Combined Beamformers for Robust Broadband Regularized Superdirective Beamforming

Reuven Berkun, Israel Cohen, and Jacob Benesty

***Abstract*—Superdirective fixed beamformers are known to attain high directivity factors, but are extremely sensitive to uncorrelated noise and slight errors in the array elements, which are modeled by the beamformer white noise gain measure. The delay-and-sum beamformer, on the other hand, manages to maximize the white noise gain, but suffers from a very low directivity factor. In this paper, we discuss the design of a broadband beamformer which controls both the directivity factor and the white noise gain. We combine a regularized version of the superdirective beamformer together with the delay-and-sum beamformer to create a robust regularized superdirective beamformer. We derive analytic closed-form expressions of the beamformer gain responses, and extend them to derive a beamformer with full control of the desired white noise gain or the directivity factor. The proposed approach offers a simple and robust broadband beamformer with controllable characteristics, shown here through persuasive simulation results.**

***Index Terms*—Beamforming, delay-and-sum beamformer, directivity factor, microphone arrays, robust superdirective beamformer, superdirective beamformer, supergain, white noise gain.**

I. Introduction

In numerous speech communication systems, the perceived signals are often distorted by reverberation and additive background noise. Consequently, microphone array broadband beamforming has been a very important topic in the field of sound acquisition and speech processing. Setting a few closely-spaced microphone elements in a line array together, enables a significant increase in the gain, given various types of noise fields. The superdirective fixed beamformer is a well-known linear filter that achieves maximum gain for a diffuse noise input [1], [2], hence enabling speech enhancement in reverberant and noisy environment. Its main disadvantage, however, is that it is very sensitive to small mismatches such as amplitude or phase errors in the sensor channels, slight position errors and uncorrelated noise, especially in low frequencies [1]–[4].

Many prior studies dealt with the design of a superdirective beamformer with robustness against different types of errors and uncorrelated noise between the sensors [1]–[20]. This can be generalized by imposing a white noise gain constraint on the relevant optimization problem. Cox *et al.* [1] introduced an optimal constrained solution, and offered a recursive algorithm to construct it. They also introduced a near-optimal solution by an oversteering approach [3]. Other works offered different formulations of the optimization problem [5], [6], or addressed specific mismatches formulations, such as errors in the steering vector [7] or in the array manifold [8].

Later studies dealt with the minimization of a mean cost function using the probability density function of the microphone characteristic errors [11]–[13]. Other popular approaches were based on a minimax design criterion for worst-case performance optimization [13]–[16], or even suggested nonlinear optimization [13], [17]. Moreover, many approaches addressed the design of a beamformer with frequency-invariant beampattern, which optimally approximates a desired frequency-invariant response in a least-square sense [11], [12], [18], [19], [21].

However, most of the existing beamforming approaches for constrained optimization involve somewhat burdensome computational optimization procedures, either by linear-programming or by sequential quadratic programming (SQP) methods [22], or by different relaxations of the optimization process [17], [20]. Furthermore, many methods require a priori knowledge of the error deviation [14] or of the probability density function of the gain, phase or position of the microphones [11]–[13], [17], [18] which are seldom available at the design phase.

Therefore, efficient and robust design of a broadband supergain beamformer with simple analytic solution is practically required. In this paper, we address the optimization problem of designing a beamformer which achieves maximum array gain given a diffuse noise input (maximum directivity factor), with constraint on the white noise gain. Unlike other approaches, which substitute the problem to convolve a multistep iterative solution, which converges step-by-step to the desired solution [1], [13], [15], [17], [20], or reformulate it in a convex linear programming problem [5], [8], [14], [20], we propose a direct and closed-form solution with simple control on both white noise gain and directivity factor. Our proposed beamformer attains an effective tradeoff between the directivity factor and the white noise gain of the microphone array, and plays the role of a regularized version of the robust superdirective beamformer. In addition, this type of beamformer enables derivation of a constant-over-frequency white noise gain response or a constant directivity factor.

The paper is organized as follows. In Section II, the signal model is presented together with some relevant terms and definitions. In Section III, we introduce two conventional fixed beamformers, which are obtained as solutions of the white

Manuscript received June 19, 2014; revised December 07, 2014; accepted February 24, 2015. Date of publication March 04, 2015; date of current version March 27, 2015. This work was supported by the Israel Science Foundation under Grant 1130/11. The associate editor coordinating the review of this manuscript and approving it for publication was Prof Nobutaka Ono.

R. Berkun and I. Cohen are with the Israel Institute of Technology Technion City, Haifa 32000, Israel (e-mail: berkun@tx.technion.ac.il; icohen@ee.technion.ac.il).

J. Benesty is with INRS-EMT, University of Quebec, Montreal, QC H5A 1K6, Canada

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TASLP.2015.2410139

noise gain and directivity factor optimization problem, respectively. Following this, we present a regularized variant of one of the conventional beamformers. In Section IV, we describe our proposed broadband beamformer, which allows control of both white noise gain and directivity factor. This is achieved by combining one conventional beamformer with the regularized second conventional beamformer. We also describe how to design a constant white noise gain or a constant directivity factor broadband beamformer, based on our proposed solution. Simulation results demonstrating the beamformer properties are presented in Section V. Finally, we conclude the work in Section VI.

II. Signal Model, Problem Formulation, and Definitions

We consider a source signal (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 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

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

where the superscript ^{T} is the transpose operator, $j = \sqrt{-1}$ 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 [1], [3], or differential beamforming [4], [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 objective is to design linear array beamformers (in the continuous frequency domain), 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

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

Therefore, our focus is on the design of such a filter.

The mth microphone signal is given by

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

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

$$\begin{aligned} \mathbf{y}(\omega) &= \begin{bmatrix} Y_1(\omega) & Y_2(\omega) & \cdots & Y_M(\omega) \end{bmatrix}^T \\ &= \mathbf{x}(\omega) + \mathbf{v}(\omega) \\ &= \mathbf{d}(\omega)X(\omega) + \mathbf{v}(\omega), \end{aligned} \tag{4}$$

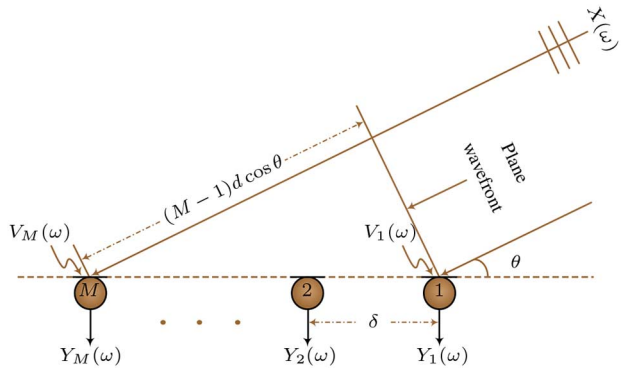

Fig. 1. Illustration of a uniformly spaced linear microphonearray for sound capture in the farfield [9].

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 $\mathbf{y}(\omega)$. This structure of the M received signals by a uniform linear microphone array is described in Fig. 1.

The beamformer output is simply [10]

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

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

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

If we take microphone 1 as the reference, we can define the input signal-to-noise ratio (SNR) with respect to this reference as

$$\text{iSNR}(\omega) = \frac{\phi_X(\omega)}{\phi_{V_1}(\omega)}, \tag{7}$$

where $\phi_X(\omega) = E[|X(\omega)|^2]$ and $\phi_{V_1}(\omega) = E[|V_1(\omega)|^2]$ are the variances of $X(\omega)$ and $V_1(\omega)$, respectively. The output SNR is defined as

$$\begin{aligned} \text{oSNR}[\mathbf{h}(\omega)] &= \phi_X(\omega)\frac{\left|\mathbf{h}^H(\omega)\mathbf{d}(\omega)\right|^2}{\mathbf{h}^H(\omega)\mathbf{\Phi}_{\mathbf{v}}(\omega)\mathbf{h}(\omega)} \\ &= \frac{\phi_X(\omega)}{\phi_{V_1}(\omega)} \times \frac{\left|\mathbf{h}^H(\omega)\mathbf{d}(\omega)\right|^2}{\mathbf{h}^H(\omega)\mathbf{\Gamma}_{\mathbf{v}}(\omega)\mathbf{h}(\omega)}, \end{aligned} \tag{8}$$

where $\mathbf{\Phi}_{\mathbf{v}}(\omega) = E[\mathbf{v}(\omega)\mathbf{v}^H(\omega)]$ and $\mathbf{\Gamma}_{\mathbf{v}}(\omega) = \frac{\mathbf{\Phi}_{\mathbf{v}}(\omega)}{\phi_{V_1}(\omega)}$ are the correlation and pseudo-coherence matrices of $\mathbf{v}(\omega)$, respectively. The definition of the gain in SNR is easily derived from the previous definitions, i.e.,

$$\begin{aligned} \mathcal{G}[\mathbf{h}(\omega)] &= \frac{\text{oSNR}[\mathbf{h}(\omega)]}{\text{iSNR}(\omega)} \\ &= \frac{\left|\mathbf{h}^H(\omega)\mathbf{d}(\omega)\right|^2}{\mathbf{h}^H(\omega)\mathbf{\Gamma}_{\mathbf{v}}(\omega)\mathbf{h}(\omega)}. \end{aligned} \tag{9}$$

