A flexible high directivity beamformer with spherical microphone arrays

Gongping Huang, Jingdong Chen, and Jacob Benesty

A flexible high directivity beamformer with spherical microphone arrays

Gongping Huang[a)] and Jingdong Chen[b)]

Center of Intelligent Acoustics and Immersive Communications, Northwestern Polytechnical University,
127 Youyi West Road, Xi'an 710072, China

Jacob Benesty

L'Institut National de la Recherche Scientifique, Énergie, Matériaux et Télécommunications,
University of Quebec, 800 de la Gauchetiere Ouest, Montreal, QC H5A 1K6, Canada

The maximum directivity (MD) beamformer with spherical microphone arrays has many salient features in processing broadband acoustic and speech signals while suppressing noise and reverberation; but it is sensitive to sensors' self-noise and mismatch among these sensors. One effective way to deal with this sensitivity is by increasing the number of microphones, thereby improving the so-called white noise gain (WNG), but this increase may lead to many other design issues in terms of cost, array aperture, and possibly other performance degradation. This paper is tackling this sensitivity problem and presents a flexible high directivity (HD) beamforming algorithm. By approximating the ideal directivity pattern and the beamformer's beampattern with spherical harmonic series, the relationship between the two is obtained. This relationship is subsequently used to deduce a flexible HD beamformer, which can improve the WNG while achieving a directivity factor (DF) between the DF of an Nth-order MD beamformer and that of an $(N-1)$th-order MD one. Also derived is the analytical link between the DF and a tuning parameter and the link between the WNG and this parameter. Based on these links, one can easily determine the optimal value of the tuning parameter once the value of the DF or the WNG is specified.
© 2018 Acoustical Society of America. https://doi.org/10.1121/1.5038275

I. INTRODUCTION

Microphone arrays, which refer to sound acquisition systems with multiple microphones to sample a sound field, are widely used in voice communications and human-machine interfaces. A core component of a microphone array system is beamforming, which takes the array observation signals as its inputs to recover a signal of interest and suppressing noise, reverberation, and interference (Brandstein and Ward, 2001; Benesty et al., 2008). A significant number of techniques have been developed, such as the superdirective beamformers (Cox et al., 1986), the robust superdirective beamformers (Doclo and Moonen, 2007; Wang et al., 2014), the tunable tradeoff superdirective beamformers (Crocco and Trucco, 2011), the subspace superdirective beamformers (Li et al., 2016; Huang et al., 2016a,b); the differential beamformers (Elko, 1996; Benesty and Chen, 2012), and adaptive beamformers (Monzingo and Miller, 1980; Gannot et al., 2004), to name but a few. Generally, the geometry of microphone arrays plays an important role on the design of the beamforming algorithm and its performance. In the literature, different geometries have been investigated, such as linear (Benesty and Chen, 2012), circular (Meyer, 2001; Benesty et al., 2015; Huang et al.,

2017a,b), spherical arrays (Meyer and Elko, 2008; Li and Duraiswami, 2007), and more general geometries (Abhayapala and Ward, 2002). Relatively, linear arrays are easy to design; but they are not very flexible in terms of beam steering, i.e., their performance varies with the look direction. In applications where full beam steering in the three-dimensional (3D) space is needed, meaning that the array response stays the same across different look directions, spherical arrays are generally preferable (Meyer and Elko, 2008; Abhayapala, 2008; Teutsch and Kellermann, 2006). As a result, beamforming with spherical microphone arrays (SMAs) has been intensively studied (Meyer and Elko, 2008; Rafaely et al., 2007a), and a number of algorithms have been developed such as sound field decomposition (Meyer and Elko, 2008; Rafaely, 2005; Zotkin et al., 2008), optimal modal beamforming (Yan et al., 2011), and eigenbeam processing (Meyer and Elko, 2008; Sun et al., 2012).

To process speech and audio signals, which are broadband in nature, it is important that the beamformer has a frequency-invariant beampattern and can perform consistently over a large frequency range of interest (Yan et al., 2007; Benesty and Chen, 2012) [note that there are also exceptional cases where frequency-independent beampatterns are desired to achieve certain spatial effects (Shabtai and Rafaely, 2014)]. One popular way to design a frequency-invariant beamformer with SMAs is by employing the spherical harmonic decomposition, which decomposes the sound field into a series of spherical harmonics.

The beamformer output is then obtained by combining spherical harmonics with proper weighting coefficients (Meyer and Elko, 2002; Rafaely and Khaykin, 2011; Teutsch, 2007). In the spherical harmonic domain, the directivity factor (DF), which is defined as the ratio between the directivity pattern at the look direction and the average of the directivity gain over other directions in the entire space, can be represented as a function of spherical harmonic series (Elko and Meyer, 2008). In this case, the maximization of the DF leads to the frequency-invariant maximum directivity (MD) beamformer, whose beampattern corresponds to the hypercardioid (Rafaely, 2015; Benesty and Chen, 2012). The value of the DF of the MD beamformer depends on the order (Elko and Meyer, 2008; Rafaely, 2015), i.e., for an Nth-order MD beamformer, its value is $(N + 1)^2$. Clearly, the higher is the order, the higher the value of the DF.

While it has a high DF and, therefore, it is effective in suppressing reverberation and spatial noise, the MD beamformer has a very low white noise gain (WNG) at low frequencies. So, it is sensitive to sensors' self-noise and mismatch among these sensors (Elko, 2000; Doclo and Moonen, 2007). This is one major problem that preventing this beamformer from being used in practice for high orders. Consequently, how to manage the WNG to a reasonable level while achieving a relative high value of DF with the MD beamformer has become an important issue. One way to do this is with the minimum-norm method by increasing the number of microphones while fixing the order of the MD beamformer (Benesty and Chen, 2012; Chen et al., 2014); but this would lead to many other design issues in terms of cost, array aperture, microphone mounting, and possibly other performance degradation as deep nulls (Rafaely et al., 2007b).

