Beamforming based on null-steering with small spacing linear microphone arrays

Changlei Li, Jacob Benesty, and Jingdong Chen

Citation: The Journal of the Acoustical Society of America **143**, 2651 (2018); doi: 10.1121/1.5035272
View online: https://doi.org/10.1121/1.5035272
View Table of Contents: https://asa.scitation.org/toc/jas/143/5
Published by the Acoustical Society of America

Beamforming based on null-steering with small spacing linear microphone arrays

Changlei Li
Center of Intelligent Acoustics and Immersive Communications, and School of Marine Science and Technology, Northwestern Polytechnical University, 127 Youyi West Road, Xi'an 710072, China

Jacob Benesty
INRS-EMT, University of Quebec, 800 de la Gauchetiere Ouest, Montreal, Quebec H5A 1K6, Canada

Jingdong Chen[a)]
Center of Intelligent Acoustics and Immersive Communications, Northwestern Polytechnical University, 127 Youyi West Road, Xi'an 710072, China

(Received 16 January 2018; revised 12 April 2018; accepted 13 April 2018; published online 3 May 2018)

This paper develops an approach to beamforming with small spacing uniform linear microphone arrays based on the null-steering (NS) principle. It first formulates the beamforming problem from the conventional mean-squared error (MSE) criterion and its normalized version. Several NS algorithms are then derived for beamforming with the constraint of placing nulls to either a single direction or multiple angles. The difference and relationships between different algorithms are discussed and their performances are evaluated. These algorithms can be used to design either fixed or adaptive beamformers. In the former case, the resulting beamformers work as differential microphone arrays (DMAs) since they exhibit frequency-invariant beampatterns and their directivity factors are high with a given number of sensors. In the latter case, the resulting beamformers can be viewed as a combination of DMAs and single-channel noise reduction since they do not only exhibit frequency-invariant beampatterns but also can achieve noise reduction based on the noise statistics.
 https://doi.org/10.1121/1.5035272

[NX] Pages: 2651–2665

I. INTRODUCTION

Microphone arrays are now widely used in a large range of applications such as smart speakers, robotics, smart home systems, automotive voice navigation, hands-free voice communication, teleconferencing, cellular phones, hearing aids, to name a few. A critical component of a microphone array system is the so-called beamforming, which is basically a multichannel filter that uses the spatial, spectral, and temporal information embedded in the array observation signals to estimate the speech signal of interest while reducing the effects of noise, reverberation, and interference (Schelkunoff, 1943; Dudgeon, 1977; Flanagan *et al.*, 1985; Haykin, 1985; Kellermann, 1991; Brandstein and Ward, 2001; Benesty *et al.*, 2008; Benesty *et al.*, 2017; Benesty and Chen, 2012; Capon, 1969; Frost, 1972; Griffiths and Jim, 1982). Great efforts have been devoted to microphone array beamforming over the last three decades, and many algorithms have been developed accordingly. Broadly, those algorithms can be divided into two categories, i.e., fixed and adaptive beamformers, based upon how the beamforming filter coefficients are designed and updated. Fixed beamformers, in either a delay-and-sum structure or a filter-and-sum form, have their beamforming coefficients designed primarily based on the array geometry and steering angle information, and those coefficients are fixed once the array system is deployed (Schelkunoff, 1943; Flanagan *et al.*, 1985; Haykin, 1985; Kellermann, 1991; Benesty *et al.*, 2008; Benesty *et al.*, 2017). These types of beamformers are generally suboptimal in terms of noise reduction performance as they do not use any information of the real ambient noise. In comparison, adaptive beamformers estimate and update their beamforming coefficients based on the noise statistics (Capon, 1969; Frost, 1972; Griffiths and Jim, 1982; Jim, 1977; Werner *et al.*, 2003). If the estimate of these statistics is accurate, adaptive beamformers may achieve better performance than fixed ones. However, how to estimate the signal and noise statistics in a reliable manner is a non-trivial task, particularly in acoustic environments in the presence of reverberation where both the desired signal and noise are generally nonstationary. Therefore, robust implementation of adaptive beamformers is always a great challenge, and it is not uncommon to have signal distortion and signal self-cancellation with adaptive beamformers in real-world applications. Consequently, fixed beamformers are often preferred in practical systems due to their performance consistency and robustness.

Fixed beamformers can be further divided into the following two subcategories: additive and differential. The former refers to those responsive to the acoustic pressure field, while the latter refers to those responsive to the spatial derivatives of different orders of the acoustic pressure field (Elko, 2000; Elko and Meyer, 2008). Generally, additive beamformers are applied to large spacing arrays with large aperture while differential beamformers are applied to small spacing and compact size microphone arrays. In comparison,

[a)]Electronic mail: jingdongchen@ieee.org

differential beamformers offer three major advantages: (1) they exhibit almost frequency-invariant beampatterns, which make them attractive for processing broadband audio and speech signals; (2) they can attain high spatial gain, which make them efficient in dealing with spatial noise and reverberation suppression in comparison with additive beamformers with the same number of sensors; and (3) they work with compact size arrays, which are easily integrated into devices with limited space such as cellular phones, robotics, hearing aids, and smart speakers. As a result, differential beamforming has attracted tremendous attention over the last two decades (Benesty and Chen, 2012; Elko and Meyer, 2008; Elko, 2000; Chen *et al.*, 2014; Benesty *et al.*, 2016; Pan *et al.*, 2015; Zhao *et al.*, 2014, 2016; Bernardini *et al.*, 2017; Wu and Chen, 2016; Sena *et al.*, 2012).

So far, there have been three major approaches to the design and implementation of differential beamformers: cascaded subtraction, null-constrained linear system solving, and series approximation. The basic principle of the first method can be dated back to the 1930s when directional sensors were invented (Olson, 1932; Olson, 1946), and this principle was then extended in the 1990s to the design of differential microphone arrays (DMAs; Elko, 2000; Elko and Meyer, 2008; Elko and Pong, 1997; Elko, 1997). In this approach, a general nth-order DMA is constructed by subtractively combining the fractionally delayed outputs of two DMAs of order $n-1$, e.g., a first-order DMA is formed by subtracting a fractionally delayed version of the output of one omnidirectional microphone (a 0th-order DMA) from the output of another adjacent omnidirectional microphone, and a second-order DMA is constructed as the difference between the outputs of two first-order DMAs (Benesty and Chen, 2012; Elko, 2000; Elko and Meyer, 2008). While straightforward, this method lacks flexibility in dealing with the problem of white noise amplification, which is intrinsic to differential beamforming, particularly when the order is high and frequency is low. In comparison, the null-constrained approach performs differential beamforming in the short-time Fourier transform (STFT) domain. It transforms the differential beamforming problem into one of linear system solving, where the linear system is constructed from some fundamental constraints (typically the steering and null directions) from the target beampattern (Benesty and Chen, 2012; Chen *et al.*, 2014; Benesty *et al.*, 2016). It has been shown that this approach is equivalent to the cascaded subtraction method if an nth-order DMA is constructed using $n+1$ physically omnidirectional microphones (Pan *et al.*, 2015). One of the prominent advantages of this method is the design of an nth-order differential beamformer using more than $n+1$ microphones to obtain the desired, target beampattern while maximizing the white noise gain (WNG), leading to the so-called robust DMA. In the series approximation approach, the exponentials in the beampattern (or steering vector) are approximated by a limited order of some expansion series (e.g., Taylor, Jacobi-Anger), and the beamforming filter is identified also by solving a linear system as in the null-constrained method, but the linear system is a natural result of the series approximation (Zhao *et al.*, 2014, 2016). Similar to the second approach, WNG can be improved by increasing the number of microphones when designing a given order of DMA.

This paper is also concerned with beamforming with small spacing uniform linear microphone arrays. We develop a null-steering (NS) approach, which designs beamformers based on the use of the array geometry and nulls information. The NS principle, which originates from the sidelobe canceller (Compton, 1988; Howells, 1976) and the generalized sidelobe canceller (Buckley and Griffiths, 1986; Buckley, 1986, 1987; Van Veen and Buckley, 1988), has been studied over a long time to cancel one or multiple competing source (interference) signals propagating from known directions (Chiba *et al.*, 1994; Friedlander and Porat, 1989; Godara and Cantoni, 1981; Godara, 1997). The nulls in the beampattern can be formed using either the amplitude (Vu, 1986) or phase information (Shore, 1984; Haupt, 1997), or more generally by adjusting the complex beamforming filter (Ng *et al.*, 1985; Steyskal *et al.*, 1986). In this paper, however, we study the NS principle from another perspective and apply it to beamforming for small spacing and compact linear microphone arrays. Different NS beamforming algorithms are derived by minimizing some mean-squared error (MSE) criterion with nulls constraints. These algorithms can be used to design either fixed or adaptive beamformers. In the former case, the resulting beamformers work as DMAs since (1) they exhibit frequency-invariant beampatterns and (2) their directivity factors (DFs) are high. In the latter case, the resulting beamformers can be viewed as a combination of DMAs and noise reduction since (1) they exhibit frequency-invariant beampatterns and (2) they can achieve adaptive noise reduction based on the noise statistics estimates.

The remainder of this paper is organized as follows. In Sec. II, we describe the signal model and the problem formulation of microphone array beamforming in the STFT domain. Section III presents some useful performance measures about beamforming and NS algorithms. In Sec. IV, we briefly describe the conventional beamforming algorithms. In Sec. V, we develop some NS algorithms under different criteria. Simulation results are presented to support our theoretical study in Sec. VI. Finally, some conclusions are drawn in Sec. VIII.

II. SIGNAL MODEL AND PROBLEM FORMULATION

We consider a uniform linear sensor array, as shown in Fig. 1, consisting of M omnidirectional microphones and for which the distance between two successive sensors is equal to δ. In a general way, the received signals at the frequency index f are expressed as (Benesty *et al.*, 2008; Benesty *et al.*, 2011; Elko and Meyer, 2008)

$$\begin{aligned} Y_m(f) &= G_m(f)S(f) + V_m(f) \\ &= X_m(f) + V_m(f), \quad m = 1, 2, \ldots, M, \end{aligned} \tag{1}$$

where $Y_m(f)$ is the mth microphone signal, $S(f)$ is the unknown speech source, which is assumed to propagate from the endfire direction (i.e., at the angle of 0°), $G_m(f)$ is the acoustic transfer function from the position of $S(f)$ to the mth microphone, and $X_m(f) = G_m(f)S(f)$ and $V_m(f)$ are, respectively, the

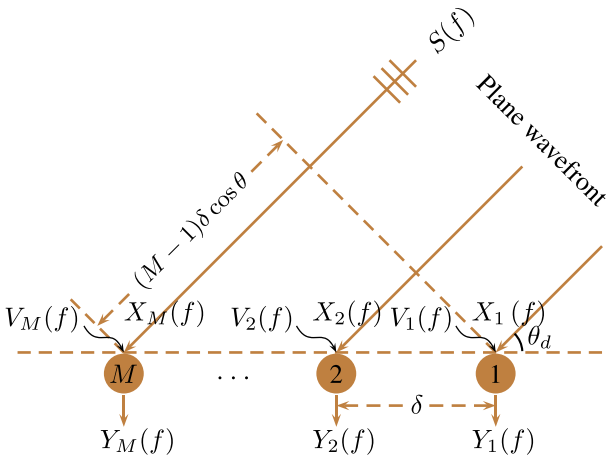


FIG. 1. (Color online) Illustration of a uniform linear microphone array for sound capture in the far field.

convolved speech signal and additive noise at the mth microphone. We assume that $X_m(f)$ and $V_m(f)$ are uncorrelated and zero mean. We further assume that δ is small in order that the array, with some processing of the sensor signals, works as a DMA (Benesty and Chen, 2012). Without loss of generality, microphone 1 is chosen as the reference. Then, the objective of beamforming or noise reduction in the frequency domain is to estimate the desired signal, $X_1(f)$, given the M observations, $Y_m(f)$, $m = 1, 2, \ldots, M$, in the best way we can.

