SUBSPACE SUPERDIRECTIVE BEAMFORMERS BASED ON JOINT DIAGONALIZATION

Changlei Li[1], *Jacob Benesty*[2], *Gongping Huang*[1], *and Jingdong Chen*[1]

[1]CIAIC and School of Marine Science and Technology
Northwestern Polytechnical University
Xi'an, Shaanxi 710072, China

[2]INRS-EMT, University of Quebec
800 de la Gauchetiere Ouest, Suite 6900
Montreal, QC H5A 1K6, Canada

ABSTRACT

Although they have been intensively studied and used in many applications due to their high directivity factor (DF), superdirective beamformers are sensitive to sensor noise and mismatch between sensors. This paper studies the problem of superdirective beamforming combined with the joint diagonalization method. We develop a subspace superdirective beamforming approach, which can achieve a good compromise between a high DF and white noise amplification. Simulations are performed to justify our theoretical analysis and demonstrate the good properties of this subspace superdirective beamforming approach.

***Index Terms*—** Microphone arrays, superdirective beamforming, robust beamforming, supergain, white noise gain, directivity factor, joint diagonalization, subspace.

1. INTRODUCTION

Microphone array beamforming has been widely used to extract signals of interest and suppress noise and interference in room acoustic environments [1–9]. Many beamforming algorithms have been developed over the last three decades, such as the delay-and-sum (DS), the filter-and-sum, the superdirective [8–13], and the differential [8, 14]. In comparison with many other beamformers, the superdirective one can achieve a high directivity factor (DF) and, as a result, it is efficient in suppressing reverberation and spherically isotropic (diffuse) noise [8, 9]. Therefore, this beamformer has the great potential to solve many important acoustic problems in voice communications and human-machine interfaces. However, superdirective beamformers are found to be very sensitive to sensor self noise, mismatch between sensors, and other array imperfections. As a matter of fact, the lack of robustness is a big hurdle that prevents superdirective beamformers to being widely deployed in practical systems [8, 9]. Consequently, how to improve the robustness of superdirective beamformers has long been an important yet challenging problem. In this paper, we develop a subspace superdirective beamformer based on the joint diagonalization method. By changing the dimension of the subspace, we can make a good compromise between a high DF and white noise amplification.

2. SIGNAL MODEL AND PROBLEM FORMULATION

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 is

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

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 $\theta = 0$.

In this work, we consider fixed directional beamformers with small values of δ, like in superdirective [3], [5] or differential beamforming [8], [11], [14], where the main lobe is at the angle $\theta = 0$ (endfire direction) and the desired signal also propagates from $\theta = 0$. The array output can then be written as

$$\begin{aligned} \mathbf{y}(\omega) &= \begin{bmatrix} Y_1(\omega) & Y_2(\omega) & \cdots & Y_M(\omega) \end{bmatrix}^T \\ &= \mathbf{x}(\omega) + \mathbf{v}(\omega) \\ &= \mathbf{d}(\omega, 0)\, X(\omega) + \mathbf{v}(\omega), \end{aligned} \quad (2)$$

where

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

is the signal received at the mth ($m = 1, 2, \ldots, M$) microphone, $X(\omega)$ is the desired source signal, $V_m(\omega)$ is the additive noise at the mth microphone, $\mathbf{x}(\omega) = \mathbf{d}(\omega, 0)\, X(\omega)$, and $\mathbf{v}(\omega)$ is defined similarly to $\mathbf{y}(\omega)$.

Linear beamforming consists of applying a complex weight at the output of each microphone and then sum all the weighted outputs together to get an estimate of the source signal [4], [15], i.e.,

$$Z(\omega) = \sum_{m=1}^{M} H_m^*(\omega)\, Y_m(\omega) = \mathbf{h}^H(\omega)\, \mathbf{y}(\omega), \quad (4)$$

where $Z(\omega)$ is the estimate of the desired signal, $X(\omega)$, $H_m(\omega)$ is a complex weighting coefficient, the superscript * denotes complex conjugation, the superscript ^{H} is the conjugate-transpose operator, and

$$\mathbf{h}(\omega) = \begin{bmatrix} H_1(\omega) & H_2(\omega) & \cdots & H_M(\omega) \end{bmatrix}^T \quad (5)$$

is the spatial filter of length M.

