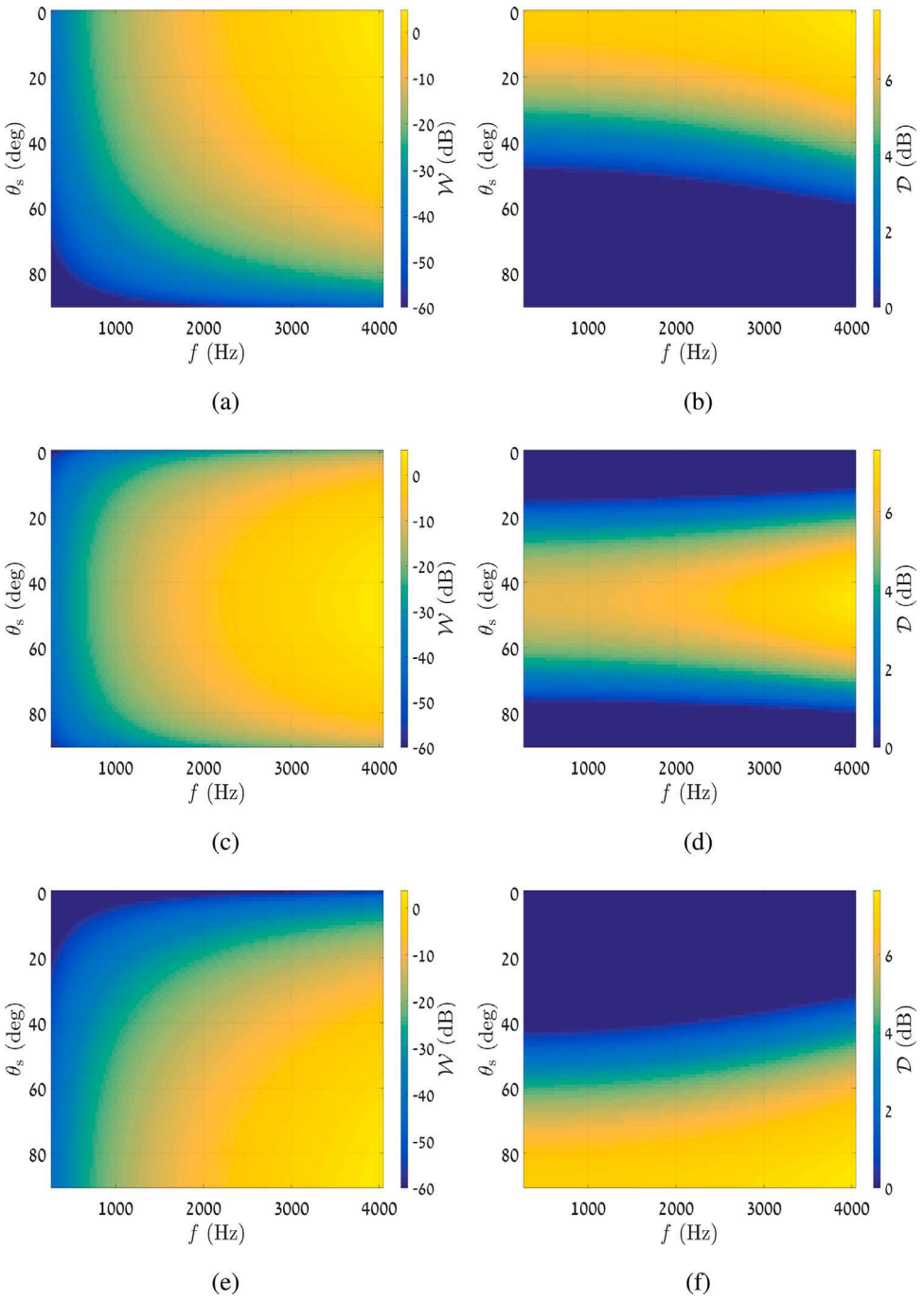


**Fig. 6.** WNG and DF measures with the MWNG rectangular differential beamformer $\mathbf{w}_{\mathrm{MWNG}(P_c,P_r)}$ and varying values of $(P_c, P_r)$ as a function of the frequency $f$ and the desired signal incident angle $\theta_s$. (a) WNG with $(P_c, P_r) = (2, 0)$, (b) DF with $(P_c, P_r) = (2, 0)$, (c) WNG with $(P_c, P_r) = (1, 1)$, (d) DF with $(P_c, P_r) = (1, 1)$, (e) WNG with $(P_c, P_r) = (0, 2)$ and (f) DF with $(P_c, P_r) = (0, 2)$, Simulation parameters: $\mathrm{M_x} = 5$, $\mathrm{M_y} = 4$, $\delta_x = 1$ cm and $\delta_y = 1.2$ cm.