It is more convenient to write the M frequency-domain microphone signals in a vector notation

$$\begin{aligned}\mathbf{y}(f) &= \mathbf{g}(f)S(f) + \mathbf{v}(f) = \mathbf{x}(f) + \mathbf{v}(f)\\ &= \mathbf{d}(f)X_1(f) + \mathbf{v}(f),\end{aligned} \tag{2}$$

where

$$\mathbf{y}(f) = \left[Y_1(f) \quad Y_2(f) \quad \cdots \quad Y_M(f)\right]^T, \tag{3}$$

$$\mathbf{x}(f) = \left[X_1(f) \quad X_2(f) \quad \cdots \quad X_M(f)\right]^T = S(f)\mathbf{g}(f), \tag{4}$$

$$\mathbf{g}(f) = \left[G_1(f) \quad G_2(f) \quad \cdots \quad G_M(f)\right]^T, \tag{5}$$

$$\mathbf{v}(f) = \left[V_1(f) \quad V_2(f) \quad \cdots \quad V_M(f)\right]^T, \tag{6}$$

$$\mathbf{d}(f) = \left[1 \quad \frac{G_2(f)}{G_1(f)} \quad \cdots \quad \frac{G_M(f)}{G_1(f)}\right]^T = \frac{\mathbf{g}(f)}{G_1(f)}, \tag{7}$$

and the superscript "^{T}" is the transpose operator. The vector $\mathbf{d}(f)$ is obviously the frequency-domain signal propagation vector, which is in the same form as the steering vector. In the anechoic and far-field context, and with the assumption that the source impinges on the array from the angle 0°, Eq. (7) degenerates to

$$\mathbf{d}(f,0) = \left[1 \quad e^{-\jmath 2\pi f\tau_0} \quad \cdots \quad e^{-\jmath 2(M-1)\pi f\tau_0}\right]^T, \tag{8}$$

where $\jmath$ is the imaginary unit with $\jmath^2 = -1$ and $\tau_0 = \delta/c$ is the delay between two successive sensors at the angle 0°, with $c \approx 340$ m/s being the speed of sound. In the rest of this work, we assume that

$$\mathbf{d}(f,0) \approx \mathbf{d}(f). \tag{9}$$

Clearly, for a given angle θ, it is known that the steering vector can be written as

$$\mathbf{d}(f,\theta) = \left[1 \quad e^{-\jmath 2\pi f\tau_0\cos\theta} \quad \cdots \quad e^{-\jmath 2(M-1)\pi f\tau_0\cos\theta}\right]^T. \tag{10}$$

The objective of beamforming is to estimate the desired signal, $X_1(f)$, given the observed vector, $\mathbf{y}(f)$. This can be achieved by applying a complex-valued filter to $\mathbf{y}(f)$, i.e.,

$$Z(f) = \mathbf{h}^H(f)\mathbf{y}(f) = \mathbf{h}^H(f)\mathbf{d}(f)X_1(f) + \mathbf{h}^H(f)\mathbf{v}(f), \tag{11}$$

where $Z(f)$ is the estimate of the desired signal, $X_1(f)$, the superscript "^{H}" is the conjugate-transpose operator, and

$$\mathbf{h}(f) = \left[H_1(f) \quad H_2(f) \quad \cdots \quad H_M(f)\right]^T \tag{12}$$

is a filter of length M containing all the complex weights applied to the microphone outputs at frequency f. With the above formulation, the objective of beamforming consists in designing an optimal spatial filter, $\mathbf{h}(f)$, such that the linear array works as a DMA (Benesty and Chen, 2012; Elko and Meyer, 2008; Elko, 2000), which has a frequency-invariant spatial response with a high spatial gain.

III. USEFUL DEFINITIONS

In this section, we give some important definitions and measures that will be used to derive and evaluate beamforming filters.

With microphone 1 as the reference, we can define the input signal-to-noise ratio (SNR) as

$$\text{iSNR}(f) = \frac{\phi_{X_1}(f)}{\phi_{V_1}(f)}, \tag{13}$$

where $\phi_{X_1}(f) = E[|X_1(f)|^2]$ and $\phi_{V_1}(f) = E[|V_1(f)|^2]$ are the variances of $X_1(f)$ and $V_1(f)$, respectively, with $E[\cdot]$ being mathematical expectation. The output SNR is written, according to Eq. (11), as

$$\begin{aligned}\text{oSNR}[\mathbf{h}(f)] &= \frac{\phi_{X_1}(f)|\mathbf{h}^H(f)\mathbf{d}(f)|^2}{\mathbf{h}^H(f)\mathbf{\Phi}_\mathbf{v}(f)\mathbf{h}(f)}\\ &= \text{iSNR}(f) \times \frac{|\mathbf{h}^H(f)\mathbf{d}(f)|^2}{\mathbf{h}^H(f)\mathbf{\Gamma}_\mathbf{v}(f)\mathbf{h}(f)},\end{aligned} \tag{14}$$

where $\mathbf{\Phi}_\mathbf{v}(f) = E[\mathbf{v}(f)\mathbf{v}^H(f)]$ and $\mathbf{\Gamma}_\mathbf{v}(f) = \mathbf{\Phi}_\mathbf{v}(f)/\phi_{V_1}(f)$ are the correlation and pseudo-coherence matrices of $\mathbf{v}(f)$, respectively. From Eqs. (13) and (14), the SNR gain, also called the array gain, is easily derived, i.e.,

$$\mathcal{G}[\mathbf{h}(f)] = \frac{\text{oSNR}[\mathbf{h}(f)]}{\text{iSNR}(f)} = \frac{|\mathbf{h}^H(f)\mathbf{d}(f)|^2}{\mathbf{h}^H(f)\mathbf{\Gamma}_\mathbf{v}(f)\mathbf{h}(f)}. \tag{15}$$

This SNR gain depends on the noise pseudo-coherence matrix, $\mathbf{\Gamma}_\mathbf{v}(f)$. Let us consider the following two typical kinds of noise.

- The temporally and spatially white noise with the same variance at all microphones. In this case, the SNR gain

is called the WNG. With this type of noise, the noise pseudo-coherence matrix is $\mathbf{\Gamma}_{\mathbf{v}}(f) = \mathbf{I}_M$, where $\mathbf{I}_M$ is the $M \times M$ identity matrix. Substituting $\mathbf{\Gamma}_{\mathbf{v}}(f) = \mathbf{I}_M$ into Eq. (15), we get the WNG

$$\mathcal{W}[\mathbf{h}(f)] = \frac{|\mathbf{h}^H(f)\mathbf{d}(f)|^2}{\mathbf{h}^H(f)\mathbf{h}(f)}. \tag{16}$$

The WNG is generally used to measure the robustness of a beamformer against random errors such as the sensors' noise and mismatch among different sensors. If we use the Cauchy-Schwarz inequality, i.e.,

$$|\mathbf{h}^H(f)\mathbf{d}(f)|^2 \leq \mathbf{h}^H(f)\mathbf{h}(f) \times \mathbf{d}^H(f)\mathbf{d}(f), \tag{17}$$

it follows immediately that

$$\mathcal{W}[\mathbf{h}(f)] \leq M, \quad \forall \mathbf{h}(f). \tag{18}$$

- The diffuse noise (this situation corresponds to the spherically isotropic noise field). In this scenario, the SNR gain is called the DF. With this type of noise, the (i,j)th element of the noise pseudo-coherence matrix is

$$\begin{aligned}[\mathbf{\Gamma}_{\mathbf{v}}(f)]_{ij} &= [\mathbf{\Gamma}_{\mathrm{si}}(f)]_{ij} = \frac{\sin[2\pi f(j-i)\tau_0]}{2\pi f(j-i)\tau_0} \\ &= \mathrm{sinc}[2\pi f(j-i)\tau_0],\end{aligned} \tag{19}$$

where the subscript "$_{\mathrm{si}}$" stands for "spherically isotropic." Substituting Eq. (19) into Eq. (15) gives the DF, i.e.,

$$\mathcal{D}[\mathbf{h}(f)] = \frac{|\mathbf{h}^H(f)\mathbf{d}(f)|^2}{\mathbf{h}^H(f)\mathbf{\Gamma}_{\mathrm{si}}(f)\mathbf{h}(f)}. \tag{20}$$

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

$$|\mathbf{h}^H(f)\mathbf{d}(f)|^2 \leq \mathbf{h}^H(f)\mathbf{\Gamma}_{\mathrm{si}}(f)\mathbf{h}(f)\mathbf{d}^H(f)\mathbf{\Gamma}_{\mathrm{si}}^{-1}(f)\mathbf{d}(f), \tag{21}$$

it can be shown from Eq. (20) that

$$\begin{aligned}\mathcal{D}[\mathbf{h}(f)] &\leq \mathbf{d}^H(f)\mathbf{\Gamma}_{\mathrm{si}}^{-1}(f)\mathbf{d}(f) \leq \mathrm{tr}\left[\mathbf{\Gamma}_{\mathrm{si}}^{-1}(f)\mathbf{d}(f)\mathbf{d}^H(f)\right] \\ &\leq M\,\mathrm{tr}\left[\mathbf{\Gamma}_{\mathrm{si}}^{-1}(f)\right], \quad \forall \mathbf{h}(f),\end{aligned} \tag{22}$$

where $\mathrm{tr}[\cdot]$ denotes the trace of a square matrix.

The beampattern, which represents the response of $\mathbf{h}(f)$ to the incidence signal as a function of the steering vector, is defined as

$$\mathcal{B}[\mathbf{h}(f)] = \mathbf{h}^H(f)\mathbf{d}(f,\theta). \tag{23}$$

It can be proved that (Benesty and Chen, 2012)

$$\mathcal{B}[\mathbf{h}(f)] \approx \sum_{n=0}^{N} \cos^n\theta \left[\frac{(\jmath 2\pi f\tau_0)^n}{n!}\sum_{m=1}^{M}(m-1)^n H_m(f)\right]. \tag{24}$$

This shows how the general definition of the beampattern is strongly related to the frequency-independent beampattern of an Nth-order DMA.

Now, we define the error signal between the estimated and desired signals at frequency f as

$$\mathcal{E}(f) = Z(f) - X_1(f) = \mathbf{h}^H(f)\mathbf{y}(f) - X_1(f). \tag{25}$$

The MSE is then

$$\begin{aligned}J[\mathbf{h}(f)] &= E\left[|\mathcal{E}(f)|^2\right] = \phi_{X_1}(f) + \mathbf{h}^H(f)\mathbf{\Phi}_{\mathbf{y}}(f)\mathbf{h}(f) \\ &\quad - \phi_{X_1}(f)\mathbf{h}^H(f)\mathbf{d}(f) - \phi_{X_1}(f)\mathbf{d}^H(f)\mathbf{h}(f) \\ &= \phi_{X_1}(f)v_{\mathrm{sd}}[\mathbf{h}(f)] + \frac{\phi_{V_1}(f)}{\xi_{\mathrm{nr}}[\mathbf{h}(f)]} \\ &= J_{\mathrm{ds}}[\mathbf{h}(f)] + J_{\mathrm{rs}}[\mathbf{h}(f)],\end{aligned} \tag{26}$$

where

$$v_{\mathrm{sd}}[\mathbf{h}(f)] = \frac{E\left[|\mathbf{h}^H(f)\mathbf{x}(f) - X_1(f)|^2\right]}{\phi_{X_1}(f)} = |\mathbf{h}^H(f)\mathbf{d}(f) - 1|^2 \tag{27}$$

is the speech distortion index (SDI),

$$\xi_{\mathrm{nr}}[\mathbf{h}(f)] = \frac{\phi_{V_1}(f)}{\mathbf{h}^H(f)\mathbf{\Phi}_{\mathbf{v}}(f)\mathbf{h}(f)} = \frac{1}{\mathbf{h}^H(f)\mathbf{\Gamma}_{\mathbf{v}}(f)\mathbf{h}(f)} \tag{28}$$

is the noise reduction factor (NRF),

$$J_{\mathrm{ds}}[\mathbf{h}(f)] = \phi_{X_1}(f)v_{\mathrm{sd}}[\mathbf{h}(f)] \tag{29}$$

is the MSE due to the speech distortion, and

$$J_{\mathrm{rs}}[\mathbf{h}(f)] = \frac{\phi_{V_1}(f)}{\xi_{\mathrm{nr}}[\mathbf{h}(f)]} \tag{30}$$

is the MSE due to the residual noise.