The objective of this work is to find a beamformer, $\mathbf{h}(\omega)$, that is able to achieve supergains at the endfire direction with a better control on the white noise gain (WNG). In our context, the distortionless constraint is desired, i.e.,

$$\mathbf{h}^H(\omega)\, \mathbf{d}(\omega, 0) = 1. \quad (6)$$

To simplify the notation, we write $\mathbf{d}(\omega, 0)$ as $\mathbf{d}(\omega)$ in the rest of this paper.

This work was supported in part by the NSFC "Distinguished Young Scientists Fund" under grant No. 61425005.

3. WHITE NOISE GAIN, DIRECTIVITY FACTOR, AND BEAMPATTERN

The performance of a beamformer is generally evaluated with the gain in signal-to-noise ratio (SNR). If we take microphone 1 as the reference, we can define the input SNR with respect to this reference as

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

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

$$\begin{aligned} \mathrm{oSNR}[\mathbf{h}(\omega)] &= \frac{E\left[\left|\mathbf{h}^H(\omega)\mathbf{x}(\omega)\right|^2\right]}{E\left[\left|\mathbf{h}^H(\omega)\mathbf{v}(\omega)\right|^2\right]} \\ &= \phi_X(\omega) \times \frac{\left|\mathbf{h}^H(\omega)\mathbf{d}(\omega)\right|^2}{\mathbf{h}^H(\omega)\mathbf{\Phi_v}(\omega)\mathbf{h}(\omega)} \\ &= \frac{\phi_X(\omega)}{\phi_{V_1}(\omega)} \times \frac{\left|\mathbf{h}^H(\omega)\mathbf{d}(\omega)\right|^2}{\mathbf{h}^H(\omega)\mathbf{\Gamma_v}(\omega)\mathbf{h}(\omega)}, \end{aligned} \tag{8}$$

where $\mathbf{\Phi_v}(\omega) = E\left[\mathbf{v}(\omega)\mathbf{v}^H(\omega)\right]$ and $\mathbf{\Gamma_v}(\omega) = \frac{\mathbf{\Phi_v}(\omega)}{\phi_{V_1}(\omega)}$ are the correlation and pseudo-coherence matrices of $\mathbf{v}(\omega)$, respectively. The definition of the SNR gain is obtained from the previous definitions, i.e.,

$$\mathcal{G}[\mathbf{h}(\omega)] = \frac{\mathrm{oSNR}[\mathbf{h}(\omega)]}{\mathrm{iSNR}(\omega)} = \frac{\left|\mathbf{h}^H(\omega)\mathbf{d}(\omega)\right|^2}{\mathbf{h}^H(\omega)\mathbf{\Gamma_v}(\omega)\mathbf{h}(\omega)}. \tag{9}$$

We consider 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 and the SNR gain is called the WNG, i.e.,

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

  It can be shown that the maximum possible WNG is equal to the number of microphones, i.e., $\mathcal{W}_{\max} = M$. We will see how the white noise is amplified with superdirective beamformers, especially at low frequencies.

- The diffuse noise[2] where

  $$\begin{aligned} [\mathbf{\Gamma_v}(\omega)]_{ij} = [\mathbf{\Gamma}_\mathrm{d}(\omega)]_{ij} &= \frac{\sin[\omega(j-i)\tau_0]}{\omega(j-i)\tau_0} \\ &= \mathrm{sinc}[\omega(j-i)\tau_0]. \end{aligned} \tag{11}$$

  In this scenario, the SNR gain is called the DF:

  $$\mathcal{D}[\mathbf{h}(\omega)] = \frac{\left|\mathbf{h}^H(\omega)\mathbf{d}(\omega)\right|^2}{\mathbf{h}^H(\omega)\mathbf{\Gamma}_\mathrm{d}(\omega)\mathbf{h}(\omega)}. \tag{12}$$

  It can be shown that the maximum possible DF is $\mathcal{D}_{\max} = M^2$ [13], which is referred to as the supergain in the literature. This gain can be achieved but at the expense of white noise amplification particularly at low frequencies.

Beampattern, also called directivity pattern, describes the beamformer's sensitivity to a planewave impinging on the array from the direction θ. It is defined as

$$\mathcal{B}[\mathbf{h}(\omega), \theta] = \mathbf{d}^H(\omega, \theta)\mathbf{h}(\omega). \tag{13}$$

[1] This noise models the sensor self noise.

