

On Robust and High Directive Beamforming With Small-Spacing Microphone Arrays for Scattered Sources

Xianghui Wang¹, Israel Cohen², *Fellow, IEEE*, Jingdong Chen¹, *Senior Member, IEEE*, and Jacob Benesty³

Abstract—This paper is devoted to beamforming with small-spacing microphone arrays for processing broadband and scattered acoustic sources. It presents a maximum diffuse noise gain (MDNG) beamformer in this context using the joint diagonalization technique, which is effective in suppressing diffuse and directional noise, but at a price of low white noise gain (WNG). We also introduce a maximum WNG (MWNG) beamformer, which is robust to the array imperfections, but paying a price of sacrificing the diffuse noise gain (DNG). To make a tradeoff between WNG and DNG so that the beamformer, on the one hand, can achieve high directivity and, on the other hand, is robust to implement, we propose a generalized MDNG beamformer, which includes both the MDNG and MWNG beamformers as particular cases. Simulations are conducted to illustrate the properties and advantages of the proposed beamformers.

Index Terms—Microphone array, scattered source, frequency-invariant beamformer, diffuse noise gain, white noise gain.

I. INTRODUCTION

MANY voice communication systems and devices, such as Voice over Internet Protocol (VoIP), hearing aids, teleconferencing and mobile phones, often work in challenging acoustic environments where the speech signal of interest is inevitably contaminated by background noise, reverberation, competing sources, and interferences. In order to deal

with sound signal acquisition in such environments, many processing techniques have been proposed over the last several decades [1]–[3], among which beamforming with microphone arrays has attracted tremendous amount of interest. A number of beamforming algorithms have been developed in the literature and the representative ones include the delay-and-sum (DS) beamformer [4]–[6], the superdirective (SD) beamformer [7]–[10], and differential microphone arrays (DMAs) [11]–[19], etc.

The DS beamforming structure, which can achieve maximum white noise gain (WNG), has been widely studied in the literature [4]. But this beamformer has frequency-dependent directivity patterns [20], which makes it unsuitable for processing broadband signals such as speech. Another drawback of the DS beamformer is its low directivity factor (DF), which limits its performance in suppressing directional noise and reverberation. To overcome these drawbacks, the SD and differential (the corresponding arrays are referred to as DMAs) beamformers have been proposed and widely studied as they exhibit frequency-invariant beampatterns and high DFs [7], [10], [12], [14], [21], [23]–[25], which have now been increasingly used in a wide range of applications.

The aforementioned beamformers were developed and studied in the context of point sources. However, in practical acoustic applications the point source model [26]–[31] is often not accurate. First of all, as small-aperture microphone arrays are used, the size of the sound source of interest, e.g., a big loudspeaker or a car engine, can be relatively large, which cannot be treated as a point source. Secondly, there may be some spatial uncertainties about the location of the source of interest, e.g., the talker moves slowly in a particular area or there are multiple talkers in the region. In these scenarios, a scattered source model is more useful than the point source model. Furthermore, the direction-of-arrival (DOA) information is generally needed in beamforming. In practice, there often exists some uncertainty about the DOA of the source of interest due to reasons such as errors in DOA estimation and tracking. In this case, considering a scattered source model can be helpful to formulate and deal with the problem. Therefore, in this paper, we study the problem of microphone array beamforming with a scattered source model in which the source of interest is assumed to be scattered in an angular range [32]–[38]. We consider three cases: the source is coherently scattered (CS), incoherently scattered (IS),

Manuscript received July 9, 2018; revised December 15, 2018; accepted February 2, 2019. Date of publication February 14, 2019; date of current version March 13, 2019. This work was supported in part by the Israel Science Foundation under Grant 576/16, in part by the National Science Foundation of China under Grant 61831019 and Grant 61425005, and in part by the ISF-NSFC joint research program under Grant 2514/17 and Grant 61761146001. The work of X. Wang was supported in part by the China Scholarship Council. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Andy W. H. Khong. (*Corresponding author: Jingdong Chen.*)

X. Wang is with the Center of Intelligent Acoustics and Immersive Communications and School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an 710072, China, and also with the Andrew and Erna Viterby Faculty of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel (e-mail: wangxh1985@126.com).

I. Cohen is with the Andrew and Erna Viterby Faculty of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel (e-mail: icohen@ee.technion.ac.il).

J. Chen is with the Center of Intelligent Acoustics and Immersive Communications, Northwestern Polytechnical University, Xi'an 710072, China (e-mail: jingdongchen@ieee.org).

J. Benesty is with the INRS-EMT, University of Quebec, Montreal, QC H5A 1K6, Canada (e-mail: benesty@emt.inrs.ca).

Digital Object Identifier 10.1109/TASLP.2019.2899517

and scattered in a partially coherent and partially incoherent manner.

Based on the scattered source model, we develop three beamformers. The first one is the maximum-diffuse-noise-gain (MDNG) beamformer, which is deduced using the joint diagonalization technique. Its beampattern is almost frequency invariant when it is applied to small spacing microphone arrays. This beamformer achieves the maximum DNG but at a price of white noise amplification, which is the same as the SD beamformer developed with the point source model. The second one is the so-called maximum-white-noise-gain (MWNG) beamformer. This beamformer, as its name indicates, achieves the maximum WNG, but at a cost of sacrificing the DNG. The third one is a generalized MDNG (GMDNG) beamformer, which can make a compromise between WNG and DNG; and the level of compromise can be controlled by adjusting a parameter related to the dimension of the desired signal subspace. When the associated parameters are properly chosen, this beamformer can be made robust to the array imperfections and effective in suppressing diffuse and directional noise. Simulation results illustrate the properties and advantages of the proposed beamformers.

The remainder of the paper is as follows. The signal model for scattered acoustic sources and problem formulation of beamforming are presented in Section II. Section III defines some useful performance measures. The MDNG and MWNG beamformers are derived in Section IV. Section V deduces the GMDNG beamformer. Simulation results are presented in Section VI. Finally, important conclusions are drawn in Section VII.

II. SIGNAL MODEL AND PROBLEM FORMULATION

Before discussing the signal model for scattered sources, let us first give the farfield, point source model. Consider a point source radiating planar waves, which propagate in an anechoic acoustic environment at the speed of sound $c = 340$ m/s. A uniform linear microphone array (ULMA) with M omnidirectional microphones is used to acquire the signal, where the spacing between two neighboring sensors is δ . We denote by θ the steering angle and, without loss of generality, we choose the first sensor as the reference. The steering vector can then be written as [20], [39], [40]

$$\mathbf{d}(\omega, \theta) \triangleq [1 e^{-j\omega\tau_0 \cos \theta} \dots e^{-j(M-1)\omega\tau_0 \cos \theta}]^T, \quad (1)$$

where $\omega = 2\pi f$ is the angular frequency, $f > 0$ is the temporal frequency, $\tau_0 = \delta/c$ is the time delay between two neighboring sensors when the source direction is $\theta = 0^\circ$, j is the imaginary unit, and $(\cdot)^T$ is the transpose operator. In this work, we focus on developing beamformers with small spacing ULMA, i.e., δ is much smaller than the smallest wavelength of the broadband signals of interest in order to avoid spatial aliasing and achieve high array gains and frequency-independent beampatterns [21], like superdirective [7], [23] and differential [21], [41] beamformers.

With a ULMA, differential and superdirective beamformers generally achieve the largest array gain at the endfire direction [42]. As a result, most study and analysis of those beamformers focus on the scenarios where the signal of interest (the desired

signal) impinges on the array from the endfire direction, i.e., $\theta = 0^\circ$. In this case, the signals observed by the microphone array are expressed as

$$\begin{aligned} \mathbf{y}(\omega) &= \mathbf{d}(\omega, 0^\circ)X(\omega) + \mathbf{v}(\omega) \\ &= \mathbf{x}(\omega) + \mathbf{v}(\omega), \end{aligned} \quad (2)$$

where $X(\omega)$ is the desired signal,

$$\mathbf{y}(\omega) = [Y_1(\omega) Y_2(\omega) \dots Y_M(\omega)]^T, \quad (3)$$

is the noisy signal vector, $\mathbf{x}(\omega)$ and $\mathbf{v}(\omega)$ are, respectively, the clean signal and additive noise vectors, which are defined in an analogous way to $\mathbf{y}(\omega)$ and are assumed to be uncorrelated with each other.

For a non-point but scattered source, the signal model in (2) is no longer valid. For simplicity, let us consider the case where the microphone array and the scattered source are in the same (assumed to be the horizontal) plane. In real situations, the source is generally scattered in an area, which is a function of not only the angle θ but also the distance. However, we focus on the farfield scenario with small-spacing microphone arrays in this work and we assume that the source is only angularly scattered. Let us denote by $p(\omega, \theta)$ the signal component from θ at frequency ω of the scattered source. By neglecting the propagation loss, we can write the array observation signals as

$$\begin{aligned} \mathbf{y}(\omega) &= \int_{-\pi}^{\pi} p(\omega, \theta)\mathbf{d}(\omega, \theta)d\theta + \mathbf{v}(\omega) \\ &= \mathbf{x}(\omega) + \mathbf{v}(\omega), \end{aligned} \quad (4)$$

where $\mathbf{x}(\omega) \triangleq \int_{-\pi}^{\pi} p(\omega, \theta)\mathbf{d}(\omega, \theta)d\theta$. Again, it is assumed that the signal $\mathbf{x}(\omega)$ and the noise $\mathbf{v}(\omega)$ are uncorrelated. The correlation matrix of $\mathbf{y}(\omega)$ is then

$$\begin{aligned} \Phi_{\mathbf{y}}(\omega) &\triangleq E[\mathbf{y}(\omega)\mathbf{y}^H(\omega)] = \Phi_{\mathbf{x}}(\omega) + \Phi_{\mathbf{v}}(\omega) \\ &= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \rho(\omega, \theta, \theta')\mathbf{d}(\omega, \theta)\mathbf{d}^H(\omega, \theta')d\theta d\theta' + \Phi_{\mathbf{v}}(\omega), \end{aligned} \quad (5)$$

where $\rho(\omega, \theta, \theta') = E[p(\omega, \theta)p^*(\omega, \theta')]$ is the source angular cross-correlation kernel, superscripts $*$ and H , denote, respectively, the complex-conjugate and conjugate-transpose operators, $E[\cdot]$ denotes the mathematical expectation, and $\Phi_{\mathbf{x}}(\omega)$ and $\Phi_{\mathbf{v}}(\omega)$ are defined in an analogous way to $\Phi_{\mathbf{y}}(\omega)$.

