2014 IEEE 28-th Convention of Electrical and Electronics Engineers in Israel

User Determined Superdirective Beamforming

Reuven Berkun, Israel Cohen
Technion, Israel Institute of Technology
Technion City, Haifa 32000, Israel

Jacob Benesty
INRS-EMT, University of Quebec
Montreal, QC H5A 1K6, Canada

***Abstract*—Superdirective beamforming is a well-known method for enhancement of reverberated speech signals. Nevertheless, it is very susceptible to errors in the sensor array characteristics, thermal noise, and white noise input, resulting in a low level of white noise gain, particularly at low frequencies. It is of great interest to develop a beamformer with superior enhancement of reverberated signals, having a high directivity factor, together with a relatively high white noise gain level.**

In this paper, a solution which controls both the directivity factor and the white noise gain is examined. We propose a linear weighted combination of two conventional beamformers, the regularized superdirective beamformer and the delay-and-sum beamformer. We analyze the beamformer gain responses, and consequently derive two user-determined frequency-dependent white noise gain and directivity factor beamformers, respectively. Simulation results approve our findings, and show of a robust user-controlled solution, with an effective tradeoff between the performance measures of the beamformer.

I. Introduction

Microphone array processing for speech signals is a challenging task of great importance. Speech signals propagating in a closed environment frequently suffer from reverberation and noise, distorting the quality and the intelligibility of the perceived signals. Beamforming represents a class of multichannel signal processing algorithms, that enable an extraction of desired source signals together with a suppression of undesired noise and reverberation [1]–[3]. Superdirective beamforming is a well-known approach that ensures high gain for reverberated signals, modeled by a diffuse noise input [4]. However, its high sensitivity to spatially white noise, significantly degrades its performance in practice [1], [4]–[6]. Slight errors between the array characteristics, such as position errors and mismatches between the sensors, pass through the beamformer like spatially white noise or uncorrelated noise. Therefore, the white noise gain is an important measure for the robustness of the beamformer.

Extensive research was conducted regarding the design of such an adaptive robust beamformer, that would have both high array gain and a satisfying level of white noise gain [1]–[9]. In [6], Cox et al. formulated the problem and introduced few approaches for the optimization problem. They presented a conventional optimal constrained solution, and proposed different methods to implement it [5], [6]. Others handled different types of beamforming mismatch errors, such as an arbitrary-type mismatch approach [10], or various models of the problem, such as worst-case optimization [7] or other optimization methods [11]. Recently, we proposed an approach of a robust beamformer with fine control of the array gain response (i.e., the directivity factor) and the white noise gain measure [8].

In this paper, we expand our solution [8], dealing with an optimization problem of the directivity factor maximization under constraint of the white noise gain. The suggested approach offers a simple optimized beamformer, as oppose to most of the familiar state-of-the-art solutions, which involve iterative solutions [6] or linear programming methods [7]. Similar to the analysis in [8], we propose a closed-form solution of a beamformer with user-defined fine control on the white noise gain and the directivity factor. We suggest a linear weighted combination of two conventional beamformers, enabling simple yet effective control of the filter response. Based on that, we derive filters which attain any desired frequency-dependent white noise gain or directivity factor.

The paper is organized as follows. In Section II, we describe the signal model and formulate the problem. In Section III, we present conventional fixed beamformers, one that maximizes the white noise gain and another that maximizes the directivity factor. In Section IV, the proposed beamformer is introduced. Based on that, we derive beamformers with user-determined white noise gain or user-determined directivity factor. This approach is based on a combination of the aforementioned conventional beamformer with the regularized adapted beamformer. This method provides a user-determined management of both white noise gain and directivity factor, and attains an effective tradeoff between the two measures. Finally, simulation results demonstrating the beamformer properties are presented in Section V.

II. Signal Model and Problem Formulation

Let us consider a source signal (plane wave), in the farfield, propagating in an anechoic acoustic environment at the speed of sound, i.e., $c = 340$ m/s, and impinging on a uniform linear sensor array consisting of M omnidirectional microphones, where the distance between two successive sensors is equal to δ. The direction of the source signal to the array is parameterized by the azimuth angle θ. In this context, the steering vector (of length M) is given by

$$\mathbf{d}(\omega,\theta) = \begin{bmatrix} 1 & e^{-\jmath\omega\tau_0\cos\theta} & \cdots & e^{-\jmath(M-1)\omega\tau_0\cos\theta} \end{bmatrix}^T, \tag{1}$$

where the superscript ^{T} is the transpose operator, $\jmath = \sqrt{-1}$ is the imaginary unit, $\omega = 2\pi f$ is the angular frequency, $f > 0$ is the temporal frequency, and $\tau_0 = \delta/c$ is the delay between two successive sensors at the angle $\theta = 0^\circ$.

We consider fixed beamformers with small values of δ, like in superdirective [5], [6] or differential beamforming [1], [9], where the main lobe is at the angle $\theta = 0^\circ$ (endfire direction) and the desired signal propagates from the same angle. Then, our goal is to design linear array beamformers, which are able to achieve supergains at the endfire with a better control on white noise amplification. For that, a complex weight, $H_m^*(\omega),\ m = 1, 2, \ldots, M$, is applied at the output of each microphone, where the superscript * denotes complex conjugation. The weighted outputs are then summed together to form the beamformer output. Putting all the gains together in a vector of length M, we get

$$\mathbf{h}(\omega) = \begin{bmatrix} H_1(\omega) & H_2(\omega) & \cdots & H_M(\omega) \end{bmatrix}^T. \tag{2}$$

This research was supported by the Israel Science Foundation (grant no. 1130/11).