When we deal with 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, $\mathbf{\Gamma_v}(\omega) = \mathbf{I}_M$, where $\mathbf{I}_M$ is the $M \times M$ identity matrix. Therefore, the white noise gain (WNG) is defined as

$$\mathcal{W}[\mathbf{h}(\omega)] = \frac{\left|\mathbf{h}^H(\omega)\mathbf{d}(\omega)\right|^2}{\mathbf{h}^H(\omega)\mathbf{h}(\omega)}. \tag{10}$$

  Using the Cauchy-Schwarz inequality, i.e.,

$$\left|\mathbf{h}^H(\omega)\mathbf{d}(\omega)\right|^2 \leq \mathbf{h}^H(\omega)\mathbf{h}(\omega) \times \mathbf{d}^H(\omega)\mathbf{d}(\omega), \tag{11}$$

  we easily deduce from (10) that

$$\mathcal{W}[\mathbf{h}(\omega)] \leq M, \ \forall \mathbf{h}(\omega). \tag{12}$$

  As a result, the maximum WNG is

$$\mathcal{W}_{\max} = M, \tag{13}$$

  which is frequency independent. On the contrary, when $\mathcal{W}[\mathbf{h}(\omega)] < 1$, the white noise is amplified at the beamformer output. 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

$$\begin{aligned}[\mathbf{\Gamma_v}(\omega)]_{ij} = [\mathbf{\Gamma}_{\rm d}(\omega)]_{ij} &= \frac{\sin\left[\omega(j-i)\tau_0\right]}{\omega(j-i)\tau_0} \\ &= \operatorname{sinc}\left[\omega(j-i)\tau_0\right].\end{aligned} \tag{14}$$

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

$$\mathcal{D}[\mathbf{h}(\omega)] = \frac{\left|\mathbf{h}^H(\omega)\mathbf{d}(\omega)\right|^2}{\mathbf{h}^H(\omega)\mathbf{\Gamma}_{\rm d}(\omega)\mathbf{h}(\omega)}. \tag{15}$$

  Again, by invoking the Cauchy-Schwarz inequality, i.e.,

$$\begin{aligned}\left|\mathbf{h}^H(\omega)\mathbf{d}(\omega)\right|^2 \leq{}& \mathbf{h}^H(\omega)\mathbf{\Gamma}_{\rm d}(\omega)\mathbf{h}(\omega) \\ &\times \mathbf{d}^H(\omega)\mathbf{\Gamma}_{\rm d}^{-1}(\omega)\mathbf{d}(\omega),\end{aligned} \tag{16}$$

  we find from (15) that

$$\mathcal{D}[\mathbf{h}(\omega)] \leq \mathbf{d}^H(\omega)\mathbf{\Gamma}_{\rm d}^{-1}(\omega)\mathbf{d}(\omega), \ \forall \mathbf{h}(\omega). \tag{17}$$

  As a result, the maximum DF is

$$\begin{aligned}\mathcal{D}_{\max}(\omega) &= \mathbf{d}^H(\omega)\mathbf{\Gamma}_{\rm d}^{-1}(\omega)\mathbf{d}(\omega) \\ &= \operatorname{tr}\left[\mathbf{\Gamma}_{\rm d}^{-1}(\omega)\mathbf{d}(\omega)\mathbf{d}^H(\omega)\right] \\ &\leq M\operatorname{tr}\left[\mathbf{\Gamma}_{\rm d}^{-1}(\omega)\right],\end{aligned} \tag{18}$$

  where $\operatorname{tr}[\cdot]$ denotes the trace of a square matrix. The maximum DF is frequency dependent. 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 $\mathcal{W}[\mathbf{h}(\omega)] \geq 1$.

[1]This noise models well the sensor noise.

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

Finally, we define the beampattern or directivity pattern as

$$\begin{aligned}\mathcal{B}[\mathbf{h}(\omega), \theta] &= \mathbf{d}^H(\omega, \theta)\mathbf{h}(\omega) \\ &= \sum_{m=1}^{M} H_m(\omega)e^{j(m-1)\omega\tau_0\cos\theta}.\end{aligned} \tag{19}$$

The distortionless constraint given in (6) can be interpreted as compulsion of a constant unit beampattern for all frequencies in the endfire direction. This single boresight constraint can be further generalized to various linear constraints which control the beam shape [1].

These definitions of the SNRs, gains, and beampattern, which are extremely useful for the evaluation of any types of beamformers, conclude this section.

III. Two Conventional Beamformers

In this section, we discuss two important conventional fixed beamformers from another perspective; one that maximizes the WNG and the other that maximizes the DF. We also relate to a regularized version of the second approach.

The simplest and the most well-known beamformer is the delay-and-sum (DS) [2], which is derived by maximizing the WNG [eq. (10)] subject to the distortionless constraint given in (6). We easily get

$$\begin{aligned}\mathbf{h}_{\rm DS}(\omega) &= \frac{\mathbf{d}(\omega)}{\mathbf{d}^H(\omega)\mathbf{d}(\omega)} \\ &= \frac{\mathbf{d}(\omega)}{M}.\end{aligned} \tag{20}$$

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

$$\mathcal{W}[\mathbf{h}_{\rm DS}(\omega)] = M = \mathcal{W}_{\max} \tag{21}$$

and

$$\mathcal{D}[\mathbf{h}_{\rm DS}(\omega)] = \frac{M^2}{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\rm d}(\omega)\mathbf{d}(\omega)} \geq 1. \tag{22}$$

Another interesting way to express (22) is

$$\mathcal{D}[\mathbf{h}_{\rm DS}(\omega)] = \mathcal{D}_{\max}(\omega)\cos^2\phi(\omega), \tag{23}$$

where

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

is the cosine of the angle between the two vectors $\mathbf{\Gamma}_{\rm d}^{1/2}(\omega)\mathbf{d}(\omega)$ and $\mathbf{\Gamma}_{\rm d}^{-1/2}(\omega)\mathbf{d}(\omega)$, with $0 \leq \cos^2\phi(\omega) \leq 1$. Let $\lambda_1(\omega)$ and $\lambda_M(\omega)$ be the maximum and minimum eigenvalues of $\mathbf{\Gamma}_{\rm d}(\omega)$, respectively. Using the Kantorovich inequality [23]:

$$\cos^2\phi(\omega) \geq \frac{4\lambda_1(\omega)\lambda_M(\omega)}{\left[\lambda_1(\omega) + \lambda_M(\omega)\right]^2}, \tag{25}$$

we deduce that

$$\frac{4\lambda_1(\omega)\lambda_M(\omega)}{\left[\lambda_1(\omega) + \lambda_M(\omega)\right]^2} \leq \frac{\mathcal{D}[\mathbf{h}_{\rm DS}(\omega)]}{\mathcal{D}_{\max}(\omega)} \leq 1. \tag{26}$$

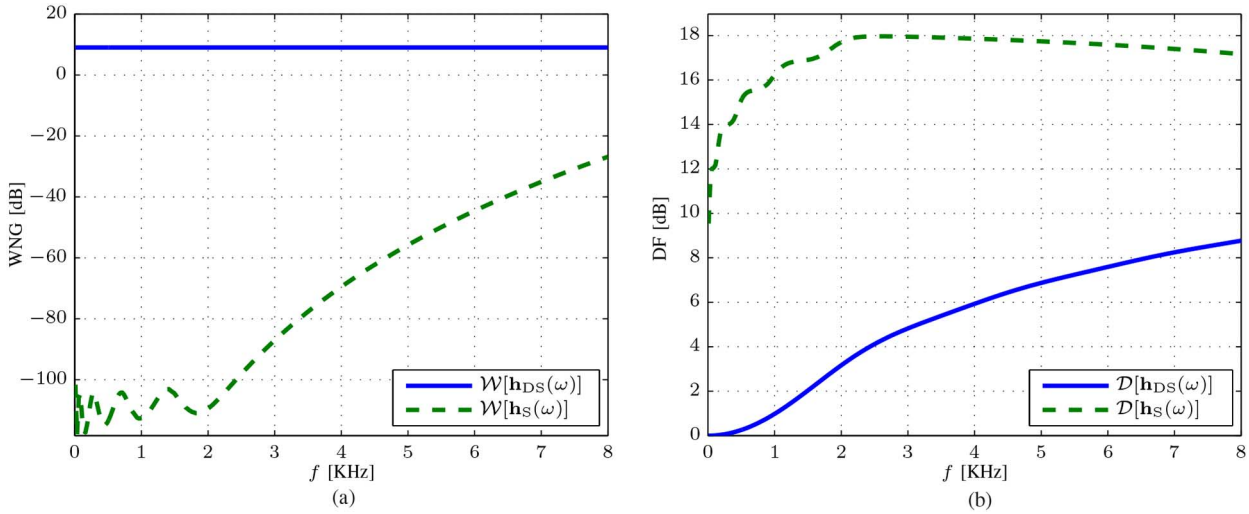

Fig. 2. Example of the array gains of the DS (solid line) and superdirective (dashed line) beamformers versus frequency, with $M = 8$ microphones and $\delta = 1$ cm. (a) WNG. (b) DF.