There are three basic classes of scattered sources: i.e., CS, IS, and a combination of partially CS and partially IS.

A. Coherently Scattered Sources

If $p(\omega, \theta)$ can be decomposed in the following form [26], [28]:

$$p(\omega, \theta) = S(\omega)g(\omega, \theta), \quad (6)$$

where $S(\omega)$ is a source signal independent of θ and $g(\omega, \theta)$ is a deterministic function that describes the spatial distribution characteristics of the source, the source is deemed as coherently scattered and we call $g(\omega, \theta)$ the signal angular density function.

It follows immediately that

$$\mathbf{x}(\omega) = \int_{-\pi}^{\pi} \mathbf{d}(\omega, \theta) S(\omega) g(\omega, \theta) d\theta. \quad (7)$$

Now, comparing the signal model defined in (2) and the model in (7), we have

$$X_1(\omega) = X(\omega) = S(\omega) \int_{-\pi}^{\pi} g(\omega, \theta) d\theta. \quad (8)$$

Then, the correlation matrix of $\mathbf{x}(\omega)$ is written as

$$\Phi_{\mathbf{x}}(\omega) = \sigma_X^2(\omega) \mathbf{b}(\omega) \mathbf{b}^H(\omega), \quad (9)$$

where $\sigma_X^2(\omega) = E[|X(\omega)|^2] = E[|X_1(\omega)|^2]$ and

$$\mathbf{b}(\omega) = \frac{S(\omega)}{X(\omega)} \int_{-\pi}^{\pi} \mathbf{d}(\omega, \theta) g(\omega, \theta) d\theta. \quad (10)$$

It is seen that $\Phi_{\mathbf{x}}(\omega)$ is a rank-1 matrix.

B. Incoherently Scattered Sources

If the source angular cross-correlation kernel satisfies

$$\rho(\omega, \theta, \theta') \neq 0, \quad \text{iff } \theta = \theta', \quad (11)$$

the source is deemed as incoherently scattered. In this case, the correlation matrix of the signal $\mathbf{x}(\omega)$ can be written as

$$\Phi_{\mathbf{x}}(\omega) = \int_{-\pi}^{\pi} \rho(\omega, \theta, \theta) \mathbf{d}(\omega, \theta) \mathbf{d}^H(\omega, \theta) d\theta. \quad (12)$$

In comparison with the CS case, the rank of $\Phi_{\mathbf{x}}(\omega)$ can be greater than 1 or even equal to M , i.e., full rank. The variance of the signal of interest at the reference sensor (Sensor 1 in this work) is then

$$\sigma_X^2(\omega) = E[|X(\omega)|^2] = \int_{-\pi}^{\pi} \rho(\omega, \theta, \theta) d\theta. \quad (13)$$

Consider a particular case where the scattered source is composed of several point sources that are distributed at θ_k , $k = 1, 2, \dots, K$ and those point sources are mutually incoherent, the source angular cross-correlation kernel becomes

$$\rho(\omega, \theta, \theta_{1:K}) = \sum_{k=1}^K \sigma_{S_k}^2(\omega) \delta(\theta - \theta_k), \quad (14)$$

where $\sigma_{S_k}^2(\omega)$ and θ_k are the variance and DOA of the k th point source, respectively, K is the total number of the point sources in the scattered region of interest, and $\delta(\cdot)$ is the Dirac delta function. In this case, we have

$$\sum_{k=1}^K \sigma_{S_k}^2(\omega) = \sigma_X^2(\omega). \quad (15)$$

Then, the signal correlation matrix can be rewritten as

$$\Phi_{\mathbf{x}}(\omega) = \mathbf{D}(\omega, \boldsymbol{\theta}) \Sigma_S(\omega) \mathbf{D}^H(\omega, \boldsymbol{\theta}), \quad (16)$$

where

$$\boldsymbol{\theta} = [\theta_1 \ \theta_2 \ \dots \ \theta_K]^T, \quad (17)$$

$$\mathbf{D}(\omega, \boldsymbol{\theta}) = [\mathbf{d}(\omega, \theta_1) \ \mathbf{d}(\omega, \theta_2) \ \dots \ \mathbf{d}(\omega, \theta_K)], \quad (18)$$

and

$$\Sigma_S(\omega) = \text{diag} [\sigma_{S_1}^2(\omega) \ \sigma_{S_2}^2(\omega) \ \dots \ \sigma_{S_K}^2(\omega)] \quad (19)$$

is a diagonal matrix.

When the source of interest is partially IS and partially CS, the signal correlation matrix is then a linear combination of the matrix in (9) and that in (12), details of which will not be presented here for brevity.

As seen, if $p(\omega, \theta) = X(\omega) \delta(\theta)$ (with the assumption that the incidence angle of the source of interest is from 0°), both the CS and IS models degenerate to the point source model. For ease of exposition, we will drop ω from the notation from now on wherever there is no ambiguity.

Given the signal model, the problem of beamforming can be described as one of estimating the desired signal, X , given the noisy signal vector \mathbf{y} . This is generally achieved by passing \mathbf{y} through a beamforming filter [21], \mathbf{h} , i.e.,

$$Z = \mathbf{h}^H \mathbf{x} + \mathbf{h}^H \mathbf{v}, \quad (20)$$

where Z is the estimate of X and

$$\mathbf{h} = [H_1 \ H_2 \ \dots \ H_M]^T \quad (21)$$

is the beamformer of length M . In the rest, the power conservation constraint of the desired signal is applied, i.e.,

$$\frac{1}{\sigma_X^2} \mathbf{h}^H \Phi_{\mathbf{x}} \mathbf{h} = 1. \quad (22)$$

III. PERFORMANCE MEASURES

To evaluate the beamforming performance, we adopt the following measures. The first one is the so-called WNG, which evaluates the sensitivity of the beamformer to the array imperfections, such as the sensors' self-noise, mismatch among microphone sensors, imprecise microphone positions, etc. In our context, the WNG can be written as [21]

$$\mathcal{W}(\mathbf{h}) = \frac{\mathbf{h}^H \Gamma_{\mathbf{x}} \mathbf{h}}{\mathbf{h}^H \mathbf{h}}, \quad (23)$$

where $\Gamma_{\mathbf{x}} = \frac{\Phi_{\mathbf{x}}}{\sigma_X^2}$ is the pseudo-coherence matrix of the desired signal vector, \mathbf{x} . Note that $\mathcal{W}(\mathbf{h}) < 1$ (0 dB) indicates that there is white noise amplification. But with today's system design technology and quality of microphone sensors, some amount of white noise amplification, e.g., 10 dB to 15 dB, is tolerable.

The second performance measure adopted in this work is the DNG, which, as its name indicates, evaluate the signal-to-noise-ratio (SNR) gain in a diffuse noise field. It is written as [21], [22]

$$\mathcal{D}(\mathbf{h}) = \frac{\mathbf{h}^H \Gamma_{\mathbf{x}} \mathbf{h}}{\mathbf{h}^H \Gamma_{\mathbf{d}} \mathbf{h}}, \quad (24)$$

where

$$(\Gamma_{\mathbf{d}})_{ij} = J_0[\omega(j-i)/c] \quad (25)$$

represents the noise pseudo-coherence matrix in cylindrically isotropic noise field and $J_0(\cdot)$ is the zero-order Bessel function of the first kind. If the source of interest is a point source, $\mathcal{D}(\mathbf{h})$

degrades to the conventional DF definition [12], [14]. It should be noted that in the literature, DF is often defined using the spherically isotropic noise field [7], [9]. But in this work, it is assumed that the microphone array and the scattered source are in the same plane with the elevation angle being 0° , so we use the cylindrically isotropic noise field in the definition of DNG.

Let us denote by $\mathcal{B}(\mathbf{h}, \theta)$ the beampattern [14], [21]. For an IS source with its power uniformly distributed in the range $[-\theta_0, \theta_0]$, i.e.,

$$\rho(\omega, \theta, \theta) = \begin{cases} \frac{\sigma_x^2(\omega)}{2\theta_0} & \text{if } \theta \in [-\theta_0, \theta_0] \\ 0 & \text{otherwise} \end{cases}, \quad (26)$$

DNG in (24) can be rewritten as

$$\mathcal{D}(\mathbf{h}) = \frac{\frac{1}{2\theta_0} \int_{-\theta_0}^{\theta_0} |\mathcal{B}(\mathbf{h}, \theta)|^2 d\theta}{\frac{1}{2\pi} \int_{-\pi}^{\pi} |\mathcal{B}(\mathbf{h}, \theta)|^2 d\theta}, \quad (27)$$

which is closely related to the front-to-back ratio [12], [14] defined as

$$\mathcal{R}_{\text{FB}}(\mathbf{h}) = \frac{\int_0^{\pi/2} |\mathcal{B}(\mathbf{h}, \theta)|^2 d\theta}{\int_{\pi/2}^{\pi} |\mathcal{B}(\mathbf{h}, \theta)|^2 d\theta} \quad (28)$$

in the cylindrically isotropic noise field.