The mth microphone signal is given by

$$Y_m(\omega) = e^{-\jmath(m-1)\omega\tau_0} X(\omega) + V_m(\omega),\ m = 1, 2, \ldots, M, \quad (3)$$

where $X(\omega)$ is the desired signal and $V_m(\omega)$ is the additive noise at the mth microphone. In a vector form, (3) becomes

$$\begin{aligned} \mathbf{y}(\omega) &= \left[\begin{array}{cccc} Y_1(\omega) & Y_2(\omega) & \cdots & Y_M(\omega) \end{array} \right]^T \\ &= \mathbf{x}(\omega) + \mathbf{v}(\omega) = \mathbf{d}(\omega)X(\omega) + \mathbf{v}(\omega), \end{aligned} \quad (4)$$

where $\mathbf{x}(\omega) = \mathbf{d}(\omega)X(\omega)$, $\mathbf{d}(\omega) = \mathbf{d}(\omega, 0^\circ)$ is the steering vector at $\theta = 0^\circ$ (direction of the source), and the noise signal vector, $\mathbf{v}(\omega)$, is defined similarly to $\mathbf{y}(\omega)$.

The beamformer output is then [2]

$$\begin{aligned} Z(\omega) &= \sum_{m=1}^{M} H_m^*(\omega) Y_m(\omega) = \mathbf{h}^H(\omega)\mathbf{y}(\omega) \\ &= \mathbf{h}(\omega)\mathbf{d}(\omega)X(\omega) + \mathbf{h}^H(\omega)\mathbf{v}(\omega), \end{aligned} \quad (5)$$

where $Z(\omega)$ is supposed to be the estimate of the desired signal, $X(\omega)$, and the superscript ^{H} is the conjugate-transpose operator. We constrain the solution to be distortionless, i.e.,

$$\mathbf{h}^H(\omega)\mathbf{d}(\omega) = 1. \quad (6)$$

If we take microphone 1 as the reference, using the definitions for input signal-to-noise ratio (SNR) and output SNR [8], the gain in SNR is defined as

$$\mathcal{G}[\mathbf{h}(\omega)] = \frac{\text{oSNR}[\mathbf{h}(\omega)]}{\text{iSNR}(\omega)} = \frac{\left|\mathbf{h}^H(\omega)\mathbf{d}(\omega)\right|^2}{\mathbf{h}^H(\omega)\mathbf{\Gamma}_{\mathbf{v}}(\omega)\mathbf{h}(\omega)}. \quad (7)$$

where $\mathbf{\Gamma}_{\mathbf{v}}(\omega) = \frac{\mathbf{\Phi}_{\mathbf{v}}(\omega)}{\phi_{V_1}(\omega)}$ and $\mathbf{\Phi}_{\mathbf{v}}(\omega) = E\left[\mathbf{v}(\omega)\mathbf{v}^H(\omega)\right]$ are the pseudo-coherence and correlation matrices of $\mathbf{v}(\omega)$, respectively, and $\phi_{V_1}(\omega) = E[|V_1(\omega)|^2]$ is the variance of $V_1(\omega)$.

In the field of superdirective beamformers, we are usually interested in two types of noise.

- The temporally and spatially white noise with the same variance at all microphones[1]. In this case, $\mathbf{\Gamma}_{\mathbf{v}}(\omega) = \mathbf{I}_M$, where $\mathbf{I}_M$ is the $M \times M$ identity matrix. Therefore, the white noise gain (WNG) is defined as

  $$\mathcal{W}[\mathbf{h}(\omega)] = \frac{\left|\mathbf{h}^H(\omega)\mathbf{d}(\omega)\right|^2}{\mathbf{h}^H(\omega)\mathbf{h}(\omega)}. \quad (8)$$

  We can easily deduce that the maximum WNG is $\mathcal{W}_{\max} = M$ which is frequency independent. The white noise amplification is the most serious problem with superdirective beamformers, which prevents them from being widely deployed in practice.
- The diffuse noise[2], where

  $$[\mathbf{\Gamma}_{\mathbf{v}}(\omega)]_{ij} = [\mathbf{\Gamma}_{\mathrm{d}}(\omega)]_{ij} = \frac{\sin\left[\omega(j-i)\tau_0\right]}{\omega(j-i)\tau_0} = \text{sinc}\left[\omega(j-i)\tau_0\right]. \quad (9)$$

  In this scenario, the gain in SNR is called the directivity factor (DF) and it is given by

  $$\mathcal{D}[\mathbf{h}(\omega)] = \frac{\left|\mathbf{h}^H(\omega)\mathbf{d}(\omega)\right|^2}{\mathbf{h}^H(\omega)\mathbf{\Gamma}_{\mathrm{d}}(\omega)\mathbf{h}(\omega)}. \quad (10)$$

  It is easy to verify that the maximum (frequency dependent) DF is $\mathcal{D}_{\max}(\omega) = \mathbf{d}^H(\omega)\mathbf{\Gamma}_{\mathrm{d}}^{-1}(\omega)\mathbf{d}(\omega)$. We refer to $\mathcal{D}_{\max}(\omega)$ as supergain when it is close to M^2, which can be achieved with a superdirective beamformer but at the expense of white noise amplification.

Then, one of the most important issues in practice is how to compromise between $\mathcal{W}[\mathbf{h}(\omega)]$ and $\mathcal{D}[\mathbf{h}(\omega)]$. Ideally, we would like $\mathcal{D}[\mathbf{h}(\omega)]$ to be as large as possible with $\mathcal{W}[\mathbf{h}(\omega)] \geq 1$.

III. Conventional Beamformers

In this section, we review two conventional fixed beamformers; one that maximizes the WNG and another that maximizes the DF. Afterwards we relate to a regularized version of the second approach.

