COMBINED DIFFERENTIAL BEAMFORMING WITH UNIFORM LINEAR MICROPHONE ARRAYS

Gongping Huang[1], *Yuzhu Wang*[2], *Jacob Benesty*[3], *Israel Cohen*[1], *and Jingdong Chen*[2]

[1]Technion–Israel Institute of Technology, Technion City, Haifa 3200003, Israel
[2]CIAIC and Shaanxi Provincial Key Laboratory of Artificial Intelligence,
Northwestern Polytechnical University, Xi'an, Shaanxi 710072, China
[3]INRS-EMT, University of Quebec, Montreal, QC H5A 1K6, Canada

ABSTRACT

While differential beamformers have been widely used in voice communication and human-machine speech interface systems to enhance speech signals of interest, how to design such beamformers that on the one hand can achieve the highest possible directivity factor (DF) and on the other hand are able to obtain a certain level of white noise gain (WNG), so that they are robust enough to sensors' self noise and array imperfections is still a challenging issue. This paper studies the problem of robust differential beamforming with small-size arrays to achieve a high DF. It presents a method for the design of differential beamformers with uniform linear arrays. We first generate differential pressure signals by applying the recently developed forward spatial difference operator to the outputs of the array with pressure sensors. The pressure microphone observation signals and the differential pressure signals are then put together, and a combined beamformer is subsequently designed, which consists of two sub-beamformers, one operates on the pressure microphone observations and the other on the differential pressure signals. A new class of combined differential beamformers are introduced, which can achieve different levels of compromises between DF and WNG using an adjustable parameter.

Index Terms— Microphone arrays, differential beamforming, white noise gain, directivity factor.

1. INTRODUCTION

Microphone array beamforming has been extensively studied and many beamforming methods have been proposed in the literature [1–6], such as superdirective beamforming [7–9], adaptive beamforming [10–12], and differential beamforming [13–19]. Among those, differential beamforming has attracted dramatic interest [20–25]. Generally, differential beamformers have two prominent properties: 1) compact sizes, so that arrays can be easily embedded into such small devices as wearable and portable ones [26–28]; 2) high directivity, so beamformers are effective in enhancing broadband acoustic signals while suppressing spatial noise and reverberation [20,29]. However, differential beamformers are also sensitive to sensors' self noise and array imperfections and, therefore, how to design such beamformers that can achieve a relatively high DF with a reasonable value of WNG is an important issue [1,30,31].

In [32], a new method of differential beamforming with uniform linear arrays (ULAs) was proposed. It introduced a forward spatial difference operator, where any order of the spatial difference of the observation signals can be represented as a product of a difference operator matrix and the microphone array observations. Then, the optimal beamforming filter was designed and applied to the differential signals. Generally, with M microphones, the Pth-order differential operator generates an $M-P$ dimensional signal. A potential way to increase the number of degrees of freedom is by considering the new observation signal as a combination of pressure and differential pressure observations. However, the combined differential beamformer that is derived from maximization of either the WNG or the DF does not give any flexibility to compromise between DF and WNG. In this paper, we analyze the limitation of this beamformer and present a new class of combined differential beamformers. The proposed beamformers offer flexibility in compromising between DF and WNG through an adjustable parameter.

2. SIGNAL MODEL, CONVENTIONAL BEAMFORMING, AND PERFORMANCE MEASURES

We consider a source signal of interest (plane wave), in the farfield, that propagates in an anechoic acoustic environment at the speed of sound, i.e., $c = 340$ m/s, and impinges on a ULA consisting of M (with $M \geq 2$) omnidirectional microphones and with an interelement spacing of δ. If we denote the steering angle (in this work we consider the two dimensional case and only the azimuth angle) as θ, the steering vector (of length M) is then written as [33]

$$\mathbf{d}_\theta(\omega) = \begin{bmatrix} 1 & e^{-\jmath\varpi_\theta(\omega)} & \cdots & e^{-\jmath(M-1)\varpi_\theta(\omega)} \end{bmatrix}^T, \quad (1)$$

where $\jmath$ is the imaginary, $\varpi_\theta(\omega) = \omega\delta\cos\theta/c$, $\omega = 2\pi f$ is the angular frequency, $f > 0$ is the temporal frequency, and the superscript ^{T} is the transpose operator.

The frequency-domain observation signal vector of length M can be expressed as [1]

$$\begin{aligned}\mathbf{y}(\omega) &= \begin{bmatrix} Y_1(\omega) & Y_2(\omega) & \cdots & Y_M(\omega) \end{bmatrix}^T \\ &= \mathbf{d}_{\theta_\mathrm{s}}(\omega) X(\omega) + \mathbf{v}(\omega),\end{aligned} \quad (2)$$

where $Y_m(\omega)$ is the mth microphone signal, $X(\omega)$ is the zero-mean desired source signal, and $\mathbf{v}(\omega)$ is the zero-mean additive noise signal vector defined similarly to $\mathbf{y}(\omega)$, $\mathbf{d}_{\theta_\mathrm{s}}(\omega)$ is the signal propagation vector (note that the relative path attenuation is neglected), which is in the same form as the steering vector, and θ_s is the incidence angle of the source signal of interest. To simplify the notation, we drop the dependence on the angular frequency, ω, in the rest of this paper. In differential beamforming, we assume the desired source signal propagates from the endfire direction ($\theta_\mathrm{s} = 0$) [14,20], so (2) becomes

$$\mathbf{y} = \mathbf{d}_0 X + \mathbf{v}. \quad (3)$$

To approximate the true acoustic pressure differentials with finite differences of the microphones' outputs, the interelement spacing, δ, should much smaller than the acoustic wavelength, $\lambda = c/f$, i.e., $\delta \ll \lambda$ [14,20].