Clearly, the DS beamformer maximizes the WNG and never amplifies the diffuse noise since $\mathcal{D}[\mathbf{h}_{\mathrm{DS}}(\omega)] \geq 1$. However, in reverberant and noisy environments, it is essential to have a large value of the DF for good speech enhancement (i.e., dereverberation and noise reduction). But, unfortunately, this does not happen, in general, with the DS beamformer, which is known to perform very poorly when the reverberation time of the room is high, even with a large number of microphones.

The second important beamformer is obtained by maximizing the DF [eq. (15)] subject to the distortionless constraint given in (6). We get the well-known superdirective beamformer [3]:

$$\mathbf{h}_{\mathrm{S}}(\omega) = \frac{\mathbf{\Gamma}_{\mathrm{d}}^{-1}(\omega)\mathbf{d}(\omega)}{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\mathrm{d}}^{-1}(\omega)\mathbf{d}(\omega)}. \tag{27}$$

This filter is a particular form of the celebrated minimum variance distortionless response (MVDR) beamformer [24], [25]. Also, (27) corresponds to the directivity pattern of the hypercardioid of order $M-1$[9]. We deduce that the WNG and the DF are, respectively,

$$\mathcal{W}[\mathbf{h}_{\mathrm{S}}(\omega)] = \frac{\left[\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\mathrm{d}}^{-1}(\omega)\mathbf{d}(\omega)\right]^2}{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\mathrm{d}}^{-2}(\omega)\mathbf{d}(\omega)} \tag{28}$$

and

$$\mathcal{D}[\mathbf{h}_{\mathrm{S}}(\omega)] = \mathbf{d}^H(\omega)\mathbf{\Gamma}_{\mathrm{d}}^{-1}(\omega)\mathbf{d}(\omega) = \mathcal{D}_{\max}(\omega). \tag{29}$$

Examples of the WNG and the DF for both DS and superdirective beamformers are shown in Fig. 2. We see that while the DS beamformer has maximal and constant WNG response, it suffers from low DF (though still greater than 0 dB). On the contrary, the superdirective beamformer maximizes the DF but has a negative WNG. It can be shown that [26]

$$\lim_{\delta \to 0} \mathcal{D}[\mathbf{h}_{\mathrm{S}}(\omega)] = M^2. \tag{30}$$

We can express the WNG as

$$\mathcal{W}[\mathbf{h}_{\mathrm{S}}(\omega)] = \mathcal{W}_{\max} \cos^2 \varphi(\omega), \tag{31}$$

where

$$\cos \varphi(\omega) = \cos\left[\mathbf{d}(\omega), \mathbf{\Gamma}_{\mathrm{d}}^{-1}(\omega)\mathbf{d}(\omega)\right] = \frac{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\mathrm{d}}^{-1}(\omega)\mathbf{d}(\omega)}{\sqrt{\mathbf{d}^H(\omega)\mathbf{d}(\omega)}\sqrt{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\mathrm{d}}^{-2}(\omega)\mathbf{d}(\omega)}} \tag{32}$$

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

$$\frac{4\lambda_1(\omega)\lambda_M(\omega)}{\left[\lambda_1(\omega) + \lambda_M(\omega)\right]^2} \leq \frac{\mathcal{W}\left[\mathbf{h}_{\mathrm{S}}(\omega)\right]}{\mathcal{W}_{\max}} \leq 1. \tag{33}$$

Examples of the DF ratio $\cos^2 \phi(\omega)$, and of the WNG ratio $\cos^2 \varphi(\omega)$ are described in Fig. 3. We observe that the DF and WNG ratios vary with frequency, and are indeed limited within the boundaries defined in eq. (26),(33), respectively.

The element spacing distance δ, is also an important factor in determining the array gain, as described thoroughly in [3].

For small values of δ ($\delta \ll \lambda$, where λ is the acoustic wavelength) and at low frequencies, $\cos^2 \varphi(\omega)$ can be very close to 0. As a result, $\mathcal{W}[\mathbf{h}_{\mathrm{S}}(\omega)]$ can be smaller than 1, which implies white noise amplification. While the superdirective beamformer gives the maximum directivity factor, which is good for speech enhancement in very reverberant rooms, it amplifies the white noise to intolerable levels, especially at low frequencies.

It is interesting to observe that

$$\frac{1}{\mathbf{h}_{\mathrm{S}}^H(\omega)\mathbf{h}_{\mathrm{DS}}(\omega)} = \mathcal{W}_{\max} \tag{34}$$

and

$$\frac{1}{\mathbf{h}_{\mathrm{S}}^H(\omega)\mathbf{\Gamma}_{\mathrm{d}}(\omega)\mathbf{h}_{\mathrm{DS}}(\omega)} = \mathcal{D}_{\max}(\omega). \tag{35}$$

We also give the obvious relationship between the two conventional beamformers:

$$\mathcal{D}_{\max}(\omega)\mathbf{\Gamma}_{\mathrm{d}}(\omega)\mathbf{h}_{\mathrm{S}}(\omega) = \mathcal{W}_{\max}\mathbf{h}_{\mathrm{DS}}(\omega). \tag{36}$$

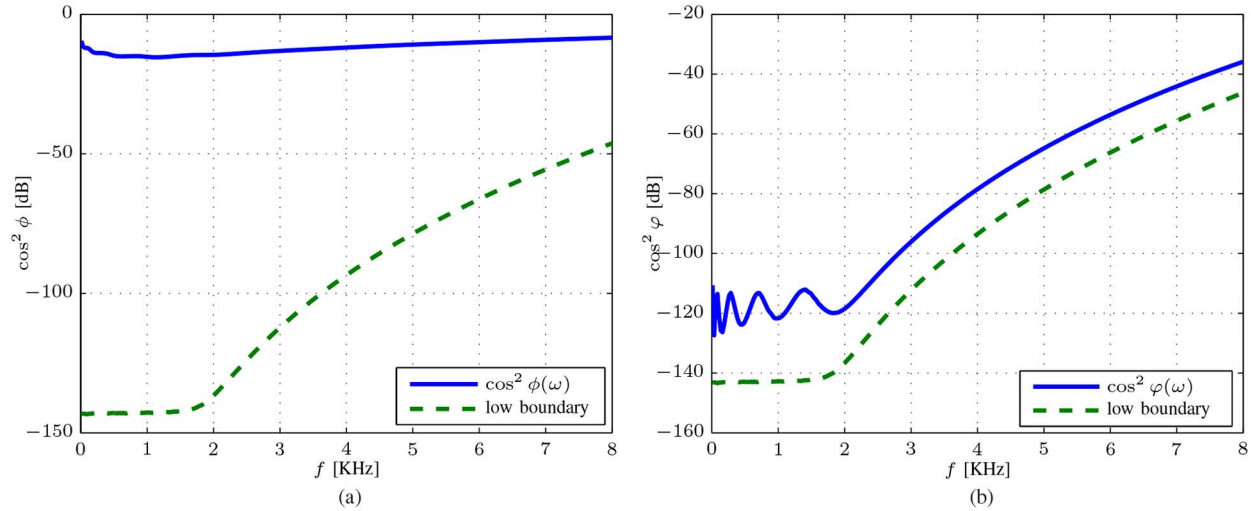

Fig. 3. (a) The DF ratio $\cos^2\phi(\omega)$ (solid line), and its low boundary (26) (dashed line) versus frequency, with $M = 8$ microphones and $\delta = 1$ cm. (b) The WNG ratio $\cos^2\varphi(\omega)$ (solid line), and its low boundary (33) (dashed line) versus frequency, with $M = 8$ microphones and $\delta = 1$ cm.