aforementioned URA to the performance of a ULA consisting of $M_x \times M_y = 20 \times 1$ microphones. We investigate the DNR and WNR with both arrays and each of the presented beamformers by taking advantage of the linearity of the beamformers and applying them to an input signal consisting of merely the specific noise component at a time. While the $(P_c, P_r)$ configurations with the URA are taken from Tables 1 and 2, $P_r$ is strictly zero for the ULA, implying that $P_c$ exclusively determines the directivity of the array. To begin with, it is shown in Table 3 that the DNR is typically maximized with $\mathbf{w}_{\mathrm{MDF}(P_c,P_r)}$, whereas the DNR with $\mathbf{w}_{\mathrm{NCMDF}(P_c,P_r)}$ is superior to the DNR with both $\mathbf{w}_{\mathrm{MWNG}(P_c,P_r)}$ and $\mathbf{w}_{\mathrm{NCMWNG}(P_c,P_r)}$. Addressing the URA, we observe that the higher the $(P_c, P_r)$ configuration, the preferable the DNR. In addition, the values of the DNR remain roughly similar regardless of the incident angle $\theta_s$. On the contrary, with the ULA, the DNR performance deteriorates as $\theta_s$ deviates from the endfire direction: when $\theta_s = 15°$, the DNR is better than the DNR with the URA, and it improves upon increasing the $(P_c, P_r)$ configuration; when $\theta_s = 45°$ the DNR is significantly worse than in the former case and compared to the URA, and when $\theta_s = 75°$ the DNR worsens even further, and might also turn negative (which implies diffuse noise amplification) with $P_c = 2$.

Let us focus on the WNR whose values with the discussed beamformers appear in Table 4. It is worth noting that when $(P_c, P_r) = (0, 0)$ white noise is typically attenuated. However, with other configurations, white noise might be significantly amplified. We note that this is in

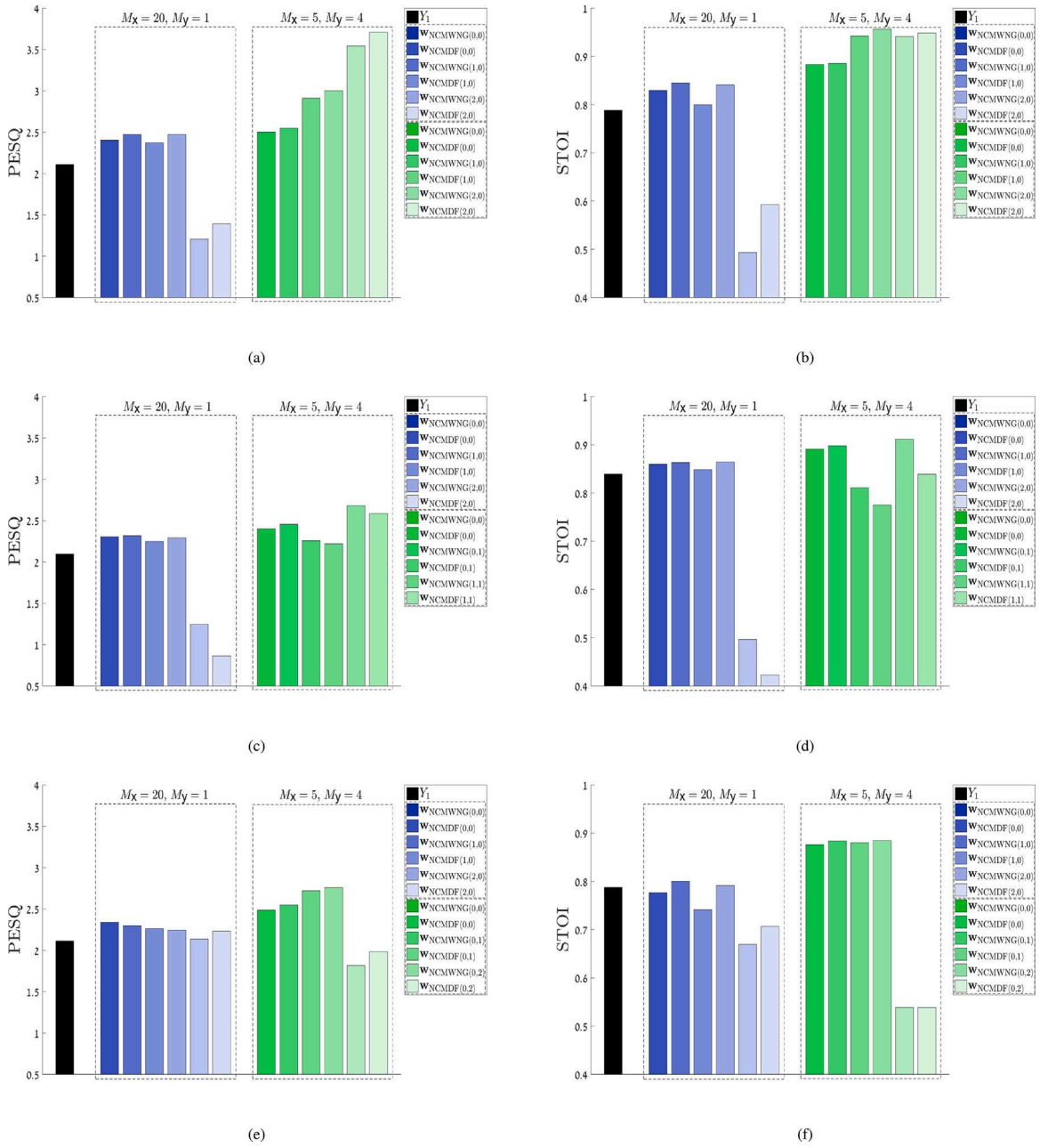


**Fig. 7.** Average PESQ and STOI scores with the NCMWNG and NCMDF rectangular differential beamformers, $\mathbf{w}_{\mathrm{NCMWNG}(P_c,P_r)}$ and $\mathbf{w}_{\mathrm{NCMDF}(P_c,P_r)}$, respectively, for three values of the incident angle $\theta_s$, two array geometries and three varying values of $(P_c, P_r)$ with each geometry. (a) Average PESQ scores and $\theta_s = 15^0$, (b) average STOI scores and $\theta_s = 15^0$, (c) average PESQ scores and $\theta_s = 45^0$, (d) average STOI scores and $\theta_s = 45^0$, (e) average PESQ scores and $\theta_s = 75^0$, and (f) average STOI scores and $\theta_s = 75^0$. Simulation parameters: $\delta_x = 1$ cm and $\delta_y = 1.2$ cm.

accordance with Section 6.1: the higher the configuration, the worse the WNR. Therefore, we infer that, in practice, whenever the microphones suffer from considerable inherent imperfections, the $(P_c, P_r)$ configuration should be kept low.