Generally, the maximum WNG and maximum DNG cannot be achieved at the same time. At low frequencies, WNG and DNG are a pair of contradictory measures, i.e., a beamformer with high WNG always has low DNG, and vice versa. Therefore, how to make a compromise between the levels of DNG and WNG is an important and challenging problem [23]–[25].

Besides WNG and DNG, beampattern will also be used as performance measure, which will become clear in Section VI.

IV. MAXIMUM DNG AND MAXIMUM WNG BEAMFORMERS

In this section, we first derive the maximum DNG (MDNG) beamformer under the power conservation constraint. Mathematically, the problem can be formulated as

$$\mathbf{h}_{\text{MDNG}} = \arg \max_{\mathbf{h}} \frac{\mathbf{h}^H \boldsymbol{\Gamma}_x \mathbf{h}}{\mathbf{h}^H \boldsymbol{\Gamma}_d \mathbf{h}} \quad \text{s.t. } \mathbf{h}^H \boldsymbol{\Gamma}_x \mathbf{h} = 1. \quad (29)$$

Let us diagonalize the two Hermitian pseudo-coherence matrices, $\boldsymbol{\Gamma}_x$ and $\boldsymbol{\Gamma}_d$, with the joint diagonalization method [43], i.e.,

$$\mathbf{P}^H \boldsymbol{\Gamma}_x \mathbf{P} = \boldsymbol{\Lambda}, \quad (30)$$

$$\mathbf{P}^H \boldsymbol{\Gamma}_d \mathbf{P} = \mathbf{I}_M, \quad (31)$$

where

$$\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_M) \quad (32)$$

and

$$\mathbf{P} = [\mathbf{p}_1 \ \mathbf{p}_2 \ \dots \ \mathbf{p}_M] \quad (33)$$

are the eigenvalue and eigenvector matrices of $\boldsymbol{\Gamma}_d^{-1} \boldsymbol{\Gamma}_x$, respectively. Here, the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_M$ are organized in a descending order, i.e., $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_M \geq 0$. Note that the

eigenvector matrix \mathbf{P} is full rank but not necessarily orthogonal. It is obvious that we have

$$\lambda_M \leq \mathcal{D}(\mathbf{h}) \leq \lambda_1, \quad \forall \mathbf{h}. \quad (34)$$

Now, we consider the beamformer \mathbf{h} of the form:

$$\mathbf{h} = \alpha \mathbf{p}_1, \quad (35)$$

where $\alpha \neq 0$ is an arbitrary scalar. Clearly, this beamformer achieves the maximum DNG for all α . Applying the power conservation constraint, we obtain

$$\alpha = \frac{1}{\sqrt{\mathbf{p}_1^H \boldsymbol{\Gamma}_x \mathbf{p}_1}}. \quad (36)$$

Substituting (36) into (35), we get the MDNG beamformer:

$$\mathbf{h}_{\text{MDNG}} = \frac{1}{\sqrt{\mathbf{p}_1^H \boldsymbol{\Gamma}_x \mathbf{p}_1}} \mathbf{p}_1. \quad (37)$$

We have the following two special cases:

- If the source of interest is a point source, \mathbf{h}_{MDNG} degenerates to the traditional superdirective beamformer [21];
- For the coherently scattered source with a pseudo-coherence matrix of $\boldsymbol{\Gamma}_x = \mathbf{b}\mathbf{b}^H$, the beamformer \mathbf{h}_{MDNG} becomes

$$\mathbf{h}_{\text{MDNG}} = \frac{\boldsymbol{\Gamma}_d^{-1} \mathbf{b}}{\mathbf{b}^H \boldsymbol{\Gamma}_d^{-1} \mathbf{b}} = \frac{\mathbf{p}_1}{\mathbf{b}^H \mathbf{p}_1}. \quad (38)$$

The MDNG beamformer based on the eigenvector corresponding to the maximum eigenvalue of the matrix $\boldsymbol{\Gamma}_d^{-1} \boldsymbol{\Gamma}_x$ maximizes the DNG; but it has a very low WNG in our context. So, this beamformer is sensitive to the array imperfections.

Another interesting beamformer, named the maximum WNG (MWNG) beamformer, can be derived as follows:

$$\mathbf{h}_{\text{MWNG}} = \arg \min_{\mathbf{h}} \mathbf{h}^H \mathbf{h} \quad \text{s.t. } \mathbf{h}^H \boldsymbol{\Gamma}_x \mathbf{h} = 1, \quad (39)$$

the solution of which is

$$\mathbf{h}_{\text{MWNG}} = \frac{1}{\sqrt{\mathcal{P}^H\{\boldsymbol{\Gamma}_x\} \boldsymbol{\Gamma}_x \mathcal{P}\{\boldsymbol{\Gamma}_x\}}} \mathcal{P}\{\boldsymbol{\Gamma}_x\}, \quad (40)$$

where $\mathcal{P}\{\cdot\}$ is the operator producing the principal eigenvector of a matrix, i.e., the eigenvector corresponding to the maximum eigenvalue.

For this beamformer, we also have two special cases as follows:

- If the source of interest is a point source, the beamformer \mathbf{h}_{MWNG} degenerates to the DS beamformer;
- For the coherently scattered source with a pseudo-coherence matrix of $\boldsymbol{\Gamma}_x = \mathbf{b}\mathbf{b}^H$, the beamformer \mathbf{h}_{MWNG} degenerates to the following form:

$$\mathbf{h}_{\text{MWNG}} = \frac{\mathbf{b}}{\mathbf{b}^H \mathbf{b}}. \quad (41)$$

The MWNG beamformer, which is based on the eigenvector corresponding to the maximum eigenvalue of the matrix $\boldsymbol{\Gamma}_x$, maximizes the WNG; but its DNG is low. This means that the MWNG beamformer is robust to the array imperfections, but it is not effective in suppressing diffuse and directional noise.

V. GENERALIZED MAXIMUM DNG BEAMFORMER

The MDNG beamformer can achieve maximum DNG, but it suffers from serious white noise amplification, especially at low frequencies. Therefore, it is important to find a beamformer that can achieve a good level of DNG with a reasonable level of WNG. To achieve this objective, let us consider beamformers of the following form:

$$\mathbf{h} = \mathbf{P}_Q \boldsymbol{\alpha}, \quad (42)$$

where

$$\mathbf{P}_Q = [\mathbf{p}_1 \ \mathbf{p}_2 \ \cdots \ \mathbf{p}_Q]. \quad (43)$$

Note that the eigenvector matrix \mathbf{P}_Q , of size $M \times Q$ ($1 \leq Q \leq M$), consists of the eigenvectors corresponding to the largest Q eigenvalues of $\Gamma_d^{-1} \Gamma_x$, and

$$\boldsymbol{\alpha} = [\alpha_1 \ \alpha_2 \ \cdots \ \alpha_Q]^T \neq \mathbf{0} \quad (44)$$

is an arbitrary vector of length Q with complex values. Here, one can view Q as the dimension of the desired signal subspace [44]–[46].

Now, let us consider the following problem:

$$\min_{\mathbf{h}} \mathbf{h}^H \mathbf{h} \quad \text{s.t.} \quad \mathbf{h}^H \Gamma_x \mathbf{h} = 1. \quad (45)$$

Using (23) and (42), we can rewrite (45) as

$$\max_{\boldsymbol{\alpha}} \frac{\boldsymbol{\alpha}^H \mathbf{P}_Q^H \Gamma_x \mathbf{P}_Q \boldsymbol{\alpha}}{\boldsymbol{\alpha}^H \mathbf{P}_Q^H \mathbf{P}_Q \boldsymbol{\alpha}} \quad \text{s.t.} \quad \boldsymbol{\alpha}^H \mathbf{P}_Q^H \Gamma_x \mathbf{P}_Q \boldsymbol{\alpha} = 1. \quad (46)$$

It can be checked that the following $\boldsymbol{\alpha}$ maximizes the ratio $\frac{\boldsymbol{\alpha}^H \mathbf{P}_Q^H \Gamma_x \mathbf{P}_Q \boldsymbol{\alpha}}{\boldsymbol{\alpha}^H \mathbf{P}_Q^H \mathbf{P}_Q \boldsymbol{\alpha}}$:

$$\begin{aligned} \boldsymbol{\alpha} &= \beta_1 \mathcal{P}\{(\mathbf{P}_Q^H \mathbf{P}_Q)^{-1} (\mathbf{P}_Q^H \Gamma_x \mathbf{P}_Q)\} \\ &= \beta_1 \mathcal{P}\{(\mathbf{P}_Q^H \mathbf{P}_Q)^{-1} \boldsymbol{\Lambda}_Q\} \\ &= \beta_1 \boldsymbol{\alpha}_{p_1}, \end{aligned} \quad (47)$$

where

$$\boldsymbol{\Lambda}_Q = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_Q), \quad (48)$$

and $\beta_1 \neq 0$ is an arbitrary scalar. Now, considering the constraint that $\boldsymbol{\alpha}^H \mathbf{P}_Q^H \Gamma_x \mathbf{P}_Q \boldsymbol{\alpha} = 1$, we obtain the generalized MDNG (GMDNG) beamformer:

$$\mathbf{h}_{\text{GMDNG},Q} = \frac{\mathbf{P}_Q \boldsymbol{\alpha}_{p_1}}{\sqrt{\boldsymbol{\alpha}_{p_1}^H \mathbf{P}_Q^H \Gamma_x \mathbf{P}_Q \boldsymbol{\alpha}_{p_1}}}. \quad (49)$$