Since (27) is sensitive to the spatially white noise, the authors in [1], [3], proposed to maximize the DF subject to a constraint on the WNG. Using the distortionless constraint, we find that the optimal solution is [1], [3]

$$\mathbf{h}_{\mathrm{S},\epsilon}(\omega) = \frac{[\mathbf{\Gamma}_{\mathrm{d}}(\omega) + \epsilon\mathbf{I}_M]^{-1}\mathbf{d}(\omega)}{\mathbf{d}^H(\omega)[\mathbf{\Gamma}_{\mathrm{d}}(\omega) + \epsilon\mathbf{I}_M]^{-1}\mathbf{d}(\omega)}, \tag{37}$$

where $\epsilon \geq 0$ is a Lagrange multiplier [27]. It is clear that (37) is a regularized (or robust) version of (27), where ϵ can be seen as the regularization parameter. This parameter tries to find a good 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 (37) 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 (37) as an ϵ-regularized superdirective beamformer:

$$\mathbf{h}_{\mathrm{S},\epsilon}(\omega) = \frac{\mathbf{\Gamma}_{\epsilon}^{-1}(\omega)\mathbf{d}(\omega)}{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\epsilon}^{-1}(\omega)\mathbf{d}(\omega)}, \tag{38}$$

where

$$\mathbf{\Gamma}_{\epsilon}(\omega) = \mathbf{\Gamma}_{\mathrm{d}}(\omega) + \epsilon\mathbf{I}_M \tag{39}$$

is a regularized version of the pseudo-coherence matrix of the diffuse noise. The corresponding WNG and DF for this beamformer are, respectively,

$$\mathcal{W}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)] = \frac{\left[\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\epsilon}^{-1}(\omega)\mathbf{d}(\omega)\right]^2}{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\epsilon}^{-2}(\omega)\mathbf{d}(\omega)} \tag{40}$$

and

$$\mathcal{D}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)] = \frac{\left[\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\epsilon}^{-1}(\omega)\mathbf{d}(\omega)\right]^2}{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\epsilon}^{-1}(\omega)\mathbf{\Gamma}_{\mathrm{d}}(\omega)\mathbf{\Gamma}_{\epsilon}^{-1}(\omega)\mathbf{d}(\omega)}. \tag{41}$$

If we define

$$\mathcal{D}_{\max,\epsilon}(\omega) = \mathbf{d}^H(\omega)\mathbf{\Gamma}_{\epsilon}^{-1}(\omega)\mathbf{d}(\omega), \tag{42}$$

from (39) we can express the DF (41) as

$$\mathcal{D}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)] = \frac{1}{\mathcal{D}_{\max,\epsilon}^{-1}(\omega) - \epsilon\mathcal{W}^{-1}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]}. \tag{43}$$

Similarly to (31), we can express the WNG (40) as

$$\mathcal{W}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)] = \mathcal{W}_{\max}\cos^2\varphi_{\epsilon}(\omega), \tag{44}$$

where

$$\begin{aligned}\cos\varphi_{\epsilon}(\omega) &= \cos\left[\mathbf{d}(\omega), \mathbf{\Gamma}_{\epsilon}^{-1}(\omega)\mathbf{d}(\omega)\right] \\ &= \frac{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\epsilon}^{-1}(\omega)\mathbf{d}(\omega)}{\sqrt{\mathbf{d}^H(\omega)\mathbf{d}(\omega)}\sqrt{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\epsilon}^{-2}(\omega)\mathbf{d}(\omega)}}\end{aligned} \tag{45}$$

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

Since the DS beamformer maximizes the WNG and the regularized superdirective beamformer enables to control the DF response, it seems natural to combine the two into the following beamformer:

$$\begin{aligned}\mathbf{h}_{\alpha,\epsilon}(\omega) &= \frac{[\mathbf{\Gamma}_{\epsilon}^{-1}(\omega) + \alpha(\omega)\mathbf{I}_M]\mathbf{d}(\omega)}{\mathbf{d}^H(\omega)[\mathbf{\Gamma}_{\epsilon}^{-1}(\omega) + \alpha(\omega)\mathbf{I}_M]\mathbf{d}(\omega)} \\ &= \frac{\mathcal{D}_{\max,\epsilon}(\omega)}{\mathcal{D}_{\max,\epsilon}(\omega) + \alpha(\omega)M} \cdot \frac{\mathbf{\Gamma}_{\epsilon}^{-1}(\omega)\mathbf{d}(\omega)}{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\epsilon}^{-1}(\omega)\mathbf{d}(\omega)} \\ &\quad + \frac{\alpha(\omega)M}{\mathcal{D}_{\max,\epsilon}(\omega) + \alpha(\omega)M} \cdot \frac{\mathbf{d}(\omega)}{M} \\ &= \frac{\mathbf{h}_{\mathrm{S},\epsilon}(\omega)}{1 + \alpha_{\epsilon}(\omega)} + \frac{\mathbf{h}_{\mathrm{DS}}(\omega)}{1 + \alpha_{\epsilon}^{-1}(\omega)},\end{aligned} \tag{46}$$

where $\alpha(\omega)$ is a real number and

$$\alpha_{\epsilon}(\omega) = \alpha(\omega)\frac{\mathcal{W}_{\max}}{\mathcal{D}_{\max,\epsilon}(\omega)}. \tag{47}$$

This beamformer allows to control both the inversion of the pseudo-coherence matrix and the regularization with ϵ, and the

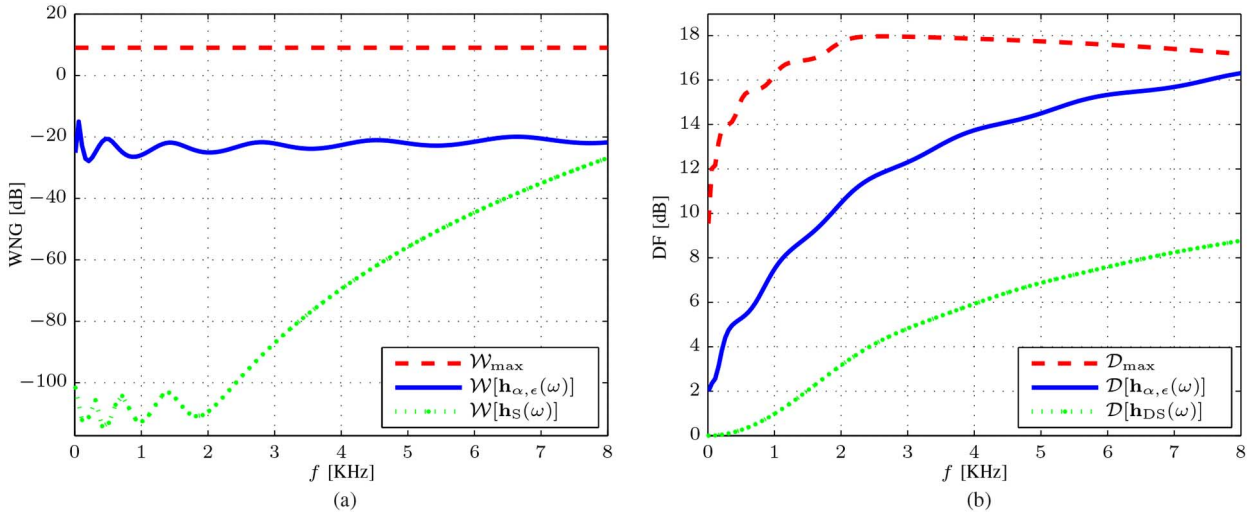

Fig. 4. Example of the array gains of $\mathbf{h}_{\alpha,\epsilon}(\omega)$ beamformer (solid line) versus frequency, with $M = 8$ microphones, $\delta = 1$ cm, $\alpha = 1$ and $\epsilon = 1 \cdot 10^{-5}$. (a) WNG. As a reference, $\mathcal{W}_{\max}$ (dashed line) and $\mathcal{W}[\mathbf{h}_{\mathrm{S}}(\omega)]$ (dotted line) are plotted. (b) DF. As a reference, $\mathcal{D}_{\max}(\omega)$ (dashed line) and $\mathcal{D}[\mathbf{h}_{\mathrm{DS}}(\omega)]$ (dotted line) are plotted.

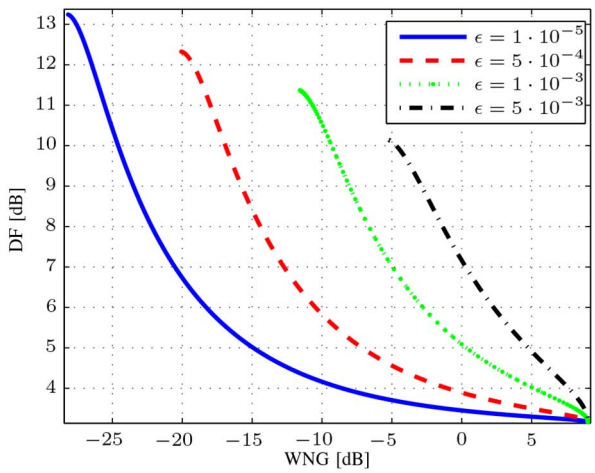

Fig. 5. The DF curve versus WNG, of $\mathbf{h}_{\alpha,\epsilon}(\omega)$ for fixed frequency and increasing α, for different values of ϵ: $\epsilon = 1 \cdot 10^{-5}$ (solid line), $\epsilon = 5 \cdot 10^{-4}$ (dashed line), $\epsilon = 1 \cdot 10^{-3}$ (dotted line), and $\epsilon = 5 \cdot 10^{-3}$ (dot-dash-dot line). $M = 8$ microphones, $\delta = 1$ cm, $f = 2$ KHz, and α is monotonically increased from 0 to ∞.