IV. CONVENTIONAL BEAMFORMING ALGORITHMS

One of the most well-known beamformers for noise reduction is the minimum variance distortionless response (MVDR), which is derived by minimizing $\mathbf{h}^H(f)\mathbf{\Phi}_{\mathbf{y}}(f)\mathbf{h}(f)$, or $\mathbf{h}^H(f)\mathbf{\Phi}_{\mathbf{v}}(f)\mathbf{h}(f)$, or $J[\mathbf{h}(f)]$, subject to the distortionless constraint, i.e., $\mathbf{h}^H(f)\mathbf{d}(f) = 1$ (Capon, 1969; Lacoss, 1971). The solution is

$$\mathbf{h}_{\mathrm{MVDR}}(f) = \frac{\mathbf{\Phi}_{\mathbf{y}}^{-1}(f)\mathbf{d}(f)}{\mathbf{d}^H(f)\mathbf{\Phi}_{\mathbf{y}}^{-1}\mathbf{d}(f)} = \frac{\mathbf{\Phi}_{\mathbf{v}}^{-1}(f)\mathbf{d}(f)}{\mathbf{d}^H(f)\mathbf{\Phi}_{\mathbf{v}}^{-1}(f)\mathbf{d}(f)}. \tag{31}$$

In the presence of the spherically isotropic noise, Eq. (31) becomes the superdirective beamformer

$$\mathbf{h}_{\mathrm{SMVDR}}(f) = \frac{\mathbf{\Gamma}_{\mathrm{si}}^{-1}(f)\mathbf{d}(f)}{\mathbf{d}^H(f)\mathbf{\Gamma}_{\mathrm{si}}^{-1}(f)\mathbf{d}(f)}, \tag{32}$$

which corresponds to the hypercardioid of order $M-1$ (Benesty and Chen, 2012). Notice that Eq. (31) is data dependent while Eq. (32) is data independent. This indicates that the former may perform better than the latter if the noise field is not spherically isotropic and the statistics of the observation or noise signals can be estimated accurately.

It can be shown that (Uzkov, 1946)

$$\lim_{\delta\to 0}\mathcal{D}[\mathbf{h}_{\mathrm{SMVDR}}(f)]=M^2. \tag{33}$$

This large gain, though, comes at a price: white noise amplification. It is well known that $\mathcal{W}[\mathbf{h}_{\mathrm{SMVDR}}(f)]$ may be much smaller than 1, especially at low frequencies. In order to deal with this problem, the authors in Cox *et al.* (1986) and Cox *et al.* (1987) proposed a robust version of the superdirective beamformer

$$\mathbf{h}_{\mathrm{SMVDR},\epsilon}(f)=\frac{[\mathbf{\Gamma}_{\mathrm{si}}(f)+\epsilon\mathbf{I}_M]^{-1}\mathbf{d}(f)}{\mathbf{d}^H(f)[\mathbf{\Gamma}_{\mathrm{si}}(f)+\epsilon\mathbf{I}_M]^{-1}\mathbf{d}(f)}, \tag{34}$$

where $\epsilon\geq 0$ is the regularization parameter. This parameter tries to find a compromise between a supergain and white noise amplification. A small value of ϵ leads to a large DF but a small WNG, while a large value of ϵ gives a small DF but a large WNG. It is clear that

$$\lim_{\epsilon\to\infty}\mathcal{W}\left[\mathbf{h}_{\mathrm{SMVDR},\epsilon}(f)\right]=M. \tag{35}$$

V. NS ALGORITHMS

In the approach developed in this section, we ignore the distortionless constraint, which is used in the development of the MVDR filter, for example, as shown in Sec. IV. Instead, our focus is to constrain to have nulls in some specific directions and, at the same time, to make sure that the output of the filter is close to the desired signal, which propagates from the endfire direction.

A. NS in one direction

In the algorithm that we want to design, we desire to have a null in the direction $\theta_1\neq 0$. Now, our problem may be stated as follows. We wish to minimize the MSE, $J[\mathbf{h}(f)]$, subject to a null in the direction θ_1. Mathematically, this is equivalent to

$$\min_{\mathbf{h}(f)}J[\mathbf{h}(f)]\quad\text{subject to}\quad\mathbf{d}^H(f,\theta_1)\mathbf{h}(f)=0, \tag{36}$$

from which we derive the NS beamformer

$$\mathbf{h}_{\mathrm{NS}}(f,\theta_1)=\phi_{X_1}(f)\left[\mathbf{I}_M-\frac{\mathbf{\Phi}_{\mathbf{y}}^{-1}\mathbf{d}(f,\theta_1)\mathbf{d}^H(f,\theta_1)}{\mathbf{d}^H(f,\theta_1)\mathbf{\Phi}_{\mathbf{y}}^{-1}\mathbf{d}(f,\theta_1)}\right]\times\mathbf{\Phi}_{\mathbf{y}}^{-1}\mathbf{d}(f). \tag{37}$$

We can express Eq. (37) as

$$\mathbf{h}_{\mathrm{NS}}(f,\theta_1)=\mathbf{P}_{\mathbf{y}}(f,\theta_1)\mathbf{h}_{\mathrm{W}}(f), \tag{38}$$

where

$$\mathbf{P}_{\mathbf{y}}(f,\theta_1)=\mathbf{I}_M-\frac{\mathbf{\Phi}_{\mathbf{y}}^{-1}\mathbf{d}(f,\theta_1)\mathbf{d}^H(f,\theta_1)}{\mathbf{d}^H(f,\theta_1)\mathbf{\Phi}_{\mathbf{y}}^{-1}\mathbf{d}(f,\theta_1)} \tag{39}$$

is a projection matrix of rank $M-1$ and

$$\mathbf{h}_W(f)=\phi_{X_1}(f)\mathbf{\Phi}_{\mathbf{y}}^{-1}\mathbf{d}(f) \tag{40}$$

is the classical Wiener beamformer, which is obtained by minimizing $J[\mathbf{h}(f)]$.

Let us define

$$\mathbf{\Gamma}_{\mathbf{y}}(f)=\frac{\mathbf{\Phi}_{\mathbf{y}}(f)}{\phi_{Y_1}(f)}=\frac{\mathrm{iSNR}(f)}{1+\mathrm{iSNR}(f)}\mathbf{d}(f)\mathbf{d}^H(f)+\frac{1}{1+\mathrm{iSNR}(f)}\mathbf{\Gamma}_{\mathbf{v}}(f), \tag{41}$$

where $\phi_{Y_1}(f)=E[|Y_1(f)|^2]$ is the variance of $Y_1(f)$. Now, $\mathbf{\Gamma}_{\mathbf{y}}(f)$ depends on the input SNR and the pseudo-coherence matrix of the noise. Hence, we can rewrite Eq. (37) as

$$\mathbf{h}_{\mathrm{NS}}(f,\theta_1)=H_W(f)\mathbf{P}_{\mathbf{y}}(f,\theta_1)\mathbf{\Gamma}_{\mathbf{y}}^{-1}(f)\mathbf{d}(f), \tag{42}$$

where

$$H_W(f)=\frac{\mathrm{iSNR(f)}}{1+\mathrm{iSNR}(f)} \tag{43}$$

is the single-channel Wiener gain.

In the presence of the spherically isotropic noise, Eq. (42) becomes the superdirective null-steering (SNS) beamformer

$$\mathbf{h}_{\mathrm{SNS}}(f,\theta_1)=H_W(f)\left[\mathbf{I}_M-\frac{\mathbf{\Gamma}_{\mathrm{si,y}}^{-1}\mathbf{d}(f,\theta_1)\mathbf{d}^H(f,\theta_1)}{\mathbf{d}^H(f,\theta_1)\mathbf{\Gamma}_{\mathrm{si,y}}^{-1}\mathbf{d}(f,\theta_1)}\right]\times\mathbf{\Gamma}_{\mathrm{si,y}}^{-1}\mathbf{d}(f), \tag{44}$$

where

$$\mathbf{\Gamma}_{\mathrm{si,y}}(f)=\frac{\mathrm{iSNR}(f)}{1+\mathrm{iSNR}(f)}\mathbf{d}(f)\mathbf{d}^H(f)+\frac{1}{1+\mathrm{iSNR}(f)}\mathbf{\Gamma}_{\mathrm{si}}(f). \tag{45}$$

Compared to $\mathbf{h}_{\mathrm{SMVDR}}(f,\theta_1)$, $\mathbf{h}_{\mathrm{SNS}}(f,\theta_1)$ is partially data dependent, since it depends on the input SNR. Nevertheless, it is possible to plot the beampattern of $\mathbf{h}_{\mathrm{SNS}}(f,\theta_1)$ for different fixed values of the input SNR. Notice that a robust version of Eq. (44) is easily obtained by replacing $\mathbf{\Gamma}_{\mathrm{si,y}}(f)$ with $\mathbf{\Gamma}_{\mathrm{si,y}}(f)+\epsilon\mathbf{I}_M$.

B. NS in multiple directions

Assume that we want N nulls in the N different directions $\theta_1\neq\theta_2\neq\cdots\neq\theta_N\neq 0$. In this case, we have N constraints of the form

$$\mathbf{d}^H(f,\theta_n)\mathbf{h}(f)=0,\quad n=1,2,\ldots,N. \tag{46}$$

Combining these constraints together, we get

$$\mathbf{D}(f,\boldsymbol{\theta})\mathbf{h}(f)=\mathbf{0}_{N\times 1}, \tag{47}$$

where

$$\mathbf{D}(f,\boldsymbol{\theta})=\begin{bmatrix}\mathbf{d}^H(f,\theta_1)\\ \mathbf{d}^H(f,\theta_2)\\ \vdots\\ \mathbf{d}^H(f,\theta_N)\end{bmatrix} \tag{48}$$

is a matrix of size $N \times M$ and $\boldsymbol{\theta}$ is a vector containing all the null information. Now, our optimization problem is as follows:

$$\min_{\mathbf{h}(f)} J[\mathbf{h}(f)] \quad \text{subject to} \quad \mathbf{D}(f,\boldsymbol{\theta})\mathbf{h}(f) = \mathbf{0}_{N\times 1}. \tag{49}$$

The solution is

$$\mathbf{h}_{\mathrm{NS}}(f,\boldsymbol{\theta}) = \mathbf{P}_{\mathbf{y}}(f,\boldsymbol{\theta})\mathbf{h}_W(f), \tag{50}$$

where

$$\mathbf{P}_{\mathbf{y}}(f,\boldsymbol{\theta}) = \mathbf{I}_M - \boldsymbol{\Phi}_{\mathbf{y}}^{-1}(f)\mathbf{D}^H(f,\boldsymbol{\theta}) \times \left[\mathbf{D}(f,\boldsymbol{\theta})\boldsymbol{\Phi}_{\mathbf{y}}^{-1}(f)\mathbf{D}^H(f,\boldsymbol{\theta})\right]^{-1}\mathbf{D}(f,\boldsymbol{\theta}) \tag{51}$$

is a projection matrix of rank $M-N$ and $\mathbf{h}_W(f)$ is defined in Eq. (40). We see from Eq. (51) that we must have $N \leq M$. If $N=M$, we have $\mathbf{h}_{\mathrm{NS}}(f,\boldsymbol{\theta}) = \mathbf{0}_{M\times 1}$, which should be avoided in practice. Therefore, we should always take $N<M$, which is the same condition to design an Nth-order DMA (Benesty and Chen, 2012).