Another way to solve (46) is as follows. Let us denote

$$\mathbf{P}_Q \boldsymbol{\alpha} = \beta_2 \mathcal{P}\{\Gamma_x\}, \quad (50)$$

where $\beta_2 \neq 0$ is an arbitrary scalar. Then, we can derive that

$$\begin{aligned} \boldsymbol{\alpha} &= \beta_2 (\mathbf{P}_Q^H \mathbf{P}_Q)^{-1} \mathbf{P}_Q^H \mathcal{P}\{\Gamma_x\} \\ &= \beta_2 \boldsymbol{\alpha}_{p_2}. \end{aligned} \quad (51)$$

Using the power conservation constraint, we obtain another form of the GMDNG beamformer as

$$\mathbf{h}_{\text{GMDNG},Q} = \frac{\mathbf{P}_Q \boldsymbol{\alpha}_{p_2}}{\sqrt{\boldsymbol{\alpha}_{p_2}^H \mathbf{P}_Q^H \Gamma_x \mathbf{P}_Q \boldsymbol{\alpha}_{p_2}}}. \quad (52)$$

If the source is CS, it can be checked that

$$\mathbf{h}_{\text{GMDNG},Q} = \frac{\mathbf{P}_Q (\mathbf{P}_Q^H \mathbf{P}_Q)^{-1} \mathbf{P}_Q^H \mathbf{b}}{\mathbf{b}^H \mathbf{P}_Q (\mathbf{P}_Q^H \mathbf{P}_Q)^{-1} \mathbf{P}_Q^H \mathbf{b}}. \quad (53)$$

We have the following three cases:

- $Q = 1$, $\mathbf{h}_{\text{GMDNG},Q}$ degenerates to the MDNG beamformer;
- $Q = M$, $\mathbf{h}_{\text{GMDNG},Q}$ degenerates to the MWNG beamformer;
- For $1 < Q < M$, the DNG value of the GMDNG beamformer is between the DNG value of the MDNG beamformer and that of the MWNG beamformer, and so is the WNG.

Substituting either (49) or (52) into (24) gives

$$\begin{aligned} \mathcal{D}(\mathbf{h}_{\text{GMDNG},Q}) &= \frac{\boldsymbol{\alpha}_{p_1}^H \mathbf{P}_Q^H \Gamma_x \mathbf{P}_Q \boldsymbol{\alpha}_{p_1}}{\boldsymbol{\alpha}_{p_1}^H \mathbf{P}_Q^H \Gamma_d \mathbf{P}_Q \boldsymbol{\alpha}_{p_1}} \\ &= \frac{\boldsymbol{\alpha}_{p_1}^H \boldsymbol{\Lambda}_Q \boldsymbol{\alpha}_{p_1}}{\boldsymbol{\alpha}_{p_1}^H \mathbf{I}_Q \boldsymbol{\alpha}_{p_1}} \\ &= \frac{\boldsymbol{\alpha}_{p_2}^H \mathbf{P}_Q^H \Gamma_x \mathbf{P}_Q \boldsymbol{\alpha}_{p_2}}{\boldsymbol{\alpha}_{p_2}^H \mathbf{P}_Q^H \Gamma_d \mathbf{P}_Q \boldsymbol{\alpha}_{p_2}}, \end{aligned} \quad (54)$$

which can be further expressed as

$$\begin{aligned} \mathcal{D}(\mathbf{h}_{\text{GMDNG},Q}) &= \frac{\mathcal{P}^H\{\Gamma_x\} \mathbf{P}_Q (\mathbf{P}_Q^H \mathbf{P}_Q)^{-1} \boldsymbol{\Lambda}_Q (\mathbf{P}_Q^H \mathbf{P}_Q)^{-1} \mathbf{P}_Q^H \mathcal{P}\{\Gamma_x\}}{\mathcal{P}^H\{\Gamma_x\} \mathbf{P}_Q (\mathbf{P}_Q^H \mathbf{P}_Q)^{-2} \mathbf{P}_Q^H \mathcal{P}\{\Gamma_x\}}, \end{aligned} \quad (55)$$

where \mathbf{I}_Q is the $Q \times Q$ identity matrix.

Similarly, substituting either (49) or (52) into (23), we obtain

$$\begin{aligned} \mathcal{W}(\mathbf{h}_{\text{GMDNG},Q}) &= \frac{\mathcal{P}^H\{\Gamma_x\} \mathbf{P}_Q (\mathbf{P}_Q^H \mathbf{P}_Q)^{-1} \boldsymbol{\Lambda}_Q (\mathbf{P}_Q^H \mathbf{P}_Q)^{-1} \mathbf{P}_Q^H \mathcal{P}\{\Gamma_x\}}{\mathcal{P}^H\{\Gamma_x\} \mathbf{P}_Q (\mathbf{P}_Q^H \mathbf{P}_Q)^{-1} \mathbf{P}_Q^H \mathcal{P}\{\Gamma_x\}}. \end{aligned} \quad (56)$$

From (54) and (56), one can obtain an interesting relation between the WNG and the DNG:

$$\frac{\mathcal{D}(\mathbf{h}_{\text{GMDNG},Q})}{\mathcal{W}(\mathbf{h}_{\text{GMDNG},Q})} = \frac{\mathcal{P}^H\{\Gamma_x\} \mathbf{P}_Q (\mathbf{P}_Q^H \mathbf{P}_Q)^{-1} \mathbf{P}_Q^H \mathcal{P}\{\Gamma_x\}}{\mathcal{P}^H\{\Gamma_x\} \mathbf{P}_Q (\mathbf{P}_Q^H \mathbf{P}_Q)^{-2} \mathbf{P}_Q^H \mathcal{P}\{\Gamma_x\}}. \quad (57)$$

It is readily to check that

$$\frac{\mathcal{D}(\mathbf{h}_{\text{GMDNG},Q})}{\mathcal{W}(\mathbf{h}_{\text{GMDNG},Q})} \geq \frac{1}{M}, \quad (58)$$

since $\mathcal{D}(\mathbf{h}_{\text{GMDNG},Q}) \geq 1$ and $\mathcal{W}(\mathbf{h}_{\text{GMDNG},Q}) \leq M$.

According to the form (42), one can expect that:

$$\mathcal{W}(\mathbf{h}_{\text{GMDNG},1}) \leq \mathcal{W}(\mathbf{h}_{\text{GMDNG},2}) \leq \dots \leq \mathcal{W}(\mathbf{h}_{\text{GMDNG},M}) \quad (59)$$

and

$$\mathcal{D}(\mathbf{h}_{\text{GMDNG},1}) \geq \mathcal{D}(\mathbf{h}_{\text{GMDNG},2}) \geq \dots \geq \mathcal{D}(\mathbf{h}_{\text{GMDNG},M}). \quad (60)$$

Consequently, by choosing a different value of Q , the GMDNG beamformer can make a compromise between the level of WNG and the value of DNG. Generally, the WNG varies with frequency and the lower the frequency, the smaller is the WNG. So, in practice, one can control the WNG by choosing a large value of Q at low frequencies and a small value of Q at high frequencies. Note that there was some study about direct and closed-form solution with explicit control on both the WNG and DF for point source model [24], which will not be discussed here.

VI. SIMULATIONS

In this section, the performances of the proposed beamformers are evaluated using beampattern, WNG, and DNG as the performance measures.

A. Simulation Setup

In the simulation, we assume that the scattering characteristics of the source is independent of the frequency.

1) *IS Source*: In this case, we assume that the angular cross-correlation kernel of the scattered source is a bell-shaped function, which is defined as

$$\rho(\omega, \theta, \theta) = \begin{cases} \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{\theta^2}{2\epsilon^2}}}{\int_{-\theta_0}^{\theta_0} \frac{1}{\sqrt{2\pi}} e^{-\frac{\theta^2}{2\epsilon^2}} d\theta} & \text{if } \theta \in [-\theta_0, \theta_0] \\ 0 & \text{otherwise} \end{cases}. \quad (61)$$

We set $\epsilon^2 = 10$ for all the simulations related to the IS case. Then, the signal correlation matrix of the IS source is computed according to (12).

2) *CS source*: In this case, we assume that there is no phase difference among the coherent point sources and

$$g(\omega, \theta) = \begin{cases} \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{\theta^2}{2\epsilon^2}}}{\int_{-\theta_0}^{\theta_0} \frac{1}{\sqrt{2\pi}} e^{-\frac{\theta^2}{2\epsilon^2}} d\theta} & \text{if } \theta \in [-\theta_0, \theta_0] \\ 0 & \text{otherwise} \end{cases}. \quad (62)$$

We set $\epsilon^2 = 1$ for all the simulations related to the CS case. Then, the signal correlation matrix in this case is computed using (9).

Note that in practice, neither the angular density function $g(\omega, \theta)$ nor the scattering parameters are directly accessible; but those parameters can be estimated using, e.g., the algorithms presented in [26]–[31]. The signal correlation matrices can subsequently be estimated according to (9) and (12).