This paper is dedicated to the noise robustness problem of the MD beamformer with SMAs. We develop a flexible high directivity (HD) beamformer, which can improve the WNG and achieve a high DF without changing the number of microphones and the array geometry. The principle we take here is to improve the WNG while slightly sacrificing the value of the DF. However, unlike the minimum-norm approach that reduces the order of the MD beamformer from N to $N - 1$ or a lower positive integer number, we attempt to achieve a DF between the DF of the order N and that of the order $N - 1$. We first approximate the ideal directivity pattern and the SMA beamformer's beampattern with spherical harmonic series of order N, thereby building a relationship between the desired, target directivity pattern and the beamformer's beampattern. We then define a new directivity pattern, which is a linear combination of the ideal directivity pattern of the MD beamformer of order N and that of the MD beamformer of order $N - 1$. A flexible HD beamformer is then deduced, which can improve the WNG while achieving a DF between the DF of an Nth-order MD beamformer and that of an $(N - 1)$th-order MD one. There is one important tuning parameter, whose value is real and in the range between 0 and 1 and affects the performance of the flexible HD beamformer. To determine the proper value of this parameter, we derive the analytical link between the DF and the tuning parameter as well as the analytical link between the WNG and this parameter. Based on these links, we show that the optimal value of the tuning parameter can be easily computed once the value of the DF or the WNG is specified. This is very useful for practical applications since we can design beamformers with a certain level of the WNG and optimize the DF without solving the optimization problem numerically, or vice versa.

The remainder of this paper is organized as follows. Section II presents the signal model, problem formulation, and performance measures. We then discuss some fundamentals of spherical harmonics in Sec. III. Sections IV and V presents, respectively, the derivation of the MD beamformer and the flexible MD beamformer. Section VI discusses how to determine the value of the tuning parameter once the value of the DF or the WNG is specified. Simulation results are presented in Sec. VII to validate and compare the performance of the MD and flexible MD beamformers. Finally, some conclusions are given in Sec. VIII.

II. SIGNAL MODEL, PROBLEM FORMULATION, AND PERFORMANCE MEASURES

A. Signal model

Consider a spherical microphone array (SMA), of radius r, consisting of M omnidirectional microphones. We assume that the center of the SMA coincides with the origin of the three-dimensional Cartesian coordinate system, azimuth angles are measured anti-clockwise from the x axis, elevation angles are measured downward from the z axis. The mth ($m = 1, 2, …, M$) microphone's position is represented as $r \times \mathbf{r}_m$, where

$$\mathbf{r}_m = \begin{bmatrix} \sin\theta_m \cos\phi_m & \sin\theta_m \sin\phi_m & \cos\theta_m \end{bmatrix}^T, \quad (1)$$

superscript ^{T} denotes the transpose of a vector or a matrix, θ_m and ϕ_m are the elevation and azimuth angles of the mth microphone, respectively.

Suppose that we want to steer the beamformer in the direction (θ, ϕ), where $\theta \in [0, \pi]$, and $\phi \in [0, 2\pi)$. In this scenario, the steering vector corresponds to plane waves of a source in the farfield of length M is defined as

$$\mathbf{d}(\omega, \theta, \phi) = \begin{bmatrix} e^{j\varpi\mathbf{p}^T\mathbf{r}_1} & e^{j\varpi\mathbf{p}^T\mathbf{r}_2} & \cdots & e^{j\varpi\mathbf{p}^T\mathbf{r}_M} \end{bmatrix}^T, \quad (2)$$

where j is the imaginary unit with $j^2 = -1$, $\varpi = \omega r/c$, $\omega = 2\pi f$ is the angular frequency, $f > 0$ is the temporal frequency, and $\mathbf{p} = \begin{bmatrix} \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \end{bmatrix}^T$.

Now, let us assume that a source signal (plane wave), located in the farfield, propagates in an anechoic acoustic environment at the speed of sound, i.e., $c = 340$ m/s, and impinges on the SMA from the direction (θ_s, ϕ_s). Then, the received signal in the frequency domain can be written as

$$\mathbf{x}(\omega) = \begin{bmatrix} X_1(\omega) & X_2(\omega) & \cdots & X_M(\omega) \end{bmatrix}^T$$
$$= \mathbf{d}(\omega, \theta_s, \phi_s)S(\omega) + \mathbf{v}(\omega), \quad (3)$$

where $\mathbf{d}(\omega, \theta_s, \phi_s)$ is the source propagation vector of length M, $S(\omega)$ is the (zero-mean) source signal of interest (also

J. Acoust. Soc. Am. **143** (5), May 2018

Huang *et al.*   3025

called the desired signal), and $\mathbf{v}(\omega)$ is the noise vector defined in a similar way to $\mathbf{x}(\omega)$. Note that in this work, we assume that the steering angle (look direction) coincides with the source incidence angle (θ_s, ϕ_s).