By following the same procedure of Sec. V A, we find the SNS beamformer as

$$\mathbf{h}_{\mathrm{SNS}}(f,\boldsymbol{\theta}) = \mathbf{P}_{\mathrm{si,y}}(f,\boldsymbol{\theta})\mathbf{h}_{\mathrm{SW}}(f), \tag{52}$$

where

$$\mathbf{P}_{\mathrm{si,y}}(f,\boldsymbol{\theta}) = \mathbf{I}_M - \boldsymbol{\Gamma}_{\mathrm{si,y}}^{-1}(f)\mathbf{D}^H(f,\boldsymbol{\theta}) \times \left[\mathbf{D}(f,\boldsymbol{\theta})\boldsymbol{\Gamma}_{\mathrm{si,y}}^{-1}(f)\mathbf{D}^H(f,\boldsymbol{\theta})\right]^{-1}\mathbf{D}(f,\boldsymbol{\theta}) \tag{53}$$

and

$$\mathbf{h}_{\mathrm{SW}}(f) = H_W(f)\boldsymbol{\Gamma}_{\mathrm{si,y}}^{-1}(f)\mathbf{d}(f) \tag{54}$$

is the superdirective Wiener beamformer.

We conclude this subsection by proposing the following beamformer:

$$\mathbf{h}_\alpha(f) = \phi_{X_1}(f)\left[\boldsymbol{\Phi}_{\mathbf{y}}(f) + \alpha\mathbf{D}^H(f,\boldsymbol{\theta})\mathbf{D}(f,\boldsymbol{\theta})\right]^{-1}\mathbf{d}(f), \tag{55}$$

where $\alpha \geq 0$ and

$$\begin{aligned}\left[\boldsymbol{\Phi}_{\mathbf{y}}(f) + \alpha\mathbf{D}^H(f,\boldsymbol{\theta})\mathbf{D}(f,\boldsymbol{\theta})\right]^{-1} &= \boldsymbol{\Phi}_{\mathbf{y}}^{-1}(f) - \boldsymbol{\Phi}_{\mathbf{y}}^{-1}(f)\mathbf{D}^H(f,\boldsymbol{\theta}) \\ &\times \left[\alpha^{-1}\mathbf{I}_N + \mathbf{D}(f,\boldsymbol{\theta})\boldsymbol{\Phi}_{\mathbf{y}}^{-1}(f)\mathbf{D}^H(f,\boldsymbol{\theta})\right]^{-1} \\ &\times \mathbf{D}(f,\boldsymbol{\theta})\boldsymbol{\Phi}_{\mathbf{y}}^{-1}(f),\end{aligned} \tag{56}$$

with $\mathbf{I}_N$ being the $N\times N$ identity matrix. We see that for $\alpha=0$ we get $\mathbf{h}_W(f)$ and for $\alpha=\infty$ we get $\mathbf{h}_{\mathrm{NS}}(f)$.

C. Modified MSE criterion

In this subsection, we propose to use a more general normalized MSE criterion

$$J_\mu[\mathbf{h}(f)] = \mu(f)\frac{J_{\mathrm{ds}}[\mathbf{h}(f)]}{\phi_{X_1}(f)} + \frac{J_{\mathrm{rs}}[\mathbf{h}(f)]}{\phi_{V_1}(f)}, \tag{57}$$

where $\mu(f) \geq 0$. Now, by minimizing $J_\mu[\mathbf{h}(f)]$ subject to $\mathbf{D}(f,\boldsymbol{\theta})\mathbf{h}(f) = \mathbf{0}_{N\times 1}$, we deduce that

$$\mathbf{h}_{\mathrm{NS},\mu}(f,\boldsymbol{\theta}) = \mu(f)\mathbf{P}_{\mathbf{y},\mu}(f,\boldsymbol{\theta})\boldsymbol{\Gamma}_{\mathbf{y},\mu}^{-1}(f)\mathbf{d}(f), \tag{58}$$

where

$$\mathbf{P}_{\mathbf{y},\mu}(f,\boldsymbol{\theta}) = \mathbf{I}_M - \boldsymbol{\Gamma}_{\mathbf{y},\mu}^{-1}(f)\mathbf{D}^H(f,\boldsymbol{\theta}) \times \left[\mathbf{D}(f,\boldsymbol{\theta})\boldsymbol{\Gamma}_{\mathbf{y},\mu}^{-1}(f)\mathbf{D}^H(f,\boldsymbol{\theta})\right]^{-1}\mathbf{D}(f,\boldsymbol{\theta}) \tag{59}$$

and

$$\boldsymbol{\Gamma}_{\mathbf{y},\mu}^{-1}(f) = \boldsymbol{\Gamma}_{\mathbf{v}}^{-1}(f) - \frac{\boldsymbol{\Gamma}_{\mathbf{v}}^{-1}(f)\mathbf{d}(f)\mathbf{d}^H(f)\boldsymbol{\Gamma}_{\mathbf{v}}^{-1}(f)}{\mu^{-1}(f) + \mathbf{d}^H(f)\boldsymbol{\Gamma}_{\mathbf{v}}^{-1}(f)\mathbf{d}(f)}. \tag{60}$$

It can be verified that for $\mu(f) = \mathrm{iSNR}(f)$, we have $\mathbf{h}_{\mathrm{NS},\mu}(f,\boldsymbol{\theta}) = \mathbf{h}_{\mathrm{NS}}(f,\boldsymbol{\theta})$ and a larger value of $\mu(f)$ implies lesser distortion to the desired signal.

D. Estimation of the input SNR

We assume that the noise field is spherically isotropic. In this case, $\boldsymbol{\Gamma}_{\mathbf{y}}(f)$ coincides with $\boldsymbol{\Gamma}_{\mathrm{si,y}}(f)$. Since $\mathbf{y}(f)$ is observable, it is easy to estimate $\boldsymbol{\Gamma}_{\mathbf{y}}(f)$. We denote this estimate by $\widehat{\boldsymbol{\Gamma}}_{\mathbf{y}}(f)$. By following the approaches developed in Zelinski (1988), Meyer and Simmer (1997), McCowan and Bourlard (2003), and Lefkimmiatis and Maragos (2007), we can write the components of the matrix $\widehat{\boldsymbol{\Gamma}}_{\mathbf{y}}(f)$ as

$$\Re\left\{\left[\widehat{\boldsymbol{\Gamma}}_{\mathbf{y}}(f)\right]_{ij}\right\} = \frac{\widehat{\mathrm{iSNR}}(f)}{1+\widehat{\mathrm{iSNR}}(f)}\Re\left[D_i(f)D_j^*(f)\right] + \frac{1}{1+\widehat{\mathrm{iSNR}}(f)}\left[\boldsymbol{\Gamma}_{\mathrm{si}}(f)\right]_{ij}, \tag{61}$$

$$\Im\left\{\left[\widehat{\boldsymbol{\Gamma}}_{\mathbf{y}}(f)\right]_{ij}\right\} = \frac{\widehat{\mathrm{iSNR}}(f)}{1+\widehat{\mathrm{iSNR}}(f)}\Im\left[D_i(f)D_j^*(f)\right], \tag{62}$$

for $i \neq j$, $i,j = 1,2,\ldots,M$, where $\Re[\cdot]$ and $\Im[\cdot]$ denote the real part and imaginary part operators, respectively, $\widehat{\mathrm{iSNR}}(f)$ is the estimate of $\mathrm{iSNR}(f)$, and $D_m(f)$ is the mth element of $\mathbf{d}(f)$. We deduce from the previous expressions that

$$\widehat{\mathrm{iSNR}}(f) = \frac{\Re\left\{\left[\widehat{\boldsymbol{\Gamma}}_{\mathbf{y}}(f)\right]_{ij}\right\} - \left[\boldsymbol{\Gamma}_{\mathrm{si}}(f)\right]_{ij}}{\Re\left[D_i(f)D_j^*(f)\right] - \Re\left\{\left[\widehat{\boldsymbol{\Gamma}}_{\mathbf{y}}(f)\right]_{ij}\right\}}, \quad i \neq j, \tag{63}$$

and

$$\widehat{\mathrm{iSNR}}(f) = \frac{\Im\left\{\left[\widehat{\boldsymbol{\Gamma}}_{\mathbf{y}}(f)\right]_{ij}\right\}}{\Im\left[D_i(f)D_j^*(f)\right] - \Im\left\{\left[\widehat{\boldsymbol{\Gamma}}_{\mathbf{y}}(f)\right]_{ij}\right\}}, \quad i \neq j. \tag{64}$$

To get a more reliable estimate, it is better to average Eqs. (63) and (64) over all possible sensor combinations (Zelinski, 1988; Meyer and Simmer, 1997; McCowan and Bourlard, 2003; Lefkimmiatis and Maragos, 2007), resulting in the estimator

$$\widehat{\mathrm{iSNR}}(f) = \frac{1}{M(M-1)} \times \sum_{i=1}^{M-1}\sum_{j=i+1}^{M} \frac{\Re\left\{\left[\widehat{\mathbf{\Gamma}}_{\mathbf{y}}(f)\right]_{ij}\right\} - [\mathbf{\Gamma}_{\mathrm{si}}(f)]_{ij}}{\Re\left[D_i(f)D_j^*(f)\right] - \Re\left\{\left[\widehat{\mathbf{\Gamma}}_{\mathbf{y}}(f)\right]_{ij}\right\}} + \frac{1}{M(M-1)} \times \sum_{i=1}^{M-1}\sum_{j=i+1}^{M} \frac{\Im\left\{\left[\widehat{\mathbf{\Gamma}}_{\mathbf{y}}(f)\right]_{ij}\right\}}{\Im\left[D_i(f)D_j^*(f)\right] - \Im\left\{\left[\widehat{\mathbf{\Gamma}}_{\mathbf{y}}(f)\right]_{ij}\right\}}. \quad (65)$$

VI. SIMULATIONS

In this section, simulations are carried out to evaluate the performance of the developed NS beamforming algorithms with the performance measures described in Sec. III. A uniform linear microphone array is configured and the spacing between two successive sensors is set to $\delta = 1.1$ cm. The steering angle is set to 0° (endfire direction). We consider to design the first-, second-, and third-order supercardioid patterns, which are given as follows (Benesty and Chen, 2012; Sena *et al.*, 2012):

$$\mathcal{B}_1(\theta) = 0.414 + 0.586\cos\theta, \quad (66)$$

$$\mathcal{B}_2(\theta) = 0.103 + 0.484\cos\theta + 0.413\cos^2\theta, \quad (67)$$

$$\mathcal{B}_3(\theta) = 0.022 + 0.217\cos\theta + 0.475\cos^2\theta + 0.286\cos^3\theta. \quad (68)$$

A. Performance of NS to one direction

Let us first consider the SNS beamformer given in Eq. (44) with two microphones. We want to design a first-order supercardioid pattern. According to Eq. (66), the single null of the first-order supercardioid pattern is at $\theta = 135°$, so the SNS beamformer is designed accordingly. We investigate three narrowband input SNR (i.e., the input SNR at the frequency bin f) conditions: −10 dB, 0 dB, and 10 dB.

Figure 2 plots the beampatterns of the SNS beamformer for the three studied input SNR conditions and at four different frequencies. As seen, the beampatterns are almost frequency invariant. When the input SNR is 10 dB, the designed beampattern is almost identical to the theoretical pattern as given in Eq. (66). As the input SNR decreases, the shape of the designed beampattern is still frequency invariant and similar to the theoretical first-order supercardioid, but the overall gain decreases. The underlying reason is that no distortionless constraint is applied to the look direction. Note, however, that this gain attenuation does not introduce signal distortion but purely a gain in signal level since the design beampattern is invariant with frequency. In practice, a simple normalization can help deal with this gain attenuation problem.