Conventional beamforming consists of applying a complex-valued linear filter, $\mathbf{h}$ of length M, to the observed signal vector to get an estimate of the source signal, i.e.,

$$Z = \mathbf{h}^H \mathbf{y}, \quad (4)$$

This work was supported in part by the Israel Science Foundation Grant 576/16, in part by ISF-NSFC joint research program Grants 2514/17 and 61761146001, in part by the National Key Research and Development Program of China under Grant No. 2018AAA0102200 and the Key Program of NSFC Grant 61831019.

where the superscript ^{H} is the conjugate-transpose operator. In our context, the distortionless constraint is desired, i.e.,

$$\mathbf{h}^H \mathbf{d}_0 = 1. \tag{5}$$

For fixed beamformers, the three commonly used performance measures are

- the beampattern, which describes the sensitivity of the beamformer to a plane wave impinging on the array from the direction θ, is defined as

$$\mathcal{B}_\theta(\mathbf{h}) = \mathbf{d}_\theta^H \mathbf{h}, \tag{6}$$

- the WNG, which evaluate the sensitivity of the beamformer to some array imperfections, is defined as [34]

$$\mathcal{W}(\mathbf{h}) = \frac{\left|\mathbf{h}^H \mathbf{d}_0\right|^2}{\mathbf{h}^H \mathbf{h}}, \tag{7}$$

- and the DF, which quantifies the ability of the beamformer to suppress spatial noise from directions other than the endfire direction, is defined as [20, 35]

$$\mathcal{D}(\mathbf{h}) = \frac{\left|\mathbf{h}^H \mathbf{d}_0\right|^2}{\mathbf{h}^H \mathbf{\Gamma}_\mathrm{d} \mathbf{h}}, \tag{8}$$

where $\mathbf{\Gamma}_\mathrm{d}$ is the pseudo-coherence matrix of the diffuse noise, whose (i, j)th $(i, j = 1, 2, \ldots, M)$ element is

$$(\mathbf{\Gamma}_\mathrm{d})_{ij} = \mathrm{sinc}\left[\varpi_0 (i-j)\right], \tag{9}$$

with $\mathrm{sinc}(x) = \sin x / x$ and $\varpi_0 = \omega\delta/c$.

3. DIFFERENTIAL BEAMFORMING THEORY

Following the framework in [32], the pth-order ($p = 0, 1, \ldots, P$, with $1 \leq P < M$) forward spatial difference of $\mathbf{y}$ is defined as

$$\begin{aligned}
\Delta Y_i &= Y_{i+1} - Y_i,\ i = 1, 2, \ldots, M-1, \\
&\vdots \\
\Delta^p Y_i &= \Delta^{p-1}(\Delta Y_i) = \Delta^{p-1} Y_{i+1} - \Delta^{p-1} Y_i \\
&= \sum_{j=0}^{p} (-1)^{p-j} \begin{pmatrix} p \\ j \end{pmatrix} Y_{i+j},\ i = 1, 2, \ldots, M-p,
\end{aligned} \tag{10}$$

where $\begin{pmatrix} p \\ j \end{pmatrix} = \frac{p!}{j!(p-j)!}$ is the binomial coefficient. It is more convenient to write (10) in a vector/matrix form as

$$\mathbf{\Delta}_{(p)} \mathbf{y} = \mathbf{y}_{(p)}, \tag{11}$$

where

$$\mathbf{\Delta}_{(p)} = \begin{bmatrix} \mathbf{c}_{(p)}^T & 0 & \cdots & 0 \\ 0 & \mathbf{c}_{(p)}^T & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{c}_{(p)}^T \end{bmatrix} \tag{12}$$

is a matrix of size $(M-p) \times M$, with

$$\mathbf{c}_{(p)} = \left[(-1)^p \begin{pmatrix} p \\ 0 \end{pmatrix} \ \cdots \ (-1)^1 \begin{pmatrix} p \\ p-1 \end{pmatrix} \ 1 \right]^T \tag{13}$$

being a vector of length $p+1$. By definition, we write $\mathbf{\Delta}_{(0)} = \mathbf{I}_M$ and $\mathbf{\Delta}_{(0)} \mathbf{y} = \mathbf{I}_M \mathbf{y} = \mathbf{y}$, where $\mathbf{I}_M$ is the $M \times M$ identity matrix.