DS influence with $\alpha(\omega)$. It is easy to verify that the beamformer $\mathbf{h}_{\alpha,\epsilon}(\omega)$ is distortionless, i.e., $\mathbf{h}_{\alpha,\epsilon}^{H}(\omega)\mathbf{d}(\omega) = 1$.

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

$$\mathcal{W}[\mathbf{h}_{\alpha,\epsilon}(\omega)] = \frac{[1+\alpha_\epsilon(\omega)]^2 \mathcal{W}[\mathbf{h}_{\mathrm{DS}}(\omega)]\mathcal{W}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]}{\mathcal{W}[\mathbf{h}_{\mathrm{DS}}(\omega)] + \left\{[1+\alpha_\epsilon(\omega)]^2 - 1\right\}\mathcal{W}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]} = \frac{[1+\alpha_\epsilon(\omega)]^2 \mathcal{W}_{\max}\cos^2\varphi_\epsilon(\omega)}{1 + \left\{[1+\alpha_\epsilon(\omega)]^2 - 1\right\}\cos^2\varphi_\epsilon(\omega)} \leq \mathcal{W}_{\max}, \tag{48}$$

which depends on the WNGs of the DS and regularized superdirective beamformers. We see that for $\alpha_\epsilon(\omega) = 0$, we have $\mathcal{W}[\mathbf{h}_{0,\epsilon}(\omega)] = \mathcal{W}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]$, and for $\alpha_\epsilon(\omega) \to \infty$, we have $\mathcal{W}[\mathbf{h}_{\infty,\epsilon}(\omega)] = \mathcal{W}[\mathbf{h}_{\mathrm{DS}}(\omega)]$. These results make obviously sense. Also, we have

$$\mathcal{W}[\mathbf{h}_{\alpha,\epsilon}(\omega)] \geq \mathcal{W}\left[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)\right], \ \forall \alpha_\epsilon(\omega) \geq 0, \tag{49}$$

suggesting that we should always choose $\alpha(\omega) \geq 0$.

It can also be verified that the inverse DF corresponding to $\mathbf{h}_{\alpha,\epsilon}(\omega)$ is

$$\begin{aligned}\mathcal{D}^{-1}[\mathbf{h}_{\alpha,\epsilon}(\omega)] &= [1+\alpha_\epsilon(\omega)]^{-2} \cdot \left\{\mathcal{D}^{-1}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]\right. \\ &\quad + 2\alpha_\epsilon(\omega)\frac{\mathbf{d}^H(\omega)\boldsymbol{\Gamma}_{\mathrm{d}}(\omega)\boldsymbol{\Gamma}_\epsilon^{-1}(\omega)\mathbf{d}(\omega)}{M \cdot \mathbf{d}^H(\omega)\boldsymbol{\Gamma}_\epsilon^{-1}(\omega)\mathbf{d}(\omega)} \\ &\quad \left. + \alpha_\epsilon^2(\omega)\mathcal{D}^{-1}[\mathbf{h}_{\mathrm{DS}}(\omega)]\right\} \\ &= [1+\alpha_\epsilon(\omega)]^{-2} \cdot \{\mathcal{D}^{-1}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)] \\ &\quad + 2\alpha_\epsilon(\omega)\left[\mathcal{D}_{\max,\epsilon}^{-1}(\omega) - \epsilon\frac{1}{M}\right] \\ &\quad + \alpha_\epsilon^2(\omega)\mathcal{D}^{-1}[\mathbf{h}_{\mathrm{DS}}(\omega)]\},\end{aligned} \tag{50}$$

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

$$\mathcal{D}[\mathbf{h}_{\alpha,\epsilon}(\omega)] \leq \mathcal{D}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)] \tag{51}$$

$$\mathcal{D}[\mathbf{h}_{\alpha,\epsilon}(\omega)] \geq \mathcal{D}[\mathbf{h}_{\mathrm{DS}}(\omega)], \ \forall \alpha_\epsilon(\omega) \geq 0, \tag{52}$$

which implies again that we should take $\alpha(\omega) \geq 0$. Examples of the WNG and the DF of $\mathbf{h}_{\alpha,\epsilon}(\omega)$ beamformer are described in Fig. 4. We observe that the received WNG and DF responses of $\mathbf{h}_{\alpha,\epsilon}(\omega)$ are obtained between the minimal and maximal values, dictated by the DS and superdirective beamformers. Our goal is to obtain a beamformer with adequate WNG level and relatively high DF. Therefore, we focus on tuning the parameters ϵ and $\alpha(\omega)$ which determine its exact response.

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 $\mathcal{W}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]$. Obviously, there are many different ways to find $\alpha(\omega)$ given a specific regularization factor ϵ, depending on what we desire. Next, we discuss two interesting possibilities.

In the first approach, we would like to find the value of $\alpha(\omega)$ in such a way that $\mathcal{W}[\mathbf{h}_{\alpha,\epsilon}(\omega)] = \mathcal{W}_0$, where $\mathcal{W}_0$ is a constant

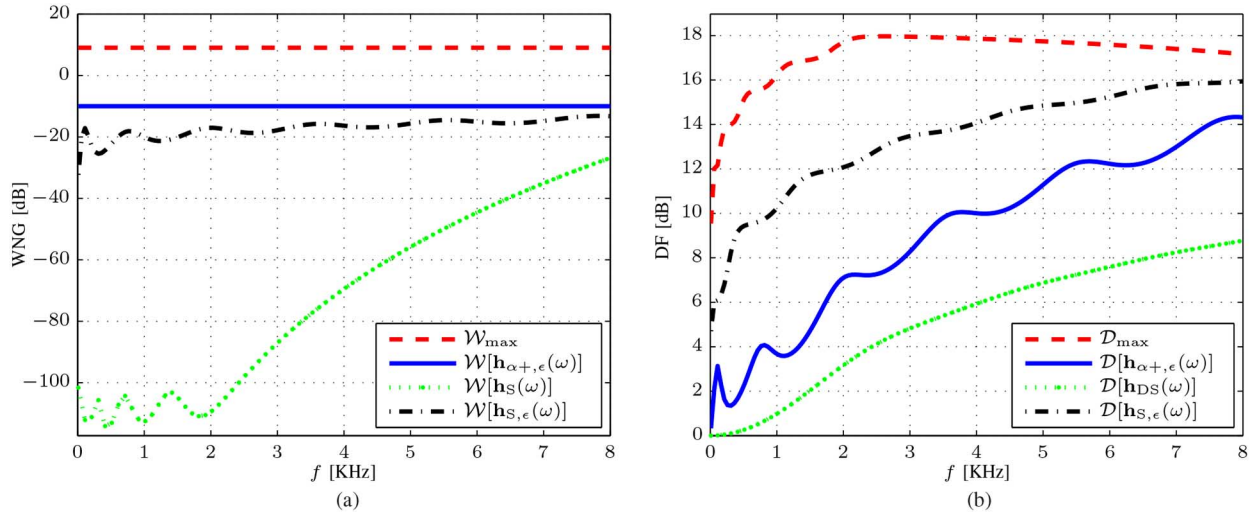

Fig. 6. Example of the array gains of $\mathbf{h}_{\alpha+,\epsilon}(\omega)$ beamformer (57) (solid line) versus frequency, with $M = 8$ microphones, $\delta = 1$ cm, and $\epsilon = 1 \cdot 10^{-4}$. The desired WNG is set to $\mathcal{W}_0 = -10$ dB. (a) WNG. As a reference, $\mathcal{W}_{\max}$ (dashed line), $\mathcal{W}[\mathbf{h}_{\mathrm{S}}(\omega)]$ (dotted line), and $\mathcal{W}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]$ (dot-dash-dot line) are plotted. (b) DF. As a reference, $\mathcal{D}_{\max}(\omega)$ (dashed line), $\mathcal{D}[\mathbf{h}_{\mathrm{DS}}(\omega)]$ (dotted line), and $\mathcal{D}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]$ (dot-dash-dot line) are plotted.