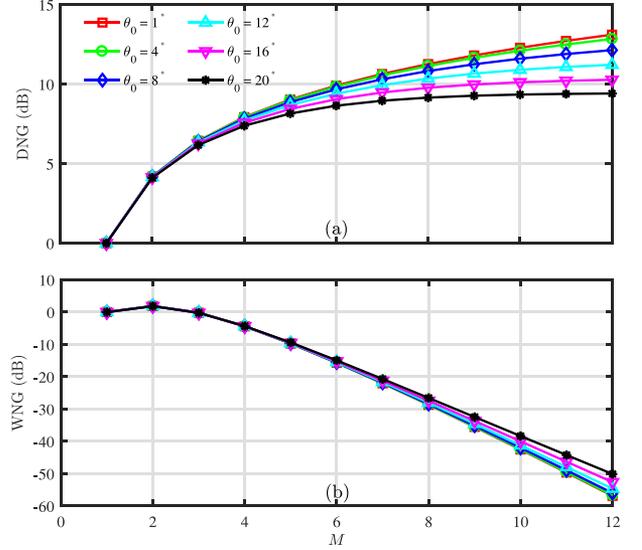


Fig. 1. DNG and WNG of the MDNG beamformer versus M for different values of θ_0 in the IS source field: (a) DNG and (b) WNG. Conditions: $\delta = 2.0$ cm and $f = 4$ kHz.

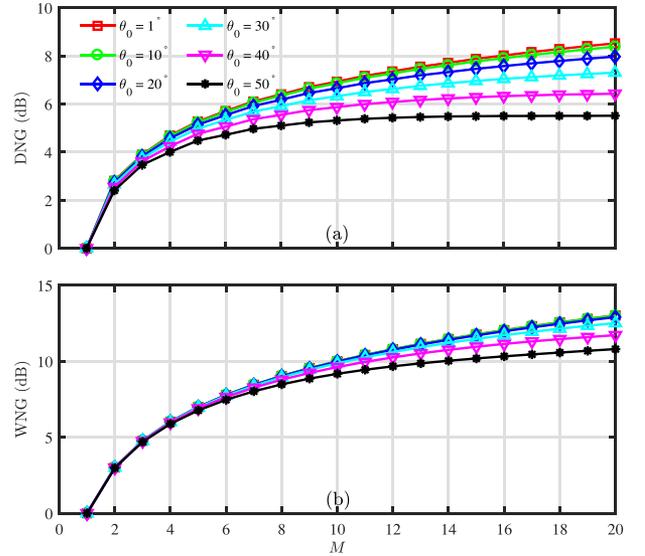


Fig. 2. DNG and WNG of the MWNG beamformer versus M for different values of θ_0 in the IS source field: (a) DNG and (b) WNG. Conditions: $\delta = 2.0$ cm and $f = 4$ kHz.

B. Performance of the MDNG and MWNG Beamformers Versus the Number of Microphones

In this simulation, we study the impact of the number of microphones on the WNG and DNG of the \mathbf{h}_{MDNG} and \mathbf{h}_{MWNG} beamformers. We consider the case with $f = 4$ kHz and different values of θ_0 . A ULMA with $\delta = 2.0$ cm is used. Figures 1 and 2 plot, respectively, the WNG and DNG of the \mathbf{h}_{MDNG} and \mathbf{h}_{MWNG} beamformers as a function of M in the IS source case while the results for the CS source case are plotted in Figs. 3 and 4.

As seen from Figs. 2 and 4, the values of DNG and WNG first increase with M . But when θ_0 becomes large, the DNG of the \mathbf{h}_{MWNG} beamformer does no longer increase with the number

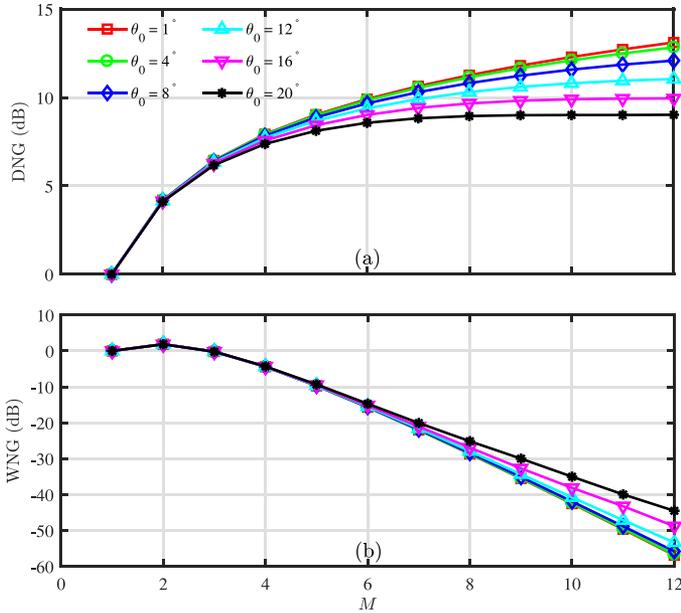


Fig. 3. DNG and WNG of the MDNG beamformer versus M for different values of θ_0 in the CS source field: (a) DNG and (b) WNG. Conditions: $\delta = 2.0$ cm and $f = 4$ kHz.

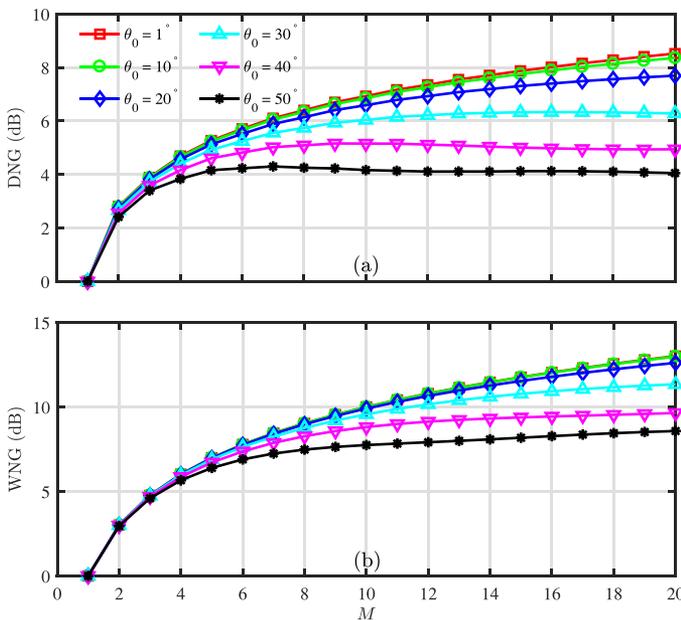


Fig. 4. DNG and WNG of the MWNG beamformer versus M for different values of θ_0 in the CS source field: (a) DNG and (b) WNG. Conditions: $\delta = 2.0$ cm and $f = 4$ kHz.

of microphones after reaching its maximum in both the IS and the CS source cases, which is different from the beamformers with the point source model. For example, in the CS source field, when $\theta_0 = 50^\circ$, the value of WNG of the MWNG beamformer does not increase much with M once $M > 10$. For $\theta_0 = 30^\circ$, 40° , and 50° , the DNG even slightly decreases as the number of microphones is large.

The values of the WNG and DNG of the \mathbf{h}_{MWNG} beamformer decrease with the increase of θ_0 in both cases. Comparing

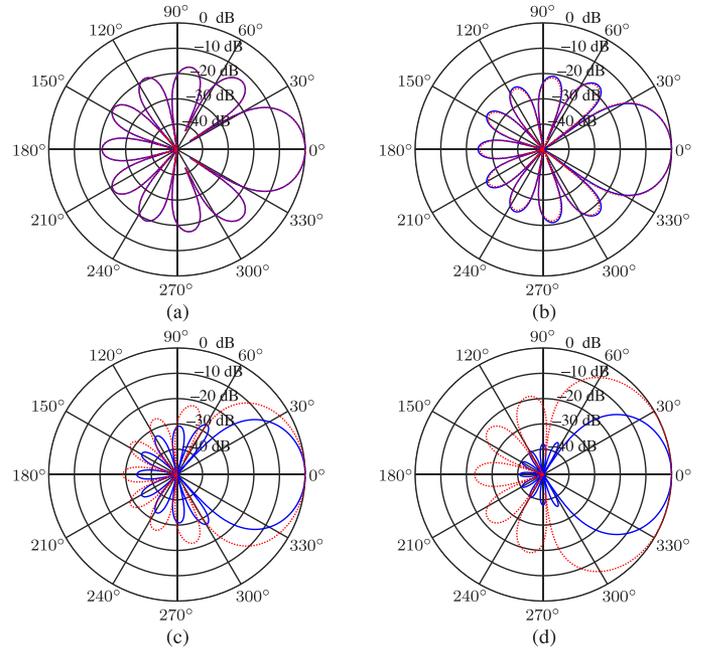


Fig. 5. Beampatterns of the MDNG (blue solid line) and LS (red dotted line) beamformers for different values of θ_0 in the IS source field: (a) $\theta_0 = 1^\circ$, (b) $\theta_0 = 20^\circ$, (c) $\theta_0 = 40^\circ$, and (d) $\theta_0 = 60^\circ$. Conditions: $M = 6$, $\delta = 1.0$ cm, and $f = 1$ kHz.

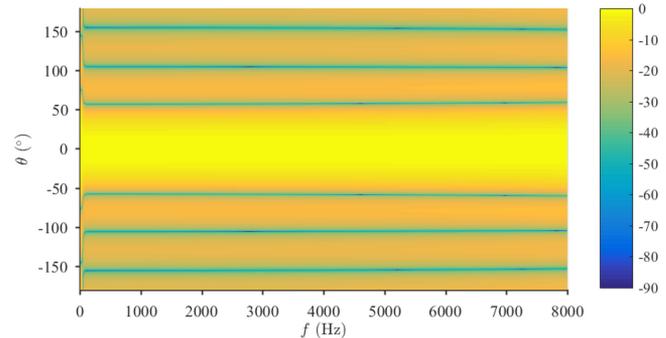


Fig. 6. Beampattern of the MDNG beamformer versus frequency in the IS source field. Conditions: $M = 4$, $\delta = 1.0$ cm, and $\theta_0 = 30^\circ$.