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

4. CONVENTIONAL SUPERDIRECTIVE BEAMFORMERS

The superdirective beamformer assumes that the noise is diffuse. Similar to the minimum variance distortionless response (MVDR) beamformer [6], [7], the superdirective beamformer is obtained by minimizing the residual noise, i.e., $\mathbf{h}^H(\omega)\mathbf{\Gamma}_\mathrm{d}(\omega)\mathbf{h}(\omega)$ subject to the distortionless constraint [eq. (6)], i.e.,

$$\min_{\mathbf{h}(\omega)} \mathbf{h}^H(\omega)\mathbf{\Gamma}_\mathrm{d}(\omega)\mathbf{h}(\omega) \text{ subject to } \mathbf{h}^H(\omega)\mathbf{d}(\omega) = 1. \tag{14}$$

The solution to this well-known problem is [3]

$$\mathbf{h}_\mathrm{S}(\omega) = \frac{\mathbf{\Gamma}_\mathrm{d}^{-1}(\omega)\mathbf{d}(\omega)}{\mathbf{d}^H(\omega)\mathbf{\Gamma}_\mathrm{d}^{-1}(\omega)\mathbf{d}(\omega)}. \tag{15}$$

Note that the superdirective beamformer is a fixed beamformer since it does not need to estimate the statistics of the signals. In fact, it can be shown that (15) is the hypercardioid of order $M-1$ [14].

It is well known that (15) is sensitive to sensor noise and array imperfections. In order to deal with this important problem, the authors in [3], [5] proposed to maximize the DF subject to a constraint on the WNG. Using the distortionless constraint, one can find the optimal solution as [3], [5]

$$\mathbf{h}_{\mathrm{R},\epsilon}(\omega) = \frac{[\mathbf{\Gamma}_\mathrm{d}(\omega) + \epsilon\mathbf{I}_M]^{-1}\mathbf{d}(\omega)}{\mathbf{d}^H(\omega)[\mathbf{\Gamma}_\mathrm{d}(\omega) + \epsilon\mathbf{I}_M]^{-1}\mathbf{d}(\omega)}, \tag{16}$$

where $\epsilon \geq 0$ is a Lagrange multiplier. It is clear that (16) is a regularized (or robust) version of (15), where ϵ is the regularization parameter. This parameter tries to find a good compromise between the DF and white noise amplification. A small ϵ leads to a large DF but a low WNG, while a large ϵ gives a large WNG but a low DF. Two interesting cases of (16) are $\mathbf{h}_{\mathrm{R},0}(\omega) = \mathbf{h}_\mathrm{S}(\omega)$ and $\mathbf{h}_{\mathrm{R},\infty}(\omega) = \mathbf{h}_\mathrm{DS}(\omega) = \mathbf{d}(\omega)/M$, which is the DS beamformer. It is, in general, very difficult to find an optimal value of this parameter.

5. JOINT DIAGONALIZATION

The joint diagonalization is going to be useful in the derivation of superdirective beamformers that can better compromise between white noise amplification and supergain. In the rest, it is assumed that we deal with the spherically isotropic noise, so that $\mathbf{\Gamma_v}(\omega) = \mathbf{\Gamma}_\mathrm{d}(\omega)$.

The correlation matrix of $\mathbf{x}(\omega)$ is $\mathbf{\Phi_x}(\omega) = \phi_X(\omega)\mathbf{d}(\omega)\mathbf{d}^H(\omega)$. Therefore, its pseudo-coherence matrix is

$$\mathbf{\Gamma_x}(\omega) = \frac{\mathbf{\Phi_x}(\omega)}{\phi_X(\omega)} = \mathbf{d}(\omega)\mathbf{d}^H(\omega), \tag{17}$$

which does not depend on $X(\omega)$.