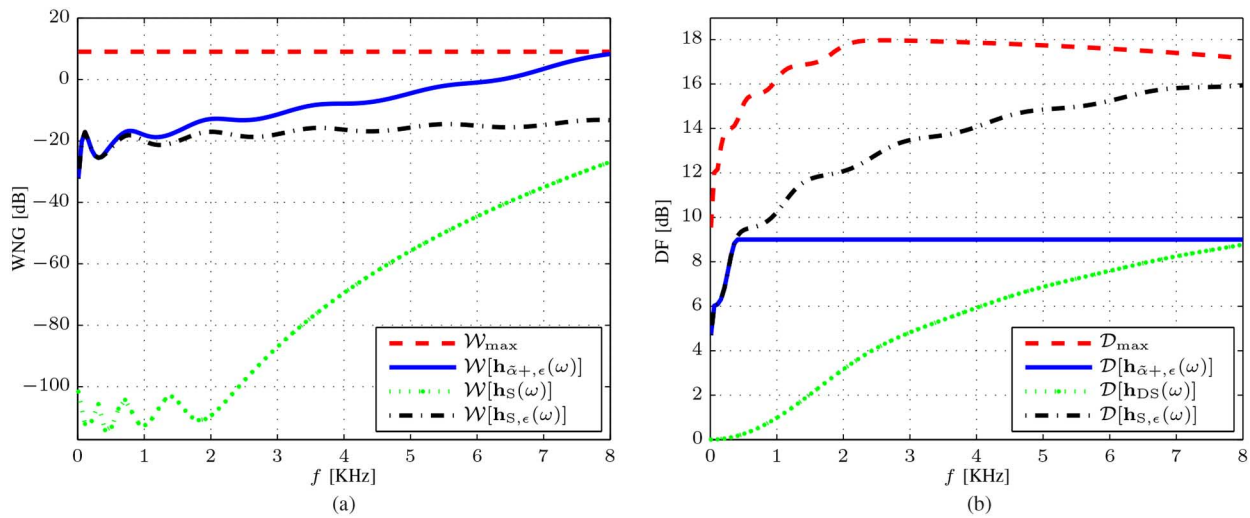

Fig. 7. Example of the array gains of $\mathbf{h}_{\bar{\alpha}+,\epsilon}(\omega)$ beamformer as a solution of (58) (solid line) versus frequency, with $M = 8$ microphones, $\delta = 1$ cm, and $\epsilon = 1 \cdot 10^{-4}$. The desired DF is set to $\mathcal{D}_0 = 9$ dB. (a) WNG. As a reference, $\mathcal{W}_{\max}$ (dashed line), $\mathcal{W}[\mathbf{h}_{\mathrm{S}}(\omega)]$ (dotted line), and $\mathcal{W}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]$ (dot-dash-dot line) are plotted. (b) DF. As a reference, $\mathcal{D}_{\max}(\omega)$ (dashed line), $\mathcal{D}[\mathbf{h}_{\mathrm{DS}}(\omega)]$ (dotted line), and $\mathcal{D}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]$ (dot-dash-dot line) are plotted.

determined by the desired white-noise robustness level, with $\mathcal{W}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)] < \mathcal{W}_0 < M,\ \forall\omega$. Using (48), we find that

$$[1+\alpha_\epsilon(\omega)]^2 = \frac{\mathcal{W}_0}{\mathcal{W}_{\max}-\mathcal{W}_0}\frac{1-\cos^2\varphi_\epsilon(\omega)}{\cos^2\varphi_\epsilon(\omega)} = \frac{\mathcal{W}_0}{\mathcal{W}_{\max}-\mathcal{W}_0}\tan^2\varphi_\epsilon(\omega), \tag{53}$$

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

$$\alpha_{\epsilon+}(\omega) = \sqrt{\frac{\mathcal{W}_0}{\mathcal{W}_{\max}-\mathcal{W}_0}}\,|\tan\varphi_\epsilon(\omega)| - 1, \tag{54}$$

$$\alpha_{\epsilon-}(\omega) = -\sqrt{\frac{\mathcal{W}_0}{\mathcal{W}_{\max}-\mathcal{W}_0}}\,|\tan\varphi_\epsilon(\omega)| - 1. \tag{55}$$

As a consequence, the corresponding values for $\alpha(\omega)$ are

$$\alpha_\pm(\omega) = \frac{\mathcal{D}_{\max,\epsilon}(\omega)}{\mathcal{W}_{\max}}\alpha_{\epsilon\pm}(\omega). \tag{56}$$

From the two solutions $\alpha_+(\omega)$ and $\alpha_-(\omega)$, we obviously choose the first one; in this case, the beamformer is

$$\mathbf{h}_{\alpha_+,\epsilon}(\omega) = \frac{[\mathbf{\Gamma}_\epsilon^{-1}(\omega)+\alpha_+(\omega)\mathbf{I}_M]\mathbf{d}(\omega)}{\mathbf{d}^H(\omega)[\mathbf{\Gamma}_\epsilon^{-1}(\omega)+\alpha_+(\omega)\mathbf{I}_M]\mathbf{d}(\omega)} = \frac{\mathbf{h}_{\mathrm{S},\epsilon}(\omega)}{1+\alpha_{\epsilon+}(\omega)} + \frac{\mathbf{h}_{\mathrm{DS}}(\omega)}{1+\alpha_{\epsilon+}^{-1}(\omega)}, \tag{57}$$

In the second approach, we would like to find the values of $\alpha(\omega)$ in such a way that $\mathcal{D}[\mathbf{h}_{\alpha,\epsilon}(\omega)] = \mathcal{D}_0$, where $\mathcal{D}_0$ is a determined diffuse-noise gain level constant, with $\mathcal{D}[\mathbf{h}_{\mathrm{DS}}(\omega)] < \mathcal{D}_0 < \mathcal{D}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)],\ \forall\omega$. We can express (50) as a second degree polynomial of $\alpha_\epsilon(\omega)$:

$$\alpha_\epsilon^2(\omega)\left\{\mathcal{D}^{-1}[\mathbf{h}_{\mathrm{DS}}(\omega)] - \mathcal{D}_0^{-1}\right\} + 2\alpha_\epsilon(\omega)[\mathcal{D}_{\max,\epsilon}^{-1}(\omega) - \epsilon\frac{1}{M} - \mathcal{D}_0^{-1}] + \mathcal{D}^{-1}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)] - \mathcal{D}_0^{-1} = 0. \tag{58}$$

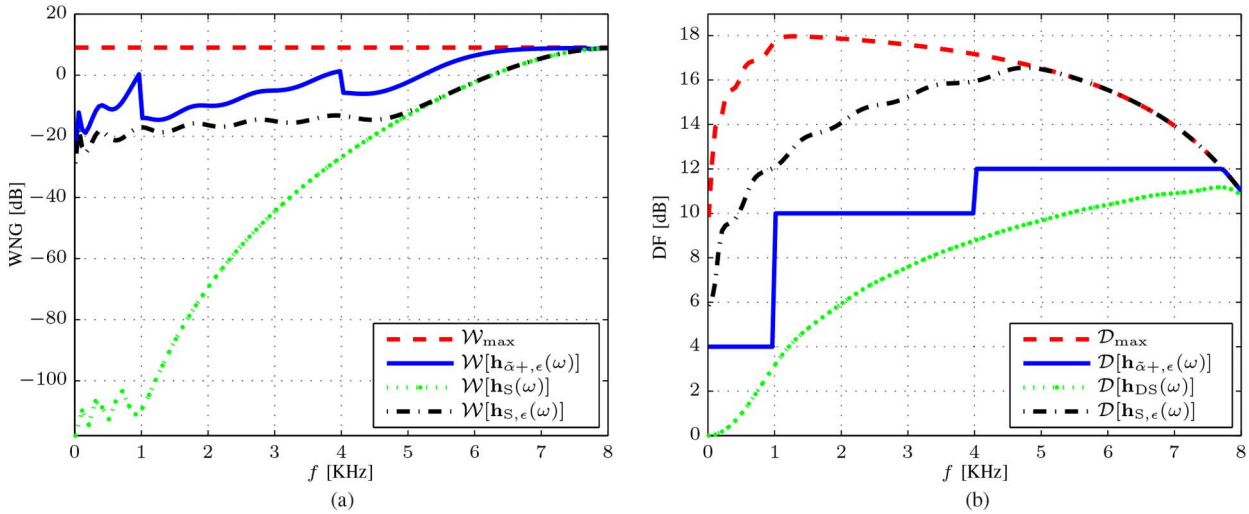

Fig. 8. The array gains of $\mathbf{h}_{\tilde{\alpha}+,\epsilon}(\omega)$ beamformer (solid line) versus frequency, for multiband $\mathcal{D}_0(\omega)$, with $M = 8$ microphones, $\delta = 2$ cm, and $\epsilon = 1 \cdot 10^{-4}$. (a) WNG. As a reference, $\mathcal{W}_{\max}$ (dashed line), $\mathcal{W}[\mathbf{h}_{\mathrm{S}}(\omega)]$ (dotted line), and $\mathcal{W}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]$ (dot-dash-dot line) are plotted. (b) DF. As a reference, $\mathcal{D}_{\max}(\omega)$ (dashed line), $\mathcal{D}[\mathbf{h}_{\mathrm{DS}}(\omega)]$ (dotted line), and $\mathcal{D}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]$ (dot-dash-dot line) are plotted.