B. Problem formulation

Beamforming is a process of applying a complex weight $H_m^*(\omega)$, where the asterisk denotes complex conjugation, to the microphone signal, $X_m(\omega)$, and then summing together the weighted outputs to form an estimate of the desired source signal (Benesty et al., 2008), i.e.,

$$
\begin{aligned}
Z(\omega) &= \sum_{m=1}^{M} H_m^*(\omega) X_m(\omega) \\
&= \mathbf{h}^H(\omega) \mathbf{x}(\omega) \\
&= \mathbf{h}^H(\omega) \mathbf{d}(\omega, \theta_s, \phi_s) S(\omega) + \mathbf{h}^H(\omega) \mathbf{v}(\omega),
\end{aligned} \tag{4}
$$

where $Z(\omega)$ is the estimate of the desired source signal, the superscript ^{H} is the conjugate-transpose operator, and

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

is the beamforming filter of length M. The objective of beamforming is then to find an optimal filter, $\mathbf{h}(\omega)$, such that $Z(\omega)$ is a good estimate of $S(\omega)$. Usually, the distortionless constraint in the look direction is needed in our context, i.e.,

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

C. Performance measures

Generally, two kinds of metrics are used to evaluate a beamformer, i.e., the beampattern (also called the directivity pattern) and the gain in signal-to-noise-ratio (SNR).

The beampattern describes the sensitivity of the beamformer to a plane wave impinging on the array from the direction (θ, ϕ) (see Fig. 1). Mathematically, the beampattern with an SMA is defined as
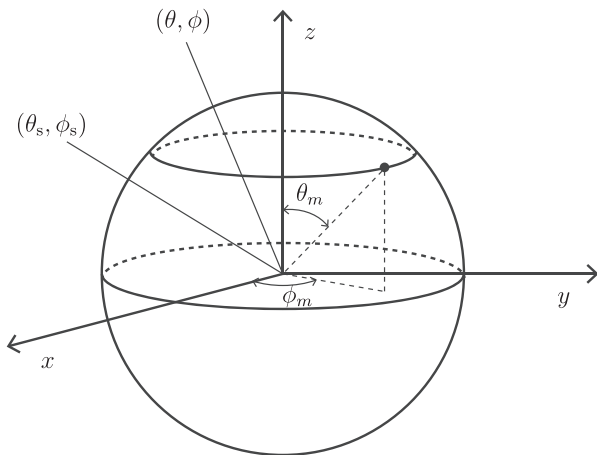

FIG. 1. Illustration of a spherical microphone arrays, where (θ, ϕ) is the array look direction, (θ_s, ϕ_s) is the source incidence angle.

$$
\mathcal{B}[\mathbf{h}(\omega), \theta, \phi] = \mathbf{h}^H(\omega) \mathbf{d}(\omega, \theta, \phi) = \sum_{m=1}^{M} H_m^*(\omega) e^{j \varpi \mathbf{p}^T \mathbf{r}_m}. \tag{7}
$$

If we choose the first microphone as the reference, the input SNR is defined according to the signal model given in Eq. (3) as

$$
\text{iSNR}(\omega) = \frac{\phi_S(\omega)}{\phi_{V_1}(\omega)}, \tag{8}
$$

where $\phi_S(\omega) = E[|S(\omega)|^2]$ and $\phi_{V_1}(\omega) = E[|V_1(\omega)|^2]$ are the variances of $S(\omega)$ and $V_1(\omega)$, respectively, with the operator $E(\cdot)$ denoting the mathematical expectation. The output SNR, according to Eq. (4), is given as

$$
\begin{aligned}
\text{oSNR}[\mathbf{h}(\omega)] &= \frac{\phi_S(\omega) |\mathbf{h}^H(\omega) \mathbf{d}(\omega, \theta_s, \phi_s)|^2}{\mathbf{h}^H(\omega) \mathbf{\Phi}_\mathbf{v}(\omega) \mathbf{h}(\omega)} \\
&= \frac{\phi_S(\omega)}{\phi_{V_1}(\omega)} \times \frac{|\mathbf{h}^H(\omega) \mathbf{d}(\omega, \theta_s, \phi_s)|^2}{\mathbf{h}^H(\omega) \mathbf{\Gamma}_\mathbf{v}(\omega) \mathbf{h}(\omega)},
\end{aligned} \tag{9}
$$

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

The SNR gain is then derived as

$$
\mathcal{G}[\mathbf{h}(\omega)] = \frac{\text{oSNR}[\mathbf{h}(\omega)]}{\text{iSNR}(\omega)} = \frac{|\mathbf{h}^H(\omega) \mathbf{d}(\omega, \theta_s, \phi_s)|^2}{\mathbf{h}^H(\omega) \mathbf{\Gamma}_\mathbf{v}(\omega) \mathbf{h}(\omega)}. \tag{10}
$$

In the literature of microphone arrays, two types of noises are often considered to evaluate a beamformer: white (both temporally and spatially) and diffuse noises. The former models the electronic and sensor noise as well as the mismatch among different sensors in a microphone array system, while the latter is often used to optimize the DF.

- The temporally and spatially white noise with the same variance across all the microphones. In this case, $\mathbf{\Gamma}_\mathbf{v}(\omega) = \mathbf{I}_M$, where $\mathbf{I}_M$ is the $M \times M$ identity matrix, and the gain given in Eq. (10) becomes

$$
\mathcal{W}[\mathbf{h}(\omega)] = \frac{|\mathbf{h}^H(\omega) \mathbf{d}(\omega, \theta_s, \phi_s)|^2}{\mathbf{h}^H(\omega) \mathbf{h}(\omega)}, \tag{11}
$$

which is called the white noise gain (WNG) (Elko and Meyer, 2008).