The two Hermitian matrices $\mathbf{\Gamma_x}(\omega)$ and $\mathbf{\Gamma}_\mathrm{d}(\omega)$ can be jointly diagonalized as follows [16]:

$$\mathbf{B}^H(\omega)\mathbf{\Gamma_x}(\omega)\mathbf{B}(\omega) = \mathbf{\Lambda}(\omega), \tag{18}$$

$$\mathbf{B}^H(\omega)\mathbf{\Gamma}_\mathrm{d}(\omega)\mathbf{B}(\omega) = \mathbf{I}_M, \tag{19}$$

where

$$\mathbf{B}(\omega) = \begin{bmatrix} \mathbf{b}_1(\omega) & \mathbf{b}_2(\omega) & \cdots & \mathbf{b}_M(\omega) \end{bmatrix} \tag{20}$$

is a full-rank square matrix (of size $M \times M$),

$$\mathbf{b}_1(\omega) = \frac{\mathbf{\Gamma}_\mathrm{d}^{-1}(\omega)\mathbf{d}(\omega)}{\sqrt{\mathbf{d}^H(\omega)\mathbf{\Gamma}_\mathrm{d}^{-1}(\omega)\mathbf{d}(\omega)}} \tag{21}$$

is the first eigenvector of the matrix $\mathbf{\Gamma}_{\mathrm{d}}^{-1}(\omega)\mathbf{\Gamma}_{\mathbf{x}}(\omega)$,

$$\mathbf{\Lambda}(\omega) = \mathrm{diag}\left[\lambda_1(\omega), 0, \ldots, 0\right] \tag{22}$$

is a diagonal matrix (of size $M \times M$), and

$$\lambda_1(\omega) = \mathbf{d}^H(\omega)\mathbf{\Gamma}_{\mathrm{d}}^{-1}(\omega)\mathbf{d}(\omega) \tag{23}$$

is the only nonnull eigenvalue of $\mathbf{\Gamma}_{\mathrm{d}}^{-1}(\omega)\mathbf{\Gamma}_{\mathbf{x}}(\omega)$, whose corresponding eigenvector is $\mathbf{b}_1(\omega)$. It is important to observe that neither $\mathbf{B}(\omega)$ nor $\lambda_1(\omega)$ depend on the statistics of the signals. It can be checked from (18) that

$$\mathbf{b}_i^H(\omega)\mathbf{d}(\omega) = 0,\ i = 2, 3, \ldots, M. \tag{24}$$

6. SUBSPACE SUPERDIRECTIVE BEAMFORMERS

In this section, we show how to derive subspace superdirective beamformers thanks to the joint diagonalization technique.

Let us define the matrix of size $M \times N$:

$$\mathbf{B}_{1:N}(\omega) = \begin{bmatrix} \mathbf{b}_1(\omega) & \mathbf{b}_2(\omega) & \cdots & \mathbf{b}_N(\omega) \end{bmatrix}, \tag{25}$$

with $1 \leq N \leq M$. We consider beamformers that have the form:

$$\mathbf{h}_N(\omega) = \mathbf{B}_{1:N}(\omega)\mathbf{a}_{1:N}(\omega), \tag{26}$$

where

$$\mathbf{a}_{1:N}(\omega) = \begin{bmatrix} A_1(\omega) & A_2(\omega) & \cdots & A_N(\omega) \end{bmatrix}^T \neq \mathbf{0} \tag{27}$$

is a vector of length N. Substituting (26) into (4), we find that

$$\begin{aligned} Z(\omega) &= \mathbf{a}_{1:N}^H(\omega)\mathbf{B}_{1:N}^H(\omega)\mathbf{d}(\omega)X(\omega) \\ &\quad + \mathbf{a}_{1:N}^H(\omega)\mathbf{B}_{1:N}^H(\omega)\mathbf{v}(\omega) \\ &= A_1^*(\omega)\sqrt{\lambda_1(\omega)}X(\omega) + \mathbf{a}_{1:N}^H(\omega)\mathbf{B}_{1:N}^H(\omega)\mathbf{v}(\omega). \end{aligned} \tag{28}$$

Since the distortionless constraint is desired, it is clear from the previous expression that we always choose

$$A_1(\omega) = \frac{1}{\sqrt{\lambda_1(\omega)}}. \tag{29}$$

Now, we need to determine the other elements of $\mathbf{a}_{1:N}(\omega)$.