We deduce two possible solutions for $\alpha_\epsilon(\omega)$ by solving this quadratic equation. These are marked as $\tilde{\alpha}_{\epsilon\pm}(\omega)$. Therefore, the corresponding values for $\alpha(\omega)$ are

$$\tilde{\alpha}_{\pm}(\omega) = \frac{\mathcal{D}_{\max,\epsilon}(\omega)}{\mathcal{W}_{\max}} \tilde{\alpha}_{\epsilon\pm}(\omega). \tag{59}$$

From the two solutions $\tilde{\alpha}_{\pm}(\omega)$, we take the positive one, and obtain a constant DF beamformer $\mathbf{h}_{\tilde{\alpha}+,\epsilon}(\omega)$.

Based on a reformulation of the DF definition (15)

$$\mathcal{D}[\mathbf{h}(\omega)] = \frac{\left|\mathcal{B}[\mathbf{h}(\omega), 0^\circ]\right|^2}{\frac{1}{2}\int_0^\pi \left|\mathcal{B}[\mathbf{h}(\omega), \theta]\right|^2 \sin\theta d\theta}, \tag{60}$$

and under the distortionless constraint (6), we find that for the fixed-DF beamformer

$$\mathcal{D}_0^{-1} = \frac{1}{2}\int_0^\pi \left|\mathcal{B}[\mathbf{h}_{\tilde{\alpha}+,\epsilon}(\omega), \theta]\right|^2 \sin\theta d\theta. \tag{61}$$

This can be interpreted as a constant mean squared beampattern constraint

$$\mathbb{E}_\theta\{\left|\mathcal{B}[\mathbf{h}(\omega), \theta]\right|^2\} = const, \ \forall\omega, \tag{62}$$

where $\mathbb{E}_\theta\{f(\theta)\} = \int_0^\pi f(\theta)\sin\theta d\theta$. Namely, the fixed-DF approach comprises a practical closed-formula of specific form of the frequency-invariant beampattern beamformer (FIBP) [11], [12], [18], [19], [12].

V. Simulations

As mentioned before, there is a major importance of setting an appropriate value for the regularization factor ϵ. It determines the range of the possible WNG and DF we can achieve. To embody this, we added the response of $\mathbf{h}_{\mathrm{S},\epsilon}(\omega)$ to the illustrated simulations. This parameter is set constant over all frequency range. Obviously, setting an adaptive frequency-dependent regularization parameter $\epsilon(\omega)$ would broaden the leeway for obtaining better WNG and DF responses for every frequency. Unfortunately, there is no solution for extracting ϵ out of the regularized superdirective WNG or DF terms [eq. (40), (41)]. Moreover, to the best of the authors knowledge, there exists no direct closed-form solution in the literature for $\epsilon(\omega)$ given a desired WNG or DF level.

Following [3], to demonstrate the influence of the filter parameters on the WNG–DF tradeoff, we show in Fig. 5 the DF curve vs. WNG of $\mathbf{h}_{\alpha,\epsilon}(\omega)$ for increasing $\alpha(\omega)$, for different ϵ values. The parameter $\alpha(\omega) = \alpha, \ \forall\omega$, varies from 0 to ∞ along the curves. 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 $\alpha \to \infty$, 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.

Obviously, the physical properties of the microphone array, such as the number of elements M and the microphone spacing δ affect the array response. In general, increasing M provides higher maximal WNG and higher maximal DF. Increasing δ provides better WNG-DF tradeoff [3], as long as $\delta \ll \lambda$. Choosing higher δ is equivalent to picking higher frequency f in the response.

We simulated the fixed-WNG beamformer $\mathbf{h}_{\alpha+,\epsilon}(\omega)$ (57). In Fig. 6, we show an example of the WNG and the DF response of such beamformer. The desired WNG is chosen such that it is always higher than $\mathcal{W}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]$. The value of $\mathcal{W}_0 = -10$ dB provides both tolerable constant white noise amplification, together with a satisfying DF, which is somewhere between the minimal DF of the superdirective beamformer and below the DF of the regularized superdirective beamformer $\mathbf{h}_{\mathrm{S},\epsilon}(\omega)$.

Next, we simulated the fixed-DF beamformer $\mathbf{h}_{\tilde{\alpha}+,\epsilon}(\omega)$ by solving (58) for fixed $\mathcal{D}_0$ beamformer. The corresponding WNG and DF responses are illustrated in Fig. 7. From Fig. 7(a), we observe that the received WNG is indeed between the WNG of the regularized superdirective beamformer and the maximal WNG of the DS beamformer. From Fig. 7(b), we note that we obtain a very nice constant DF, except at very low frequencies.

$\mathcal{D}_0$ is limited within the range of $\mathcal{D}[\mathbf{h}_{\mathrm{DS}}(\omega)] < \mathcal{D}_0 < \mathcal{D}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)], \ \forall\omega$. We can hold this by designing a multi-

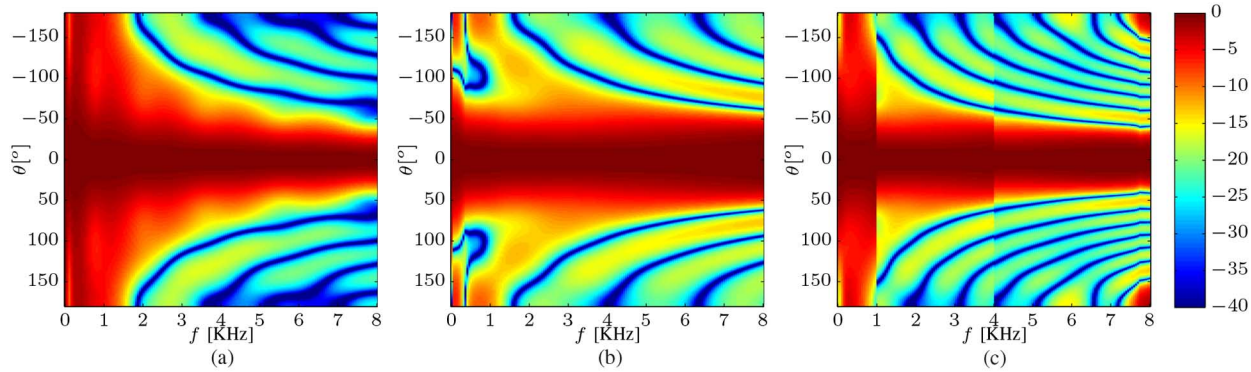

Fig. 9. The squared beampattern [dB] versus frequency and θ, with $M = 8$ microphones and $\epsilon = 1 \cdot 10^{-4}$. (a) $|\mathcal{B}[\mathbf{h}_{\alpha+,\epsilon}(\omega), \theta]|^2$ of the fixed-WNG beamformer, with $\delta = 1$ cm. (b) $|\mathcal{B}[\mathbf{h}_{\tilde{\alpha}+,\epsilon}(\omega), \theta]|^2$ of the fixed-DF beamformer ($\delta = 1$ cm). (c) $|\mathcal{B}[\mathbf{h}_{\tilde{\alpha}+,\epsilon}(\omega), \theta]|^2$ of the multiband-fixed DF beamformer ($\delta = 2$ cm).

band-fixed DF, where for every frequency band we set a different $\mathcal{D}_0$, i.e., $\mathcal{D}_0 = \mathcal{D}_0(\omega)$. The values of $\alpha_\epsilon(\omega)$ are still obtained by (58), but now we replace $\mathcal{D}_0$ with $\mathcal{D}_0(\omega)$. An example of such beamformer is described in Fig. 8. $\mathcal{D}_0(\omega)$ is set here such that it is (much) lower than $\mathcal{D}[\mathbf{h}_{\mathrm{S},\epsilon}(\omega)]$ at low frequencies. Note that δ is set here to 2 cm, twice of its value so far. A similar analysis can be followed for designing a multiband-fixed WNG beamformer, by setting $\mathcal{W}_0 = \mathcal{W}_0(\omega)$.

In addition, we simulated the beampattern of the above beamformers. In Fig. 9(a)–(c), the squared beampattern responses of the fixed-WNG beamformer $\mathbf{h}_{\alpha+,\epsilon}(\omega)$, the fixed-DF beamformer $\mathbf{h}_{\tilde{\alpha}+,\epsilon}(\omega)$, and the multiband-fixed DF beamformer are illustrated, respectively. Examining Fig. 9(b), we note that its significant beampattern level is roughly constant over frequency, as oppose to Fig. 9(a). From Fig. 9(b)–(c), we deduce that by the (multiband) fixed-DF solution we get a practical beamformer with (intermittent) frequency-invariant beampattern.

VI. Conclusions

We have proposed a new approach to robust broadband regularized beamforming as a combination of a regularized version of the superdirective beamformer and the DS beamformer. Given a user-determined regularization factor ϵ, it enables control of the WNG–DF tradeoff by simple terms for the parameter $\alpha(\omega)$, that determines the beamformer frequency response. In addition, two closed-form beamformers have been derived for a constant-WNG and a constant-DF beamformer, which approximates a frequency-invariant mean beampattern as well. Finally, we saw that by using this approach one can design and obtain any desired frequency-dependent WNG or DF beamformer, within the allowed range.