Next, we define the desired signal Reverberations Reduction (RR) factor by:

$$\mathrm{RR} = \frac{E\left[x_{\mathrm{r},1}^2\right]}{E\left[x_{\mathrm{r},f}^2\right]}, \tag{66}$$

and the Interference Reduction (IR) factor by:

$$\mathrm{IR} = \frac{E\left[v_{\mathrm{r},1,1}^2\right] + E\left[v_{\mathrm{r},2,1}^2\right]}{E\left[v_{\mathrm{r},1,f}^2\right] + E\left[v_{\mathrm{r},2,f}^2\right]}. \tag{67}$$

We note that both factors are a function of the RIRs as opposed to the DNR and WNR; nevertheless, they are computed similarly. The RR and IR with the discussed URA, ULA, and their corresponding beamformers are shown in Tables 5 and 6, respectively. We begin by addressing the RRs. It is clear that the RR with the ULA is roughly slightly better than with the URA for $\theta_s = 15°$. However, for $\theta_s = 45°$, and in particular for $\theta_s = 75°$- the URA is superior. In addition, unlike with the WNR and DNR, the RR performance does not monotonically improve or worsen upon increasing the $(P_c, P_r)$ configuration as it highly depends on the beampatterns and the simulated scenario. However, we note that with the URA, it is highly likely for the RR to improve by modifying from $(P_c, P_r) = (0, 0)$ to the successive configuration. This is true for all examined values of $\theta_s$. On the contrary, with the ULA, when $\theta_s = 15°$, the RR does improve by such a modification;

**Table 3**
The DNR in dB units for the three investigated values of $\theta_s$, three configurations of the ULA with $M_x = 20$ and $M_y = 1$, and three configurations of the URA with $M_x = 5$ and $M_y = 4$. Simulation parameters: $\delta_x = 1$ cm and $\delta_y = 1.2$ cm.

| | $\theta_s = 15^o$ | | | | | | $\theta_s = 45^o$ | | | | | | $\theta_s = 75^o$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $M_x = 20,\ M_y = 1$ | | | $M_x = 5,\ M_y = 4$ | | | $M_x = 20,\ M_y = 1$ | | | $M_x = 5,\ M_y = 4$ | | | $M_x = 20,\ M_y = 1$ | | | $M_x = 5,\ M_y = 4$ | | |
| $(P_c, P_r) =$ | (0,0) | (1,0) | (2,0) | (0,0) | (1,0) | (2,0) | (0,0) | (1,0) | (2,0) | (0,0) | (0,1) | (1,1) | (0,0) | (1,0) | (2,0) | (0,0) | (0,1) | (0,2) |
| $\mathbf{w}_{\mathrm{MWNG}(P_c,P_r)}$ | 6.8 | 8.5 | 9.8 | 3.8 | 6.3 | 7.6 | 5.5 | 6.1 | 5.4 | 3.8 | 4.7 | 7.5 | 5.0 | 0.0 | −8.9 | 3.7 | 6.2 | 7.4 |
| $\mathbf{w}_{\mathrm{MDF}(P_c,P_r)}$ | **11.7** | **12.6** | **12.9** | **10.2** | **12.0** | **12.6** | 7.5 | **8.0** | **8.1** | **10.3** | **11.5** | **13.0** | **6.6** | 4.1 | −3.1 | **10.1** | **11.6** | **11.9** |
| $\mathbf{w}_{\mathrm{NCMWNG}(P_c,P_r)}$ | 8.7 | 9.2 | 10.2 | 6.3 | 8.5 | 10.0 | 6.7 | 6.6 | 5.8 | 6.4 | 8.0 | 9.9 | 1.8 | 0.3 | −7.5 | 5.9 | 8.9 | 10.9 |
| $\mathbf{w}_{\mathrm{NCMDF}(P_c,P_r)}$ | 10.5 | 11.3 | 12.0 | 8.5 | 10.2 | 11.6 | **7.8** | 7.7 | 7.7 | 8.6 | 9.8 | 11.6 | 6.4 | **4.8** | **6.5** | 8.3 | 10.3 | 11.8 |

**Table 4**
The WNR in dB units for the three investigated values of $\theta_s$, three configurations of the ULA with $M_x = 20$ and $M_y = 1$, and three configurations of the URA with $M_x = 5$ and $M_y = 4$. Simulation parameters: $\delta_x = 1$ cm and $\delta_y = 1.2$ cm.