Substituting (3) into (10), we get

$$\Delta^p Y_i = \tau_0^p e^{-\jmath(i-1)\varpi_0} X + \Delta^p V_i,\ i = 1, 2, \ldots, M-p, \tag{14}$$

where $\tau_0 = e^{-\jmath\varpi_0} - 1$. In a vector form, (14) becomes

$$\mathbf{\Delta}_{(p)} \mathbf{y} = \tau_0^p \mathbf{d}_{0, M-p} X + \mathbf{v}_{(p)} = \mathbf{y}_{(p)}, \tag{15}$$

where

$$\mathbf{d}_{0, M-p} = \left[\begin{array}{cccc} 1 & e^{-\jmath\varpi_0} & \cdots & e^{-\jmath(M-p-1)\varpi_0} \end{array}\right]^T \tag{16}$$

is the steering vector of length $M-p$ at $\theta = 0$ and $\mathbf{v}_{(p)} = \mathbf{\Delta}_{(p)} \mathbf{v}$.

The Pth-order ($P < M$) differential beamformer is designed by applying a complex-valued linear filter, $\mathbf{h}_{(P)}$ of length $M-P$, to the differential observed signal $\mathbf{y}_{(P)}$, i.e.,

$$Z_{(P)} = \mathbf{h}_{(P)}^H \mathbf{y}_{(P)} = X_{\mathrm{fd},(P)} + V_{\mathrm{rn},(P)}, \tag{17}$$

where $Z_{(P)}$ is the estimate of X, $X_{\mathrm{fd},(P)} = X \tau_0^P \mathbf{h}_{(P)}^H \mathbf{d}_{0, M-P}$ is the filtered desired signal, and $V_{\mathrm{rn},(P)} = \mathbf{h}_{(P)}^H \mathbf{v}_{(P)}$ is the residual noise.

Now, the WNG and DF are, respectively,

$$\mathcal{W}\left(\mathbf{h}_{(P)}\right) = \frac{|\tau_0|^{2P} \left|\mathbf{h}_{(P)}^H \mathbf{d}_{0, M-P}\right|^2}{\mathbf{h}_{(P)}^H \mathbf{\Delta}_{(P)} \mathbf{\Delta}_{(P)}^T \mathbf{h}_{(P)}} \tag{18}$$

and

$$\mathcal{D}\left(\mathbf{h}_{(P)}\right) = \frac{|\tau_0|^{2P} \left|\mathbf{h}_{(P)}^H \mathbf{d}_{0, M-P}\right|^2}{\mathbf{h}_{(P)}^H \mathbf{\Delta}_{(P)} \mathbf{\Gamma}_\mathrm{d} \mathbf{\Delta}_{(P)}^T \mathbf{h}_{(P)}}, \tag{19}$$

and the power beampattern is

$$\left|\mathcal{B}_\theta\left(\mathbf{h}_{(P)}\right)\right|^2 = |\tau_\theta|^{2P} \left|\mathbf{h}_{(P)}^H \mathbf{d}_{\theta, M-P}\right|^2, \tag{20}$$

where $\tau_\theta = e^{-\jmath\varpi_0 \cos\theta} - 1$.

4. COMBINED DIFFERENTIAL BEAMFORMERS

Generally, the optimal Pth-order differential beamformers can be derived from the maximization of the WNG or the DF. A potential way to achieve compromises between high DF and robustness is by considering the observed signal vector as a combination of pressure and difference pressure observations. For example, by taking the first P components of $\mathbf{y}$, a new observation signal vector of length M can be constructed as

$$\begin{aligned}
\overrightarrow{\mathbf{y}} &= \left[\begin{array}{ccccc} Y_1 & Y_2 & \cdots & Y_P & \mathbf{y}_{(P)}^T \end{array}\right]^T \\
&= \left[\begin{array}{cc} \mathbf{y}_P^T & \mathbf{y}_{(P)}^T \end{array}\right]^T \\
&= \overrightarrow{\mathbf{d}}_0 X + \overrightarrow{\mathbf{v}},
\end{aligned} \tag{21}$$

where

$$\overrightarrow{\mathbf{d}}_0 = \left[\begin{array}{cc} \mathbf{d}_{0,P}^T & \tau_0^P \mathbf{d}_{0, M-P}^T \end{array}\right]^T \tag{22}$$

is the steering vector of length P at $\theta = 0$ and $\overrightarrow{\mathbf{v}} = [\mathbf{v}_P^T \ \mathbf{v}_{(P)}^T]^T$ is defined in a similar way to $\overrightarrow{\mathbf{y}}$.

Consequently, the proposed beamformer output is

$$Z_{\overrightarrow{\mathbf{h}}} = \overrightarrow{\mathbf{h}}^H \overrightarrow{\mathbf{y}}, \tag{23}$$

where $\overrightarrow{\mathbf{h}}$ is a beamforming filter of length M. In this case, the WNG is

$$\mathcal{W}\left(\overrightarrow{\mathbf{h}}\right) = \frac{\left|\mathbf{h}_{(P)}^H \overrightarrow{\mathbf{d}}_0\right|^2}{\mathbf{h}_{(P)}^H \overrightarrow{\mathbf{\Delta}}_{(P)} \overrightarrow{\mathbf{\Delta}}_{(P)}^T \mathbf{h}_{(P)}}, \tag{24}$$

where

$$\overrightarrow{\mathbf{\Delta}}_{(P)} = \begin{bmatrix} \mathbf{I}_P & \mathbf{0} \\ & \mathbf{\Delta}_{(P)} \end{bmatrix} \tag{25}$$

is an $M \times M$ matrix with $\mathbf{I}_P$ being the $P \times P$ identity matrix and $\mathbf{0}$ being the $P \times (M-P)$ zero matrix. The DF is

$$\mathcal{D}\left(\overrightarrow{\mathbf{h}}\right) = \frac{\left|\mathbf{h}_{(P)}^H \overrightarrow{\mathbf{d}}_0\right|^2}{\mathbf{h}_{(P)}^H \overrightarrow{\mathbf{\Delta}}_{(P)} \mathbf{\Gamma}_{\mathrm{d}} \overrightarrow{\mathbf{\Delta}}_{(P)}^T \mathbf{h}_{(P)}}. \tag{26}$$