- The diffuse noise. In this case, the elements of $\mathbf{\Gamma}_\mathbf{v}(\omega)$ are given by (Teal et al., 2002)

$$
[\mathbf{\Gamma}_\mathbf{v}(\omega)]_{ij} = [\mathbf{\Gamma}_d(\omega)]_{ij} = \text{sinc}\left(\frac{\omega d_{ij}}{c}\right), \tag{12}
$$

with $i, j = 1, 2, \ldots, M$, and d_{ij} being the distance between microphones i and j. Now, the gain is

$$
\mathcal{D}[\mathbf{h}(\omega)] = \frac{|\mathbf{h}^H(\omega) \mathbf{d}(\omega, \theta_s, \phi_s)|^2}{\mathbf{h}^H(\omega) \mathbf{\Gamma}_d(\omega) \mathbf{h}(\omega)}, \tag{13}
$$

which is called the directivity factor (DF) (Elko and Meyer, 2008; Beranek, 1986).

III. SPHERICAL HARMONICS

The spherical harmonics of order n ($n = 0, 1, 2, \ldots, +\infty$) and degree l ($l = 0, \pm 1, \ldots, \pm n$) are defined as (Williams, 1999; Rafaely, 2015)

$$Y_n^l(\theta, \phi) = \sqrt{\frac{2n+1}{4\pi} \frac{(n-l)!}{(n+l)!}} \mathcal{P}_n^l(\cos \theta) e^{jl\phi}, \tag{14}$$

where $(\cdot)!$ represents the factorial function, and $\mathcal{P}_n^l(\cos \theta)$ is the associated Legendre function of order n and degree l.

The spherical harmonics satisfy the following orthogonality (Williams, 1999; Rafaely, 2015):

$$\int_0^{2\pi} \int_0^{\pi} Y_n^l(\theta, \phi) \left[Y_{n'}^{l'}(\theta, \phi) \right]^* \sin \theta d\theta d\phi = \delta_{nn'} \delta_{ll'}, \tag{15}$$

where $\delta_{nn'}$ and $\delta_{ll'}$ are the Kronecker delta functions. Consequently, a proper sampling method should be chosen such that when we approximate the integral with a sum of samples, the orthogonality still holds approximately, i.e.,

$$\sum_{m=1}^{M} a_m Y_n^l(\theta_m, \phi_m) \left[Y_{n'}^{l'}(\theta_m, \phi_m) \right]^* \approx \delta_{nn'} \delta_{ll'}, \tag{16}$$

where a_m are sampling weights (Rafaely, 2015). Several sampling methods have been developed, such as the equal-angle sampling, Gaussian sampling, uniform and nearly uniform sampling, which have been fully summarized in the literature (Rafaely, 2015). In this work, we consider the uniform and nearly uniform sampling for simplicity, where the sampling weights are chosen as a constant, i.e., $a_m = 4\pi/M$.

Without loss of generality, we consider an open sphere. The unit-amplitude plane wave, which comes from the direction (θ, ϕ), can be represented as a series of spherical harmonics (Rafaely, 2015), i.e.,

$$e^{j\varpi \mathbf{p}^T \mathbf{r}_m} = \sum_{n=0}^{\infty} \sum_{l=-n}^{n} \beta_n(\varpi) \left[Y_n^l(\theta_m, \phi_m) \right]^* Y_n^l(\theta, \phi), \tag{17}$$

where

$$\beta_n(\varpi) = 4\pi j^n \mathcal{J}_n(\varpi), \tag{18}$$

with $\mathcal{J}_n(\cdot)$ being the nth-order spherical Bessel function of the first kind.

IV. MAXIMUM DIRECTIVITY BEAMFORMER

In beamforming, we expect the directivity pattern to be as sharp as possible. Ideally, the beampattern is expected to have a high spike at $(\theta, \phi) = (\theta_s, \phi_s)$ and zero elsewhere, i.e.,

$$\mathcal{B}(\theta, \phi) = \delta(\cos \theta - \cos \theta_s) \delta(\phi - \phi_s). \tag{19}$$

In this case, the DF may be infinitely large.

The ideal directivity pattern (19) can be represented in terms of a spherical harmonic series as (Rafaely, 2015)

$$\mathcal{B}(\theta, \phi) = \sum_{n=0}^{\infty} \sum_{l=-n}^{n} \left[Y_n^l(\theta_s, \phi_s) \right]^* Y_n^l(\theta, \phi). \tag{20}$$

Limiting the order in Eq. (20) up to N gives

$$\mathcal{B}'_N(\theta, \phi) = \sum_{n=0}^{N} \sum_{l=-n}^{n} \left[Y_n^l(\theta_s, \phi_s) \right]^* Y_n^l(\theta, \phi). \tag{21}$$

Using the spherical harmonic addition theorem (Rafaely, 2015), we can write Eq. (21) as

$$\mathcal{B}'_N(\theta, \phi) = \sum_{n=0}^{N} \frac{2n+1}{4\pi} P_n(\cos \Theta), \tag{22}$$

where $P_n(\cos \Theta)$ is the nth-order Legendre function, and Θ is the angle difference between (θ, ϕ) and (θ_s, ϕ_s), with $\cos \Theta = \cos \theta_s \cos \theta + \cos(\phi_s - \phi) \sin \theta_s \sin \theta$. When $(\theta, \phi) = (\theta_s, \phi_s)$, i.e., $\Theta = 0$, we have

$$\kappa_N = \mathcal{B}'_N(\theta_s, \phi_s) = \sum_{n=0}^{N} \frac{2n+1}{4\pi} P_n(1) = \frac{(N+1)^2}{4\pi}. \tag{23}$$

From Eq. (23), we define the normalized Nth-order ideal directivity pattern as

$$\mathcal{B}_N(\theta, \phi) = \frac{1}{\kappa_N} \sum_{n=0}^{N} \sum_{l=-n}^{n} \left[Y_n^l(\theta_s, \phi_s) \right]^* Y_n^l(\theta, \phi). \tag{24}$$

Figure 2 plots the ideal directivity pattern (as a function of Θ) approximated by the spherical harmonics series of order $N = 3$ and $N = 10$. It is clearly seen that the Nth-order ideal directivity pattern have N distinct nulls and it corresponds to the Nth-order hypercardioid (Elko and Meyer, 2008, Benesty and Chen, 2012).

