Robust Differential Beamforming with Rectangular Arrays

Gal Itzhak
Faculty of Electrical Engineering
Technion –Israel Institute of Technology
Haifa, Israel
galitz@campus.technion.ac.il

Israel Cohen
Faculty of Electrical Engineering
Technion –Israel Institute of Technology
Haifa, Israel
icohen@ee.technion.ac.il

Jacob Benesty
INRS-EMT, University of Quebec
Montreal, Canada
benesty@emt.inrs.ca

***Abstract*—In this paper, we present a robust approach for rectangular differential beamforming. At first, we propose to employ a 2-D multistage mean which operates independently on the columns and rows of the observation signals of a URA. The number of mean stages along the columns and rows of the URA is controlled by two parameters, Q_c and Q_r, respectively. Then, the resulting matrix is (column-wise) concatenated to a vector form to serve as a modified version of the observations. Finally, we design a rectangular differential beamformer and apply it to the latter. We show that the first mean operation improves the white noise robustness of the resulting beamformer. We focus on the maximum directivity factor (MDF) and null-constrained maximum directivity factor (NCMDF) differential beamformers and analyze their performances in terms of both the white noise gain (WNG) and directivity factor (DF) measures. We show that the configuration (Q_c, Q_r) constitutes a useful mean to mitigate the white noise amplification of differential beamformers in low frequencies.**

***Index Terms*—Microphone arrays, uniform rectangular arrays (URAs), differential beamforming, robust beamforming, two-dimensional (2-D) arrays.**

I. Introduction

Communication and speech signals are often degraded by undesired noise which may severely deteriorate the functionality of systems that involve such signals. To attenuate the undesired noise, sensor arrays, or beamformers, are often employed. That is, an array consisting of multiple microphones is used to simultaneously capture samples in different locations in space. Among the extensively-studied field of array signal processing, differential microphone arrays (DMAs) are known to be particularly suitable for practical applications for the two following reasons: their physical size is small and their beampatterns tend to be frequency-invariant [1], [2]. As a result, DMAs have been widely studied and optimized over the years [3]–[6].

Recently, there has been a growing interest in differential uniform rectangular arrays (URAs) [7], [8]. Taking advantage of the rectangular geometry, such DMAs exhibit a better beam steering performance than uniform linear arrays (ULAs) and better directivity than uniform circular arrays (UCAs). In addition, URA beamformers may be decomposed into sub-beamformers by employing the Kronecker-product (KP) decomposition. This allows a significant design flexibility: the KP decomposition is not unique and each of the sub-beamformers may be independently designed with respect to a different criterion [7]. Nevertheless, high-directivity rectangular differential beamformers tend to be sensitive to white noise, in particular in low frequencies [9]–[12].

In this paper, we present a robust approach for rectangular differential beamforming. First, we propose to employ a 2-D multistage mean which operates independently on the columns and rows of the observation signals of a URA. The number of mean stages along the columns and rows of the URA is controlled by two parameters, Q_c and Q_r, respectively. Then, the resulting matrix is (column-wise) concatenated to a vector form to serve as a modified version of the observations. Finally, we design a rectangular differential beamformer and apply it to the latter. We show that the first mean operation improves the white noise sensitivity of the resulted beamformer.

While there is some partial technical resemblance between this work and the approach presented in [8], the two greatly differ. That is, the objective of the multistage differentials in [8] is to improve the array directivity in a controlled manner, which is attained at the expense of an increased sensitivity to white noise. In contrast, in this work, our objective is the opposite as we aim to improve the white noise robustness. We focus on the MDF and NCMDF differential beamformers and analyze their performances in terms of both the WNG and DF measures. We show that the configuration (Q_c, Q_r) constitutes a useful mean to mitigate the white noise amplification of differential beamformers in the low frequency range.

II. Signal Model

Consider a two-dimensional (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 δ_{x} and M_{y} omnidirectional sensors along the y (negative) axis with a uniform interelement spacing equal to δ_{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, θ, in an anechoic acoustic environment at the

This research was supported by the ISF-NSFC joint research program (grant No. 2514/17), and the Pazy Research Foundation.