4.1. Direct Optimization

A straightforward way to derive the optimal beamformers is maximizing the WNG in (24), which gives the maximum WNG (MWNG) differential beamformer [32]:

$$\overrightarrow{\mathbf{h}}_{\mathrm{MWNG}} = \frac{\left(\overrightarrow{\mathbf{\Delta}}_{(P)} \overrightarrow{\mathbf{\Delta}}_{(P)}^T\right)^{-1} \overrightarrow{\mathbf{d}}_0}{\overrightarrow{\mathbf{d}}_0^H \left(\overrightarrow{\mathbf{\Delta}}_{(P)} \overrightarrow{\mathbf{\Delta}}_{(P)}^T\right)^{-1} \overrightarrow{\mathbf{d}}_0}, \tag{27}$$

and maximizing the DF in (26), which gives the maximum DF (MDF) beamformer:

$$\overrightarrow{\mathbf{h}}_{\mathrm{MDF}} = \frac{\left(\overrightarrow{\mathbf{\Delta}}_{(P)} \mathbf{\Gamma}_{\mathrm{d}} \overrightarrow{\mathbf{\Delta}}_{(P)}^T\right)^{-1} \overrightarrow{\mathbf{d}}_0}{\overrightarrow{\mathbf{d}}_0^H \left(\overrightarrow{\mathbf{\Delta}}_{(P)} \mathbf{\Gamma}_{\mathrm{d}} \overrightarrow{\mathbf{\Delta}}_{(P)}^T\right)^{-1} \overrightarrow{\mathbf{d}}_0}. \tag{28}$$

We show that the MWNG beamformer is identical to the delay-and-sum (DS) beamformer while the MDF beamformer is equal to the well-known superdirective beamformer.

Proof. From (21) and (22), we have

$$\overrightarrow{\mathbf{y}} = \overrightarrow{\mathbf{\Delta}}_{(P)} \mathbf{y}, \tag{29}$$

$$\overrightarrow{\mathbf{d}}_0 = \overrightarrow{\mathbf{\Delta}}_{(P)} \mathbf{d}_0. \tag{30}$$

Since $\overrightarrow{\mathbf{\Delta}}_{(P)}$ is a full-rank matrix, we have

$$\left(\overrightarrow{\mathbf{\Delta}}_{(P)} \overrightarrow{\mathbf{\Delta}}_{(P)}^T\right)^{-1} = \left(\overrightarrow{\mathbf{\Delta}}_{(P)}^T\right)^{-1} \overrightarrow{\mathbf{\Delta}}_{(P)}^{-1}, \tag{31}$$

$$\left(\overrightarrow{\mathbf{\Delta}}_{(P)} \mathbf{\Gamma}_{\mathrm{d}} \overrightarrow{\mathbf{\Delta}}_{(P)}^T\right)^{-1} = \left(\overrightarrow{\mathbf{\Delta}}_{(P)}^T\right)^{-1} \mathbf{\Gamma}_{\mathrm{d}}^{-1} \overrightarrow{\mathbf{\Delta}}_{(P)}^{-1}. \tag{32}$$

Substituting (30) and (31) into (27), we get

$$\overrightarrow{\mathbf{h}}_{\mathrm{MWNG}} = \frac{\left(\overrightarrow{\mathbf{\Delta}}_{(P)}^T\right)^{-1} \mathbf{d}_0}{\mathbf{d}_0^H \mathbf{d}_0} = \frac{\left(\overrightarrow{\mathbf{\Delta}}_{(P)}^T\right)^{-1} \mathbf{d}_0}{M}. \tag{33}$$

Then, by substituting (29) and (33) into (23), the MWNG beamformer output can be written as

$$Z_{\overrightarrow{\mathbf{h}}} = \overrightarrow{\mathbf{h}}_{\mathrm{MWNG}}^H \overrightarrow{\mathbf{y}} = \frac{\mathbf{d}_0^H \overrightarrow{\mathbf{\Delta}}_{(P)}^{-1}}{M} \overrightarrow{\mathbf{\Delta}}_{(P)} \mathbf{y}$$
$$= \frac{1}{M} \mathbf{d}_0^H \mathbf{y} = \mathbf{h}_{\mathrm{DS}}^H \mathbf{y}, \tag{34}$$

where $\mathbf{h}_{\mathrm{DS}} = \mathbf{d}_0 / M$ is the DS beamformer [1]. So, the MWNG differential beamformer is identical to the DS beamformer.