The most well-known beamformer is the delay-and-sum (DS) [4], which is derived by maximizing the WNG (8) subject to the distortionless constraint (6). We get

$$\mathbf{h}_{\mathrm{DS}}(\omega) = \frac{\mathbf{d}(\omega)}{\mathbf{d}^H(\omega)\mathbf{d}(\omega)} = \frac{\mathbf{d}(\omega)}{M}. \quad (11)$$

Therefore, with this filter, the WNG and the DF are, respectively,

$$\mathcal{W}[\mathbf{h}_{\mathrm{DS}}(\omega)] = M = \mathcal{W}_{\max} \quad (12)$$

and

$$\mathcal{D}[\mathbf{h}_{\mathrm{DS}}(\omega)] = \frac{M^2}{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\mathrm{d}}(\omega)\mathbf{d}(\omega)} \geq 1. \quad (13)$$

This beamformer maximizes the WNG and never amplifies the diffuse noise since $\mathcal{D}\left[\mathbf{h}_{\mathrm{DS}}(\omega)\right] \geq 1$. However, in reverberant and noisy environments, our aim is to obtain high DF for good speech enhancement (i.e., dereverberation and noise reduction). This unfortunately does not happen with the DS beamformer, that malfunctions in reverberant rooms, even with a large number of microphones.

The second important beamformer is obtained by maximizing the DF (10) subject to the distortionless constraint (6). We get the well-known superdirective beamformer [5]:

$$\mathbf{h}_{\mathrm{S}}(\omega) = \frac{\mathbf{\Gamma}_{\mathrm{d}}^{-1}(\omega)\mathbf{d}(\omega)}{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\mathrm{d}}^{-1}(\omega)\mathbf{d}(\omega)}. \quad (14)$$

Its WNG and DF are, respectively,

$$\mathcal{W}[\mathbf{h}_{\mathrm{S}}(\omega)] = \frac{\left[\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\mathrm{d}}^{-1}(\omega)\mathbf{d}(\omega)\right]^2}{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\mathrm{d}}^{-2}(\omega)\mathbf{d}(\omega)} \quad (15)$$

and

$$\mathcal{D}[\mathbf{h}_{\mathrm{S}}(\omega)] = \mathbf{d}^H(\omega)\mathbf{\Gamma}_{\mathrm{d}}^{-1}(\omega)\mathbf{d}(\omega) = \mathcal{D}_{\max}(\omega). \quad (16)$$

While the DS beamformer has maximal and constant WNG response, but suffers from low DF, the superdirective beamformer, on the other hand, maximizes the DF but has a negative WNG.

We can express the WNG as

$$\mathcal{W}[\mathbf{h}_{\mathrm{S}}(\omega)] = \mathcal{W}_{\max}\cos^2\varphi(\omega), \quad (17)$$

where

$$\begin{aligned} \cos\varphi(\omega) &= \cos\left[\mathbf{d}(\omega), \mathbf{\Gamma}_{\mathrm{d}}^{-1}(\omega)\mathbf{d}(\omega)\right] \\ &= \frac{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\mathrm{d}}^{-1}(\omega)\mathbf{d}(\omega)}{\sqrt{\mathbf{d}^H(\omega)\mathbf{d}(\omega)}\sqrt{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\mathrm{d}}^{-2}(\omega)\mathbf{d}(\omega)}} \end{aligned} \quad (18)$$

is the cosine of the angle between the two vectors $\mathbf{d}(\omega)$ and $\mathbf{\Gamma}_{\mathrm{d}}^{-1}(\omega)\mathbf{d}(\omega)$, with $0 \leq \cos^2\varphi(\omega) \leq 1$.

At low frequencies, $\cos^2\varphi(\omega)$ can be very close to 0. As a result, $\mathcal{W}\left[\mathbf{h}_{\mathrm{S}}(\omega)\right]$ can be smaller than 1, which implies white noise amplification. While the superdirective beamformer attains maximum

[1]This noise models appropriately the sensor noise.

[2]This situation corresponds to the spherically isotropic noise field.

directivity factor, which is good for speech enhancement in reverberant rooms, it amplifies the white noise to intolerable levels, especially at low frequencies.

Since (14) is sensitive to the spatially white noise, Cox et al. [5], [6] proposed to maximize the DF subject to a constraint on the WNG. Under the distortionless constraint (6), the obtained optimal solution is [5], [6]

$$\mathbf{h}_{\mathrm{S},\epsilon}(\omega) = \frac{[\mathbf{\Gamma}_{\mathrm{d}}(\omega) + \epsilon \mathbf{I}_M]^{-1}\,\mathbf{d}(\omega)}{\mathbf{d}^H(\omega)\,[\mathbf{\Gamma}_{\mathrm{d}}(\omega) + \epsilon \mathbf{I}_M]^{-1}\,\mathbf{d}(\omega)}, \tag{19}$$

where $\epsilon \geq 0$ is a Lagrange multiplier. This is a regularized version of (14), where ϵ can be seen as the regularization parameter. This parameter tries to find a compromise between a supergain and white noise amplification. A small ϵ leads to a large DF and a low WNG, while a large ϵ gives a low DF and a large WNG. Two interesting cases of (19) are $\mathbf{h}_{\mathrm{S},0}(\omega) = \mathbf{h}_{\mathrm{S}}(\omega)$ and $\mathbf{h}_{\mathrm{S},\infty}(\omega) = \mathbf{h}_{\mathrm{DS}}(\omega)$.