Figure 3 plots the DF, WNG, NRF, and SDI of the SNS beamformer with respect to frequency at different SNR levels. One can see that the SNS beamformer has a constant DF, which is invariant with frequency and input SNR. With two microphones, the DF is approximately 5.3 dB, which is the same as the theoretical value of the DF of the first-order supercardioid DMA. The WNG also does not change with the input SNR. The value of WNG increases as frequency, which is the same as the first-order supercardioid DMA designed with either the traditional cascaded subtraction or the null-constrained methods (Benesty and Chen, 2012). The fact that both the DF and the WNG do not change with the input SNR indicates that the SNS beamformer works as a first-order DMA with the supercardioid pattern regardless of what input SNR value is used. From Figs. 3(c) and 3(d), one can see that both the SDI and the NRF increases as the SNR decreases; so, more noise is reduced in a lower SNR condition, which is the same as the classical noise reduction techniques (Benesty *et al.*, 2009). Note that given an SNR level, both the SDI and the NRF are constant over frequencies in the studied frequency range. The fact that the SDI does not vary with frequency corroborates with the observation that the SNS beamformer does not introduce distortion to the signal from the look direction but purely an attenuation of the signal level as long as the subband SNRs from different frequency bands stay the same.

B. Performance of NS to multiple directions

Now, we evaluate the SNS beamformer with multiple nulls, which was given in Eq. (50). We consider two cases: (1) designing a second-order supercardioid pattern as given in Eq. (67) with three microphones, and (2) designing a third-order supercardioid pattern as given in Eq. (68) with four microphones.

In the first scenario, the two nulls are at, respectively, $\theta_1 = 106°$ and $\theta_2 = 153°$. Again, we consider three narrowband input SNR conditions: −10 dB, 0 dB, and 10 dB, respectively. The results are plotted in Figs. 4 and 5. Again, at 10-dB input SNR, the designed beampattern is identical to the theoretical pattern as given in Eq. (67). As the input SNR decreases, there is a gain attenuation; but the shape of the designed beampattern is similar to the theoretical second-order supercardioid and is frequency invariant. We again, observe that both the DF and the WNG are independent of the input SNR level and are same as those of the second-order supercardioid DMA designed by the traditional cascaded subtraction or the null-constrained methods (Benesty and Chen, 2012).

In the second case, the three nulls are at, respectively, $\theta_1 = 98°$, $\theta_2 = 125°$, and $\theta_3 = 160°$. The results for the second case are plotted in Figs. 6 and 7. Similar observations to the previous case can be made from the results.

C. Performance with the modified MSE criterion

In this subsection, we evaluate the NS beamformer derived from the more general normalized MSE criterion

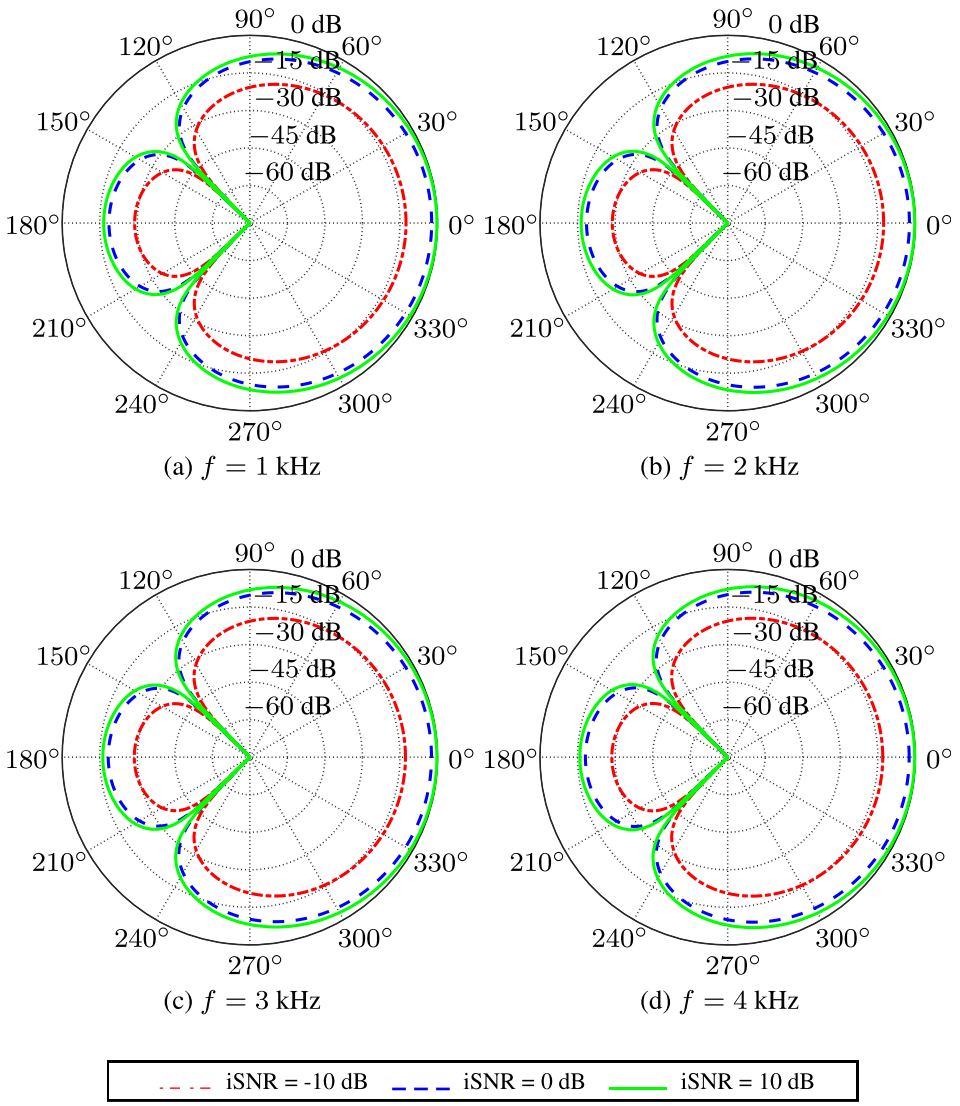


FIG. 2. (Color online) Beampatterns of the SNS beamformer with one null at $\theta = 135°$. $M = 2$ and $\delta = 1.1$ cm.

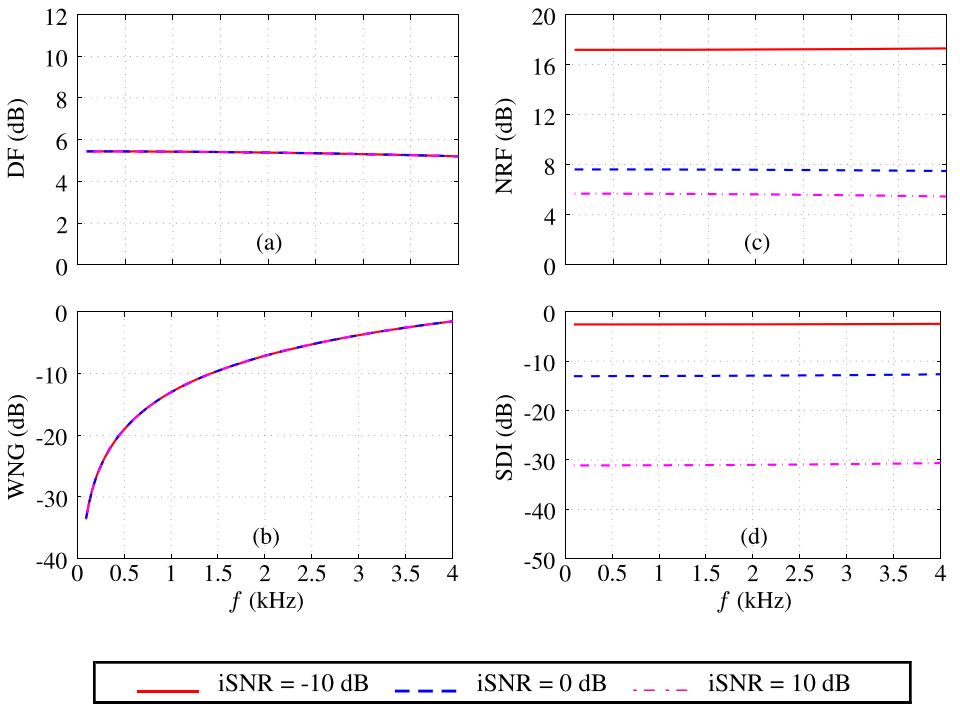


FIG. 3. (Color online) Performance of the SNS beamformer as a function of frequency and with one null at $\theta = 135°$: (a) DF, (b) WNG, (c) noise reduction factor, and (d) SDI. $M = 2$ and $\delta = 1.1$ cm.

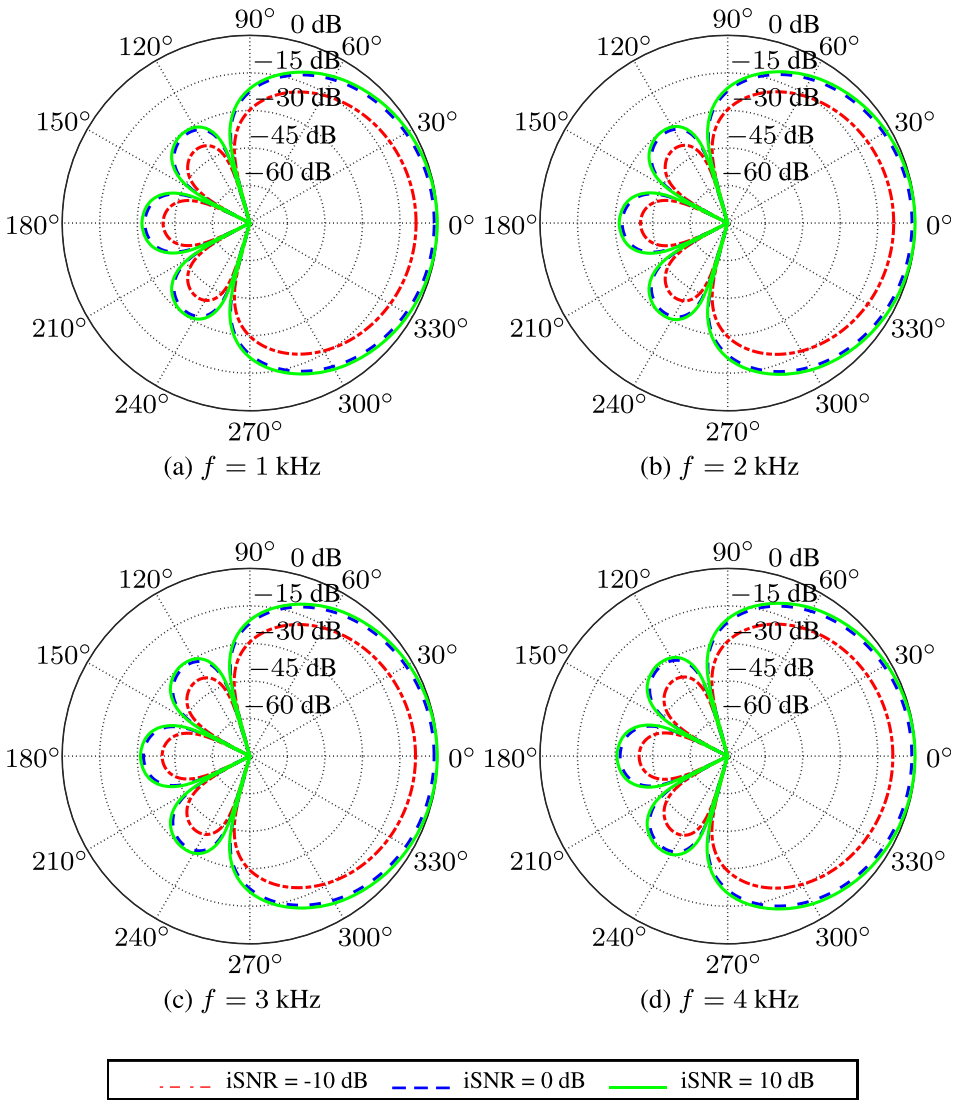


FIG. 4. (Color online) Beampatterns of the SNS beamformer with two nulls at $\theta_1 = 106°$ and $\theta_2 = 153°$. $M = 3$ and $\delta = 1.1$ cm.