Substituting (30) and (32) into (28), we get

$$\overrightarrow{\mathbf{h}}_{\mathrm{MDF}} = \frac{\left(\overrightarrow{\mathbf{\Delta}}_{(P)}^T\right)^{-1} \mathbf{\Gamma}_{\mathrm{d}}^{-1} \mathbf{d}_0}{\mathbf{d}_0^H \mathbf{\Gamma}_{\mathrm{d}}^{-1} \mathbf{d}_0}. \tag{35}$$

Then, by substituting (30) and (35) into (23), the MDF beamformer output can be written as

$$Z_{\overrightarrow{\mathbf{h}}} = \overrightarrow{\mathbf{h}}_{\mathrm{MDF}}^H \overrightarrow{\mathbf{y}} = \frac{\mathbf{d}_0^H \mathbf{\Gamma}_{\mathrm{d}}^{-1} \overrightarrow{\mathbf{\Delta}}_{(P)}^{-1}}{\mathbf{d}_0^H \mathbf{\Gamma}_{\mathrm{d}}^{-1} \mathbf{d}_0} \overrightarrow{\mathbf{\Delta}}_{(P)} \mathbf{y}$$
$$= \frac{\mathbf{d}_0^H \mathbf{\Gamma}_{\mathrm{d}}^{-1}}{\mathbf{d}_0^H \mathbf{\Gamma}_{\mathrm{d}}^{-1} \mathbf{d}_0} \mathbf{y} = \mathbf{h}_{\mathrm{SD}}^H \mathbf{y}, \tag{36}$$

where $\mathbf{h}_{\mathrm{SD}} = \frac{\mathbf{\Gamma}_{\mathrm{d}}^{-1} \mathbf{d}_0}{\mathbf{d}_0^H \mathbf{\Gamma}_{\mathrm{d}}^{-1} \mathbf{d}_0}$ is the superdirective beamformer [34]. So, the MDF differential beamformer is equal to the superdirective beamformer. □

This shows that the combined differential beamformer that is derived from the maximization of the WNG (resp. DF) is theoretically equal to the conventional DS (resp. superdirective) beamformer, which does not offer any extra flexibility in compromising between DF and WNG.

4.2. Separate Optimization

A certainly better way to achieve compromises is to optimize separately the two beamformers corresponding to the pressure and difference pressure observations. In this case, we set

$$\overrightarrow{\mathbf{h}} = \begin{bmatrix} \mathbf{h}_P^T & \mathbf{h}_{(P)}^T \end{bmatrix}^T, \tag{37}$$

where $\mathbf{h}_P$ is a beamformer of length P applied to $\mathbf{y}_P$ and $\mathbf{h}_{(P)}$ is a beamformer of length $M-P$ applied to $\mathbf{y}_{(P)}$. The beamformer's output can be written as

$$Z_{\overrightarrow{\mathbf{h}}} = \overrightarrow{\mathbf{h}}^H \overrightarrow{\mathbf{y}} \tag{38}$$
$$= \mathbf{h}_P^H \mathbf{y}_P + \mathbf{h}_{(P)}^H \mathbf{y}_{(P)}$$
$$= \left(\mathbf{h}_P^H \mathbf{d}_{0,P} + \mathbf{h}_{(P)}^H \tau_0^P \mathbf{d}_{0,M-P}\right) X + \mathbf{h}_P^H \mathbf{v}_P + \mathbf{h}_{(P)}^H \mathbf{v}_{(P)}.$$

To satisfy the distortionless constraint at the desired direction, we require

$$\mathbf{h}_P^H \mathbf{d}_{0,P} + \tau_0^P \mathbf{h}_{(P)}^H \mathbf{d}_{0,M-P} = 1. \tag{39}$$

We can set

$$\mathbf{h}_P^H \mathbf{d}_{0,P} = \alpha,$$
$$\mathbf{h}_{(P)}^H \mathbf{d}_{0,M-P} = \frac{1}{\tau_0^P}(1-\alpha),$$

where $0 \leq \alpha \leq 1$ is a real coefficient that determines the level of compromise.

For the beamformer $\mathbf{h}_{(P)}$ applied to the difference pressure observations $\mathbf{y}_{(P)}$, we attempt to maximize the corresponding WNG in (18), which is obtained from the following optimization:

$$\min_{\mathbf{h}_{(P)}} \mathbf{h}_{(P)}^H \mathbf{\Delta}_{(P)} \mathbf{\Delta}_{(P)}^T \mathbf{h}_{(P)}$$
$$\text{s. t. } \mathbf{h}_{(P)}^H \mathbf{d}_{0,M-P} = \frac{1}{\tau_0^P}(1-\alpha). \tag{40}$$

The solution is the Pth-order MWNG differential beamformer:

$$\mathbf{h}_{(P),\mathrm{MWNG}} = \frac{(1-\alpha)\left(\mathbf{\Delta}_{(P)} \mathbf{\Delta}_{(P)}^T\right)^{-1} \mathbf{d}_{0,M-P}}{\left(\tau_0^*\right)^P \mathbf{d}_{0,M-P}^H \left(\mathbf{\Delta}_{(P)} \mathbf{\Delta}_{(P)}^T\right)^{-1} \mathbf{d}_{0,M-P}}, \tag{41}$$

where the superscript * is the complex-conjugate operator. The beamformer $\mathbf{h}_{(P)}$ can also be derived from maximization of the DF in (19), which is equivalent to

$$\min_{\mathbf{h}_{(P)}} \mathbf{h}_{(P)}^H \mathbf{\Delta}_{(P)} \mathbf{\Gamma}_{\mathrm{d}} \mathbf{\Delta}_{(P)}^T \mathbf{h}_{(P)}$$
$$\text{s. t. } \mathbf{h}_{(P)}^H \mathbf{d}_{0,M-P} = \frac{1}{\tau_0^P}(1-\alpha), \tag{42}$$