We can express (19) as an ϵ-regularized superdirective beamformer:

$$\mathbf{h}_{\mathrm{S},\epsilon}(\omega) = \frac{\mathbf{\Gamma}_{\epsilon}^{-1}(\omega)\mathbf{d}(\omega)}{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\epsilon}^{-1}(\omega)\mathbf{d}(\omega)}, \tag{20}$$

where $\mathbf{\Gamma}_{\epsilon}(\omega) = \mathbf{\Gamma}_{\mathrm{d}}(\omega) + \epsilon \mathbf{I}_M$ is a regularized version of the pseudo-coherence matrix of the diffuse noise. The corresponding WNG and DF for this beamformer are, respectively,

$$\mathcal{W}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)] = \frac{\left[\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\epsilon}^{-1}(\omega)\mathbf{d}(\omega)\right]^2}{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\epsilon}^{-2}(\omega)\mathbf{d}(\omega)} \tag{21}$$

and

$$\mathcal{D}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)] = \frac{\left[\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\epsilon}^{-1}(\omega)\mathbf{d}(\omega)\right]^2}{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\epsilon}^{-1}(\omega)\mathbf{\Gamma}_{\mathrm{d}}(\omega)\mathbf{\Gamma}_{\epsilon}^{-1}(\omega)\mathbf{d}(\omega)}. \tag{22}$$

Similarly to (17), we can express the WNG (21) as

$$\mathcal{W}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)] = \mathcal{W}_{\max}\cos^2\varphi_{\epsilon}(\omega), \tag{23}$$

where

$$\begin{aligned}\cos\varphi_{\epsilon}(\omega) &= \cos\left[\mathbf{d}(\omega), \mathbf{\Gamma}_{\epsilon}^{-1}(\omega)\mathbf{d}(\omega)\right] \\ &= \frac{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\epsilon}^{-1}(\omega)\mathbf{d}(\omega)}{\sqrt{\mathbf{d}^H(\omega)\mathbf{d}(\omega)}\sqrt{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\epsilon}^{-2}(\omega)\mathbf{d}(\omega)}}\end{aligned} \tag{24}$$

is the cosine of the angle between the two vectors $\mathbf{d}(\omega)$ and $\mathbf{\Gamma}_{\epsilon}^{-1}(\omega)\mathbf{d}(\omega)$, with $0 \leq \cos^2\varphi_{\epsilon}(\omega) \leq 1$. For small ϵ, $\cos\varphi_{\epsilon}(\omega)$ would be similar to $\cos\varphi(\omega)$. Large ϵ would enlarge $\mathcal{W}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]$, so that $\cos^2\varphi_{\epsilon}(\omega)$ would be closer to 1.

While $\mathbf{h}_{\mathrm{S},\epsilon}(\omega)$ has some control on white noise amplification, it is certainly not easy to find a closed-form expression for ϵ given a desired value of the WNG.

IV. Proposed Beamformer

A. Derivation

As we saw, the DS and the regularized superdirective beamformers achieve maximum WNG and high DF, respectively. Therefore, we suggest to linearly combine the aforementioned beamformers into the following beamformer:

$$\begin{aligned}\mathbf{h}_{\alpha,\beta,\epsilon}(\omega) &= \alpha(\omega)\mathbf{h}_{\mathrm{S},\epsilon}(\omega) + \beta(\omega)\mathbf{h}_{\mathrm{DS}}(\omega) \\ &= \alpha(\omega)\frac{\mathbf{\Gamma}_{\epsilon}^{-1}(\omega)\mathbf{d}(\omega)}{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\epsilon}^{-1}(\omega)\mathbf{d}(\omega)} + \beta(\omega)\frac{\mathbf{d}(\omega)}{M} \\ &= \left[\alpha(\omega)\mathbf{I}_M + \beta(\omega)\mathbf{h}_{\mathrm{DS}}(\omega)\mathbf{d}^H(\omega)\right]\mathbf{h}_{\mathrm{S},\epsilon} \\ &= \left[\beta(\omega)\mathbf{I}_M + \alpha(\omega)\mathbf{h}_{\mathrm{S},\epsilon}(\omega)\mathbf{d}^H(\omega)\right]\mathbf{h}_{\mathrm{DS}},\end{aligned} \tag{25}$$

where $\alpha(\omega)$ and $\beta(\omega)$ are two real numbers with

$$\alpha(\omega) + \beta(\omega) = 1. \tag{26}$$

It is easy to verify that with the condition (26), this beamformer (25) is distortionless, i.e., $\mathbf{h}_{\alpha,\beta,\epsilon}^H(\omega)\mathbf{d}(\omega) = 1$. This beamformer controls the regularization with ϵ, and the DS or regularized superdirective influence with $\alpha(\omega)$ and $\beta(\omega)$.

It is not hard to show that the WNG of $\mathbf{h}_{\alpha,\beta,\epsilon}(\omega)$ is

$$\begin{aligned}\mathcal{W}[\mathbf{h}_{\alpha,\beta,\epsilon}(\omega)] &= \frac{\mathcal{W}[\mathbf{h}_{\mathrm{DS}}(\omega)]\mathcal{W}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]}{\alpha^2(\omega)\mathcal{W}[\mathbf{h}_{\mathrm{DS}}(\omega)] + [1-\alpha^2(\omega)]\,\mathcal{W}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]} \\ &= \frac{\mathcal{W}_{\max}\cos^2\varphi_{\epsilon}(\omega)}{\alpha^2(\omega) + [1-\alpha^2(\omega)]\cos^2\varphi_{\epsilon}(\omega)} \leq \mathcal{W}_{\max},\end{aligned} \tag{27}$$

which depends on the WNGs of the DS and regularized superdirective beamformers. We see that for $\alpha(\omega) = 0$ [in this case $\beta(\omega) = 1$], we have $\mathcal{W}[\mathbf{h}_{0,1,\epsilon}(\omega)] = \mathcal{W}[\mathbf{h}_{\mathrm{DS}}(\omega)]$, and for $\alpha(\omega) = 1$ [so $\beta(\omega) = 0$], we have $\mathcal{W}[\mathbf{h}_{1,0,\epsilon}(\omega)] = \mathcal{W}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]$. Also, we have

$$\mathcal{W}[\mathbf{h}_{\alpha,\beta,\epsilon}(\omega)] \geq \mathcal{W}\left[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)\right],\ \forall\alpha^2(\omega) \leq 1, \tag{28}$$

suggesting that we should always choose $-1 \leq \alpha(\omega) \leq 1$.