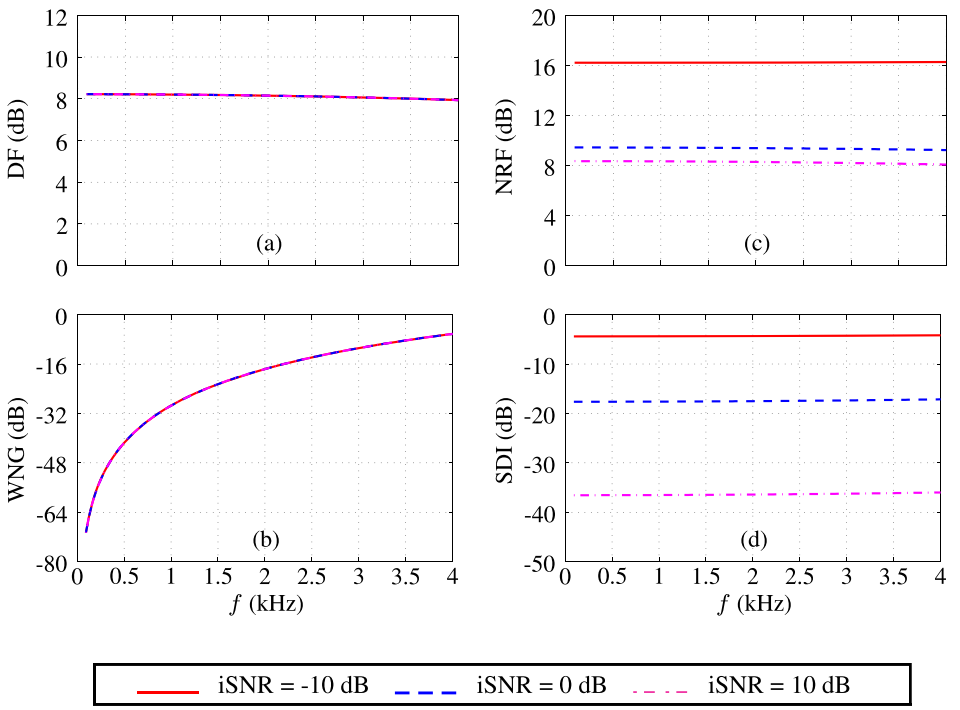


FIG. 5. (Color online) Performance of the SNS beamformer (with two nulls at $\theta_1 = 106°$ and $\theta_2 = 153°$), as a function of frequency: (a) DF, (b) WNG, (c) noise reduction factor, and (d) SDI. $M = 3$ and $\delta = 1.1$ cm.

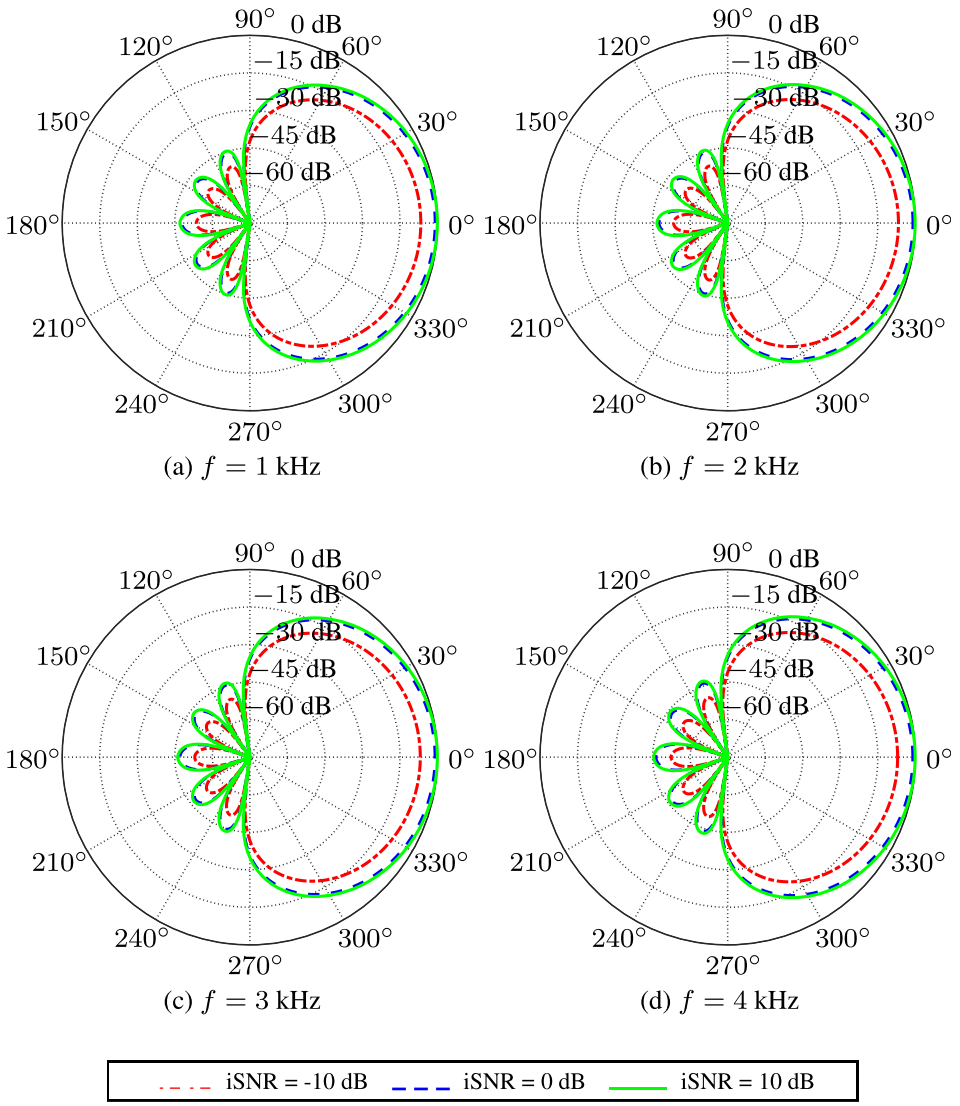


FIG. 6. (Color online) Beampatterns of the SNS beamformer with three nulls at $\theta_1 = 98°$, $\theta_2 = 125°$, and $\theta_3 = 160°$. $M = 4$ and $\delta = 1.1$ cm.

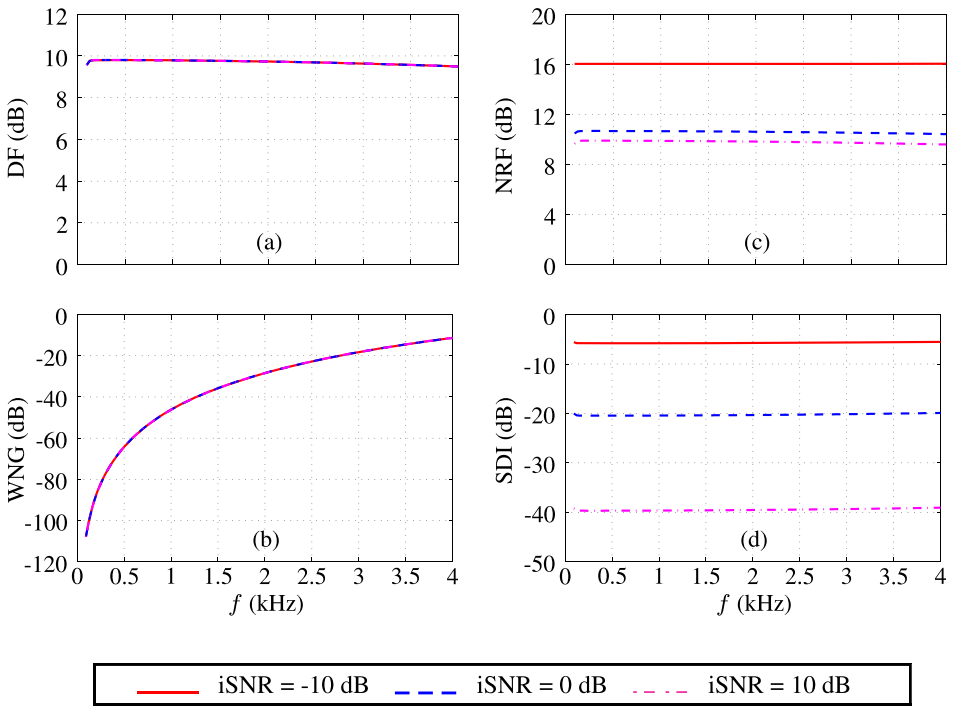


FIG. 7. (Color online) Performance of the SNS beamformer (with three nulls at $\theta_1 = 98°$, $\theta_2 = 125°$, and $\theta_3 = 160°$), as a function of frequency: (a) DF, (b) WNG, (c) noise reduction factor, and (d) SDI. $M = 4$ and $\delta = 1.1$ cm.

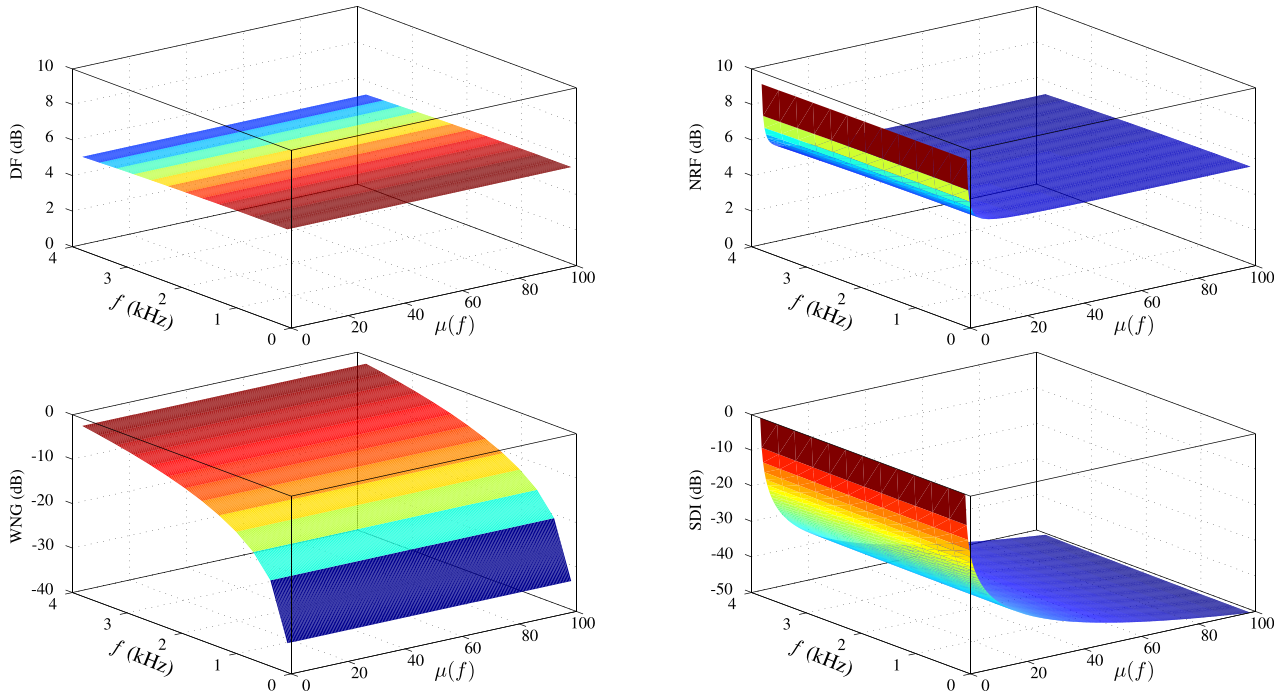


FIG. 8. (Color online) Performance of the SNS beamformer derived in Eq. (58) as a function of frequency and the parameter $\mu(f)$, and with one null at $\theta = 135^\circ$. $M = 2$ and $\delta = 1.1$ cm.

given in Eq. (58). Note that this NS beamformer is independent of the input SNR. We consider to design the first-order supercardiod pattern as given in Eq. (66) with two microphones. The results, as a function of the frequency and the value of the parameter $\mu(f)$, are plotted in Fig. 8. As seen, the DF is a constant and it does not change with either frequency or the value of $\mu(f)$. The WNG decreases with frequency, but at a given frequency, the WNG of this beamformer does not vary with $\mu(f)$. Therefore, this beamformer works as a DMA with the supercardiod pattern. According to Eq. (57), there is a tradeoff between the NRF and the SDI by setting different values of $\mu(f)$, which is verified from Fig. 8. Generally, the larger the value of $\mu(f)$, the smaller are the SDI and the NRF, which is, again, similar to the classical noise reduction techniques. Note if we set $\mu(f) = \text{iSNR}$, the NS algorithm under the normalized MSE criterion degenerates to the NS beamformer derived from the MSE criterion. However, with the normalized MSE criterion, the NS beamformer does not depend on the input SNR. Therefore, normalization is not needed to deal with the gain attenuation problem.