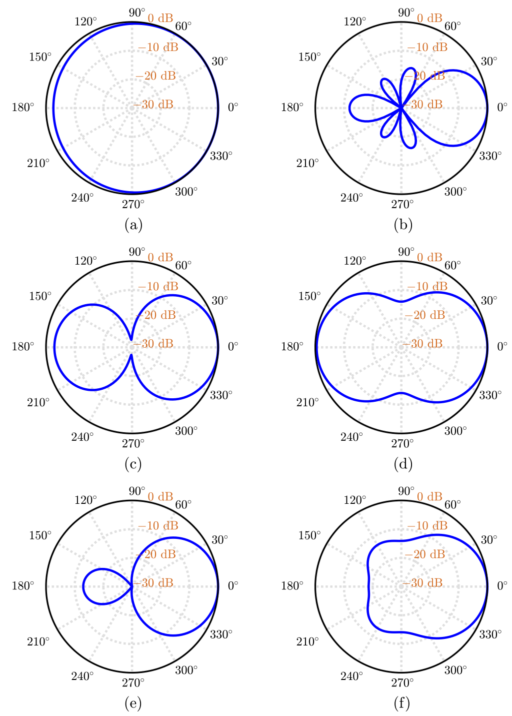

Fig. 1. Beampatterns of different kinds of differential beamformers with a ULA: (a) $\overrightarrow{\mathbf{h}}_{\mathrm{MWNG}}$, (b) $\overrightarrow{\mathbf{h}}_{\mathrm{MDF}}$, (c) $\overrightarrow{\mathbf{h}}_{\mathrm{C},1}$, (d) $\overrightarrow{\mathbf{h}}_{\mathrm{C},2}$, (e) $\overrightarrow{\mathbf{h}}_{\mathrm{C},3}$, and (f) $\overrightarrow{\mathbf{h}}_{\mathrm{C},4}$. Conditions: $M = 4$, $\delta = 1.5$ cm, $\alpha = 0.2$, and $f = 2$ kHz.

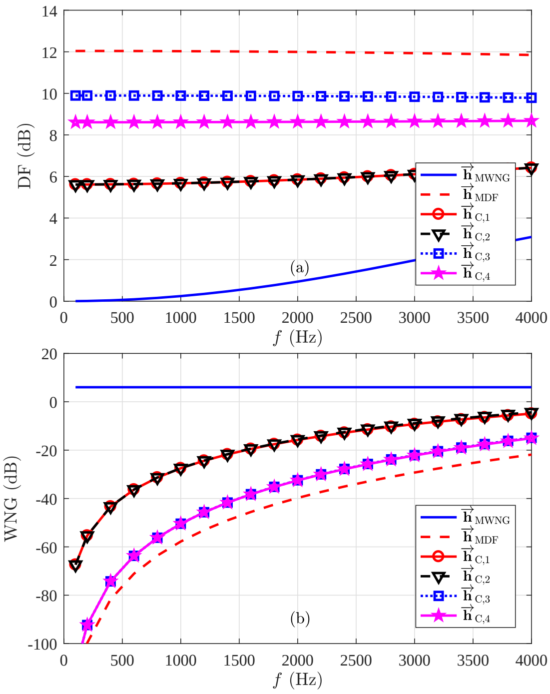

Fig. 2. WNGs and DFs of different kinds of differential beamformers as a function of the frequency: (a) DF and (b) WNG. Conditions: $M = 4$, $\delta = 1.5$ cm, and $\alpha = 0.2$.

from which we deduce the Pth-order MDF differential beamformer:

$$\mathbf{h}_{(P),\mathrm{MDF}} = \frac{(1-\alpha)\left(\mathbf{\Delta}_{(P)}\mathbf{\Gamma}_{\mathrm{d}}\mathbf{\Delta}_{(P)}^T\right)^{-1}\mathbf{d}_{0,M-P}}{(\tau_0^*)^P\,\mathbf{d}_{0,M-P}^H\left(\mathbf{\Delta}_{(P)}\mathbf{\Gamma}_{\mathrm{d}}\mathbf{\Delta}_{(P)}^T\right)^{-1}\mathbf{d}_{0,M-P}}. \quad (43)$$

For the beamformer $\mathbf{h}_P$ applied to the pressure observations $\mathbf{y}_P$, it can be derived from maximization of the WNG, which gives the MWNG (also the DS) beamformer:

$$\mathbf{h}_{P,\mathrm{MWNG}} = \frac{\alpha}{P}\mathbf{d}_{0,P}. \quad (44)$$

The beamformer $\mathbf{h}_P$ can also be derived from maximization of the DF, which leads to the MDF beamformer:

$$\mathbf{h}_{P,\mathrm{MDF}} = \frac{\alpha\mathbf{\Gamma}_{\mathrm{d},P}^{-1}\mathbf{d}_{0,P}}{\mathbf{d}_{0,P}^H\mathbf{\Gamma}_{\mathrm{d},P}^{-1}\mathbf{d}_{0,P}}. \quad (45)$$

where $\mathbf{\Gamma}_{\mathrm{d},P}$ is the pseudo-coherence matrix of the diffuse noise corresponding to a ULA consisting of P sensors, which is defined in a similar way to (9).