If we define $\mathcal{D}_{\max,\epsilon}(\omega) = \mathbf{d}^H(\omega)\mathbf{\Gamma}_{\epsilon}^{-1}(\omega)\mathbf{d}(\omega)$, it can be verified that the inverse DF corresponding to $\mathbf{h}_{\alpha,\beta,\epsilon}(\omega)$ is

$$\begin{aligned}&\mathcal{D}^{-1}[\mathbf{h}_{\alpha,\beta,\epsilon}(\omega)] = \alpha^2(\omega)\mathcal{D}^{-1}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)] + \\ &2\alpha(\omega)\beta(\omega)\tfrac{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\mathrm{d}}(\omega)\mathbf{\Gamma}_{\epsilon}^{-1}(\omega)\mathbf{d}(\omega)}{M\cdot\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\epsilon}^{-1}(\omega)\mathbf{d}(\omega)} + \beta^2(\omega)\mathcal{D}^{-1}[\mathbf{h}_{\mathrm{DS}}(\omega)] \\ &= \alpha^2(\omega)\mathcal{D}^{-1}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)] + 2\alpha(\omega)\beta(\omega)\left[\mathcal{D}_{\max,\epsilon}^{-1}(\omega) - \epsilon\tfrac{1}{M}\right] \\ &\quad + \beta^2(\omega)\mathcal{D}^{-1}[\mathbf{h}_{\mathrm{DS}}(\omega)],\end{aligned} \tag{29}$$

which depends on the DFs of the DS and regularized superdirective beamformers. We observe that for $\beta(\omega) = 0$, [i.e., $\alpha(\omega) = 1$] we have $\mathcal{D}[\mathbf{h}_{1,0,\epsilon}(\omega)] = \mathcal{D}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]$, and for $\beta(\omega) = 1$ [so $\alpha(\omega) = 0$], we have $\mathcal{D}[\mathbf{h}_{0,1,\epsilon}(\omega)] = \mathcal{D}[\mathbf{h}_{\mathrm{DS}}(\omega)]$. These results are consistent with the ones obtained for the WNGs. Also, we have

$$\mathcal{D}[\mathbf{h}_{\alpha,\beta,\epsilon}(\omega)] \leq \mathcal{D}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)] \tag{30}$$

$$\mathcal{D}[\mathbf{h}_{\alpha,\beta,\epsilon}(\omega)] \geq \mathcal{D}[\mathbf{h}_{\mathrm{DS}}(\omega)],\ \forall\beta^2(\omega) \leq 1, \tag{31}$$

suggesting that we should always take $-1 \leq \beta(\omega) \leq 1$. From all of the above we deduce that $0 \leq \alpha(\omega), \beta(\omega) \leq 1$.

Examples of the WNG and the DF of $\mathbf{h}_{\alpha,\beta,\epsilon}(\omega)$ beamformer are described in Fig. 1(a)-(b). Our goal is to design a beamformer with adequate WNG level and relatively high DF. When we design the filter parameters, first we set the regularization factor ϵ. It will determine the maximal DF $\mathcal{D}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]$, and the minimal WNG $\mathcal{W}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]$. Setting the parameters $\alpha(\omega)$ and $\beta(\omega)$ is depended on what we desire. Next, we discuss two interesting approaches.

In the first approach, we would like to find the value of $\alpha(\omega)$ in such a way that $\mathcal{W}[\mathbf{h}_{\alpha,\beta,\epsilon}(\omega)] = \mathcal{W}_0(\omega)$, where $\mathcal{W}_0(\omega)$ is a user-determined frequency-dependent WNG response, with $\mathcal{W}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)] < \mathcal{W}_0(\omega) < M,\ \forall\omega$. Using (27), we find that

$$\alpha^2(\omega) = \left(\frac{\mathcal{W}_{\max}}{\mathcal{W}_0(\omega)} - 1\right)\frac{\cos^2\varphi_{\epsilon}(\omega)}{1-\cos^2\varphi_{\epsilon}(\omega)} \tag{32}$$

from which we deduce two possible solutions for $\alpha(\omega)$:

$$\alpha_{\pm}(\omega) = \pm\sqrt{\frac{\mathcal{W}_{\max}}{\mathcal{W}_0(\omega)} - 1}\,\left|\cot\varphi_{\epsilon}(\omega)\right|. \tag{33}$$

The corresponding values for $\beta(\omega)$ are

$$\beta_1(\omega) = 1 - \alpha_+(\omega),\ \ \beta_2(\omega) = 1 - \alpha_-(\omega). \tag{34}$$

From the two pairs of solutions $\{\alpha_+(\omega), \beta_1(\omega)\}$ and $\{\alpha_-(\omega), \beta_2(\omega)\}$, we obviously choose the first one; therefore the obtained beamformer is

$$\mathbf{h}_{\alpha_+,\beta_1,\epsilon}(\omega) = \alpha_+(\omega)\mathbf{h}_{\mathrm{S},\epsilon}(\omega) + \beta_1(\omega)\mathbf{h}_{\mathrm{DS}}(\omega). \tag{35}$$

In the second approach, we would like to design the parameters in such a way that $\mathcal{D}[\mathbf{h}_{\alpha,\beta,\epsilon}(\omega)] = \mathcal{D}_0(\omega)$, where $\mathcal{D}_0(\omega)$ is a desired frequency-dependent DF with $\mathcal{D}[\mathbf{h}_{\mathrm{DS}}(\omega)] < \mathcal{D}_0(\omega) < \mathcal{D}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]$, $\forall\omega$. We can express (29) as a second degree polynomial of $\beta(\omega)$:

$$\begin{aligned} &\beta^2(\omega)\left\{\mathcal{D}^{-1}[\mathbf{h}_{\mathrm{DS}}(\omega)] + \mathcal{D}^{-1}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)] - 2\left[\mathcal{D}^{-1}_{\max,\epsilon}(\omega) - \epsilon\tfrac{1}{M}\right]\right\} \\ &\quad +2\beta(\omega)\left\{\left[\mathcal{D}^{-1}_{\max,\epsilon}(\omega) - \epsilon\tfrac{1}{M}\right] - \mathcal{D}^{-1}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]\right\} \\ &\quad\quad +\mathcal{D}^{-1}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)] - \mathcal{D}_0^{-1}(\omega) = 0, \end{aligned} \tag{36}$$