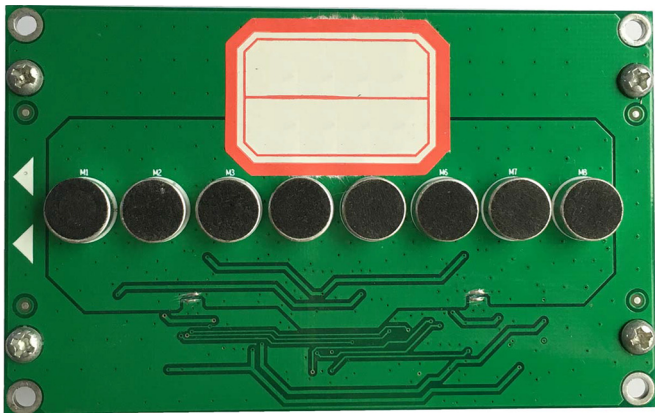
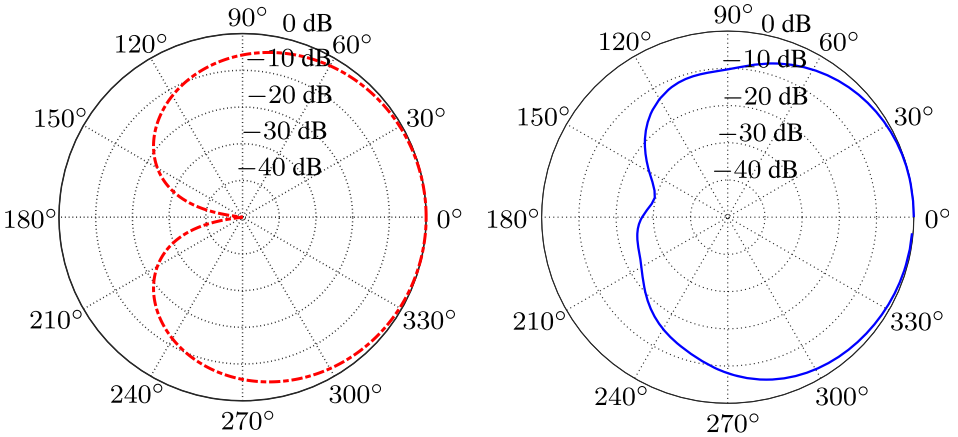


FIG. 9. (Color online) A photo of the designed eight-microphone array.

VII. EXPERIMENTS

To further validate the developed NS beamforming algorithms, we carried out some experiments. A microphone array is designed and fabricated, which consists of eight electret microphones. The spacing between two neighboring microphones is 1.15 cm. The microphone outputs are passed through a preamplifier that has an adjustable gain of 26–36 dB and a low-pass filter with a cutoff frequency of 8 kHz. The signals are then fed to a 24-bit analog-to-digital converter (ADC) with a sampling rate of 16 kHz and the digitized signals are subsequently passed on to a TI floating-point digital signal

FIG. 10. (Color online) The experimental setup for measuring the real beampatterns.



(a) Theoretical beampattern (b) Measured beampattern

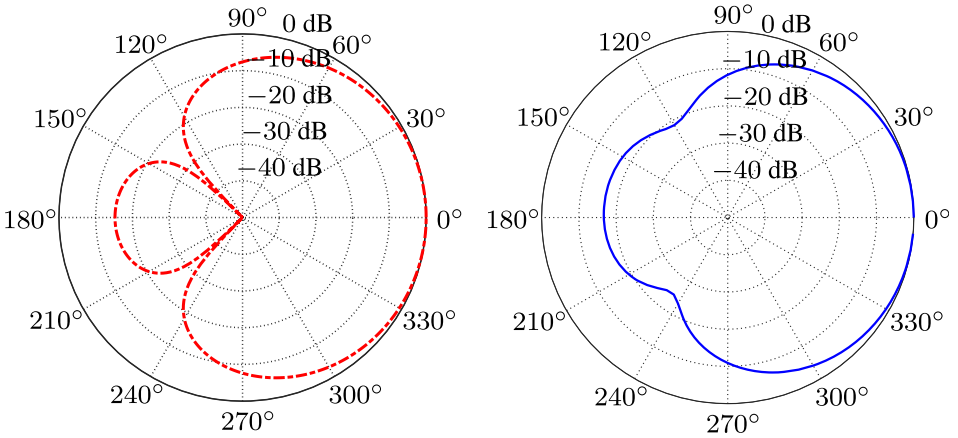
FIG. 11. (Color online) Designed and measured beampatterns of the NS beamformer with one null at $\theta = 180°$. $M = 2$, $\delta = 1.15$ cm, and $f = 2.1$ kHz.

processor (TMS320C6726, Texas Instruments, Dallas, TX) for processing. A photo of the designed microphone array is shown in Fig. 9.

The null beamforming algorithms were implemented using the framework presented in Benesty and Chen (2012) and Chen *et al*. (2014). The eight-channel microphone signals are partitioned into short frames with a frame length of 8 ms and each frame is transformed into the STFT domain using the fast Fourier transform (FFT) with a length of 128. In each STFT subband, NS beamformers are designed and applied. The subband beamforming outputs are then transformed into the time domain to produce the broadband beamforming output using the inverse fast Fourier transform (IFFT). To avoid aliasing, the overlap add technique is used, where the overlap factor is 75%.

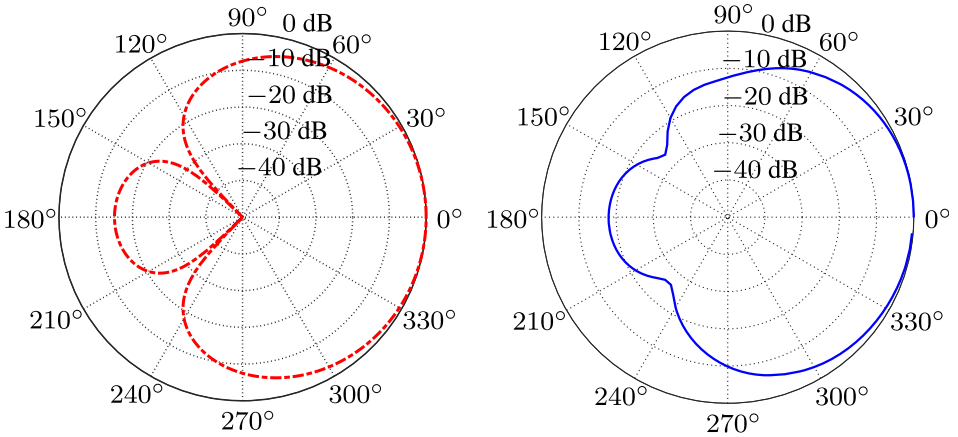
To measure the real beampatterns, the microphone array is placed on a rotating platform in the anechoic chamber at the Center of Intelligent Acoustic and Immersive Communication, Northwestern Polytechnic University. The chamber is 11.8 m long × 4.2 m wide × 3.8 m high. The rotating platform is placed in the center of the floating floor of the chamber, as shown in Fig. 10. To simulate a source, a loudspeaker is placed on the same horizontal plane as the microphone array. The distance between the loudspeaker and the center of the array is 1 m. The loudspeaker is configured to play back a narrowband sinusoidal signal with a specified frequency for 5 s and then it keeps silence for 2 s. This process is repeated till the measurement is completed. The rotating platform is configured to rotate the microphone array clockwise with respect to the array center from 0° to 360° with an increment of 5°. The array stays at every direction for 5 s before moving to the next direction. The rotating platform is strictly synchronized with the loudspeaker signal so that the platform rotation noise only happens during loudspeaker silence periods and does not affect the performance measurement. The beampattern is finally obtained by computing the array again at different angles based on the loudspeaker signal and the array beamforming output.

We measured the beampatterns for different orders of NS beamformers at various frequencies and report only a few cases in this section. In the first experiment, we consider the design of a cardioid pattern using two microphones, where there is one null at 180°. Both the designed (theoretical) and measured beampatterns at 2.1 kHz are plotted in Fig. 11. In general, the measured beampattern is close to the designed beampattern, but there are some discrepancies: (1) the null of the measured beampattern is slightly away from 180°, and (2) the null of the measured beampattern is not as deep as that in the designed beampattern. The underlying major reasons could be as follows. First, to make the processing delay small for real-time communication, we used a 128-point frame length



(a) Theoretical beampattern (b) Measured beampattern

FIG. 12. (Color online) Designed and measured beampatterns of the NS beamformer with one null at $\theta = 135°$. $M = 2$, $\delta = 1.15$ cm, and $f = 1.1$ kHz.



(a) Theoretical beampattern (b) Measured beampattern

FIG. 13. (Color online) Designed and measured beampatterns of the NS beamformer with one null at $\theta = 135°$. $M = 2$, $\delta = 1.15$ cm, and $f = 2.1$ kHz.

and FFT. So, each STFT band is not a single frequency, but has a bandwidth of 123 Hz. Second, we used a loudspeaker to simulate the source. The size of the loudspeaker is approximately 10 cm wide and 12 cm high, which is not a point source as assumed in the theoretical model. Third, there is some mismatch between the microphone frequency responses (within 2 B in our experiments). Finally, the measurements were made every 5°. Nevertheless, even with the aforementioned reasons, the difference between the designed and measured beampatterns is negligible. The directivity index (DI), i.e., DF expressed in a dB scale, of the designed cardioid beampattern is 4.2 dB. The DI of the measured beampattern is 4.9 dB, which again indicates that the measured beampattern is similar to the designed one.

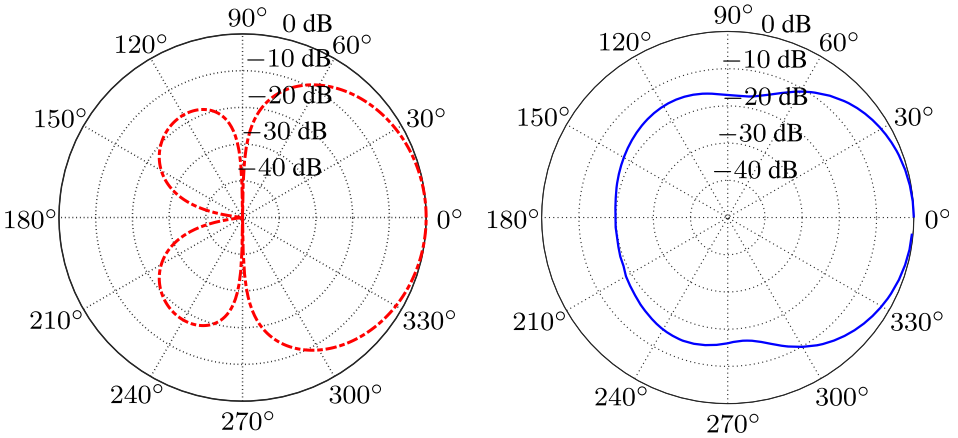
Note that the DI value computed from the measured beampattern is slightly higher than its theoretical counterparts. The reason could be that the measurements were made every 5°, which may lead to some error in computing the DI. We are working to get a rotating platform that can enable finer measurements at every 1°.