| | $\theta_s = 15^o$ | | | | | | $\theta_s = 45^o$ | | | | | | $\theta_s = 75^o$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $M_x = 20,\ M_y = 1$ | | | $M_x = 5,\ M_y = 4$ | | | $M_x = 20,\ M_y = 1$ | | | $M_x = 5,\ M_y = 4$ | | | $M_x = 20,\ M_y = 1$ | | | $M_x = 5,\ M_y = 4$ | | |
| $(P_c, P_r) =$ | (0,0) | (1,0) | (2,0) | (0,0) | (1,0) | (2,0) | (0,0) | (1,0) | (2,0) | (0,0) | (0,1) | (1,1) | (0,0) | (1,0) | (2,0) | (0,0) | (0,1) | (0,2) |
| $\mathbf{w}_{\mathrm{MWNG}(P_c,P_r)}$ | **12.9** | **5.0** | **−34.9** | **12.9** | **−6.3** | **−53.1** | **12.9** | **2.6** | **−40.3** | **12.9** | **−9.3** | **−51.6** | **12.9** | **−5.8** | −57.7 | **12.9** | **−6.6** | **−54.5** |
| $\mathbf{w}_{\mathrm{MDF}(P_c,P_r)}$ | 1.1 | −25.9 | −67.6 | −2.5 | −33.1 | −71.8 | 0.4 | −28.0 | −71.7 | −2.5 | −33.7 | −69.6 | −0.5 | −35.0 | −83.6 | −2.5 | −32.4 | −70.7 |
| $\mathbf{w}_{\mathrm{NCMWNG}(P_c,P_r)}$ | 7.1 | −19.7 | −63.0 | 4.2 | −27.2 | −65.7 | 5.9 | −22.7 | −68.1 | 4.2 | −29.1 | −66.1 | 2.5 | −26.4 | −60.1 | 3.7 | −27.0 | −65.2 |
| $\mathbf{w}_{\mathrm{NCMDF}(P_c,P_r)}$ | 10.2 | −17.9 | −63.9 | 7.3 | −20.5 | −60.7 | 10.1 | −20.0 | −67.0 | 7.3 | −20.0 | −58.1 | 9.7 | −19.9 | **−54.2** | 7.0 | −19.2 | −57.7 |

**Table 5**
The RR in dB units for the three investigated values of $\theta_s$, three configurations of the ULA with $M_x = 20$ and $M_y = 1$, and three configurations of the URA with $M_x = 5$ and $M_y = 4$. Simulation parameters: $\delta_x = 1$ cm, $\delta_y = 1.2$ cm, and $T_{60} = 600$ ms.

| | $\theta_s = 15^o$ | | | | | | $\theta_s = 45^o$ | | | | | | $\theta_s = 75^o$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $M_x = 20,\ M_y = 1$ | | | $M_x = 5,\ M_y = 4$ | | | $M_x = 20,\ M_y = 1$ | | | $M_x = 5,\ M_y = 4$ | | | $M_x = 20,\ M_y = 1$ | | | $M_x = 5,\ M_y = 4$ | | |
| $(P_c, P_r) =$ | (0,0) | (1,0) | (2,0) | (0,0) | (1,0) | (2,0) | (0,0) | (1,0) | (2,0) | (0,0) | (0,1) | (1,1) | (0,0) | (1,0) | (2,0) | (0,0) | (0,1) | (0,2) |
| $\mathbf{w}_{\mathrm{MWNG}(P_c,P_r)}$ | 1.4 | 5.5 | **7.3** | 0.1 | 4.6 | **4.6** | −0.3 | 3.0 | **1.8** | 0.0 | 3.6 | 6.2 | −0.3 | −3.7 | −14.2 | 0.1 | 4.6 | 4.5 |
| $\mathbf{w}_{\mathrm{MDF}(P_c,P_r)}$ | **8.2** | **8.8** | −0.2 | **3.8** | 6.6 | −10.7 | 2.7 | 3.5 | −3.0 | 3.7 | 6.5 | −1.4 | 1.4 | **4.0** | −12.9 | **3.7** | 6.8 | 4.4 |
| $\mathbf{w}_{\mathrm{NCMWNG}(P_c,P_r)}$ | 6.3 | 6.4 | −4.8 | 2.4 | 6.5 | −2.4 | 3.3 | 3.3 | −9.4 | 2.9 | 5.9 | 2.1 | −3.5 | −5.6 | −14.4 | 1.7 | 6.5 | 4.4 |
| $\mathbf{w}_{\mathrm{NCMDF}(P_c,P_r)}$ | 6.7 | 7.4 | 2.9 | 3.4 | **7.2** | 0.9 | **4.1** | **4.1** | 0.0 | **3.9** | **7.3** | **7.1** | **2.6** | −0.1 | **4.1** | 3.0 | **7.5** | **8.9** |

**Table 6**
The IR in dB units for the three investigated values of $\theta_s$, three configurations of the ULA with $M_x = 20$ and $M_y = 1$, and three configurations of the URA with $M_x = 5$ and $M_y = 4$. Simulation parameters: $\delta_x = 1$ cm, $\delta_y = 1.2$ cm, and $T_{60} = 600$ ms.