Now, substituting Eq. (17) into Eq. (7), we can rewrite the beampattern as

$$\begin{aligned} \mathcal{B}[\mathbf{h}(\omega), \theta, \phi] &= \sum_{m=1}^{M} H_m^*(\omega) e^{j\varpi \mathbf{p}^T \mathbf{r}_m} \\ &= \sum_{m=1}^{M} H_m^*(\omega) \sum_{n=0}^{\infty} \sum_{l=-n}^{n} \beta_n(\varpi) \\ &\quad \times \left[Y_n^l(\theta_m, \phi_m) \right]^* Y_n^l(\theta, \phi) \\ &= \sum_{n=0}^{\infty} \sum_{l=-n}^{n} Y_n^l(\theta, \phi) \beta_n(\varpi) \\ &\quad \times \sum_{m=1}^{M} H_m^*(\omega) \left[Y_n^l(\theta_m, \phi_m) \right]^*. \end{aligned} \tag{25}$$

Limiting the expansion to the order N, we obtain an approximation of the beampattern,

$$\begin{aligned} \mathcal{B}_N[\mathbf{h}(\omega), \theta, \phi] &= \sum_{n=0}^{N} \sum_{l=-n}^{n} Y_n^l(\theta, \phi) \beta_n(\varpi) \\ &\quad \times \sum_{m=1}^{M} H_m^*(\omega) \left[Y_n^l(\theta_m, \phi_m) \right]^*. \end{aligned} \tag{26}$$

Equating the beampattern in Eq. (26) to the Nth-order ideal directivity pattern in Eq. (24), we find the following relationship:
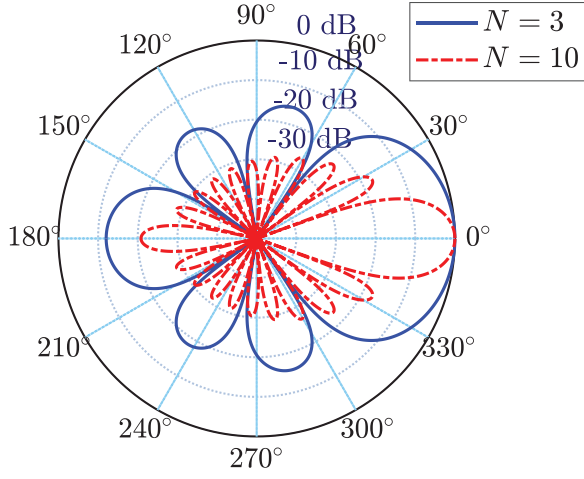
J. Acoust. Soc. Am. **143** (5), May 2018

Huang *et al.* 3027

FIG. 2. (Color online) Ideal directivity pattern (as a function of Θ) approximated by the spherical harmonic series of order $N = 3$ and 10.

$$\beta_n(\varpi) \sum_{m=1}^{M} H_m^*(\omega) \left[Y_n^l(\theta_m, \phi_m) \right]^* = \frac{1}{\kappa_N} \left[Y_n^l(\theta_s, \phi_s) \right]^*$$

(27)

or, equivalently,

$$\beta_n^*(\varpi) \sum_{m=1}^{M} H_m(\omega) Y_n^l(\theta_m, \phi_m) = \frac{1}{\kappa_N} Y_n^l(\theta_s, \phi_s).$$

(28)

We observe from Eq. (28) that the filter coefficients $H_m(\omega)$ can be determined given the coefficients $Y_n^l(\theta_s, \phi_s)$ and $\beta_n(\varpi)$, i.e.,

$$\mathbf{Y}\mathbf{h}(\omega) = \boldsymbol{\eta}(\varpi),$$

(29)

where

$$\mathbf{Y} = \begin{bmatrix} Y_0^0(\theta_1, \phi_1) & Y_0^0(\theta_2, \phi_2) & \cdots & Y_0^0(\theta_M, \phi_M) \\ Y_1^{-1}(\theta_1, \phi_1) & Y_1^{-1}(\theta_2, \phi_2) & \cdots & Y_1^{-1}(\theta_M, \phi_M) \\ Y_1^0(\theta_1, \phi_1) & Y_1^0(\theta_2, \phi_2) & \cdots & Y_1^0(\theta_M, \phi_M) \\ Y_1^1(\theta_1, \phi_1) & Y_1^1(\theta_2, \phi_2) & \cdots & Y_1^1(\theta_M, \phi_M) \\ \vdots & \vdots & \ddots & \vdots \\ Y_N^N(\theta_1, \phi_1) & Y_N^N(\theta_2, \phi_2) & \cdots & Y_N^N(\theta_M, \phi_M) \end{bmatrix}$$

(30)

is a matrix of size $(N + 1)^2 \times M$, and

$$\boldsymbol{\eta}(\varpi) = \frac{1}{\kappa_N} \left[\frac{Y_0^0(\theta_s, \phi_s)}{\beta_0^*(\varpi)} \quad \frac{Y_1^{-1}(\theta_s, \phi_s)}{\beta_1^*(\varpi)} \quad \cdots \quad \frac{Y_N^N(\theta_s, \phi_s)}{\beta_N^*(\varpi)} \right]^T$$

(31)

is a vector of length $(N + 1)^2$.