In future work, we would test this approach of robust regularized beamforming over other types of noise fields. We would inspect a frequency-dependent regularization for a more versatile performance. Additionally, we would examine generalization of the proposed solution to a problem with general linear constraints, such as representations of correlated noise, spatial limitations, side-lobe requirements, etc. We would also discuss practical design considerations for implementation in the discrete frequency domain.

Acknowledgment

The authors thank the anonymous reviewers for their constructive comments and useful suggestions.

References

[1] H. Cox, R. M. Zeskind, and M. M. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, no. 10, pp. 1365–1376, Oct. 1987.

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

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

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

[5] 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.

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

[7] J. Li, P. Stoica, and Z. Wang, "On robust capon beamforming and diagonal loading," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1702–1715, Jul. 2003.

[8] R. G. Lorenz and S. P. Boyd, "Robust minimum variance beamforming," *IEEE Trans. Signal Process.*, vol. 53, no. 5, pp. 1684–1696, May 2005.

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

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

[11] 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, May 2011.

[12] M. Crocco and A. Trucco, "Stochastic and analytic optimization of sparse aperiodic arrays and broadband beamformers with robust superdirective patterns," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 20, no. 9, pp. 2433–2447, Nov. 2012.

[13] S. Doclo and M. Moonen, "Design of broadband beamformers robust against gain and phase errors in the microphone array characteristics," *IEEE Trans. Signal Process.*, vol. 51, no. 10, pp. 2511–2526, Oct. 2003.

[14] R. C. Nongpiur, "Design of minimax broadband beamformers that are robust to microphone gain, phase, and position errors," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 22, no. 6, pp. 1013–1022, Jun. 2014.

[15] R. C. Nongpiur, "Design of minimax robust broadband beamformers with optimized microphone positions," *Digital Signal Process.*, vol. 32, pp. 100–108, Sep. 2014.

[16] H. Chen, W. Ser, and Z. L. Yu, "Optimal design of nearfield wideband beamformers robust against errors in microphone array characteristics," *IEEE Trans. Circuits Syst.*, vol. 54, no. 9, pp. 1950–1959, Sep. 2007.
[17] M. Crocco and A. Trucco, "A computationally efficient procedure for the design of robust broadband beamformers," *IEEE Trans. Signal Process.*, vol. 58, no. 20, pp. 5420–5424, Oct. 2010.
[18] F. Traverso, M. Crocco, and A. Trucco, "Design of frequency-invariant robust beam patterns by the oversteering of end-fire arrays," *Signal Process.*, vol. 99, pp. 129–135, Jun. 2014.
[19] E. Mabande, A. Schad, and W. Kellermann, "A time-domain implementation of data-independent robust broadband beamformers with low filter order," in *Proc. Joint Workshop Hands-Free Speech Commun. Microphone Arrays (HSCMA)*, May 2011, pp. 81–85.
[20] R. C. Nongpiur and D. J. Shpak, "L -infinity norm design of linear-phase robust broadband beamformers using constrained optimization," *IEEE Trans. Signal Process.*, vol. 61, no. 23, pp. 6034–6046, Dec. 2013.
[21] D. B. Ward, R. A. Kennedy, and R. C. Williamson, "Theory and design of broadband sensor arrays with frequency invariant far-field beam patterns," *J. Acoust. Soc. Amer.*, vol. 97, no. 2, pp. 1023–1034, 1995.
[22] R. Fletcher, *Practical Methods of Optimization*. New York, NY, USA: Wiley, 1987.
[23] G. A. F. Seber, *A Matrix Handbook for Statisticians*. Hoboken, NJ, USA: Wiley, 2008.
[24] J. Capon, "High resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, no. 8, pp. 1408–1418, Aug. 1969.
[25] R. T. Lacoss, "Data adaptive spectral analysis methods," *Geophysics*, vol. 36, pp. 661–675, Aug. 1971.
[26] A. I. Uzkov, "An approach to the problem of optimum directive antenna design," *Comptes Rendus (Doklady) de l'Academie des Sciences de l'URSS*, vol. LIII, no. 1, pp. 35–38, 1946.
[27] G. A. F. Seber and A. J. Lee, "Linear Regression: Estimation and Distribution Theory," in *Linear Regression Analysis*. Hoboken, NJ, USA: Wiley, 2003, ch. 3, pp. 33–95.

Reuven Berkun received the B.Sc. degrees (*Cum Laude*) in electrical engineering and physics in 2011 from the Technion–Israel Institute of Technology, Haifa, Israel. He is currently pursuing the M.Sc. degree in electrical engineering at the Technion. From 2010 to 2014, he was a video algorithms and architecture engineer in CSR, Haifa, Israel. Since 2014, he has been a Teaching Assistant at the Electrical Engineering Department, Technion. His research interests include multi-channel signal processing, speech enhancement and microphone arrays.

Israel Cohen (M'01–SM'03–F'15) received the B.Sc. (*Summa Cum Laude*), M.Sc. and Ph.D. degrees in electrical engineering from the Technion–Israel Institute of Technology, in 1990, 1993 and 1998, respectively.

He is a Professor of electrical engineering at the Technion–Israel Institute of Technology, Haifa, Israel. From 1990 to 1998, he was a Research Scientist with RAFAEL Research Laboratories, Haifa, Israel Ministry of Defense. From 1998 to 2001, he was a Postdoctoral Research Associate with the Computer Science Department, Yale University, New Haven, CT, USA. In 2001 he joined the Electrical Engineering Department of the Technion. He is a coeditor of the Multichannel Speech Processing Section of the *Springer Handbook of Speech Processing* (Springer, 2008), a coauthor of *Noise Reduction in Speech Processing* (Springer, 2009), a Coeditor of *Speech Processing in Modern Communication: Challenges and Perspectives* (Springer, 2010), and a General Cochair of the 2010 International Workshop on Acoustic Echo and Noise Control (IWAENC). He served as Guest Editor of the *European Association for Signal Processing Journal on Advances in Signal Processing* Special Issue on Advances in Multimicrophone Speech Processing and the *Elsevier Speech Communication Journal* a Special Issue on Speech Enhancement. His research interests are statistical signal processing, analysis and modeling of acoustic signals, speech enhancement, noise estimation, microphone arrays, source localization, blind source separation, system identification, and adaptive filtering. Dr. Cohen was a recipient of the Alexander Goldberg Prize for Excellence in Research, and the Muriel and David Jacknow Award for Excellence in Teaching. He serves as a member of the IEEE Audio and Acoustic Signal Processing Technical Committee and the IEEE Speech and Language Processing Technical Committee. He served as Associate Editor of the IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING and IEEE SIGNAL PROCESSING LETTERS.

Jacob Benesty was born in 1963. He received a Master degree in microwaves from Pierre & Marie Curie University, France, in 1987, and a Ph.D. degree in control and signal processing from Orsay University, France, in April 1991. During his Ph.D. (from Nov. 1989 to Apr. 1991), he worked on adaptive filters and fast algorithms at the Centre National d'Etudes des Telecomunications (CNET), Paris, France. From January 1994 to July 1995, he worked at Telecom Paris University on multichannel adaptive filters and acoustic echo cancellation. From October 1995 to May 2003, he was first a Consultant and then a Member of the Technical Staff at Bell Laboratories, Murray Hill, NJ, USA. In May 2003, he joined the University of Quebec, INRS-EMT, in Montreal, Quebec, Canada, as a Professor. He is also a Visiting Professor at the Technion, in Haifa, Israel, an Adjunct Professor at Aalborg University, in Denmark, and a Guest Professor at Northwestern Polytechnical University, in Xi'an, Shaanxi, China. His research interests are in signal processing, acoustic signal processing, and multimedia communications. He is the inventor of many important technologies. In particular, he was the lead researcher at Bell Labs who conceived and designed the world-first real-time hands-free full-duplex stereophonic teleconferencing system. Also, he conceived and designed the world-first PC-based multi-party hands-free full-duplex stereo conferencing system over IP networks. He was the co-chair of the 1999 International Workshop on Acoustic Echo and Noise Control and the general co-chair of the 2009 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. He is the recipient, with Morgan and Sondhi, of the IEEE Signal Processing Society 2001 Best Paper Award. He is the recipient, with Chen, Huang, and Doclo, of the IEEE Signal Processing Society 2008 Best Paper Award. He is also the co-author of a paper for which Huang received the IEEE Signal Processing Society 2002 Young Author Best Paper Award. In 2010, he received the Gheorghe Cartianu Award from the Romanian Academy. In 2011, he received the Best Paper Award from the IEEE WASPAA for a paper that he co-authored with Chen.