| | $\theta_s = 15^o$ | | | | | | $\theta_s = 45^o$ | | | | | | $\theta_s = 75^o$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $M_x = 20,\ M_y = 1$ | | | $M_x = 5,\ M_y = 4$ | | | $M_x = 20,\ M_y = 1$ | | | $M_x = 5,\ M_y = 4$ | | | $M_x = 20,\ M_y = 1$ | | | $M_x = 5,\ M_y = 4$ | | |
| $(P_c, P_r) =$ | (0,0) | (1,0) | (2,0) | (0,0) | (1,0) | (2,0) | (0,0) | (1,0) | (2,0) | (0,0) | (0,1) | (1,1) | (0,0) | (1,0) | (2,0) | (0,0) | (0,1) | (0,2) |
| $\mathbf{w}_{\mathrm{MWNG}(P_c,P_r)}$ | 7.8 | 11.6 | **10.0** | 4.6 | 10.8 | 1.9 | 7.0 | 10.2 | **5.7** | 5.1 | 2.5 | 7.7 | 4.5 | 5.6 | −10.3 | 5.0 | 4.3 | −7.1 |
| $\mathbf{w}_{\mathrm{MDF}(P_c,P_r)}$ | **12.7** | 12.7 | −0.9 | **13.2** | **15.7** | 0.2 | 10.4 | 10.8 | −4.3 | **13.6** | 13.3 | 3.4 | 6.4 | 9.4 | −14.5 | 11.5 | 12.1 | −10.7 |
| $\mathbf{w}_{\mathrm{NCMWNG}(P_c,P_r)}$ | 11.6 | 11.6 | −5.6 | 10.1 | 11.9 | 4.8 | 10.6 | 9.9 | −10.7 | 10.7 | 12.0 | 6.4 | 7.0 | 4.6 | −5.1 | 9.8 | 11.8 | −4.1 |
| $\mathbf{w}_{\mathrm{NCMDF}(P_c,P_r)}$ | **12.7** | **13.2** | 2.2 | 12.2 | 13.8 | **5.5** | **11.2** | **11.2** | −0.3 | 13.2 | **13.7** | **9.1** | **11.7** | **10.7** | **10.1** | **11.7** | **12.9** | **−1.8** |

when $\theta_s = 45°$ the RR improvement depends on the type of the beamformer, and when $\theta_s = 75°$, the RR typically worsens. We infer that to obtain a reduction in the reverberations of the desired signal with considerable beam steering requirements, a multistage differential URA should be preferred over a multistage differential ULA and its configuration should not be set to $(P_c, P_r) = (0, 0)$. Addressing the IR, we note that a similar analysis applies for this measure as well. However, it is shown that except for with $\mathbf{w}_{\mathrm{MWNG}(P_c,P_r)}$, the IR with the URA is typically superior to the IR with the ULA — even for $\theta_s = 15°$.

We end this section by analyzing the average PESQ (Rix et al., 2001) and STOI (Taal et al., 2011) scores of the time-domain enhanced signals with $\mathbf{w}_{\mathrm{NCMWNG}(P_c,P_r)}$ and $\mathbf{w}_{\mathrm{NCMDF}(P_c,P_r)}$ with both the URA and ULA, and with all three analyzed incident angles. The corresponding scores are depicted in Fig. 7. To begin with, we observe that in terms of both scores and with all three incident angles — superior performance is obtained with the URA, that is, by applying a multistage rectangular differential beamformer. Moreover, considering the multistage linear differential beamformers of the ULA, it is evident that high values of $P_c$ degrade the PESQ and STOI scores. In contrast, with the URA, higher $(P_c, P_r)$ configurations improve these scores. For example, for $\theta_s = 15°$, with $(P_c, P_r) = (2, 0)$ the PESQ score is significantly higher with both rectangular beamformers with respect to with $(P_c, P_r) = (0, 0)$, and for $\theta_s = 45°$ both scores are maximized with $\mathbf{w}_{\mathrm{NCMWNG}(1,1)}$. On the contrary, for $\theta_s = 75°$, both rectangular beamformers considerably degrade the speech signal with $(P_c, P_r) = (0, 2)$. This is an artifact of the inability to attenuate the undesirable directional interferences and their reverberations, as reflected in the IR measure. Considering the power beampatterns, this may be embodied in the shape of amplified directions from which reverberations are impinging on the array. We also note that in this case, the rectangular differential beamformers are relatively short: they are of length 10, implying that $P_c$ and $P_r$ should be set carefully, as the increase in either of them results in a decrease in the differential beamformer's length and its degrees of freedom. We infer that considering beam steering, and by appropriately choosing the $(P_c, P_r)$ configuration, the multistage rectangular differential approach outperforms the multistage linear differential approach in terms of the enhanced speech quality and intelligibility.

### 6.3. Comparison to existing rectangular differential beamformers

In this part, we compare the proposed rectangular differential beamforming approach to two existing rectangular differential beamformers from the literature.