Finally, we get four kinds of combined differential beamformers:

$$\overrightarrow{\mathbf{h}}_{\mathrm{C},1} = \left[\begin{array}{cc}\mathbf{h}_{P,\mathrm{MDF}}^T & \mathbf{h}_{(P),\mathrm{MWNG}}^T\end{array}\right]^T, \quad (46)$$

$$\overrightarrow{\mathbf{h}}_{\mathrm{C},2} = \left[\begin{array}{cc}\mathbf{h}_{P,\mathrm{MWNG}}^T & \mathbf{h}_{(P),\mathrm{MWNG}}^T\end{array}\right]^T, \quad (47)$$

$$\overrightarrow{\mathbf{h}}_{\mathrm{C},3} = \left[\begin{array}{cc}\mathbf{h}_{P,\mathrm{MDF}}^T & \mathbf{h}_{(P),\mathrm{MDF}}^T\end{array}\right]^T, \quad (48)$$

$$\overrightarrow{\mathbf{h}}_{\mathrm{C},4} = \left[\begin{array}{cc}\mathbf{h}_{P,\mathrm{MWNG}}^T & \mathbf{h}_{(P),\mathrm{MDF}}^T\end{array}\right]^T. \quad (49)$$

5. SIMULATIONS

In this section, we study the performance of the proposed differential beamformers. We use a ULA consisting of four closely spaced microphones, with $\delta = 1.5$ cm. The desired source signal propagates from the endfire direction, i.e., $\theta_{\mathrm{s}} = 0°$.

Figure 1 shows plots of the beampatterns (at $f = 2$ kHz) of the MWNG beamformer, MDF beamformer, and the proposed four kinds of combined differential beamformers (with a chosen parameter $\alpha = 0.2$). Figure 2 plots the WNG and DF of the aforementioned beamformers as a function of the frequency, f. One can see that the MWNG beamformer (which is equal to the DS beamformer) has a low yet frequency dependent DF, which limits its use in practice; but it achieves the largest WNG among all the studied beamformers. The MDF beamformer (which is equal to the superdirective beamformer) has three nulls in the range between 0° and 180°, which corresponds to the third-order hypercardioid [20]. While the MDF beamformer can achieve the maximum DF, it suffers from significant white noise amplification, particularly at low frequencies. In comparison, the combined differential beamformers can achieve a tradeoff performance between the MDF and MWNG beamformers, and their DFs are frequency invariant. Consequently, by choosing a proper value of the parameter α, we can achieve a good compromise between a large value of DF and high value of WNG.

6. CONCLUSIONS

This paper studied the problem of robust differential beamforming with small-size microphone arrays to achieve a high directivity. It presented a differential beamforming method, which combines the pressure microphone observations and differential pressure signals obtained using the so-called forward spatial difference operator. To achieve a good compromise between the contradicting performance metrics of large DF values and high WNG, the two sub-beamformers that operate on the pressure microphone observations and differential pressure signals are optimized individually, each of which is derived either from the maximization of WNG or DF. A new class of four different combined differential beamformers were then introduced, which can achieve different levels of compromises between DF and WNG using an adjustable parameter.

7. REFERENCES

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

[2] G. W. Elko, "Superdirectional microphone arrays," in *Acoustic Signal Processing for Telecommunication*, pp. 181–237, Springer, 2000.

[3] J. Benesty, I. Cohen, and J. Chen, *Array Processing: Kronecker Product Beamforming*, vol. 18. Berlin, Germany: Springer-Verlag, 2019.

[4] B. Rafaely, *Fundamentals of Spherical Array Processing*. Berlin, Germany: Springer-Verlag, 2015.

[5] S. Yan, Y. Ma, and C. Hou, "Optimal array pattern synthesis for broadband arrays," *J. Acoust. Soc. Am.*, vol. 122, no. 5, pp. 2686–2696, 2007.

[6] S. Yan, "Optimal design of modal beamformers for circular arrays," *J. Acoust. Soc. Am.*, vol. 138, no. 4, pp. 2140–2151, 2015.

[7] R. M. Derkx and K. Janse, "Theoretical analysis of a first-order azimuth-steerable superdirective microphone array," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 17, no. 1, pp. 150–162, 2009.

[8] Y. Wang, Y. Yang, Y. Ma, and Z. He, "Robust high-order superdirectivity of circular sensor arrays," *J. Acoust. Soc. Am.*, vol. 136, no. 4, pp. 1712–1724, 2014.

[9] M. Crocco and A. Trucco, "Design of robust superdirective arrays with a tunable tradeoff between directivity and frequency-invariance," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2169–2181, 2011.