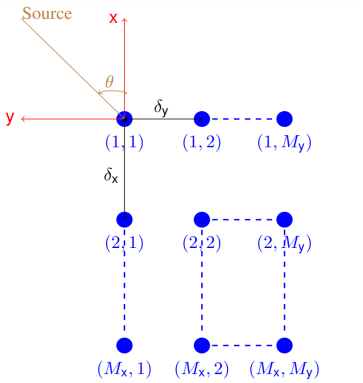


Fig. 1: Illustration of the studied rectangular microphone array.

speed of sound, i.e., $c = 340$ m/s, and impinges on the above described array. Then, the corresponding steering matrix (of size $M_{\mathsf{x}} \times M_{\mathsf{y}}$) is [1]:

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

where

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

is the steering vector associated with the x axis,

$$\begin{aligned}\mathbf{b}_\theta(\omega) &= \left[\begin{array}{cccc} B_{\theta,1}(\omega) & B_{\theta,2}(\omega) & \cdots & B_{\theta,M_{\mathsf{y}}}(\omega)\end{array}\right]^T \\ &= \left[\begin{array}{cccc} 1 & e^{-\jmath\varpi_{\theta,\mathsf{y}}(\omega)} & \cdots & e^{-\jmath(M_{\mathsf{y}}-1)\varpi_{\theta,\mathsf{y}}(\omega)}\end{array}\right]^T \end{aligned} \tag{3}$$

is the steering vector associated with the y axis,

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

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 $M_{\mathsf{x}} \times M_{\mathsf{y}}$ of the URA can be expressed in the frequency domain as [2]:

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

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 $M_{\mathsf{x}} \times M_{\mathsf{y}}$, 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

$$\begin{aligned}\mathbf{y}(\omega) &= \left[\begin{array}{cccc} \mathbf{y}_1^T(\omega) & \mathbf{y}_2^T(\omega) & \cdots & \mathbf{y}_{M_{\mathsf{y}}}^T(\omega)\end{array}\right]^T \\ &= \mathbf{d}_\theta(\omega)\,X(\omega) + \mathbf{v}(\omega), \end{aligned} \tag{6}$$

where

$$\begin{aligned}\mathbf{y}_{m_{\mathsf{y}}}(\omega) &= \left[\begin{array}{cccc} Y_{m_{\mathsf{y}},1}(\omega) & Y_{m_{\mathsf{y}},2}(\omega) & \cdots & Y_{m_{\mathsf{y}},M_{\mathsf{x}}}(\omega)\end{array}\right]^T \\ &= B_{\theta,m_{\mathsf{y}}}(\omega)\,\mathbf{a}_\theta(\omega)\,X(\omega) + \mathbf{v}_{m_{\mathsf{y}}}(\omega), \end{aligned} \tag{7}$$

for $m_{\mathsf{y}} = 1, 2, \ldots, M_{\mathsf{y}}$. Dropping the dependence on ω to simplify the notation, we define the covariance matrix of $\mathbf{y}$ by:

$$\mathbf{\Phi}_{\mathbf{y}} = E\left(\mathbf{y}\mathbf{y}^H\right) = \phi_X\mathbf{d}_\theta\mathbf{d}_\theta^H + \mathbf{\Phi}_{\mathbf{v}}, \tag{8}$$

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 (8) as:

$$\mathbf{\Phi}_{\mathbf{y}} = \phi_X\mathbf{d}_\theta\mathbf{d}_\theta^H + \phi_V\mathbf{\Gamma}_{\mathbf{v}}, \tag{9}$$

where ϕ_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 (9), we deduce that the input signal-to-noise ratio (SNR) is:

$$\text{iSNR} = \frac{\text{tr}\left(\phi_X\mathbf{d}_\theta\mathbf{d}_\theta^H\right)}{\text{tr}\left(\phi_V\mathbf{\Gamma}_{\mathbf{v}}\right)} = \frac{\phi_X}{\phi_V}, \tag{10}$$

where $\text{tr}(\cdot)$ denotes the trace of a square matrix.

III. Robust Differential Beamforming

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

$$\Sigma Y_{m_{\mathsf{y}},i} = Y_{m_{\mathsf{y}},i+1} + Y_{m_{\mathsf{y}},i} = Y_{m_{\mathsf{y}},(1),i}, \;\; i = 1, 2, \ldots, M_{\mathsf{x}} - 1, \tag{11}$$

where Σ is the forward spatial mean operator. Clearly, the forward spatial mean operator may be applied multiple times. in general, let $q = 0, 1, \ldots, Q_c$, with $1 \leq Q_c < M_{\mathsf{x}}$. Let us represent Σ in a vector/matrix form. By definition, we write $\mathbf{\Sigma}_{(0)} = \mathbf{I}_{M_{\mathsf{x}}}$,where $\mathbf{I}_{M_{\mathsf{x}}}$ is the $M_{\mathsf{x}} \times M_{\mathsf{x}}$ identity matrix. Therefore,

$$\mathbf{\Sigma}_{(0)}\mathbf{y}_{m_{\mathsf{y}}} = \mathbf{I}_{M_{\mathsf{x}}}\mathbf{y}_{m_{\mathsf{y}}} = \mathbf{y}_{m_{\mathsf{y}}}. \tag{12}$$