With the proposed beamformer, the WNG and the DF are, respectively,

$$\begin{aligned} \mathcal{W}\left[\mathbf{h}_N(\omega)\right] &= \frac{\left|\mathbf{h}_N^H(\omega)\mathbf{d}(\omega)\right|^2}{\mathbf{h}_N^H(\omega)\mathbf{h}_N(\omega)} \\ &= \frac{\left|\mathbf{a}_{1:N}^H(\omega)\mathbf{B}_{1:N}^H(\omega)\mathbf{d}(\omega)\right|^2}{\mathbf{a}_{1:N}^H(\omega)\mathbf{B}_{1:N}^H(\omega)\mathbf{B}_{1:N}(\omega)\mathbf{a}_{1:N}(\omega)} \end{aligned} \tag{30}$$

and

$$\begin{aligned} \mathcal{D}\left[\mathbf{h}_N(\omega)\right] &= \frac{\left|\mathbf{h}_N^H(\omega)\mathbf{d}(\omega)\right|^2}{\mathbf{h}_N^H(\omega)\mathbf{\Gamma}_{\mathrm{d}}(\omega)\mathbf{h}_N(\omega)} \\ &= \frac{\left|\mathbf{a}_{1:N}^H(\omega)\mathbf{B}_{1:N}^H(\omega)\mathbf{d}(\omega)\right|^2}{\mathbf{a}_{1:N}^H(\omega)\mathbf{B}_{1:N}^H(\omega)\mathbf{\Gamma}_{\mathrm{d}}(\omega)\mathbf{B}_{1:N}(\omega)\mathbf{a}_{1:N}(\omega)} \\ &= \frac{\left|\mathbf{a}_{1:N}^H(\omega)\mathbf{B}_{1:N}^H(\omega)\mathbf{d}(\omega)\right|^2}{\mathbf{a}_{1:N}^H(\omega)\mathbf{a}_{1:N}(\omega)}. \end{aligned} \tag{31}$$

The maximization of the DF or, equivalently, the minimization of $\mathbf{h}_N^H(\omega)\mathbf{\Gamma}_{\mathrm{d}}(\omega)\mathbf{h}_N(\omega)$ subject to $\mathbf{h}_N^H(\omega)\mathbf{d}(\omega) = 1$, leads to the conventional superdirective beamformer:

$$\begin{aligned} \mathbf{h}_{\mathrm{S}}(\omega) &= \frac{\mathbf{B}_{1:N}(\omega)\mathbf{B}_{1:N}^H(\omega)\mathbf{d}(\omega)}{\mathbf{d}^H(\omega)\mathbf{B}_{1:N}(\omega)\mathbf{B}_{1:N}^H(\omega)\mathbf{d}(\omega)} \\ &= \frac{\mathbf{b}_1(\omega)\mathbf{b}_1^H(\omega)\mathbf{d}(\omega)}{\left|\mathbf{b}_1^H(\omega)\mathbf{d}(\omega)\right|^2} = \frac{\mathbf{\Gamma}_{\mathrm{d}}^{-1}(\omega)\mathbf{d}(\omega)}{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\mathrm{d}}^{-1}(\omega)\mathbf{d}(\omega)}. \end{aligned} \tag{32}$$

The most interesting subspace beamformer is derived by maximizing the WNG. This is equivalent to minimizing $\mathbf{h}_N^H(\omega)\mathbf{h}_N(\omega)$ subject to $\mathbf{h}_N^H(\omega)\mathbf{d}(\omega) = 1$. We find

$$\mathbf{h}_N(\omega) = \frac{\mathbf{P}_{\mathbf{B}_{1:N}}(\omega)\mathbf{d}(\omega)}{\mathbf{d}^H(\omega)\mathbf{P}_{\mathbf{B}_{1:N}}(\omega)\mathbf{d}(\omega)}, \tag{33}$$

where

$$\mathbf{P}_{\mathbf{B}_{1:N}}(\omega) = \mathbf{B}_{1:N}(\omega)\left[\mathbf{B}_{1:N}^H(\omega)\mathbf{B}_{1:N}(\omega)\right]^{-1}\mathbf{B}_{1:N}^H(\omega). \tag{34}$$

For $N = 1$, we get

$$\mathbf{h}_1(\omega) = \frac{\mathbf{b}_1(\omega)}{\mathbf{d}^H(\omega)\mathbf{b}_1(\omega)} = \mathbf{h}_{\mathrm{S}}(\omega), \tag{35}$$

which is the conventional superdirective beamformer, and for $N = M$, we obtain

$$\mathbf{h}_M(\omega) = \frac{\mathbf{d}(\omega)}{\mathbf{d}^H(\omega)\mathbf{d}(\omega)} = \mathbf{h}_{\mathrm{DS}}(\omega), \tag{36}$$

which is the DS beamformer. Therefore, by playing with N, we obtain different beamformers whose performances are in between the performances of $\mathbf{h}_{\mathrm{S}}(\omega)$ and $\mathbf{h}_{\mathrm{DS}}(\omega)$.