Figs. 2 and 4, one can see that the values of the WNG and DNG of the \mathbf{h}_{MWNG} beamformer in the IS source case are larger than those in the CS source case. From Figs. 1 and 3, one can see that, in both the IS and CS source scenarios, the WNG of the beamformer \mathbf{h}_{MDNG} increases while the DNG decreases with the value of θ_0 . When the value of θ_0 is large, the DNG of the \mathbf{h}_{MDNG} beamformer does not always increase with M in both the IS and CS source fields.

C. Beampattern of the MDNG Beamformer

In this simulation, we first consider a ULMA with $M = 6$ and $\delta = 1.0$ cm. The beampatterns of the \mathbf{h}_{MDNG} beamformer at $f = 1$ kHz and with different values of θ_0 in both the IS and CS cases are plotted in Figs. 5 and 7. For comparison, we also plot the beampatterns obtained by the least-squares (LS)

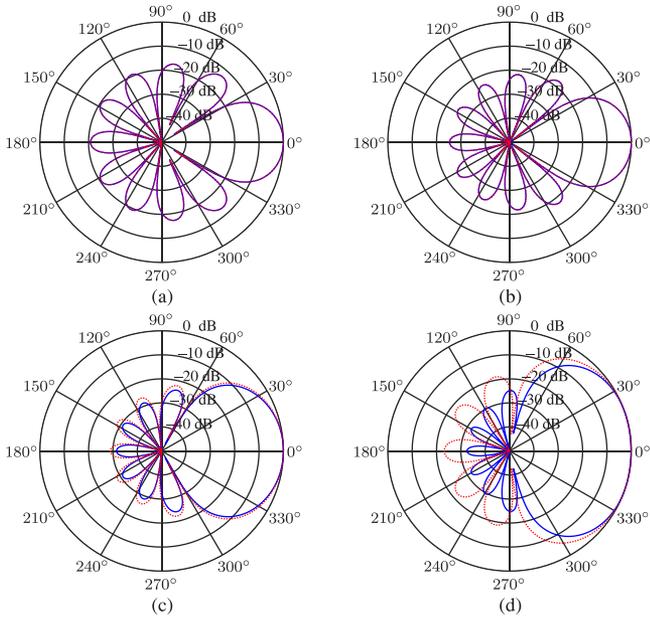


Fig. 7. Beampatterns of the MDNG (blue solid line) and LS (red dotted line) beamformers for different values of θ_0 in the CS source field: (a) $\theta_0 = 1^\circ$, (b) $\theta_0 = 20^\circ$, (c) $\theta_0 = 40^\circ$, and (d) $\theta_0 = 60^\circ$. Conditions: $M = 6$, $\delta = 1.0$ cm, and $f = 1$ kHz.

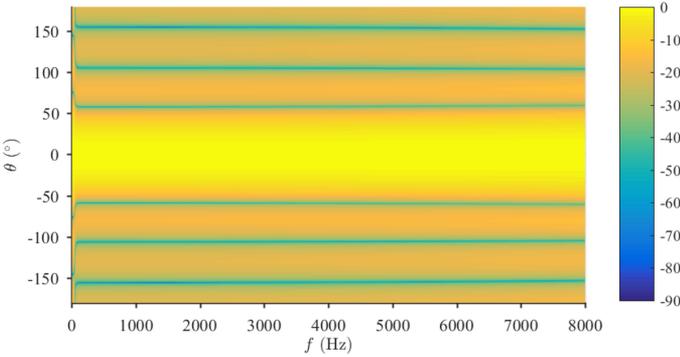


Fig. 8. Beampattern of the MDNG beamformer versus frequency in the CS source field. Conditions: $M = 4$, $\delta = 1.0$ cm, and $\theta_0 = 30^\circ$.

approach [20], [47]. The LS beamformer is designed by forcing the beampattern to be 1 for $\theta \in [-\theta_0, \theta_0]$ and 0 otherwise.

As seen, the beampatterns of both the MDNG and LS beamformers vary greatly with θ_0 . The beamwidth of the beampattern increases with θ_0 in both the IS and CS cases. It should be noted that for the given microphone array ($M = 6$, $\delta = 1$ cm), the beamwidth of the designed beampattern cannot be very narrow even when $\theta_0 \rightarrow 0^\circ$, which is limited by the aperture of the array. The beampatterns of the MDNG beamformer in Figs. 5(a) and 7(a) are the same, since when $\theta_0 \rightarrow 0^\circ$, \mathbf{h}_{MDNG} degenerates to the superdirective (fifth-order hypercardioid) beamformer in both the IS and CS cases.

We can see that, for both the IS and CS cases, when θ_0 is small, the beampatterns of both the MDNG and LS beamformers are similar. But when θ_0 is relatively large, the beamwidth of the MDNG beamformer is narrower because this beamformer takes the source scattering property into consideration for better signal estimation.

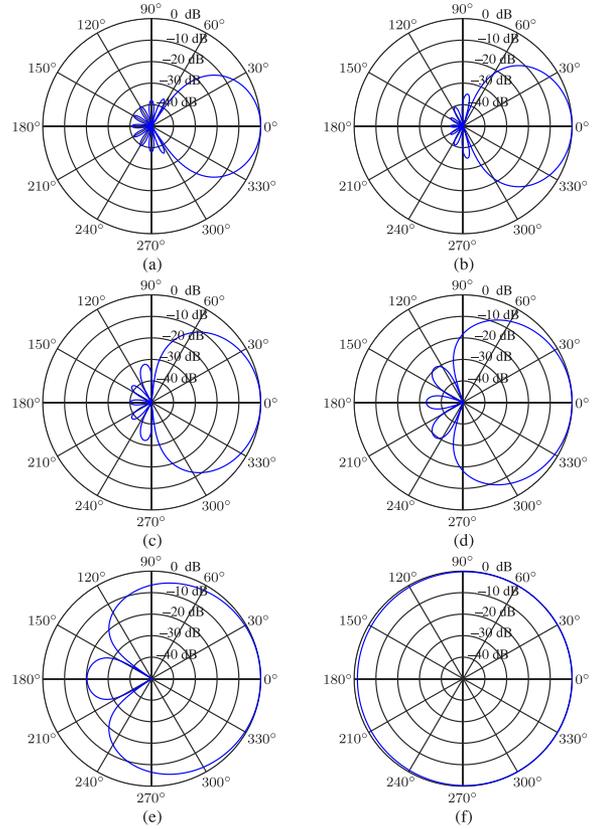


Fig. 9. Beampatterns of the GMDNG beamformer for different values of Q in the IS source field: (a) $Q = 1$, (b) $Q = 2$, (c) $Q = 3$, (d) $Q = 4$, (e) $Q = 5$, and (f) $Q = 6$. Conditions: $M = 6$, $\delta = 1.0$ cm, $\theta_0 = 60^\circ$, and $f = 1$ kHz.

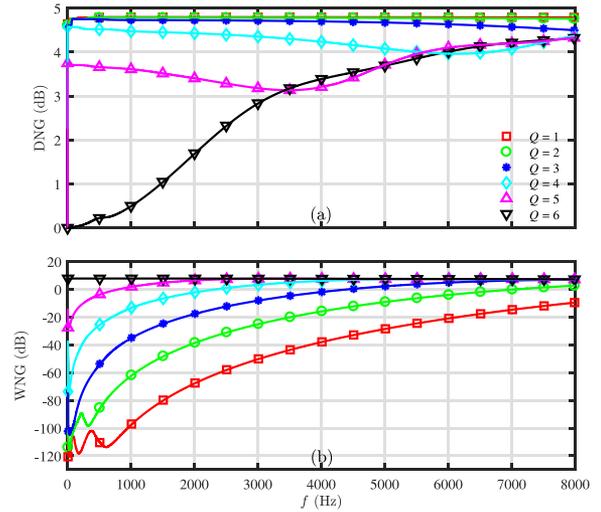


Fig. 10. DNG and WNG of the GMDNG beamformer versus frequency for different values of Q in the IS source field: (a) DNG and (b) WNG. Conditions: $M = 6$, $\delta = 1.0$ cm, and $\theta_0 = 60^\circ$.

Figures 6 and 8 plot the beampatterns of the designed MDNG beamformers versus frequency for $\theta_0 = 30^\circ$ with a ULMA of $M = 4$ and $\delta = 1.0$ cm. As seen, the obtained beampatterns are almost the same across the studied frequency range except the very low-frequency cases where the beampatterns deforms slightly, which is due to the numerical problem.