We define the qth-order forward spatial (unnormalized) mean of $\mathbf{y}_{m_{\mathsf{y}}}$ as:

$$\begin{aligned}\Sigma^q Y_{m_{\mathsf{y}},i} &= \Sigma^{q-1}\left(\Sigma Y_{m_{\mathsf{y}},i}\right) = \Sigma^{q-1}Y_{m_{\mathsf{y}},i+1} + \Sigma^{q-1}Y_{m_{\mathsf{y}},i} \\ &= \sum_{j=0}^{q}\left(\begin{array}{c} q \\ j \end{array}\right) Y_{m_{\mathsf{y}},i+j}, \end{aligned} \tag{13}$$

where $i = 1, 2, \ldots, M_{\mathsf{x}} - q$ and $\left(\begin{array}{c} q \\ j \end{array}\right)$ is the binomial coefficient. In a vector/matrix form, (13) is:

$$\mathbf{\Sigma}_{(Q_c)}\mathbf{y}_{m_{\mathsf{y}}} = \mathbf{y}_{m_{\mathsf{y}},(Q_c)}, \tag{14}$$

where

$$\mathbf{\Sigma}_{(Q_c)} = \begin{bmatrix} \mathbf{c}_{(Q_c)}^T & 0 & \cdots & 0 \\ 0 & \mathbf{c}_{(Q_c)}^T & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{c}_{(Q_c)}^T \end{bmatrix} \tag{15}$$

is a matrix of size $(M_{\mathsf{x}} - Q_c) \times M_{\mathsf{x}}$, with

$$\mathbf{c}_{(Q_c)} = \begin{bmatrix} \begin{pmatrix} Q_c \\ 0 \end{pmatrix} & \begin{pmatrix} Q_c \\ 1 \end{pmatrix} & \cdots & \begin{pmatrix} Q_c \\ Q_c - 1 \end{pmatrix} & 1 \end{bmatrix}^T \tag{16}$$

being a vector of length $Q_c + 1$.

Now, substituting (7) into (13), it can be shown that:

$$\begin{aligned} \mathbf{\Sigma}_{(Q_c)} \mathbf{y}_{m_{\mathsf{y}}} &= B_{\theta,m_{\mathsf{y}}} \mathbf{\Sigma}_{(Q_c)} \mathbf{a}_\theta X + \mathbf{\Sigma}_{(Q_c)} \mathbf{v}_{m_{\mathsf{y}}} \\ &= B_{\theta,m_{\mathsf{y}}} \mu_{\theta,\mathsf{x}}^{Q_c} \mathbf{a}_{\theta,(Q_c)} X + \mathbf{v}_{m_{\mathsf{y}},(Q_c)} \\ &= \mathbf{y}_{m_{\mathsf{y}},(Q_c)}, \end{aligned} \tag{17}$$

where

$$\mu_{\theta,\mathsf{x}} = e^{-\jmath \varpi_{\theta,\mathsf{x}}} + 1, \tag{18}$$

$$\mathbf{a}_{\theta,(Q_c)} = \begin{bmatrix} 1 & e^{-\jmath \varpi_{\theta,\mathsf{x}}} & \cdots & e^{-\jmath (M_{\mathsf{x}} - Q_c - 1) \varpi_{\theta,\mathsf{x}}} \end{bmatrix}^T \tag{19}$$

is the steering vector of length $M_{\mathsf{x}} - Q_c$, and $\mathbf{v}_{m_{\mathsf{y}},(Q_c)} = \mathbf{\Sigma}_{(Q_c)} \mathbf{v}_{m_{\mathsf{y}}}$. In an analogous manner, equations (11)-(19) can be rewritten with the roles of x and y axes interchanged. That is, we may average over the rows of $\mathbf{Y}$ instead of over its columns. Define:

$$\mu_{\theta,\mathsf{y}} = e^{-\jmath \varpi_{\theta,\mathsf{y}}} + 1, \tag{20}$$

$$\mathbf{b}_{\theta,(Q_r)} = \begin{bmatrix} 1 & e^{-\jmath \varpi_{\theta,\mathsf{y}}} & \cdots & e^{-\jmath (M_{\mathsf{y}} - Q_r - 1) \varpi_{\theta,\mathsf{y}}} \end{bmatrix}^T, \tag{21}$$

with $1 \leq Q_r < M_{\mathsf{y}}$. Recalling the matrix form in (4), we may define:

$$\begin{aligned} \mathbf{Y}_{(Q_c,Q_r)} &= \mathbf{\Sigma}_{(Q_c)} \mathbf{Y} \mathbf{\Sigma}_{(Q_r)}^T \\ &= \mu_{\theta,\mathsf{x}}^{Q_c} \mu_{\theta,\mathsf{y}}^{Q_r} \left(\mathbf{b}_{\theta,(Q_r)}^T \otimes \mathbf{a}_{\theta,(Q_c)} \right) X \\ &+ \mathbf{\Sigma}_{(Q_c)} \mathbf{V} \mathbf{\Sigma}_{(Q_r)}^T. \end{aligned} \tag{22}$$