with two possible solutions for $\beta(\omega)$, marked by $\beta_\pm(\omega)$. Therefore, the corresponding values for $\alpha(\omega)$ are

$$\alpha_1(\omega) = 1 - \beta_+(\omega),\ \alpha_2(\omega) = 1 - \beta_-(\omega), \tag{37}$$

from which we take the positive one, and obtain the user-determined DF beamformer:

$$\mathbf{h}_{\alpha_1,\beta_+,\epsilon}(\omega) = \alpha_1(\omega)\mathbf{h}_{\mathrm{S},\epsilon}(\omega) + \beta_+(\omega)\mathbf{h}_{\mathrm{DS}}(\omega). \tag{38}$$

B. Justification

First, we can rewrite the robust superdirective beamformer (19) as

$$\mathbf{h}_{\mathrm{S},\epsilon}(\omega) = \mathcal{S}_\epsilon(\omega)[\mathbf{\Gamma}_{\mathrm{d}}(\omega) + \epsilon\mathbf{I}_M]^{-1}\mathbf{d}(\omega), \tag{39}$$

where

$$\mathcal{S}_\epsilon(\omega) = \frac{1}{\mathbf{d}^H(\omega)[\mathbf{\Gamma}_{\mathrm{d}}(\omega) + \epsilon\mathbf{I}_M]^{-1}\mathbf{d}(\omega)} \tag{40}$$

is a scaling factor which ensures that $\mathbf{h}_{\mathrm{S},\epsilon}(\omega)$ is distortionless. Therefore, $\mathcal{S}_\epsilon(\omega)$ has no effect on the robustness of the filter and only the term $\epsilon\mathbf{I}_M$ in $[\mathbf{\Gamma}_{\mathrm{d}}(\omega) + \epsilon\mathbf{I}_M]^{-1}\mathbf{d}(\omega)$ has, since it is the linear system that we want to solve.

Let us assume that $\alpha(\omega) \neq 0$; Using Woodbury identity [12] we can express the proposed beamformer as

$$\begin{aligned} \mathbf{h}_{\alpha,\beta,\epsilon}(\omega) &= \left[\alpha(\omega)\mathbf{I}_M + \beta(\omega)\mathbf{h}_{\mathrm{DS}}(\omega)\mathbf{d}^H(\omega)\right]\mathbf{h}_{\mathrm{S},\epsilon}(\omega) \\ &= \alpha(\omega)\left[\mathbf{I}_M - \beta(\omega)\tfrac{\mathbf{d}(\omega)\mathbf{d}^H(\omega)}{\mathbf{d}^H(\omega)\mathbf{d}(w)}\right]^{-1}\tfrac{\mathbf{\Gamma}_\epsilon^{-1}(\omega)\mathbf{d}(\omega)}{\mathbf{d}^H(\omega)\mathbf{\Gamma}_\epsilon^{-1}(\omega)\mathbf{d}(\omega)} \\ &= \mathcal{S}_\alpha(\omega)\left[\mathbf{\Gamma}_{\mathrm{d}}(\omega) - \beta(\omega)\tfrac{\mathbf{\Gamma}_\epsilon(\omega)\mathbf{d}(\omega)\mathbf{d}^H(\omega)}{\mathbf{d}^H(\omega)\mathbf{d}(w)}\right]^{-1}\mathbf{d}(\omega), \end{aligned} \tag{41}$$

where

$$\mathcal{S}_\alpha(\omega) = \frac{\alpha(\omega)}{\mathbf{d}^H(\omega)\mathbf{\Gamma}_\epsilon^{-1}(\omega)\mathbf{d}(\omega)} \tag{42}$$

is a scaling factor which ensures that $\mathbf{h}_{\alpha,\beta,\epsilon}(\omega)$ is distortionless. We can say that (41) is also a regularized superdirective beamformer with a rank-one complex matrix.

Let us assume that $\beta(\omega) \neq 0$. In this case, the proposed beamformer can be rewritten as

$$\begin{aligned} \mathbf{h}_{\alpha,\beta,\epsilon}(\omega) &= \left[\beta(\omega)\mathbf{I}_M + \alpha(\omega)\mathbf{h}_{\mathrm{S},\epsilon}(\omega)\mathbf{d}^H(\omega)\right]\mathbf{h}_{\mathrm{DS}}(\omega) \\ &= \beta(\omega)\left[\mathbf{I}_M - \alpha(\omega)\tfrac{\mathbf{\Gamma}_\epsilon^{-1}(\omega)\mathbf{d}(\omega)\mathbf{d}^H(\omega)}{\mathbf{d}^H(\omega)\mathbf{\Gamma}_\epsilon^{-1}(\omega)\mathbf{d}(w)}\right]^{-1}\tfrac{\mathbf{d}(\omega)}{\mathbf{d}^H(\omega)\mathbf{d}(\omega)} \\ &= \mathcal{S}_\beta(\omega)\left[\mathbf{I}_M - \alpha(\omega)\tfrac{\mathbf{\Gamma}_\epsilon^{-1}(\omega)\mathbf{d}(\omega)\mathbf{d}^H(\omega)}{\mathbf{d}^H(\omega)\mathbf{\Gamma}_\epsilon^{-1}(\omega)\mathbf{d}(w)}\right]^{-1}\mathbf{d}(\omega), \end{aligned} \tag{43}$$

where

$$\mathcal{S}_\beta(\omega) = \frac{\beta(\omega)}{\mathbf{d}^H(\omega)\mathbf{d}(\omega)} \tag{44}$$

is a scaling factor which ensures that $\mathbf{h}_{\alpha,\beta,\epsilon}(\omega)$ is distortionless. Again, we can consider (43) as a regularized form of the superdirective beamformer.