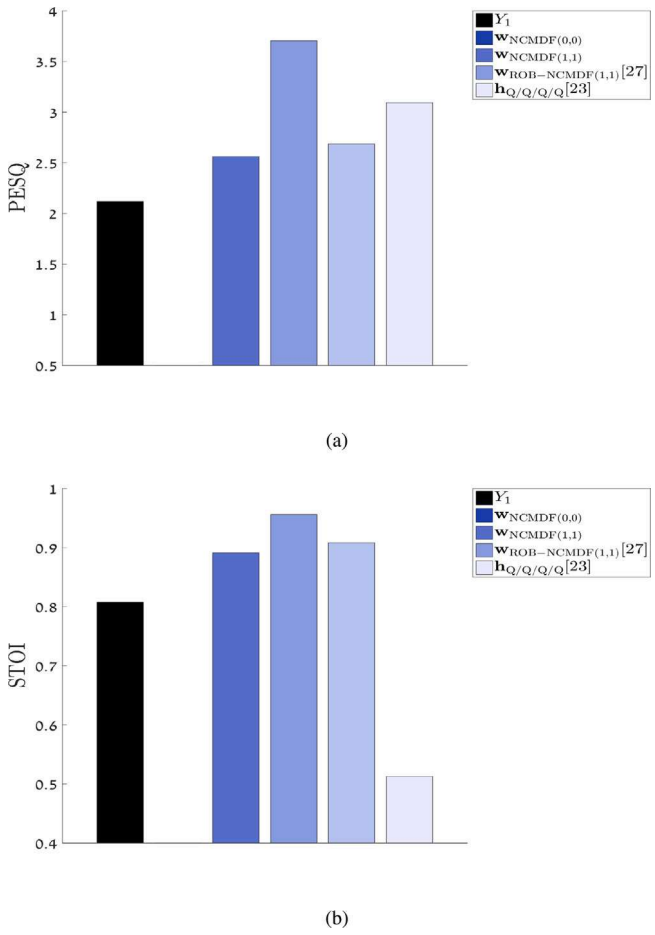


**Fig. 8.** Average PESQ and STOI scores comparison between the NCMDF rectangular differential beamformer $\mathbf{w}_{\mathrm{NCMDF}(P_c,P_r)}$ and two existing rectangular differential beamformers. (a) Average PESQ scores and (b) average STOI scores. Simulation parameters: $\theta_s = 45^0$, $M_x = M_y = 5$ and $\delta_x = \delta_y = 1$ cm.

To allow a fair comparison, we modify the reverberant settings from the previous part as follows. The URA is changed to an $M_x \times M_y = 5 \times 5$ microphone array whose interelement spacing is $\delta_x = \delta_y = 1$ cm. In addition, we set the positions of the directional interferences, which are now equally powerful, to $180^0$ and 270°, respectively, implying each is approximately 6 dB weaker than the desired signal. Finally, we configure $T_{60} = 200$ ms, $\theta_s = 45°$, and maintain all other settings, including the desired signal, the spatially white noise power, and the location of the array as before.

We compare the NCMDF rectangular differential beamformer $\mathbf{w}_{\mathrm{NCMDF}(P_c,P_r)}$ with $(P_c, P_r) = (0,0)$ and $(P_c, P_r) = (1,1)$ to the robust NCMDF rectangular differential beamformer of Itzhak et al. (2021b) with its scenario-optimal design parameters, denoted by $\mathbf{w}_{\mathrm{ROB-NCMDF}(1,1)}$, and the square-differential quadrupole beamformer of Zhao et al. (2021). More specifically, we extend the proposed $\mathbf{h}_{\mathrm{Q/Q}}$ beamformer of Zhao et al. (2021) which fits a squared $3 \times 3$ differential array to a $\mathbf{h}_{\mathrm{Q/Q/Q/Q}}$ beamformer which fits our case. The extension is carried out as suggested in the original paper.

The average PESQ and STOI scores with the four beamformers are depicted in Fig. 8. We observe that $\mathbf{w}_{\mathrm{NCMDF}(1,1)}$ outperforms the rest of the beamformers by a great deal, exhibiting a considerable performance gap in terms of both the quality and intelligibility of the enhanced signal. In addition, $\mathbf{h}_{\mathrm{Q/Q/Q/Q}}$ achieves the second-best PESQ score. However, it significantly degrades the intelligibility of the observed signal. This could be an artifact of the significant white noise amplification reported in Zhao et al. (2021). Finally, we observe that with $\mathbf{w}_{\mathrm{ROB-NCMDF}(1,1)}$ both scores are higher with respect to the base configuration of $\mathbf{w}_{\mathrm{NCMDF}(0,0)}$, resulted by the former's mitigated level of white noise amplification, as reported in Itzhak et al. (2021b).

We end this section by comparing the power beampatterns of the four beamformers shown in Fig. 9. Firstly, we note that the improved directivity of $\mathbf{w}_{\mathrm{NCMDF}(1,1)}$ with respect to $\mathbf{w}_{\mathrm{NCMDF}(0,0)}$ is embodied in the shape of a considerably reduced main lobe width and an attenuated back lobe. On the contrary, the main lobe of $\mathbf{w}_{\mathrm{ROB-NCMDF}(1,1)}$ is the widest, implying lower array directivity. Finally, we note that the power beampattern of $\mathbf{h}_{\mathrm{Q/Q/Q/Q}}$ zeros the directions of the interferences, and its directivity is high. On the contrary, its general shape greatly differs from the beampattern of the WNG-optimal DS beamformer, whose main lobe is typically wide, and its side and back lobes are attenuated with respect to the main lobe. This implies that $\mathbf{h}_{\mathrm{Q/Q/Q/Q}}$ exhibits a considerable level of spatially white noise amplification.

## 7. Conclusions

In this paper, we have presented a multistage rectangular approach for steerable differential beamforming. As a first step, we have proposed to employ a 2-D differentiation scheme that operates independently on the columns and rows of the observed signals of the rectangular array. This yields a differentials matrix of the noisy observations. Then, as a second step, a rectangular differential beamformer is designed and applied to the vector form of the differentials matrix. We have shown that the proposed differentiation scheme may significantly improve the directivity of the resulting beamformer at the expense of white noise amplification. This improvement depends on an appropriate selection of the $(P_c, P_r)$ configuration, which is done per the desired signal incident angle. For example, when the incident angle is close to the endfire direction, $P_c$ should be larger than $P_r$; when the incident angle is close to the broadside direction, $P_r$ should be larger than $P_c$; and