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

$$\begin{aligned} \mathbf{y}_{(Q_c,Q_r)} &= \mathrm{vec}\left[\mathbf{Y}_{(Q_c,Q_r)} \right] \\ &= \mu_{\theta,\mathsf{x}}^{Q_c} \mu_{\theta,\mathsf{y}}^{Q_r} \mathbf{d}_{\theta,(Q_c,Q_r)} X \\ &+ \left(\Sigma_{(Q_r)} \otimes \Sigma_{(Q_c)} \right) \mathbf{v}, \end{aligned} \tag{23}$$

where $\mathbf{d}_{\theta,(Q_c,Q_r)} = \mathbf{b}_{\theta,(Q_r)} \otimes \mathbf{a}_{\theta,(Q_c)}$ is a 2-D differential steering vector of length $(M_{\mathsf{x}} - Q_c)(M_{\mathsf{y}} - Q_r)$. We deduce that the $(M_{\mathsf{x}} - Q_c)(M_{\mathsf{y}} - Q_r) \times (M_{\mathsf{x}} - Q_c)(M_{\mathsf{y}} - Q_r)$ covariance matrix of $\mathbf{y}_{(Q_c,Q_r)}$ is:

$$\begin{aligned} \mathbf{\Phi}_{\mathbf{y}_{(Q_c,Q_r)}} &= \phi_X \left| \mu_{\theta,\mathsf{x}} \right|^{2Q_c} \left| \mu_{\theta,\mathsf{y}} \right|^{2Q_r} \\ &\times \mathbf{d}_{\theta,(Q_c,Q_r)} \mathbf{d}_{\theta,(Q_c,Q_r)}^H \\ &+ \phi_V \left(\Sigma_{(Q_r)} \otimes \Sigma_{(Q_c)} \right) \mathbf{\Gamma}_{\mathbf{v}} \left(\Sigma_{(Q_r)} \otimes \Sigma_{(Q_c)} \right)^T. \end{aligned} \tag{24}$$

We immediately obtain the WNG and DF between $\mathbf{y}_{(Q_c,Q_r)}$ and $\mathbf{y}$:

$$\begin{aligned} \mathcal{W}_{(Q_c,Q_r)} &= \frac{\left| \mu_{\theta,\mathsf{x}} \right|^{2Q_c} \left| \mu_{\theta,\mathsf{y}} \right|^{2Q_r} (M_{\mathsf{x}} - Q_c)(M_{\mathsf{y}} - Q_r)}{\mathrm{tr}\left(\left(\Sigma_{(Q_r)} \otimes \Sigma_{(Q_c)} \right) \left(\Sigma_{(Q_r)} \otimes \Sigma_{(Q_c)} \right)^T \right)} \\ &= \frac{\left| \mu_{\theta,\mathsf{x}} \right|^{2Q_c}}{\begin{pmatrix} 2Q_c \\ Q_c \end{pmatrix}} \times \frac{\left| \mu_{\theta,\mathsf{y}} \right|^{2Q_r}}{\begin{pmatrix} 2Q_r \\ Q_r \end{pmatrix}}, \end{aligned} \tag{25}$$

$$\mathcal{D}_{(Q_c,Q_r)} = \frac{\left| \mu_{\theta,\mathsf{x}} \right|^{2Q_c} \left| \mu_{\theta,\mathsf{y}} \right|^{2Q_r} (M_{\mathsf{x}} - Q_c)(M_{\mathsf{y}} - Q_r)}{\mathrm{tr}\left(\left(\Sigma_{(Q_r)} \otimes \Sigma_{(Q_c)} \right) \mathbf{\Gamma}_{\mathrm{d}} \left(\Sigma_{(Q_r)} \otimes \Sigma_{(Q_c)} \right)^T \right)}. \tag{26}$$

where $\mathbf{\Gamma}_{\mathrm{d}}$ is the pseudo-coherence matrix of the spherically-isotropic diffuse noise [8].