With the proposed beamformer, the WNG is

$$\begin{aligned} \mathcal{W}\left[\mathbf{h}_N(\omega)\right] &= \mathbf{d}^H(\omega)\mathbf{P}_{\mathbf{B}_{1:N}}(\omega)\mathbf{d}(\omega) \\ &= \lambda_1(\omega)\mathbf{i}^T\left[\mathbf{B}_{1:N}^H(\omega)\mathbf{B}_{1:N}(\omega)\right]^{-1}\mathbf{i}, \end{aligned} \tag{37}$$

where $\mathbf{i}$ is the first column of the $N \times N$ identity matrix, $\mathbf{I}_N$, with

$$\mathcal{W}\left[\mathbf{h}_1(\omega)\right] = \frac{\left|\mathbf{b}_1^H(\omega)\mathbf{d}(\omega)\right|^2}{\mathbf{b}_1^H(\omega)\mathbf{b}_1(\omega)} = \frac{\lambda_1(\omega)}{\mathbf{b}_1^H(\omega)\mathbf{b}_1(\omega)} \leq M \tag{38}$$

and

$$\mathcal{W}\left[\mathbf{h}_M(\omega)\right] = M. \tag{39}$$

The DF is

$$\begin{aligned} \mathcal{D}\left[\mathbf{h}_N(\omega)\right] &= \frac{\left[\mathbf{d}^H(\omega)\mathbf{P}_{\mathbf{B}_{1:N}}(\omega)\mathbf{d}(\omega)\right]^2}{\mathbf{d}^H(\omega)\mathbf{P}_{\mathbf{B}_{1:N}}(\omega)\mathbf{\Gamma}_{\mathrm{d}}(\omega)\mathbf{P}_{\mathbf{B}_{1:N}}(\omega)\mathbf{d}(\omega)} \\ &= \lambda_1(\omega)\frac{\left\{\mathbf{i}^T\left[\mathbf{B}_{1:N}^H(\omega)\mathbf{B}_{1:N}(\omega)\right]^{-1}\mathbf{i}\right\}^2}{\mathbf{i}^T\left[\mathbf{B}_{1:N}^H(\omega)\mathbf{B}_{1:N}(\omega)\right]^{-2}\mathbf{i}}, \end{aligned} \tag{40}$$

with

$$\mathcal{D}\left[\mathbf{h}_1(\omega)\right] = \lambda_1(\omega) \leq M^2 \tag{41}$$

and

$$\begin{aligned} \mathcal{D}\left[\mathbf{h}_M(\omega)\right] &= \frac{\left[\mathbf{d}^H(\omega)\mathbf{d}(\omega)\right]^2}{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\mathrm{d}}(\omega)\mathbf{d}(\omega)} \\ &= \frac{M^2}{\mathbf{d}^H(\omega)\mathbf{\Gamma}_{\mathrm{d}}(\omega)\mathbf{d}(\omega)} \geq 1. \end{aligned} \tag{42}$$

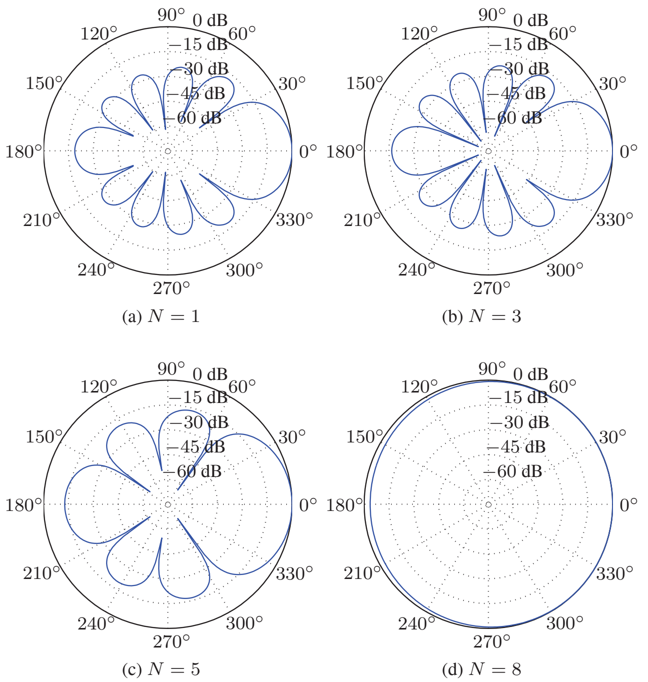

Fig. 1. Beampatterns of the subspace superdirective beamformer with different values of N at $f = 1$ kHz. $\delta = 1$ cm and $M = 8$.