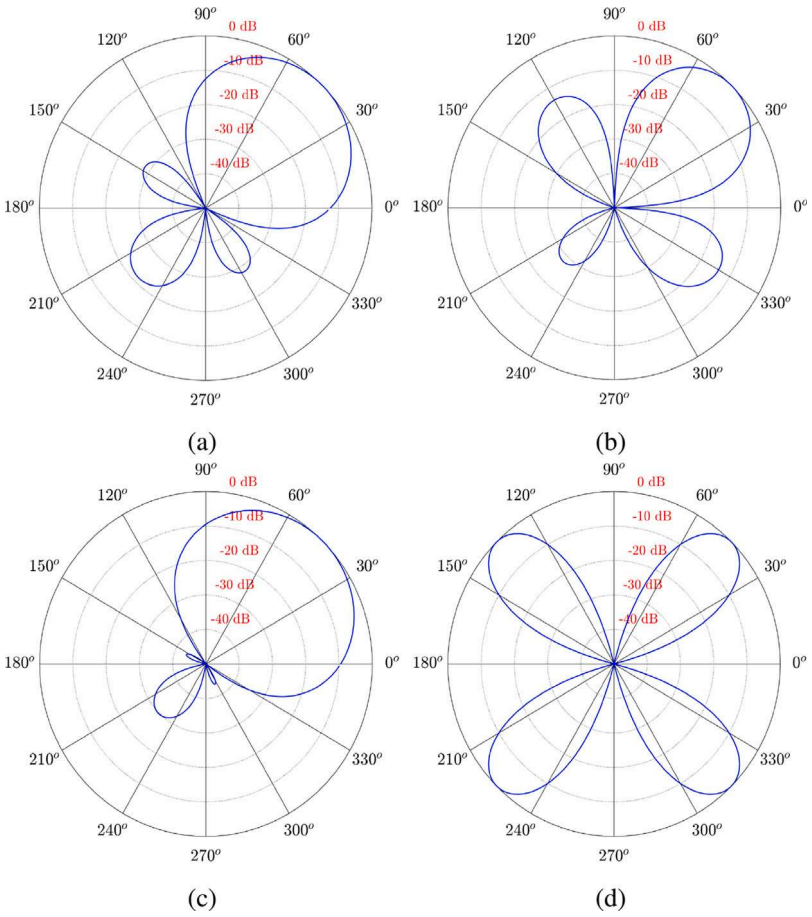


**Fig. 9.** Power beampatterns of the proposed NCMDF rectangular differential beamformer $\mathbf{w}_{\mathrm{NCMDF}(P_c,P_r)}$ and two other existing rectangular differential beamformers. (a) $\mathbf{w}_{\mathrm{NCMDF}(0,0)}$, (b) $\mathbf{w}_{\mathrm{NCMDF}(1,1)}$, (c) $\mathbf{w}_{\mathrm{ROB-NCMDF}(1,1)}$[27] and (d) $\mathbf{h}_{\mathrm{Q/Q/Q/Q}}$[23]. Simulation parameters: $\theta_s = 45^0$, $M_x = M_y = 5$, $\delta_x = \delta_y = 1$ cm and $f = 1$ kHz.

when the incident angle is $45^0$, $P_c$ and $P_r$ should be roughly equal. We have proposed four rectangular differential beamforming: two maximum SNR gain beamformers, $\mathbf{w}_{\mathrm{MWNG}(P_c,P_r)}$ and $\mathbf{w}_{\mathrm{MDF}(P_c,P_r)}$, as well as their two null-constrained versions, $\mathbf{w}_{\mathrm{NCMWNG}(P_c,P_r)}$ and $\mathbf{w}_{\mathrm{NCMDF}(P_c,P_r)}$. We have analyzed their performance in terms of both the WNG and DF measures for different $(P_c, P_r)$ configurations and demonstrated that the configuration should be set according to the desired signal incident angle and the desired level of array directivity. Moreover, addressing reverberant scenarios with sound sources located on the URA plane, we have analyzed and compared the WNR, DNR, RR, and IR of all the presented beamformers. We have shown that when the desired signal incident angle is not close to the endfire direction, the DNR, RR, and IR are highly preferable with the proposed rectangular differential approach compared to the former linear differential approach. Considering average PESQ and STOI scores of time-domain enhanced signals, we have demonstrated the null-constrained rectangular differential beamformers to outperform their linear counterparts for all three analyzed incident angles. Finally, we have compared our proposed approach to two existing rectangular differential beamformers from the literature. We have shown the proposed approach to be preferable — considering both the quality and intelligibility of the enhanced signals. In future work, we may analyze the performance of the proposed approach concerning elevated sound sources not necessarily located on the URA plane.

## CRediT authorship contribution statement

**Gal Itzhak:** Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **Jacob Benesty:** Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. **Israel Cohen:** Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Writing – review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

No data was used for the research described in the article.

## References

Allen, J.B., Berkley, D.A., 1979. Image method for efficiently simulating small-room acoustics. J. Acoust. Soc. Am. 65 (4), 943–950.
Anon, 1993. DARPA TIMIT Acoustic Phonetic Continuous Speech Corpus CDROM. NIST.
Bai, M., I.H., J., Benesty, J., 2014. Acoustic Array Systems: Theory, Implementation, and Application. Wiley-IEEE Press, New York.
Benesty, J., Chen, J., 2013. Study and Design of Differential Microphone Arrays. Springer-Verlag Berlin Heidelberg, Berlin.
Benesty, J., Chen, J., Habets, E., 2012. Speech Enhancement in the STFT Domain. Springer-Verlag Berlin Heidelberg, Berlin.
Benesty, J., Chen, J., Huang, Y., Cohen, I., 2009. Noise Reduction in Speech Processing, first ed. Springer-Verlag Berlin Heidelberg, Berlin.
Benesty, J., Cohen, I., Chen, J., 2017. Fundamentals of Signal Enhancement and Array Signal Processing, first ed. Wiley-IEEE Press, New York.
Benesty, J., Cohen, I., Chen, J., 2019. Array Processing - Kronecker Product Beamforming. Springer-Verlag, Switzerland.