Generally, it is assumed that $M \geq (N + 1)^2$. If $M > (N + 1)^2$, the solution to Eq. (29) is not unique. In this case, the beamforming filter, $\mathbf{h}(\omega)$, is derived from the following optimization:

$$\min_{\mathbf{h}(\omega)} \mathbf{h}^H(\omega) \mathbf{h}(\omega) \ \text{s. t.} \ \mathbf{Y}\mathbf{h}(\omega) = \boldsymbol{\eta}(\varpi).$$

(32)

The solution of Eq. (32), which is the minimum-norm solution of Eq. (29), gives the Nth-order MD beamformer

$$\mathbf{h}_N(\omega) = \mathbf{Y}^H (\mathbf{Y}\mathbf{Y}^H)^{-1} \boldsymbol{\eta}(\varpi),$$

(33)

where the subscript $_N$ is used to indicate the Nth order.

V. FLEXIBLE HIGH DIRECTIVITY BEAMFORMER

The MD beamformer may suffer from serious white noise amplification, particularly at low frequencies. One possible way to improve the WNG is by using more microphones while fixing the order of the beamformer. However, this would lead to many issues in terms of cost, microphone mounting, array aperture, interelement spacing, and possibly other performance degradation as deep nulls. Now, we consider taking another approach, i.e., we improve the WNG by slightly sacrificing the value of the DF.

To begin, we define the Nth-order flexible directivity pattern as a tradeoff between the Nth-order and $(N-1)$th-order ideal directivity patterns,

$$\mathcal{B}_{N,\alpha}(\theta, \phi) = \frac{\alpha}{\kappa'_N} \sum_{n=0}^{N} \sum_{l=-n}^{n} \left[Y_n^l(\theta_s, \phi_s) \right]^* Y_n^l(\theta, \phi)$$
$$+ \frac{1 - \alpha}{\kappa'_N} \sum_{n=0}^{N-1} \sum_{l=-n}^{n} \left[Y_n^l(\theta_s, \phi_s) \right]^* Y_n^l(\theta, \phi),$$

(34)

where α is a real coefficient, with $\alpha \in [0, 1]$, and

$$\kappa'_N = \frac{1}{4\pi} \left[N^2 + \alpha(2N + 1) \right]$$

(35)

is a normalization factor.

To simplify the derivation, we rewrite Eq. (34) into an equivalent form as

$$\mathcal{B}_{N,\alpha}(\theta, \phi) = \frac{1}{\kappa'_N} \sum_{n=0}^{N} \sum_{l=-n}^{n} \xi_n \left[Y_n^l(\theta_s, \phi_s) \right]^* Y_n^l(\theta, \phi),$$

(36)

where

$$\xi_n = \begin{cases} 1, & n = 0, 1, \ldots, N - 1, \\ \alpha, & n = N. \end{cases}$$

(37)

Equating the beampattern in Eq. (26) to the Nth-order flexible directivity pattern, we obtain the following relationship:

$$\beta_n^*(\varpi) \sum_{m=1}^{M} H_m(\omega) Y_n^l(\theta_m, \phi_m) = \frac{\xi_n}{\kappa'_N} Y_n^l(\theta_s, \phi_s).$$

(38)

It follows immediately that

$$\mathbf{Y}\mathbf{h}(\omega) = \boldsymbol{\eta}'(\varpi),$$

(39)

where **Y** has been defined in Eq. (30) and

$$\boldsymbol{\eta}'(\varpi) = \frac{1}{\kappa'_N}\left[\frac{\xi_0 Y_0^0(\theta_s,\phi_s)}{\beta_0^*(\varpi)} \frac{\xi_1 Y_1^{-1}(\theta_s,\phi_s)}{\beta_1^*(\varpi)} \cdots \frac{\xi_N Y_N^N(\theta_s,\phi_s)}{\beta_N^*(\varpi)}\right]^T \tag{40}$$

is a vector of length $(N+1)^2$.

The minimum-norm solution of Eq. (39) gives the Nth-order flexible HD beamformer,

$$\mathbf{h}_{N,\alpha}(\omega) = \mathbf{Y}^H(\mathbf{Y}\mathbf{Y}^H)^{-1}\boldsymbol{\eta}'(\varpi). \tag{41}$$

Depending on the value of α, we have the following three cases.

- $\alpha = 1$: $\mathbf{h}_{N,\alpha}(\omega)$ becomes the Nth-order MD beamformer, i.e., $\mathbf{h}_{N,1}(\omega) = \mathbf{h}_N(\omega)$.
- $\alpha = 0$: $\mathbf{h}_{N,\alpha}(\omega)$ becomes the MD beamformer of order $N-1$, i.e., $\mathbf{h}_{N,0}(\omega) = \mathbf{h}_{N-1}(\omega)$.
- $0 < \alpha < 1$: we obtain a flexible HD beamformer whose performance is between the performances of the MD beamformers of order N and order $N-1$.

VI. DETERMINATION OF THE TUNING PARAMETER IN THE FLEXIBLE HD BEAMFORMER

The value of the tuning parameter, α, plays an important role on the performance of the flexible HD beamformer. In this section, we discuss the DF and the WNG of the proposed beamformer and show how to determine the value of α based on a specified value of the DF or the WNG.