[10] E. A. P. Habets, J. Benesty, I. Cohen, S. Gannot, and J. Dmochowski, "New insights into the MVDR beamformer in room acoustics," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 1, p. 158, 2010.

[11] T. Higuchi, N. Ito, S. Araki, T. Yoshioka, M. Delcroix, and T. Nakatani, "Online MVDR beamformer based on complex gaussian mixture model with spatial prior for noise robust ASR," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 25, no. 4, pp. 780–793, 2017.

[12] V. M. Tavakoli, J. R. Jensen, M. G. Christenseny, and J. Benesty, "Pseudo-coherence-based MVDR beamformer for speech enhancement with ad hoc microphone arrays," in *Proc. IEEE ICASSP*, pp. 2659–2663, IEEE, 2015.

[13] G. W. Elko, "Microphone array systems for hands-free telecommunication," *Speech Commun.*, vol. 20, no. 3, pp. 229–240, 1996.

[14] G. W. Elko, "Differential microphone arrays," in *Audio Signal Processing for Next-Generation Multimedia Communication Systems*, pp. 11–65, Springer, 2004.

[15] M. Kolundzija, C. Faller, and M. Vetterli, "Spatiotemporal gradient analysis of differential microphone arrays," *Journal Audio Eng. Soc.*, vol. 59, no. 1/2, pp. 20–28, 2011.

[16] G. Huang, J. Chen, and J. Benesty, "On the design of differential beamformers with arbitrary planar microphone array," *J. Acoust. Soc. Am.*, vol. 144, no. 1, pp. EL66–EL70, 2018.

[17] E. D. Sena, H. Hacihabiboglu, and Z. Cvetkovic, "On the design and implementation of higher-order differential microphones," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 20, pp. 162–174, Jan. 2012.

[18] J. Byun, Y. C. Park, and S. W. Park, "Continuously steerable second-order differential microphone arrays," *J. Acoust. Soc. Am.*, vol. 143, no. 3, pp. EL225–EL230, 2018.

[19] G. Huang, I. Cohen, J. Chen, and J. Benesty, "Continuously steerable differential beamformers with null constraints for circular microphone arrays," *J. Acoust. Soc. Am.*, vol. 146, no. 1, pp. 1248–1258, 2020.

[20] J. Benesty and J. Chen, *Study and Design of Differential Microphone Arrays*. Berlin, Germany: Springer-Verlag, 2012.

[21] G. Huang, J. Benesty, and J. Chen, "On the design of frequency-invariant beampatterns with uniform circular microphone arrays," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 25, no. 5, pp. 1140–1153, 2017.

[22] F. Borra, A. Bernardini, F. Antonacci, and A. Sarti, "Uniform linear arrays of first-order steerable differential microphones," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 27, no. 12, pp. 1906–1918, 2019.

[23] F. Borra, A. Bernardini, F. Antonacci, and A. Sarti, "Efficient implementations of first-order steerable differential microphone arrays with arbitrary planar geometry," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, 2020.

[24] G. Huang, J. Benesty, I. Cohen, and J. Chen, "Differential beamforming on graphs," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 28, no. 1, pp. 901–913, 2020.

[25] I. Cohen, J. Benesty, and J. Chen, "Differential Kronecker product beamforming," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 27, pp. 892–902, May. 2019.

[26] H. Teutsch and G. W. Elko, "First-and second-order adaptive differential microphone arrays," in *Proc. IEEE IWAENC*, pp. 35–38, 2001.

[27] T. D. Abhayapala and A. Gupta, "Higher order differential-integral microphone arrays," *J. Acoust. Soc. Am.*, vol. 136, pp. 227–233, May 2010.

[28] E. Messner, H. Pessentheiner, J. A. Morales-Cordovilla, and M. Hagmüller, "Adaptive differential microphone arrays used as a front-end for an automatic speech recognition system," in *Proc. IEEE ICASSP*, pp. 2689–2693, IEEE, 2015.

[29] G. Huang, J. Chen, and J. Benesty, "Insights into frequency-invariant beamforming with concentric circular microphone arrays," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 26, no. 12, pp. 2305–2318, 2018.

[30] A. Bernardini, F. Antonacci, and A. Sarti, "Wave digital implementation of robust first-order differential microphone arrays," *IEEE Signal Process. Lett.*, vol. 25, no. 2, pp. 253–257, 2017.

[31] G. Huang, J. Chen, and J. Benesty, "Design of planar differential microphone arrays with fractional orders," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 28, pp. 116–130, 2019.

[32] G. Huang, J. Benesty, I. Cohen, and J. Chen, "A simple theory and new method of differential beamforming with uniform linear microphone arrays," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 28, no. 1, pp. 1079–1093, 2020.

[33] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques*. Upper Saddle River, NJ: Prentice-Hall, 1993.

[34] H. Cox, R. M. Zeskind, and T. Kooij, "Practical supergain," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, no. 3, pp. 393–398, 1986.

[35] G. W. Elko and J. Meyer, "Microphone arrays," in *Springer Handbook of Speech Processing* (J. Benesty, M. M. Sondhi, and Y. Huang, eds.), ch. 48, pp. 1021–1041, Berlin, Germany: Springer-Verlag, 2008.