We also deduce an interesting relationship between the WNG and the DF:

$$\frac{\mathcal{D}\left[\mathbf{h}_N(\omega)\right]}{\mathcal{W}\left[\mathbf{h}_N(\omega)\right]} = \frac{\mathbf{i}^T\left[\mathbf{B}_{1:N}^H(\omega)\mathbf{B}_{1:N}(\omega)\right]^{-1}\mathbf{i}}{\mathbf{i}^T\left[\mathbf{B}_{1:N}^H(\omega)\mathbf{B}_{1:N}(\omega)\right]^{-2}\mathbf{i}}, \tag{43}$$

where

$$\frac{1}{M} \leq \frac{\mathcal{D}\left[\mathbf{h}_N(\omega)\right]}{\mathcal{W}\left[\mathbf{h}_N(\omega)\right]} < \infty. \tag{44}$$

We should always have

$$M^2 \geq \mathcal{D}\left[\mathbf{h}_1(\omega)\right] \geq \mathcal{D}\left[\mathbf{h}_2(\omega)\right] \geq \cdots \geq \mathcal{D}\left[\mathbf{h}_M(\omega)\right] \tag{45}$$

and

$$M = \mathcal{W}\left[\mathbf{h}_M(\omega)\right] \geq \mathcal{W}\left[\mathbf{h}_{M-1}(\omega)\right] \geq \cdots \geq \mathcal{W}\left[\mathbf{h}_1(\omega)\right]. \tag{46}$$

Clearly, the beamformer $\mathbf{h}_N(\omega)$ is able to control white noise amplification while giving a reasonably good DF.

7. SIMULATIONS

In this section, we briefly study the performance of the subspace superdirective beamformer given in (33) through simulations. We use a uniform linear microphone array consisting of eight closely spaced microphones, with $\delta = 1$ cm. The beampatterns, DFs, and WNGs are plotted in Figs. 1 and 2. When $N = 1$, we get the conventional superdirective beamformer, which has a high DF but suffers from significant white noise amplification as seen in Fig. 2. When $N = M$, we get the DS beamformer, which gives the maximum WNG but its DF is low. As seen, the WNG increases with N while the DF decreases with N. So, a tradeoff between a high DF and good robustness can be obtained by choosing a proper value of N.

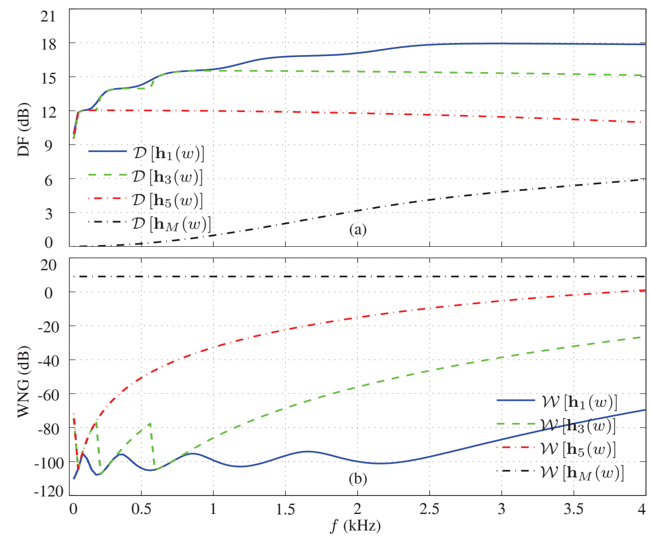

Fig. 2. SNR gains of the subspace superdirective beamformer: (a) DF and (b) WNG. $\delta = 1$ cm and $M = 8$.