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

$$Z_{(Q_c,Q_r)} = \mathbf{w}_{(Q_c,Q_r)}^H \mathbf{y}_{(Q_c,Q_r)} = X_{\mathrm{fd},(Q_c,Q_r)} + V_{\mathrm{rn},(Q_c,Q_r)}, \tag{27}$$

where $Z_{(Q_c,Q_r)}$ is the estimate of X,

$$X_{\mathrm{fd},(Q_c,Q_r)} = \mu_{\theta,\mathsf{x}}^{Q_c} \mu_{\theta,\mathsf{y}}^{Q_r} \left(\mathbf{w}^H \mathbf{d}_{\theta,(Q_c,Q_r)} \right) X \tag{28}$$

is the filtered desired signal, and:

$$V_{\mathrm{rn},(Q_c,Q_r)} = \mathbf{w}_{(Q_c,Q_r)}^H \mathbf{v}_{(Q_c,Q_r)} \tag{29}$$

is the residual noise, where $\mathbf{v}_{(Q_c,Q_r)} = \left(\Sigma_{(Q_r)} \otimes \Sigma_{(Q_c)} \right) \mathbf{v}$. Consequently, the variance of $Z_{(Q_c,Q_r)}$ is:

$$\begin{aligned} \phi_{Z_{(Q_c,Q_r)}} &= \mathbf{w}_{(Q_c,Q_r)}^H \mathbf{\Phi}_{\mathbf{y}_{(Q_c,Q_r)}} \mathbf{w}_{(Q_c,Q_r)} \\ &= \phi_{X_{\mathrm{fd},(Q_c,Q_r)}} + \phi_{V_{\mathrm{rn},(Q_c,Q_r)}}, \end{aligned} \tag{30}$$

where

$$\begin{aligned} \phi_{X_{\mathrm{fd},(Q_c,Q_r)}} &= \phi_X \left| \mu_{\theta,\mathsf{x}} \right|^{2Q_c} \left| \mu_{\theta,\mathsf{y}} \right|^{2Q_r} \\ &\times \left| \mathbf{w}_{(Q_c,Q_r)}^H \mathbf{d}_{\theta,(Q_c,Q_r)} \right|^2, \end{aligned} \tag{31}$$

$$\phi_{V_{\mathrm{rn},(Q_c,Q_r)}} = \mathbf{w}_{(Q_c,Q_r)}^H \mathbf{\Phi}_{\mathbf{v}_{(Q_c,Q_r)}} \mathbf{w}_{(Q_c,Q_r)}, \tag{32}$$

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

$$\begin{aligned} \mathbf{\Phi}_{\mathbf{v}_{(Q_c,Q_r)}} &= \phi_V \left(\Sigma_{(Q_r)} \otimes \Sigma_{(Q_c)} \right) \mathbf{\Gamma}_{\mathbf{v}} \left(\Sigma_{(Q_r)} \otimes \Sigma_{(Q_c)} \right)^T \\ &= \phi_V \mathbf{\Gamma}_{\mathbf{v}_{(Q_c,Q_r)}}. \end{aligned} \tag{33}$$

Ultimately, it is clear that the distortionless constraint is given by:

$$\mathbf{w}_{(Q_c,Q_r)}^H \mathbf{d}_{\theta,(Q_c,Q_r)} = \mu_{\theta,\mathsf{x}}^{-Q_c} \mu_{\theta,\mathsf{y}}^{-Q_r}. \tag{34}$$

Now, let us relate the SNR gains corresponding to $\mathbf{w}_{(Q_c,Q_r)}$. It is clear from (30)-(32) that the WNG and DF are given by:

$$\begin{aligned} \mathcal{W}\left(\mathbf{w}_{(Q_c,Q_r)} \right) &= \left| \mu_{\theta,\mathsf{x}} \right|^{2Q_c} \left| \mu_{\theta,\mathsf{y}} \right|^{2Q_r} \\ &\times \frac{\left| \mathbf{w}_{(Q_c,Q_r)}^H \mathbf{d}_{\theta,(Q_c,Q_r)} \right|^2}{\mathbf{w}_{(Q_c,Q_r)}^H \mathbf{\Xi}_{(Q_c,Q_r)} \mathbf{w}_{(Q_c,Q_r)}}, \end{aligned} \tag{35}$$

and the DF:

$$\mathcal{D}\left(\mathbf{w}_{(Q_c,Q_r)}\right) = \left|\mu_{\theta,\mathsf{x}}\right|^{2Q_c}\left|\mu_{\theta,\mathsf{y}}\right|^{2Q_r} \times \frac{\left|\mathbf{w}^H_{(Q_c,Q_r)}\mathbf{d}_{\theta,(Q_c,Q_r)}\right|^2}{\mathbf{w}^H_{(Q_c,Q_r)}\mathbf{\Gamma}_{\mathrm{d},(Q_c,Q_r)}\mathbf{w}_{(Q_c,Q_r)}}, \quad (36)$$

where

$$\mathbf{\Xi}_{(Q_c,Q_r)} = \left(\Sigma_{(Q_r)}\otimes\Sigma_{(Q_c)}\right)\left(\Sigma_{(Q_r)}\otimes\Sigma_{(Q_c)}\right)^T, \quad (37)$$

$$\mathbf{\Gamma}_{\mathrm{d},(Q_c,Q_r)} = \left(\Sigma_{(Q_r)}\otimes\Sigma_{(Q_c)}\right)\mathbf{\Gamma}_{\mathrm{d}}\left(\Sigma_{(Q_r)}\otimes\Sigma_{(Q_c)}\right)^T. \quad (38)$$