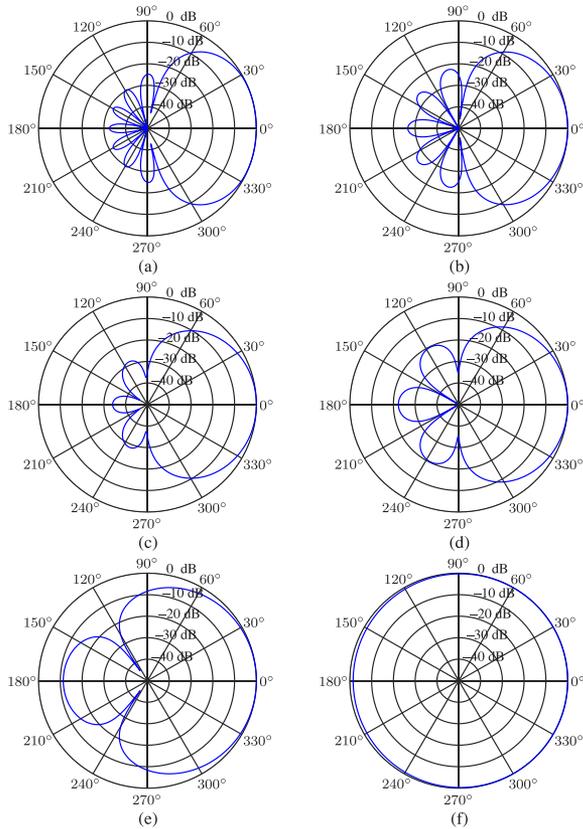


Fig. 11. Beam patterns of the GMDNG beamformer for different values of Q in the CS source field: (a) $Q = 1$, (b) $Q = 2$, (c) $Q = 3$, (d) $Q = 4$, (e) $Q = 5$, and (f) $Q = 6$. Conditions: $M = 6$, $\delta = 1.0$ cm, $\theta_0 = 60^\circ$, and $f = 1$ kHz.

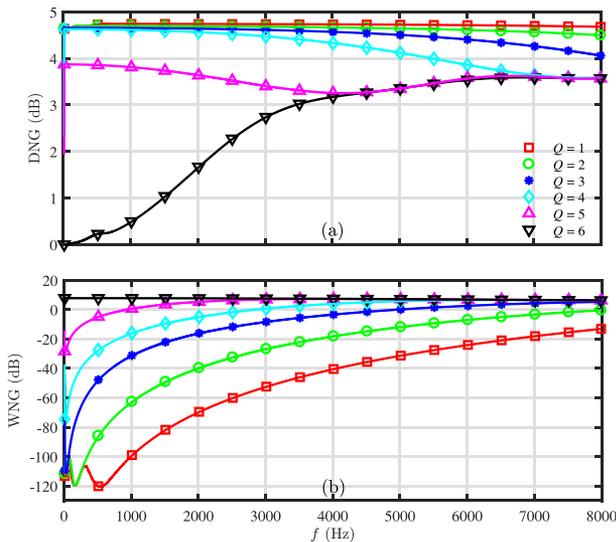


Fig. 12. DNG and WNG of the GMDNG beamformer versus frequency for different values of Q in the CS source field: (a) DNG and (b) WNG. Conditions: $M = 6$, $\delta = 1.0$ cm, and $\theta_0 = 60^\circ$.

D. Performance of the GMDNG Beamformer

In this simulation, we consider a ULMA with $M = 6$ and $\delta = 1.0$ cm. Figures 9 and 11 plot the beam patterns of the GMDNG beamformer in both the IS and CS situations for $\theta_0 = 60^\circ$, $f = 1$ kHz, and several values of Q . The beam pat-

terns of the GMDNG beamformer vary greatly with Q , from the beam pattern corresponding to the MDNG beamformer when $Q = 1$ to the one corresponding to the MWNG beamformer when $Q = 6$. The WNG and DNG of the GMDNG beamformer as functions of frequency for different Q values are plotted in Figs. 10 and 12. In both the CS and IS cases, the MDNG beamformer ($Q = 1$) achieved high DNG, but it suffers from serious white noise amplification at low frequencies. The MWNG beamformer ($Q = 6$), however, obtained high WNG, but the DNG is rather limited. It is seen that the WNG of the GMDNG beamformer increases while the DNG decreases as the value of Q increases from 1 to 6 in both cases. These results agree well with the analysis presented in the previous sections and show that the GMDNG beamformer can achieve a tradeoff between the DNG and WNG by adjusting the value of Q .

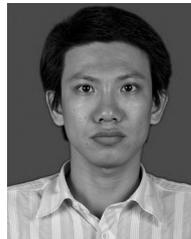
VII. CONCLUSIONS

In this paper, we have studied the problem of beamforming for scattered acoustic sources. We discussed the signal models for coherently and incoherently scattered sources. Three beamformers were then developed, namely, MDNG, MWNG, and GMDNG, for estimating the signal of interest from scattered sources and the beamwidth of their beam patterns can cover the scattered range of the source of interest. Simulations based on the use of a ULMA show that increasing the number of microphones can help improve the DNG of the MDNG and MWNG beamformers but only when the source scattered range is not too large, which is different from what was observed for beamformers with point sources. The GMDNG beamformer is basically a generalization of the MDNG and MWNG beamformers. Instead of maximizing either DNG or WNG, this beamformer can achieve a compromise between the value of DNG and the level of WNG through adjusting one parameter, which was validated by simulations.

REFERENCES

- [1] J. Benesty, S. Makino, and J. Chen, *Speech Enhancement*. Berlin, Germany: Springer-Verlag, 2005.
- [2] P. C. Loizou, *Speech Enhancement: Theory Practice*. Boca Raton, FL, USA: CRC, 2007.
- [3] P. A. Naylor and N. D. Gaubitch, *Speech Dereverberation*. Berlin, Germany: Springer Science & Business Media, 2010.
- [4] S. A. Schelkunoff, "A mathematical theory of linear arrays," *Bell Syst. Tech. J.*, vol. 22, pp. 80–107, Jan. 1943.
- [5] B. Rafaely, "Phase-mode versus delay-and-sum spherical microphone array processing," *IEEE Signal Process. Lett.*, vol. 12, no. 10, pp. 713–716, Oct. 2005.
- [6] Y. Zeng and R. C. Hendriks, "Distributed delay and sum beamformer for speech enhancement via randomized gossip," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 22, no. 1, pp. 260–273, Jan. 2014.
- [7] 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.
- [8] J. M. Kates, "Superdirective arrays for hearing aids," *J. Acoust. Soc. Amer.*, vol. 94, no. 4, pp. 1930–1933, 1993.
- [9] Y. Wang, Y. Yang, Y. Ma, and Z. He, "Robust high-order superdirectivity of circular sensor arrays," *J. Acoust. Soc. Amer.*, vol. 136, no. 4, pp. 1712–1724, 2014.
- [10] S. Doclo and M. Moonen, "Superdirective beamforming robust against microphone mismatch," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 15, no. 2, pp. 617–631, Feb. 2007.