Bernardini, A., Antonacci, F., Sarti, A., 2018. Wave digital implementation of robust first-order differential microphone arrays. IEEE Signal Process. Lett. 25 (2), 253–257.

Borra, F., Bernardini, A., Antonacci, F., Sarti, A., 2019. Uniform linear arrays of first-order steerable differential microphones. IEEE/ACM Trans. Audio Speech Lang. Process. 27 (12), 1906–1918.

Buchris, Y., Cohen, I., Benesty, J., 2018. Frequency-domain design of asymmetric circular differential microphone arrays. IEEE/ACM Trans. Audio Speech Lang. Process. 26 (4), 760–773.

Buck, M., 2002. Aspects of first-order differential microphone arrays in the presence of sensor imperfections. Eur. Trans. Telecommun. 13, 115–122.

Chen, J., Benesty, J., Pan, C., 2014. On the design and implementation of linear differential microphone arrays. J. Acoust. Soc. Am. 136 (6), 3097–3113.

Cohen, I., Benesty, J., Chen, J., 2019. Differential kronecker product beamforming. IEEE/ACM Trans. Audio Speech Lang. Process. 27 (5), 892–902.

Elko, G., 2004. Differential microphone arrays. In: Huang, Y., Benesty, J. (Eds.), Audio Signal Processing for Next-Generation Multimedia Communication Systems. Springer US, Boston, MA, pp. 11–65.

Habets, E., 2008. Room impulse response (RIR) generator.

Huang, G., Benesty, J., Chen, J., 2017. Design of robust concentric circular differential microphone arrays. J. Acoust. Soc. Am. 141 (5), 3236–3249.

Huang, G., Benesty, J., Chen, J., Cohen, I., 2020a. Robust and steerable kronecker product differential beamforming with rectangular microphone arrays. In: ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing. ICASSP, pp. 211–215.

Huang, G., Benesty, J., Cohen, I., Chen, J., 2020b. A simple theory and new method of differential beamforming with uniform linear microphone arrays. IEEE/ACM Trans. Audio Speech Lang. Process. 28, 1079–1093.

Huang, G., Chen, J., Benesty, J., 2018. On the design of differential beamformers with arbitrary planar microphone array geometry. J. Acoust. Soc. Am. 144 (1), EL66–EL70.

Huang, G., Chen, J., Benesty, J., 2020c. Design of planar differential microphone arrays with fractional orders. IEEE/ACM Trans. Audio Speech Lang. Process. 28, 116–130.

Huang, G., Chen, J., Benesty, J., Cohen, I., Zhao, X., 2020d. Steerable differential beamformers with planar microphone arrays. EURASIP J. Audio Speech Music Proc. 2020.

Itzhak, G., Benesty, J., Cohen, I., 2021a. On the design of differential kronecker product beamformers. IEEE/ACM Trans. Audio Speech Lang. Process. 29, 1397–1410.

Itzhak, G., Cohen, I., Benesty, J., 2021b. Robust differential beamforming with rectangular arrays. In: Proc. 29th European Signal Processing Conference, EUSIPCO-2021.

Johnson, D.H., Dudgeon, D.E., 1992. Array Signal Processing: Concepts and Techniques. Simon and Schuster, Inc., USA.

Kolundzija, M., Faller, C., Vetterli, M., 2011. Spatiotemporal gradient analysis of differential microphone arrays. J. Audio Eng. Soc. 59 (1), 20–28.

Olson, H., 1946. Gradient microphones. J. Acoust. Soc. Am. 17 (3), 192–198.

Pierce, A., 2019. Acoustics: An Introduction to Its Physical Principles and Applications. Springer International Publishing, Switzerland.

Rix, A.W., Beerends, J.G., Hollier, M.P., Hekstra, A.P., 2001. Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs. In: 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings, vol. 2. pp. 749–752 vol.2.

Sena, E.D., Hacihabiboglu, H., Cvetkovic, Z., 2012. On the design and implementation of higher order differential microphones. IEEE Trans. Audio Speech Lang. Process. 20 (1), 162–174.

Taal, C.H., Hendriks, R.C., Heusdens, R., Jensen, J., 2011. An algorithm for intelligibility prediction of time–frequency weighted noisy speech. IEEE Trans. Audio Speech Lang. Process. 19 (7), 2125–2136.

Van Trees, H., 2004. Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory. In: Detection, Estimation, and Modulation Theory, Wiley, New York.

Weinberger, J., Olson, H., Massa, F., 1933. A uni-directional ribbon microphone. J. Acoust. Soc. Am. 5 (2), 139–147.

Wu, X., Chen, H., 2016. Directivity factors of the first-order steerable differential array with microphone mismatches: Deterministic and worst-case analysis. IEEE/ACM Trans. Audio Speech Lang. Process. 24 (2), 300–315.

Zhao, X., Huang, G., Benesty, J., Chen, J., Cohen, I., 2021. On the design of square differential microphone arrays with a multistage structure. In: ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing. ICASSP, pp. 746–750.