In the second experiment, we consider to design a supercardioid pattern using two microphones, which has a null at 135°. Figures 12 and 13 plot the designed and measured beampatterns at, respectively, 1.1 kHz and 2.1 kHz. Again, one can see that the measured beampatterns are close to the designed ones even though there exists some discrepancy. The DI for the designed supercardioid is 4.6 dB, while the DI of the measured beampatterns at 1.1 kHz and 2.1 kHz are, respectively, 5.2 dB and 5.3 dB, which, again, indicates that the measured beampattern is similar to the designed beampattern.

In the last experiment, we present the design of a second-order cardioid pattern using three microphones, which has two nulls, one at 120° and the other at 180°. Figure 14 plots the designed and measured beampatterns at 2.1 kHz. One can see that there are two nulls appearing in the measured beampattern at 120° and 180°, but the nulls are not as deep as those in the designed beampattern. Again, this happens mainly due to the short FFT length, the size of the loudspeaker, and the mismatch between sensors. The DI for the designed beampattern is 6.6 dB, and that for the measured beampattern is 6.8 dB. The two DI values agree fairly well.

VIII. CONCLUSIONS

In this paper, we developed a NS approach to beamforming with small-spacing uniform linear microphone arrays. Two criteria were defined, i.e., the MSE criterion and the normalized MSE criterion. Under the MSE criterion, we derived two NS algorithms, one considering only a null in one direction and the other considering nulls in multiple directions. Similarly, two NS algorithms were developed with the normalized MSE criterion. These algorithms can be used to



(a) Theoretical beampattern (b) Measured beampattern

FIG. 14. (Color online) Designed and measured beampatterns of the NS beamformer with two nulls at, respectively, $\theta_1 = 120°$ and $\theta_2 = 180°$. $M = 3$, $\delta = 1.15$ cm, and $f = 2.1$ kHz.

design either fixed or adaptive beamformers. To design fixed beamformers, the resulting beamformers work like DMAs as they exhibit frequency-invariant beampatterns. To design adaptive beamformers, the resulting beamformers can be viewed as a combination of DMAs and single-channel noise reduction since they do not only exhibit frequency-invariant beampatterns but also can achieve noise reduction and compromise between noise reduction and speech distortion. The algorithms with the MSE criterion are a function of the narrowband input SNR, therefore, a normalization procedure may be needed in order to keep the desired signal from the look direction not attenuated. In comparison, the NS algorithm derived from the normalized MSE criterion is independent of the input SNR and as a result there will be no signal attenuation at the look direction. Simulations and experiments were carried out and the results validated the deduced algorithms.

ACKNOWLEDGMENT

This work was supported in part by the NSFC (National Natural Science Foundation of China) "Distinguished Young Scientists Fund" under Grant No. 61425005 and the NSFC-ISF (National Natural Science Foundation of China – Israel Science Foundation) joint research program under Grant No. 61761146001.

Benesty, J., and Chen, J. (**2012**). *Study and Design of Differential Microphone Arrays* (Springer-Verlag, Berlin, Germany), 182 pp.
Benesty, J., Chen, J., and Habets, E. (**2011**). *Speech Enhancement in the STFT Domain* (Springer, Berlin, Germany), 111 pp.
Benesty, J., Chen, J., and Huang, Y. (**2008**). *Microphone Array Signal Processing* (Springer, Berlin, Germany), 232 pp.
Benesty, J., Chen, J., Huang, Y., and Cohen, I. (**2009**). *Noise Reduction in Speech Processing* (Springer, Berlin, Germany), 229 pp.
Benesty, J., Chen, J., and Pan, C. (**2016**). *Fundamentals of Differential Beamforming* (Springer-Verlag, Berlin, Germany), 122 pp.
Benesty, J., Cohen, I., and Chen, J. (**2017**). *Fundamentals of Signal Enhancement and Array Signal Processing* (Wiley, Hoboken, NJ), 451 pp.
Bernardini, A., Aria, M. D., Sannino, R., and Sarti, A. (**2017**). "Efficient continuous beam steering for planar arrays of differential microphones," IEEE Signal Process. Lett. **24**, 794–798.
Brandstein, M., and Ward, D. B. (**2001**). *Microphone Arrays: Signal Processing Techniques and Applications* (Springer, Berlin, Germany), 398 pp.
Buckley, K. M. (**1986**). "Broad-band beamforming and the generalized sidelobe canceller," IEEE Trans. Acoust., Speech, Signal Process. **34**, 1322–1323.
Buckley, K. M. (**1987**). "Spatial/spectral filtering with linearly constrained minimum variance beamformers," IEEE Trans. Acoust., Speech, Signal Process. **35**, 249–266.
Buckley, K. W., and Griffiths, L. J. (**1986**). "An adaptive generalized sidelobe canceller with derivative constraints," IEEE Trans. Antennas Propag. **34**, 311–319.
Capon, J. (**1969**). "High-resolution frequency-wavenumber spectrum analysis," Proc. IEEE **57**, 1408–1418.
Chen, J., Benesty, J., and Pan, C. (**2014**). "On the design and implementation of linear differential microphone arrays," J. Acoust. Soc. Am. **136**, 3097–3113.
Chiba, I., Takahashi, T., and Karasawa, Y. (**1994**). "Transmitting null beam forming with beam space adaptive array antennas," in *Proc. IEEE 44th VTC*, pp. 1498–1502.
Compton, R. T. (**1988**). *Adaptive Antennas: Concepts and Performance* (Prentice-Hall, Englewood Cliffs, NJ), 400 pp.
Cox, H., Zeskind, R. M., and Kooij, T. (**1986**). "Practical supergain," IEEE Trans. Acoust., Speech, Signal Process. **34**, 393–398.
Cox, H., Zeskind, R. M., and Owen, M. M. (**1987**). "Robust adaptive beamforming," IEEE Trans. Acoust., Speech, Signal Process. **35**, 1365–1376.
Dudgeon, D. E. (**1977**). "Fundamentals of digital array processing," Proc. IEEE **65**, 898–904.
Elko, G. W. (**1997**). "Steerable and variable first order differential microphone array," U.S. patent 6,041,127 (March 21, 2000).
Elko, G. W. (**2000**). "Superdirectional microphone arrays," in *Acoustic Signal Processing for Telecommunication*, edited by S. L. Gay, and J. Benesty (Kluwer Academic Publishers, Boston, MA), Chap. 10, pp. 181–237.
Elko, G. W., and Meyer, J. (**2008**). "Microphone arrays," in *Springer Handbook of Speech Processing*, edited by J. Benesty, M. M. Sondhi, and Y. Huang (Springer-, Berlin, Germany), Chap. 50, pp. 1021–1041.
Elko, G. W., and Pong, A. T. N. (**1997**). "A steerable and variable first-order differential microphone array," in *Proc. IEEE ICASSP*, pp. 223–226.
Flanagan, J. L., Johnston, J. D., Zahn, R., and Elko, G. W. (**1985**). "Computer-steered microphone arrays for sound transduction in large rooms," J. Acoust. Soc. Am. **78**, 1508–1518.
Friedlander, B., and Porat, B. (**1989**). "Performance analysis of a null-steering algorithm based on direction-of-arrival estimation," IEEE Trans. Acoust., Speech, Signal Process. **37**, 461–466.
Frost, O. L. (**1972**). "An algorithm for linearly constrained adaptive array processing," Proc. IEEE **60**, 926–935.
Godara, L. C. (**1997**). "Application of antenna arrays to mobile communications, part II: Beam-forming and direction-of-arrival considerations," Proc. IEEE **85**, 1195–1245.
Godara, L. C., and Cantoni, A. (**1981**). "Uniqueness and linear independence of steering vectors in array space," J. Acoust. Soc. Am. **70**, 467–475.
Griffiths, L. J., and Jim, C. W. (**1982**). "An alternative approach to linearly constrained adaptive beamforming," IEEE Trans. Antennas Propag. **30**, 27–34.
Haupt, R. L. (**1997**). "Phase-only adaptive nulling with a genetic algorithm," IEEE Trans. Antennas Propag. **45**, 1009–1015.
Haykin, S. (**1985**). *Array Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ).
Howells, P. W. (**1976**). "Explorations in fixed and adaptive resolution at GE and SURC," IEEE Trans. Antennas Propag. **24**, 575–584.
Jim, C. W. (**1977**). "A comparison of two LMS constrained optimal array structures," Proc. IEEE **65**, 1730–1731.
Kellermann, W. (**1991**). "A self-steering digital microphone array," in *Proc. IEEE ICASSP*, Vol. 5, pp. 3581–3584.
Lacoss, R. T. (**1971**). "Data adaptive spectral analysis methods," Geophysics **36**, 661–675.
Lefkimmiatis, S., and Maragos, P. (**2007**). "A generalized estimation approach for linear and nonlinear microphone array post-filters," Speech Commun. **49**, 657–666.
Meyer, J., and Simmer, K. U. (**1997**). "Multi-channel speech enhancement in a car environment using wiener filtering and spectral subtraction," in *Proc. IEEE ICASSP*, Vol. 2, pp. 1167–1170.
McCowan, I. A., and Bourlard, H. (**2003**). "Microphone array post-filter based on noise field coherence," IEEE Trans. Speech, Audio Process. **11**, 709–716.
Ng, T., Cheah, J., and Paoloni, F. J. (**1985**). "Optimization with controlled null placement in antenna array pattern synthesis," IEEE Trans. Antennas Propag. **33**, 215–217.
Olson, H. F. (**1932**). "A uni-directional ribbon microphone," J. Acoust. Soc. Am. **3**, 315–316.
Olson, H. F. (**1946**). "Gradient microphones," J. Acoust. Soc. Am. **17**, 192–198.
Pan, C., Chen, J., and Benesty, J. (**2015**). "Theoretical analysis of differential microphone array beamforming and an improved solution," IEEE/ACM Trans. Audio, Speech, Lang. Process. **23**, 2093–2105.
Schelkunoff, S. A. (**1943**). "A mathematical theory of linear arrays," Bell Syst. Tech. J. **22**, 80–107.
Sena, E. De, Hacihabiboglu, H., and Cvetković, Z. (**2012**). "On the design and implementation of higher order differential microphones," IEEE Trans. Audio, Speech, Lang. Process. **20**, 162–174.
Shore, R. A. (**1984**). "Nulling at symmetric pattern location with phase-only weight control," IEEE Trans. Antennas Propag. **32**, 530–533.
Steyskal, H., Shore, R. A., and Haupt, R. L. (**1986**). "Methods for null control and their effects on the radiation pattern," IEEE Trans. Antennas Propag. **34**, 404–409.
Uzkov, A. L. (**1946**). "An approach to the problem of optimum directive antenna design," in *Comptes Rendus (Doklady) de lAcademie des Sciences de lURSS*, Vol. 53, pp. 35–38.

Van Veen, B. D., and Buckley, K. M. (**1988**). "Beamforming: A versatile approach to spatial filtering," IEEE ASSP Mag. **5**, 4–24.
Vu, T. (**1986**). "Simultaneous nulling in sum and difference patterns by amplitude control," IEEE Trans. Antennas Propag. **34**, 214–218.
Werner, S., Apolinario, J. A., and Campos, M. L. (**2003**). "On the equivalence of RLS implementations of LCMV and GSC processors," IEEE Signal Process. Lett. **10**, 356–359.
Wu, X., and Chen, H. (**2016**). "Directivity factors of the first-order steerable differential array with microphone mismatches: Deterministic and worst-case analysis," IEEE/ACM Trans. Audio, Speech, Lang. Process. **24**, 300–315.
Zelinski, R. (**1988**). "A microphone array with adaptive post-filtering for noise reduction in reverberant rooms," in *Proc. IEEE ICASSP*, pp. 2578–2581.
Zhao, L., Benesty, J., and Chen, J. (**2014**). "Design of robust differential microphone arrays," IEEE/ACM Trans. Audio, Speech, Lang. Process. **22**, 1455–1466.
Zhao, L., Benesty, J., and Chen, J. (**2016**). "Design of robust differential microphone arrays with the Jacobi-Anger expansion," Appl. Acoust. **110**, 194–206.