8. CONCLUSIONS

In this paper, we studied the problem of robust superdirective beamforming by jointly diagonalizing the desired and noise signals correlation matrices. Unlike the regularized superdirective beamformer, which introduces a constraint on the WNG to improve the robustness of the beamformer against the microphone sensor noise, we developed a subspace superdirective beamformer. By properly choosing the dimension of the subspace, the developed beamformer can find a good compromise between a high DF and a low WNG. Simulation results demonstrated the good property of this subspace superdirective beamformer.

9. RELATION TO PRIOR WORK

Microphone arrays are widely used in many speech communication applications. The most important component of a microphone array system is beamforming, which plays a critical role on the array performance in desired signal estimation and noise suppression. Various beamforming algorithms have been developed over the past several decades [1–20]. Among those, the superdirective beamforming method has been intensively studied for its ability to achieve a high DF [8–13]. However, this beamformer is very sensitive to sensor self noise and other array imperfections. How to deal with the robustness is always the utmost issue in the design of a superdirective beamformer. Many efforts have been devoted to circumvent this issue. The most popular approach so far is the so-called regularized superdirective beamformer [3, 9, 10, 12], in which the WNG is controlled by a regularization parameter. The value of this parameter has to be carefully chosen. On the one hand, there is not much improvement in the WNG if the value is too small and, on the other hand, the DF of the beamformer may suffer significant degradation and the beampattern becomes more frequency dependent if the regularization parameter is too large. In practice, it is not easy to find the proper value of this parameter that fits to different application scenarios. In this paper, we developed a subspace superdirective beamformer based on the joint diagonalization, which provides an easy way to control the tradeoff between a high DF and robustness of the beamformer.

10. REFERENCES

[1] S. A. Schelkunoff, "A mathematical theory of linear arrays," *Bell Syst. Tech. J.*, vol. 22, pp. 80–107, Jan. 1943.

[2] J. L. Flanagan, J. D. Johnston, R. Zahn, and G. W. Elko, "Computer-steered microphone arrays for sound transduction in large rooms," *J. Acoust. Soc. Amer.*, vol. 78, pp. 1508–1518, Nov. 1985.

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

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

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

[6] J. Capon, "High resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, pp. 1408–1418, Aug. 1969.

[7] R. T. Lacoss, "Data adaptive spectral analysis methods," *Geophysics*, vol. 36, pp. 661–675, Aug. 1971.

[8] 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.

[9] S. Doclo and M. Moonen, "Superdirective beamforming robust against microphone mismatch," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 15, pp. 617–631, Feb. 2007.

[10] M. Crocco and A. Trucco, "Design of robust superdirective arrays with a tunable tradeoff between directivity and frequency-invariance," *IEEE Trans. Sig. Process.*, vol. 59, pp. 2169–2181, May 2011.

[11] G. W. Elko, "Superdirectional microphone arrays," in *Acoustic Signal Processing for Telecommunication*, S. L. Gay and J. Benesty, Eds. Boston, MA: Kluwer Academic Publishers, 2000, Chapter 10, pp. 181–237.

[12] R. Berkun, I. Cohen, and J. Benesty, "Combined beamformers for robust broadband regularized superdirective beamforming," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 23, pp. 877–886, May 2015.

[13] 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.

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

[15] J. Benesty, J. Chen, and E. Habets, *Speech Enhancement in the STFT Domain*. Springer Briefs in Electrical and Computer Engineering, 2011.

[16] J. N. Franklin, *Matrix Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1968.

[17] W. Liao, M. Hong, I. Merks, T. Zhang, and Z. Luo, "Incorporating spatial information in binaural beamforming for noise suppression in hearing aids," in *Proc. IEEE ICASSP*, 2015, pp. 5733–5737.

[18] E. Mabande, M. Buerger, and W. Kellermann, "Design of robust polynomial beamformers for symmetric arrays," in *Proc. IEEE ICASSP*, 2012, pp. 1–4.

[19] M. Zhang, F. Rubio, D.P. Palomar, and X. Mestre, "Robust adaptive beamforming with imprecise steering vector and noise covariance matrix due to finite sample size," in *Proc. IEEE ICASSP*, 2013, pp. 3786–3790.

[20] L. Bianchi, R. Magalotti, F. Antonacci, A. Sarti, and S. Tubaro, "Robust beamforming under uncertainties in the loudspeakers directivity pattern," in *Proc. IEEE ICASSP*, 2014, pp. 4448–4452.