V. Simulations

In general, both the filter-design parameters and the physical properties of the microphone array determine the filter response. First, the regularization factor ϵ controls the range of the WNG and DF we can get, as noted in (28) and (30). To embody this, we added the response of $\mathbf{h}_{\mathrm{S},\epsilon}(\omega)$ to the illustrated simulations. Clearly, the array physical properties, such as the number of elements M and the microphone spacing δ, affect the response as well [5], [8].

First, we simulated the user-determined WNG beamformer $\mathbf{h}_{\alpha_+,\beta_1,\epsilon}(\omega)$ (35). In Fig. 1(c)-(d), we show an example of the WNG and the DF response of such beamformer. The determined WNG example of $\mathcal{W}_0(\omega) = -8 + 5\sin(\frac{1}{2\pi}\omega)$ dB provides a tolerable white noise amplification, above the minimal WNG of $\mathcal{W}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]$, together with a satisfying DF, between the DF of the superdirective beamformer and the DF of the regularized superdirective beamformer. Of course, any desired WNG can be set, within the allowed range.

Next, we simulated the user-determined DF beamformer $\mathbf{h}_{\alpha_1,\beta_+,\epsilon}(\omega)$ by solving (36) for a linear equation example: $\mathcal{D}_0(\omega) = 4 + \frac{1}{2\pi}\cdot 10^{-3}\omega$ dB. The corresponding WNG and DF responses are illustrated in Fig. 1(e)-(f). We observe that the received WNG is between the WNG levels of the regularized superdirective and the DS beamformers, as expected. Likewise, the determined $\mathcal{D}_0(\omega)$ is limited within the range of $\mathcal{D}[\mathbf{h}_{\mathrm{DS}}(\omega)] < \mathcal{D}_0(\omega) < \mathcal{D}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]$, $\forall\omega$.

Finally, following [5], to demonstrate the influence of the filter parameters on the WNG–DF tradeoff, we show in Fig. 2 the DF curve vs. WNG for increasing $\alpha(\omega)$, for different ϵ values. The parameter $\alpha(\omega) = \alpha$, $\forall\omega$, varies from 0 to 1 along the curves, where $\beta(\omega) = 1 - \alpha$ correspondingly. This example indicates of a monotonic relationship between α (given a specific regularization factor ϵ) and the gains of the beamformer. Increase of the WNG from its minimal value $\mathcal{W}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]$ at $\alpha = 0$ to its maximum, at $\alpha = 1$, causes monotonic decrease in the DF from its maximal value (of $\mathcal{D}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]$) to the low DF of the DS beamformer. One can see that setting different regularization factor ϵ changes the WNG–DF tradeoff vastly, hence there is a major importance of choosing an appropriate value for this parameter as well.

VI. Conclusion

We have extended our previously proposed approach [8] to robust regularized beamforming by using a linear combination of a regularized version of the superdirective beamformer and the DS beamformer. The proposed solution allows the user to adjust each of these beamformers influence, by setting the beamformer parameters $\alpha(\omega)$ and $\beta(\omega)$, in favor of the desired application. Using the proposed beamformer enables us to achieve any user-determined WNG or DF (such as the sine and linear equation shown here), within a wide allowed range. We examined the WNG–DF tradeoff, and analyzed the influence of the filter parameters α, β and ϵ on the WNG–DF relationship.

The presented approach opens a window to a wide family of combinations of known beamformers, with management of the beamformer frequency response. By setting the filter parameters, the user determines the weight of each factor, hence controlling the behavior of the beamformer WNG and DF.