- [11] R. N. Marshall and W. R. Harry, "A new microphone providing uniform directivity over an extended frequency range," *J. Acoust. Soc. Amer.*, vol. 12, pp. 481–497, Apr. 1941.
- [12] G. W. Elko, S. L. Gay, and J. Benesty, *Acoustic Signal Processing for Telecommunication*. Boston, MA, USA: Kluwer, 2000.
- [13] 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, pp. 1021–1041.
- [14] J. Chen, J. Benesty, and C. Pan, "On the design and implementation of linear differential microphone arrays," *J. Acoust. Soc. Amer.*, vol. 136, pp. 3097–3113, Dec. 2014.
- [15] C. Pan, J. Chen, and J. Benesty, "Theoretical analysis of differential microphone array beamforming and an improved solution," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 23, no. 11, pp. 2093–2105, Nov. 2015.
- [16] T. D. Abhayapala and A. Gupta, "Higher order differential-integral microphone arrays," *J. Acoust. Soc. Amer.*, vol. 127, pp. 227–233, May 2010.
- [17] J. Weinberger, H. F. Olson, and F. Massa, "A uni-directional ribbon microphone," *J. Acoust. Soc. Amer.*, vol. 5, pp. 139–147, Oct. 1933.
- [18] H. F. Olson, "Gradient microphones," *J. Acoust. Soc. Amer.*, vol. 17, pp. 192–198, Jan. 1946.
- [19] G. M. Sessler and J. E. West, "Directional transducers," *IEEE Trans. Audio Electroacoust.*, vol. 19, no. 1, pp. 19–23, Mar. 1971.
- [20] J. Benesty, J. Chen, and Y. Huang, *Microphone Array Signal Processing*. Berlin, Germany: Springer-Verlag, 2008.
- [21] J. Benesty and J. Chen, *Study Design Differential Microphone Arrays*. Berlin, Germany: Springer-Verlag, 2012.
- [22] L. L. Beranek, *Acoustics*. Woodbury, NY, USA: Acoustic Society of America, 1986.
- [23] H. Cox, R. Zeskind, and M. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, no. 10, pp. 1365–1376, Oct. 1987.
- [24] R. Berkun, I. Cohen, and J. Benesty, "Combined beamformers for robust broadband regularized superdirective beamforming," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 23, no. 5, pp. 877–886, May 2015.
- [25] C. Li, J. Benesty, G. Huang, and J. Chen, "Subspace superdirective beamformers based on joint diagonalization," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2016, pp. 400–404.
- [26] S. Valaee, B. Champagne, and P. Kabal, "Parametric localization of distributed sources," *IEEE Trans. Signal Process.*, vol. 43, no. 9, pp. 2144–2153, Sep. 1995.
- [27] Y. U. Lee, J. Choi, I. Song, and S. R. Lee, "Distributed source modeling and direction-of-arrival estimation techniques," *IEEE Trans. Signal Process.*, vol. 45, no. 4, pp. 960–969, Apr. 1997.
- [28] S. Shahbazpanahi, S. Valaee, and M. H. Bastani, "Distributed source localization using ESPRIT algorithm," *IEEE Trans. Signal Process.*, vol. 49, no. 10, pp. 2169–2178, Oct. 2001.
- [29] O. Besson and P. Stoica, "Decoupled estimation of DOA and angular spread for a spatially distributed source," *IEEE Trans. Signal Process.*, vol. 48, no. 7, pp. 1872–1882, Jul. 2000.
- [30] M. Bengtsson and B. Ottersten, "Low-complexity estimators for distributed sources," *IEEE Trans. Signal Process.*, vol. 48, no. 8, pp. 2185–2194, Aug. 2000.
- [31] M. Souden, S. Affes, and J. Benesty, "A two-stage approach to estimate the angles of arrival and the angular spreads of locally scattered sources," *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1968–1983, May 2008.
- [32] K. L. Bell and H. L. V. Trees, "Robust adaptive beamforming for spatially spread sources," in *Proc. 9th IEEE Workshop Statist. Signal Array Process.*, Portland, OR, USA, 1998, pp. 1–4.
- [33] 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.
- [34] L. Zhang and W. Liu, "Robust forward backward based beamformer for a general-rank signal model with real-valued implementation," *Signal Process.*, vol. 92, pp. 163–169, Jan. 2012.
- [35] J.-S. Hu and M.-T. Lee, "Norm-constrained capon beamforming using multirank signal models with kalman filter implementation," *IEEE Trans. Antennas Propag.*, vol. 62, no. 9, pp. 4574–4583, Sep. 2014.
- [36] N. Q. K. Duong, E. Vincent, and R. Gribonval, "Under-determined reverberant audio source separation using a full-rank spatial covariance model," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 7, pp. 1830–1840, Sep. 2010.
- [37] H. Chen and A. B. Gershman, "Robust adaptive beamforming for general-rank signal models using positive semi-definite covariance constraint," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2008, pp. 2341–2344.
- [38] S.-J. Kim, A. Magnani, A. Mutapcic, S. P. Boyd, and Z.-Q. Luo, "Robust beamforming via worst-case SINR maximization," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1539–1547, Apr. 2008.
- [39] M. Brandstein and D. B. Ward, Eds., *Microphone Arrays: Signal Processing Techniques and Applications*. Berlin, Germany: Springer-Verlag, 2001.
- [40] J. Benesty, I. Cohen, and J. Chen, *Fundamentals of Signal Enhancement and Array Signal Processing*. Hoboken, NJ, USA: Wiley, 2017.
- [41] G. W. Elko, J. Meyer, J. Benesty, M. M. Sondhi, and Y. Huang, *Springer Handbook of Speech Processing*. Berlin, Germany: Springer-Verlag, 2008.
- [42] C. Pan, J. Chen, and J. Benesty, "Performance study of the MVDR beamformer as a function of the source incident angle," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 22, no. 1, pp. 67–79, Jan. 2014.
- [43] J. N. Franklin, *Matrix Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1968.
- [44] X. Wang, J. Benesty, and J. Chen, "A single-channel noise cancellation filter in the short-time-fourier-transform domain," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2016, pp. 5235–5239.
- [45] S. M. Nørholm, J. Benesty, J. R. Jensen, and M. G. Christensen, "Single-channel noise reduction using unified joint diagonalization and optimal filtering," *EURASIP J. Adv. Signal Process.*, vol. 2014, Dec. 2014, Art. no. 11.
- [46] J. R. Jensen, J. Benesty, and M. G. Christensen, "Noise reduction with optimal variable span linear filters," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 24, no. 4, pp. 631–644, Apr. 2016.
- [47] S. Doclo and M. Moonen, "Design of far-field and near-field broadband beamformers using eigenfilters," *Signal Process.*, vol. 83, pp. 2641–2673, 2003.



Xianghui Wang received the bachelor's degree in electronics and information engineering from the Xi'an University, Xi'an, China, in 2008, and the master's degree in signal and information processing from the Northwestern Polytechnical University, Xi'an, China, in 2011, where he is currently working toward the Ph.D. degree in information and communication engineering. He is also a Visiting Ph.D. Student at the Andrew and Erna Viterby Faculty of Electrical Engineering, Technion – Israel Institute of Technology, Haifa, Israel. His research interests include speech enhancement and microphone array signal processing.



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, Haifa, Israel, in 1990, 1993, and 1998, respectively.

From 1990 to 1998, he was a Research Scientist with the RAFAEL Research Laboratories, Haifa, Israel Ministry of Defense. From 1998 to 2001, he was a Postdoctoral Research Associate with the Department of Computer Science, Yale University, New Haven, CT, USA. In 2001 he joined the Department

of Electrical Engineering, Technion. He is a Professor of electrical engineering with the Technion–Israel Institute of Technology, Haifa, Israel. He is also a Visiting Professor with the Northwestern Polytechnical University, Xi'an, China. He is a coeditor of the Multichannel Speech Processing Section of the *Springer Handbook of Speech Processing* (Springer, 2008), and a coauthor of *Fundamentals of Signal Enhancement and Array Signal Processing* (Wiley-IEEE Press, 2018). His research interests include array processing, 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 received the Norman Seiden Prize for Academic Excellence (2017), the SPS Signal Processing Letters Best Paper Award (2014), the Alexander Goldberg Prize for Excellence in Research (2010), and the Muriel and David Jacknow Award for Excellence in Teaching (2009). He is an Associate Member of the IEEE Audio and Acoustic Signal Processing Technical Committee, and as Distinguished Lecturer of the IEEE Signal Processing Society. He was an Associate Editor of the IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, IEEE SIGNAL PROCESSING LETTERS, as a member of the IEEE Audio and Acoustic Signal Processing Technical Committee, and the IEEE Speech and Language Processing Technical Committee.



Jingdong Chen (M'99–SM'09) received the Ph.D. degree in pattern recognition and intelligence control from the Chinese Academy of Sciences, Beijing, China, in 1998.

From 1998 to 1999, he was with ATR Interpreting Telecommunications Research Laboratories, Kyoto, Japan, where he conducted research on speech synthesis, speech analysis, as well as objective measurements for evaluating speech synthesis. He then joined the Griffith University, Brisbane, Australia, where he engaged in research on robust speech recognition and signal processing.

From 2000 to 2001, he worked with ATR Spoken Language Translation Research Laboratories on robust speech recognition and speech enhancement. From 2001 to 2009, he was a member of Technical Staff at Bell Laboratories, Murray Hill, New Jersey, working on acoustic signal processing for telecommunications. He subsequently joined WeVoice Inc. in New Jersey, as the Chief Scientist. He is currently a Professor with the Northwestern Polytechnical University, Xi'an, China. He has coauthored 12 monograph books including *Fundamentals of Signal Enhancement and Array Signal Processing*, (Wiley, 2017), *Design of Circular Differential Microphone Arrays* (Springer, 2015), *Study and Design of Differential Microphone Arrays* (Springer, 2013), *Noise Reduction in Speech Processing* (Springer, 2009), *Microphone Array Signal Processing* (Springer, 2008), and *Acoustic MIMO Signal Processing* (Springer, 2006), etc. His research interests include acoustic signal processing, adaptive signal processing, speech enhancement, adaptive noise/echo control, microphone array signal processing, signal separation, and speech communication.

Dr. Chen was an Associate Editor of the IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING from 2008 to 2014 and as a technical committee (TC) member of the IEEE Signal Processing Society (SPS) TC on Audio and Electroacoustics from 2007 to 2009. He is currently a member of the IEEE SPS TC on Audio and Acoustic Signal Processing, and a member of the editorial advisory board of the Open Signal Processing Journal. He was the General Co-Chair of ACM WUWNET 2018 and IWAENC 2016, the Technical Program Chair of the IEEE TENCON 2013, a Technical Program Co-Chair of IEEE WASPAA 2009, IEEE ChinaSIP 2014, IEEE ICSPCC 2014, and IEEE ICSPCC 2015, and helped organize many other conferences. He received the 2008 Best Paper Award from the IEEE Signal Processing Society (with Benesty, Huang, and Doclo), the Best Paper Award from the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics in 2011 (with Benesty), the Bell Labs Role Model Teamwork Award twice, respectively, in 2009 and 2007, the NASA Tech Brief Award twice, respectively, in 2010 and 2009, and the Young Author Best Paper Award from the 5th National Conference on Man-Machine Speech Communications in 1998. He has coauthored a paper for which his Ph.D. student, Chao Pan, received the IEEE R10 (Asia-Pacific Region) Distinguished Student Paper Award (First Prize) in 2016. He was also a recipient of the Japan Trust International Research Grant from the Japan Key Technology Center in 1998 and the "Distinguished Young Scientists Fund" from the National Natural Science Foundation of China in 2014.



Jacob Benesty received a master's degree in microwaves from Pierre and Marie Curie University, Paris, France, in 1987, and the Ph.D. degree in control and signal processing from Orsay University, Orsay, France, in April 1991. During his Ph.D. (from November 1989 to April 1991), he worked on adaptive filters and fast algorithms at the Centre National d'Etudes des Telecommunications, 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 with the Technion, Haifa, Israel, an Adjunct Professor at Aalborg University, Denmark, and a Guest Professor at Northwestern Polytechnical University, Xi'an, China. 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 multiparty hands-free full-duplex stereo conferencing system over IP networks. He has coauthored and coedited/coauthored numerous books in the area of acoustic signal processing. He is the editor of the book series Springer Topics in Signal Processing. His research interests include signal processing, acoustic signal processing, and multimedia communications.

Prof. Benesty was the General Chair and Technical Chair of many international conferences and a member of several IEEE technical committees. Four of his journal papers were awarded by the IEEE Signal processing Society and in 2010 he received the Gheorghe Cartianu Award from the Romanian Academy.