IV. Optimal Robust Differential Beamformers

Let us start by considering equation (36). The maximum DF (MDF) beamformer is derived from:

$$\min_{\mathbf{w}_{(Q_c,Q_r)}} \mathbf{w}^H_{(Q_c,Q_r)}\mathbf{\Gamma}_{\mathrm{d},(Q_c,Q_r)}\mathbf{w}_{(Q_c,Q_r)} \quad \text{s. t.} \quad \mathbf{w}^H_{(Q_c,Q_r)}\mathbf{d}_{\theta,(Q_c,Q_r)} = \mu^{-Q_c}_{\theta,\mathsf{x}}\mu^{-Q_r}_{\theta,\mathsf{y}}, \quad (39)$$

in which we considered the distortionless constraint. The solution is therefore given by:

$$\mathbf{w}_{\mathrm{MDF}(Q_c,Q_r)} = \frac{1}{\left(\mu^{Q_c}_{\theta,\mathsf{x}}\mu^{Q_r}_{\theta,\mathsf{y}}\right)^*} \times \frac{\mathbf{\Gamma}^{-1}_{\mathrm{d},(Q_c,Q_r)}\mathbf{d}_{\theta,(Q_c,Q_r)}}{\mathbf{d}^H_{\theta,(Q_c,Q_r)}\mathbf{\Gamma}^{-1}_{\mathrm{d},(Q_c,Q_r)}\mathbf{d}_{\theta,(Q_c,Q_r)}}. \quad (40)$$

We now turn to null-constrained version of $\mathbf{w}_{\mathrm{MDF}(Q_c,Q_r)}$. In practice, in order to give a desired shape to a beampattern or attenuate directional interferences, spatial null constraints may be required. For example, with $N = 2$ distinct null constraints (39) is transformed into:

$$\min_{\mathbf{w}} \mathbf{w}^H\mathbf{\Gamma}_{\mathrm{d},(Q_c,Q_r)}\mathbf{w} \quad \text{s. t.} \quad \mathbf{C}^H\left(\Sigma_{(Q_r)}\otimes\Sigma_{(Q_c)}\right)^T\mathbf{w}_{(Q_c,Q_r)} = \begin{bmatrix}1\\0\\0\end{bmatrix}. \quad (41)$$

where $\mathbf{C}$ is a constraint matrix of size $M_{\mathsf{x}}M_{\mathsf{y}}\times 3$:

$$\mathbf{C} = \begin{bmatrix}\mathbf{d}_\theta & \mathbf{d}_{\theta_1} & \mathbf{d}_{\theta_2}\end{bmatrix}, \quad (42)$$

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 MDF (NCMDF) beamformer is given by:

$$\mathbf{w}_{\mathrm{NCMDF}(Q_c,Q_r)} = \mathbf{\Gamma}^{-1}_{\mathrm{d},(Q_c,Q_r)}\left(\Sigma_{(Q_r)}\otimes\Sigma_{(Q_c)}\right)\mathbf{C} \times\left[\mathbf{C}^H\left(\Sigma_{(Q_r)}\otimes\Sigma_{(Q_c)}\right)^T\mathbf{\Gamma}^{-1}_{\mathrm{d},(Q_c,Q_r)} \times\left(\Sigma_{(Q_r)}\otimes\Sigma_{(Q_c)}\right)\mathbf{C}\right]^{-1}\begin{bmatrix}1\\0\\0\end{bmatrix}. \quad (43)$$

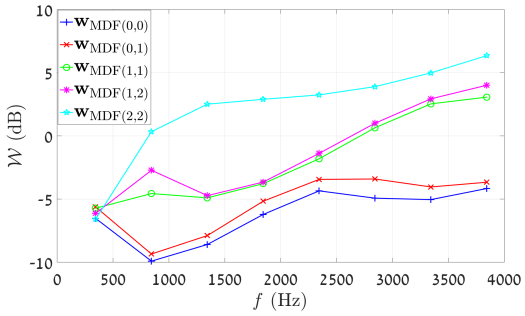
TABLE I: The WNG and DF in dB units between $\mathbf{y}_{(Q_c,Q_r)}$ and $\mathbf{y}$ for varying values of (Q_c,Q_r) and $f = 3$ kHz. Gray background color indicates optimal configurations further discussed in the paper. Simulation parameters: $\theta = 0^o$,$M_{\mathsf{x}} = 5$, $M_{\mathsf{y}} = 5$, $\delta_{\mathsf{x}} = 1$ cm and $\delta_{\mathsf{y}} = 1.5$ cm.