A. Determination of the tuning parameter with a specified DF value

Following the definition of the DF, the frequency-independent DF of the Nth-order flexible HD beamformer can be defined as

$$\mathcal{D}_{N,\alpha} = \frac{|\mathcal{B}_{N,\alpha}(\theta_s,\phi_s)|^2}{\frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi}|\mathcal{B}_{N,\alpha}(\theta,\phi)|^2 \sin\theta d\theta d\phi}. \tag{42}$$

From Eq. (36), we have

$$|\mathcal{B}_{N,\alpha}(\theta,\phi)|^2 = \frac{1}{\kappa_N'^2}\sum_{n=0}^{N}\sum_{l=-n}^{n}\xi_n\left[Y_n^l(\theta_s,\phi_s)\right]^* Y_n^l(\theta,\phi)$$
$$\times \sum_{n'=0}^{N}\sum_{l'=-n'}^{n'}\xi_{n'}Y_{n'}^{l'}(\theta_s,\phi_s)\left[Y_{n'}^{l'}(\theta,\phi)\right]^*. \tag{43}$$

Using Eq. (43) and also the orthogonality property of spherical harmonics (Rafaely, 2015), we can write the denominator of the right-hand side of Eq. (42) as

$$\frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi}|\mathcal{B}_{N,\alpha}(\theta,\phi)|^2 \sin\theta d\theta d\phi$$
$$= \frac{1}{4\pi\kappa_N'^2}\sum_{n=0}^{N}\sum_{l=-n}^{n}\xi_n^2\left[Y_n^l(\theta_s,\phi_s)\right]^* Y_n^l(\theta_s,\phi_s)$$
$$= \frac{1}{4\pi\kappa_N'^2}\sum_{n=0}^{N}\xi_n^2\frac{2n+1}{4\pi}$$
$$= \frac{[N^2 + \alpha^2(2N+1)]}{(4\pi\kappa_N')^2}. \tag{44}$$

Substituting Eqs. (37) and (44) into Eq. (42) and using the fact that $|\mathcal{B}_{N,\alpha}(\theta_s,\phi_s)|^2 = 1$, we obtain an explicit form of the DF,

$$\mathcal{D}_{N,\alpha} = \frac{[N^2 + \alpha(2N+1)]^2}{N^2 + \alpha^2(2N+1)}. \tag{45}$$

To clearly see the changing trend of $\mathcal{D}_{N,\alpha}$, we compute the gradient of $\mathcal{D}_{N,\alpha}$ with respect to α, i.e.,

$$\frac{\partial\mathcal{D}_{N,\alpha}}{\partial\alpha} = \frac{N^2(2N+1)[N^2 + \alpha(2N+1)](1-\alpha)}{[N^2 + \alpha^2(2N+1)]^2}. \tag{46}$$

Clearly, for $0 \le \alpha \le 1$, we always have

$$\frac{\partial\mathcal{D}_{N,\alpha}}{\partial\alpha} \ge 0, \tag{47}$$

where the equality holds if and only if $\alpha = 1$. Consequently, the DF of the Nth-order flexible HD beamformer is an increasing function of α. From Eq. (45), it is seen that the DF of the Nth-order flexible HD beamformer satisfies

$$N^2 \le \mathcal{D}_{N,\alpha} \le (N+1)^2. \tag{48}$$

From Eq. (45), one can compute the value of the tuning parameter once the value of the DF is specified. As shown in Appendix A, with a specified DF, $\mathcal{D}$, the parameter α is given as

$$\alpha = \frac{N^2(2N+1) - N\sqrt{(2N+1)\mathcal{D}[(N+1)^2 - \mathcal{D}]}}{(2N+1)(\mathcal{D} - 2N - 1)}. \tag{49}$$

Table I lists some examples of the tuning parameter for different orders of the flexible HD beamformer.

B. Determination of the tuning parameter with a specified WNG value

If we neglect the approximation error on the distortionless constraint in the look direction (we will come back to this point in the simulation part), the WNG of the Nth-order flexible HD beamformer is in the following form:

$$\mathcal{W}[\mathbf{h}_{N,\alpha}(\omega)] = \frac{1}{\mathbf{h}_{N,\alpha}(\omega)\mathbf{h}_{N,\alpha}^H(\omega)}. \tag{50}$$

J. Acoust. Soc. Am. **143** (5), May 2018

Huang *et al.* 3029

TABLE I. Examples of computing the parameter α with a specified DF, $\mathcal{D}$. The special case of $\mathcal{D} = 2N + 1$ is computed according to Eq. (A3), i.e., for $N = 1$ and $\mathcal{D} = 3$, $\alpha = 1/3$, for $N = 2$ and $\mathcal{D} = 5$, $\alpha = 0.1$.

N	$\mathcal{D}_{min}$	$\mathcal{D}_{max}$	α
1	1	4	$\dfrac{3 - \sqrt{3\mathcal{D}(4 - \mathcal{D})}}{3(\mathcal{D} - 3)}, 1 \leq \mathcal{D} \leq 4, \mathcal{D} \neq 3$
2	4	9	$\dfrac{20 - 2\sqrt{5\mathcal{D}(9 - \mathcal{D})}}{5(\mathcal{D} - 5)}, 4 \leq \mathcal{D} \leq 9, \mathcal{D} \neq 5$
3	9	16	$\dfrac{63 - 3\sqrt{7\mathcal{D}(16 - \mathcal{D})}}{7(\mathcal{D} - 7)}, 9 \leq \mathcal{D} \leq 16$
4	16	25	$\dfrac{48 - 4\sqrt{\mathcal{D}(25 - \mathcal{D})}}{3(\mathcal{D} - 9)}, 16 \leq \mathcal{D} \leq 25$