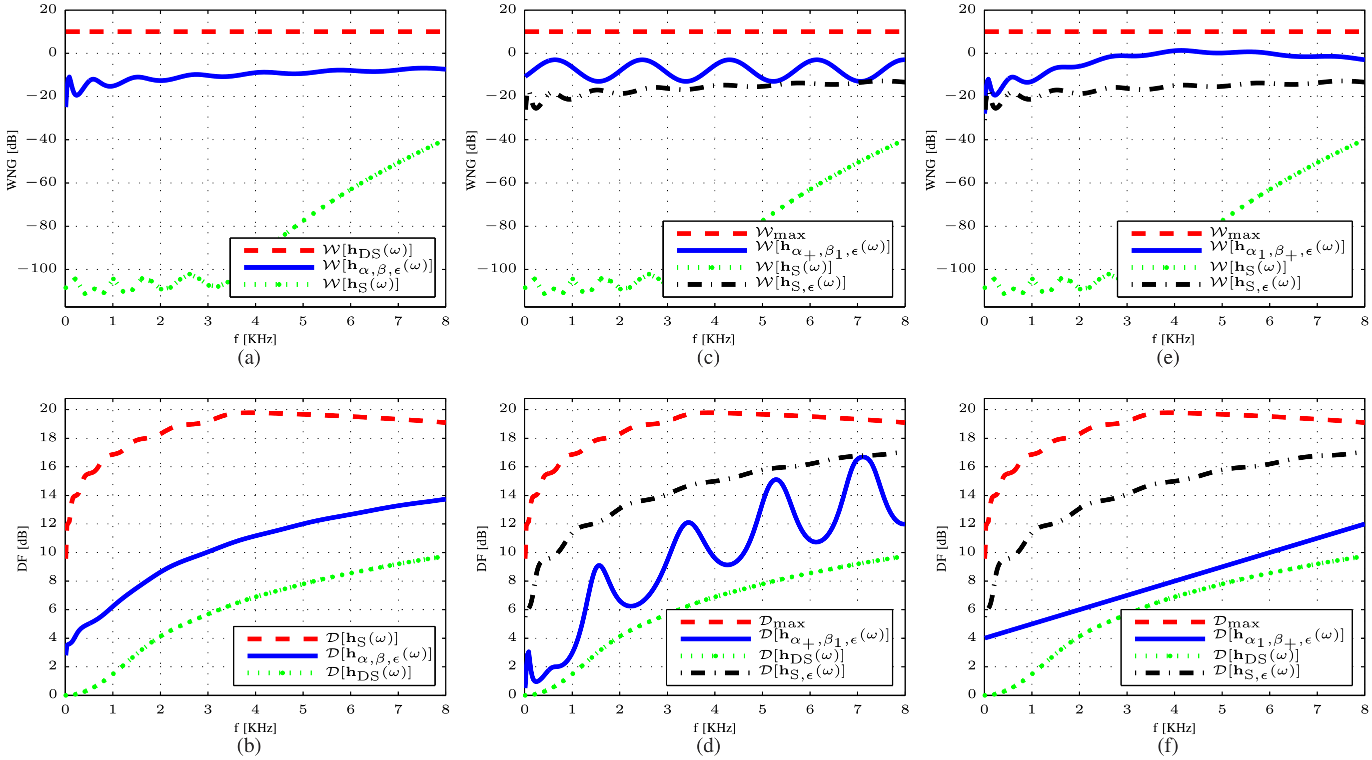

Fig. 1: Examples of the array gain of each beamformer, with $M = 10$ microphones, $\delta = 1$ cm, and $\epsilon = 1 \cdot 10^{-4}$. The top figures illustrate the proposed beamformer WNG (solid line) versus frequency. As a reference, $\mathcal{W}_{\max}$ (dashed line), $\mathcal{W}[\mathbf{h}_{\rm S}(\omega)]$ (dotted line), and $\mathcal{W}[\mathbf{h}_{{\rm S},\epsilon}(\omega)]$ (dot-dash-dot line) are plotted. The bottom figures illustrate the proposed beamformer DF (solid line). As a reference, $\mathcal{D}_{\max}(\omega)$ (dashed line), $\mathcal{D}[\mathbf{h}_{\rm DS}(\omega)]$ (dotted line), and $\mathcal{D}[\mathbf{h}_{{\rm S},\epsilon}(\omega)]$ (dot-dash-dot line) are plotted. (a)-(b) WNG and DF of $\mathbf{h}_{\alpha,\beta,\epsilon}(\omega)$, with $\alpha = \beta = 0.5$. (c)-(d) WNG and DF of $\mathbf{h}_{\alpha_+,\beta_1,\epsilon}(\omega)$ (35). The desired WNG is set to $\mathcal{W}_0(\omega) = -8 + 5\sin(\frac{1}{2\pi}\omega)$ dB. (e)-(f) WNG and DF of $\mathbf{h}_{\alpha_1,\beta_+,\epsilon}(\omega)$ as a solution of (36). The desired DF is set to $\mathcal{D}_0(\omega) = 4 + \frac{1}{2\pi} \cdot 10^{-3} \cdot \omega$ dB.

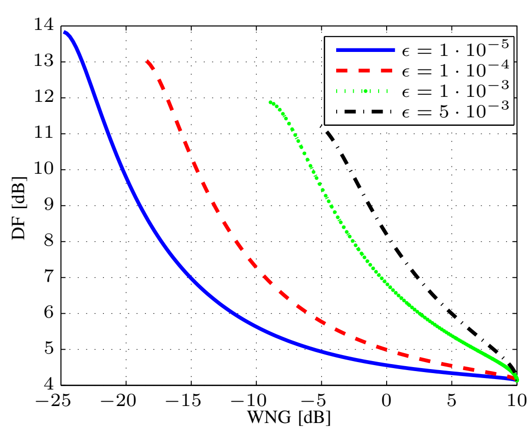

Fig. 2: The DF curve versus WNG, of $\mathbf{h}_{\alpha,\beta,\epsilon}(\omega)$ for increasing α, for different values of ϵ: $\epsilon = 1 \cdot 10^{-5}$ (solid line), $\epsilon = 1 \cdot 10^{-4}$ (dashed line), $\epsilon = 1 \cdot 10^{-3}$ (dotted line), and $\epsilon = 5 \cdot 10^{-3}$ (dot-dash-dot line). $M = 10$ microphones, $\delta = 1$ cm, and α is monotonically increased from 0 to 1 (with $\beta = 1 - \alpha$ correspondingly).

References

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

[2] J. Benesty, J. Chen, and Y. Huang, *Microphone Array Signal Processing*. Berlin, Germany: Springer-Verlag, 2008.

[3] B. D. Van Veen, and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Magazine*, vol. 5, no. 2, pp. 4–24, April 1988.

[4] M. Brandstein, and D. Ward, *Microphone arrays: Signal processing techniques and applications*. Berlin, Germany: Springer-Verlag, 2001.

[5] H. Cox, R. M. Zeskind, and T. Kooij, "Practical supergain," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-34, pp. 393–398, June 1986.

[6] H. Cox, R. M. Zeskind, and M. M. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, pp. 1365–1376, Oct. 1987.

[7] S. A. Vorobyov, A. B. Gershman, and Z.-Q. Lou, "Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 313–324, Feb. 2003.

[8] R. Berkun, I. Cohen, and J. Benesty, "Combined beamformers for robust regularized superdirective beamforming,", submitted to *IEEE Trans. Ausio, Speech, and Language Process.*, July 2014.

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

[10] S. Shahbazpanahi, A. B. Gershman, Z. Q. Luo, and K. M. Wong, "Robust adaptive beamforming for general-rank signal models," *IEEE Trans. Signal Process.*, vol. 51, no. 9 pp. 2257–2269, Sep. 2003.

[11] S. Doclo, and M. Moonen, "Superdirective beamforming robust against microphone mismatch," *IEEE Trans. Audio, Speech, and Language Process.*, vol. 15, no. 2 pp. 617–631, Feb. 2007.

[12] G. A. F. Seber, *A Matrix Handbook for Statisticians*. Hoboken, NJ: John Wiley & Sons, Inc., 2008.