$\mathcal{W}_{(Q_c,Q_r)}$		Q_r					$\mathcal{D}_{(Q_c,Q_r)}$		Q_r				
		0	1	2	3	4			0	1	2	3	4
Q_c	0	0.0	3.0	4.3	5.1	5.6	Q_c	0	0.0	0.2	0.5	0.7	0.9
	1	2.7	5.7	6.9	7.7	8.3		1	-0.2	0.0	0.3	0.5	0.7
	2	3.6	6.6	7.9	8.6	9.2		2	-0.5	-0.2	0.0	0.3	0.5
	3	4.0	7.0	8.3	9.1	9.7		3	-0.7	-0.4	-0.2	0.0	0.3
	4	4.3	7.3	8.5	9.3	9.9		4	-0.9	-0.7	-0.4	-0.2	0.0

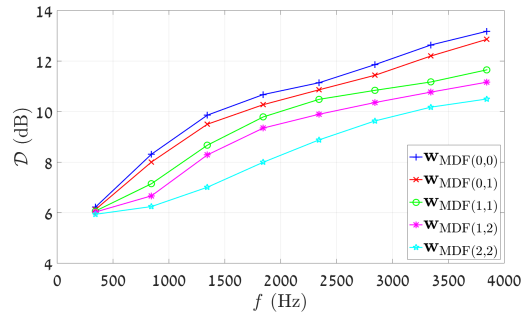
V. Simulations

For the purpose of the simulative part of the paper, let us assume $\theta = 0^o$ as well as the following URA: $M_{\mathsf{x}} = 5$, $M_{\mathsf{y}} = 5$, $\delta_{\mathsf{x}} = 1$ cm and $\delta_{\mathsf{y}} = 1.5$ cm. To begin with, it is valuable to evaluate the WNG and DF between $\mathbf{y}_{(Q_c,Q_r)}$ and $\mathbf{y}$. Recalling equations (25) and (26), we realize that the optimal configurations are scenario-dependent and are a function of the rectangular array structure and the desired signal incident angle. For example, $\mathcal{W}_{(Q_c,Q_r)}$ and $\mathcal{D}_{(Q_c,Q_r)}$ in our scenario are elaborated in Table I for $f = 3$ kHz. Analyzing the results, it is intuitively clear that averaging along the rows would be more beneficial than averaging along the columns. That is, the desired signal is, in fact, in broadside with respect to each ULA along the y axis. Therefore, applying the mean along the rows resembles the application of a series of delay-and-sum beamformers (which are known to be optimal in terms of the WNG) of length 2 to each two adjacent samples. In addition, we note that in case $Q_r \geq Q_c$ averaging over the rows yields either zero or small positive values of $\mathcal{D}_{(Q_c,Q_r)}$, in contrast to the the complementary case. Consequently, we will next focus on the five configurations of (Q_c,Q_r) which are marked in gray in Table I.

Next we investigate the WNG and DF performance of $\mathbf{w}_{\mathrm{MDF}(Q_c,Q_r)}$ and $\mathbf{w}_{\mathrm{NCMDF}(Q_c,Q_r)}$. We note that the latter is designed with two distinct nulls in $\theta_1 = 90^o$ and $\theta_2 = -70^o$. The results are depicted in Fig 2 and Fig 3, respectively. We observe that with both beamformers the WNG is improved upon a (Q_c,Q_r) configuration change, with a great accordance to the values of $\mathcal{W}_{(Q_c,Q_r)}$. In particular, it is important to accentuate the performance gap with $f = 1$ kHz, a relatively low frequency to which the human ear is highly sensitive, but in which high-directivity DMAs tend to exhibit significant white noise amplification. We observe that $\mathbf{w}_{\mathrm{MDF}(2,2)}$ is better than $\mathbf{w}_{\mathrm{MDF}(0,0)}$ by roughly 10 dB, whereas $\mathbf{w}_{\mathrm{NCMDF}(2,2)}$ is better than $\mathbf{w}_{\mathrm{NCMDF}(0,0)}$ by roughly 5 dB. On the contrary, we note that as the WNG improves the DF deteriorates, even though in the selected configurations $\mathcal{D}_{(Q_c,Q_r)}$ is always non-negative. The reason for that is clear- as the values of Q_c or Q_r increase, the length of $\mathbf{y}_{(Q_c,Q_r)}$ decreases. This results in smaller beamformers with less degrees of freedom which are optimized with respect to the array directivity.

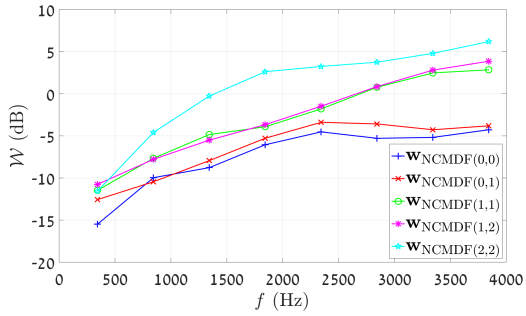


(a)

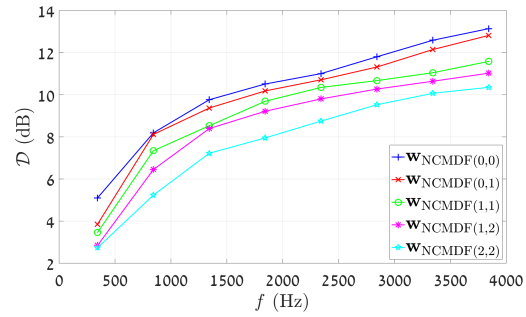


(b)

Fig. 2: WNG and DF measures with the robust MDF differential beamformers, $\mathbf{w}_{\mathrm{MDF}(Q_c,Q_r)}$, with varying values of (Q_c, Q_r). Simulation parameters: $M_{\mathsf{x}} = 5$, $M_{\mathsf{y}} = 5$, $\delta_{\mathsf{x}} = 1$ cm and $\delta_{\mathsf{y}} = 1.5$ cm. (a) WNG and (b) DF.



(a)



(b)

Fig. 3: WNG and DF measures with the robust NCMDF differential beamformers, $\mathbf{w}_{\mathrm{NCMDF}(Q_c,Q_r)}$, with varying values of (Q_c, Q_r). Simulation parameters: $M_{\mathsf{x}} = 5$, $M_{\mathsf{y}} = 5$, $\delta_{\mathsf{x}} = 1$ cm and $\delta_{\mathsf{y}} = 1.5$ cm. (a) WNG and (b) DF.

Nevertheless, even a mild selection of (Q_c, Q_r), for instance, $(1, 1)$, improves the WNG by approximately 5 dB, at the expanse of about 1 dB degradation in the DF.

VI. Conclusions

We have presented a robust approach for rectangular differential beamforming. We proposed to employ a 2-D multistage mean which operates independently on the columns and rows of the observation signals of a URA. The number of mean stages along the columns and rows of the URA is controlled by two parameters, Q_c and Q_r, respectively. Then, the resulting matrix is (column-wise) concatenated to a vector form to serve as a modified version of the observations. We showed that the first mean operation improves the robustness of the beamformer to white noise. We focused on the MDF and NCMDF differential beamformers and analyzed their performances in terms of both the WNG and DF measures. We showed that the configuration (Q_c, Q_r) constitutes a useful mean to mitigate the white noise amplification of rectangular differential beamformers in low frequencies.

References

[1] H. Van Trees, *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*, ser. Detection, Estimation, and Modulation Theory. Wiley, 2004.

[2] J. Benesty, I. Cohen, and J. Chen, *Fundamentals of Signal Enhancement and Array Signal Processing*. Singapore: Wiley-IEEE Press, 2018.

[3] ——, *Array Processing - Kronecker Product Beamforming*. Switzerland: Springer-Verlag, 2019.

[4] I. Cohen, J. Benesty, and J. Chen, "Differential kronecker product beamforming," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 5, pp. 892–902, 2019.

[5] G. Itzhak, J. Benesty, and I. Cohen, "On the design of differential kronecker product beamformers," *Submitted to IEEE/ACM Transactions on Audio, Speech, and Language Processing*.

[6] J. Benesty, I. Cohen, and J. Chen, *Array Beamforming with Linear Difference Equations*. Springer, 2021.

[7] G. Huang, J. Benesty, J. Chen, and I. Cohen, "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)*, 2020, pp. 211–215.

[8] G. Itzhak, J. Benesty, and I. Cohen, "Multistage approach for rectangular differential beamforming," *Submitted to IEEE/ACM Transactions on Audio, Speech, and Language Processing*.

[9] X. Wu and H. Chen, "Directivity Factors of the First-Order Steerable Differential Array With Microphone Mismatches: Deterministic and Worst-Case Analysis," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 2, pp. 300–315, 2016.

[10] G. Huang, J. Benesty, and J. Chen, "Design of robust concentric circular differential microphone arrays," *The Journal of the Acoustical Society of America*, vol. 141, no. 5, pp. 3236–3249, 2017.

[11] Y. Buchris, I. Cohen, and J. Benesty, "Frequency-Domain Design of Asymmetric Circular Differential Microphone Arrays," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 4, pp. 760–773, 2018.

[12] J. Jin, G. Huang, X. Wang, J. Chen, J. Benesty, and I. Cohen, "Steering study of linear differential microphone arrays," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 158–170, 2021.