With the uniform or nearly uniform sampling (Rafaely, 2015), we have

$$\mathbf{Y}\mathbf{Y}^H = \frac{M}{4\pi}\mathbf{I}_M. \tag{51}$$

Using Eqs. (40) and (51), we can write the denominator in Eq. (50) as

$$
\begin{aligned}
\mathbf{h}_{N,\alpha}^H(\omega)\mathbf{h}_{N,\alpha}(\omega) &= \frac{4\pi}{M}\boldsymbol{\eta}'^H(\varpi)\boldsymbol{\eta}'(\varpi) \\
&= \frac{4\pi}{M\kappa\iota_N^2}\sum_{n=0}^{N}\sum_{l=-n}^{n}\frac{\xi_n^2}{|\beta_n(\varpi)|^2}|Y_n^l(\theta_s, \phi_s)|^2 \\
&= \frac{4\pi}{M\kappa\iota_N^2}\sum_{n=0}^{N}\frac{\xi_n^2}{|\beta_n(\varpi)|^2}\frac{2n+1}{4\pi}.
\end{aligned}
\tag{52}
$$

Substituting Eqs. (18) and (52) into Eq. (50), we can rewrite the WNG as

$$\mathcal{W}[\mathbf{h}_{N,\alpha}(\omega)] = \frac{M\kappa\iota_N^2}{\displaystyle\sum_{n=0}^{N}\frac{\xi_n^2(2n+1)}{|\beta_n(\varpi)|^2}} = \frac{(4\pi)^2 M\kappa\iota_N^2}{\displaystyle\sum_{n=0}^{N}\frac{\xi_n^2(2n+1)}{\mathcal{J}_n^2(\varpi)}}, \tag{53}$$

where $\mathcal{J}_n(\omega)$, again, is the nth-order spherical Bessel function. Finally, by substituting Eqs. (35) and (37) into Eq. (53), we derive

$$
\begin{aligned}
\mathcal{W}[\mathbf{h}_{N,\alpha}(\omega)] &= \frac{M[N^2 + (2N+1)\alpha]^2}{\displaystyle\sum_{n=0}^{N-1}\frac{(2n+1)}{\mathcal{J}_n^2(\varpi)} + \frac{(2N+1)}{\mathcal{J}_N^2(\varpi)}\alpha^2} \\
&= \frac{M\left[(2N+1)^2\alpha^2 + 2N^2(2N+1)\alpha + N^4\right]}{\dfrac{(2N+1)}{\mathcal{J}_N^2(\varpi)}\alpha^2 + \displaystyle\sum_{n=0}^{N-1}\frac{(2n+1)}{\mathcal{J}_n^2(\varpi)}}.
\end{aligned}
\tag{54}
$$

From Eq. (54), it is seen that the WNG is frequency dependent, which is different from the DF. When the value of ϖ approaches 0 (i.e., at low frequencies), the value of $\mathcal{J}_n(\varpi)$ becomes very small and so is the value of the WNG, which leads to serious white noise amplification. We also see that the WNG is an increasing function of M, which indicates that, for a specified order N, increasing the number of microphones can improve the WNG.

Generally, for a small size array, the value of $\mathcal{J}_N(\varpi)$ is very small, and $(2N+1)/\mathcal{J}_N^2(\varpi)$ is much larger than the other coefficients. Consequently, the WNG is mainly affected by $[(2N+1)/\mathcal{J}_N^2(\varpi)]\alpha^2$ in the denominator of Eq. (54), and it is a decreasing function of α.

From Eq. (54), it can be shown that the WNG of the Nth-order flexible HD beamformer at a given frequency satisfies

$$\frac{M(N+1)^4}{\displaystyle\sum_{n=0}^{N}\frac{(2n+1)}{\mathcal{J}_n^2(\varpi)}} \leq \mathcal{W}[\mathbf{h}_{N,\alpha}(\omega)] \leq \frac{MN^4}{\displaystyle\sum_{n=0}^{N-1}\frac{(2n+1)}{\mathcal{J}_n^2(\varpi)}}. \tag{55}$$

According to Eq. (54), one can compute the value of the tuning parameter once the level of the WNG is specified. As shown in Appendix B, with a specified value of the WNG, $\mathcal{W}$, the parameter α can be computed as

$$\alpha = \frac{2MN^2(2N+1) + \sqrt{\Delta}}{2\mathcal{W}\frac{(2N+1)}{\mathcal{J}_N^2(\varpi)} - 2M(2N+1)^2}, \tag{56}$$

where Δ is given in Eq. (B8) of Appendix B. This is very useful for practical applications. Once we know the level of the sensors' self-noise and electronic noise of our system, we can determine the minimum value of the WNG, which can subsequently be used to compute the tuning parameter and the DF without going through an experimental or numerical optimization process.

VII. SIMULATIONS

In this section, we study the performance of the developed flexible HD beamformer through simulations. We consider an open SMA with a radius of 3 cm, consisting of 32 omnidirectional microphones. While the designed beampattern can be perfectly steered to any directions, we only consider here the case where the signal of interest comes from the direction $(\theta_s, \phi_s) = (90°, 115°)$.

Figure 3 plots the beampatterns of the 1st-, 2nd-, 3rd-, and 4th-order flexible HD beamformers for $\alpha \in \{1, 0.8, 0.5, 0.2, 0.1\}$. The designed beampatterns are symmetric with respect to the look direction, so we only plot the two-dimensional beampattern as a function of the azimuth angle ϕ. It is seen that the beampattern of the Nth-order flexible HD beamformer changes with the value of α. As the value of α decreases from 1 to 0, the designed beampattern achieves a compromise between the beampattern of the Nth-order MD beamformer and that of the $(N-1)$th-order MD one.

Figure 4 plots the DF and the WNG of the MD beamformer, both as a function of frequency, for three different orders, i.e., $N \in \{1, 2, 3, 4\}$. It is clearly seen from Fig. 4(a) that the MD beamformer can achieve a DF of $10\log_{10}(N+1)^2$ dB over the studied frequency range. It is also seen that the WNG is an increasing function of frequency. The value of the WNG is very small at low frequencies, particularly for higher orders, which leads to significant white noise amplification.
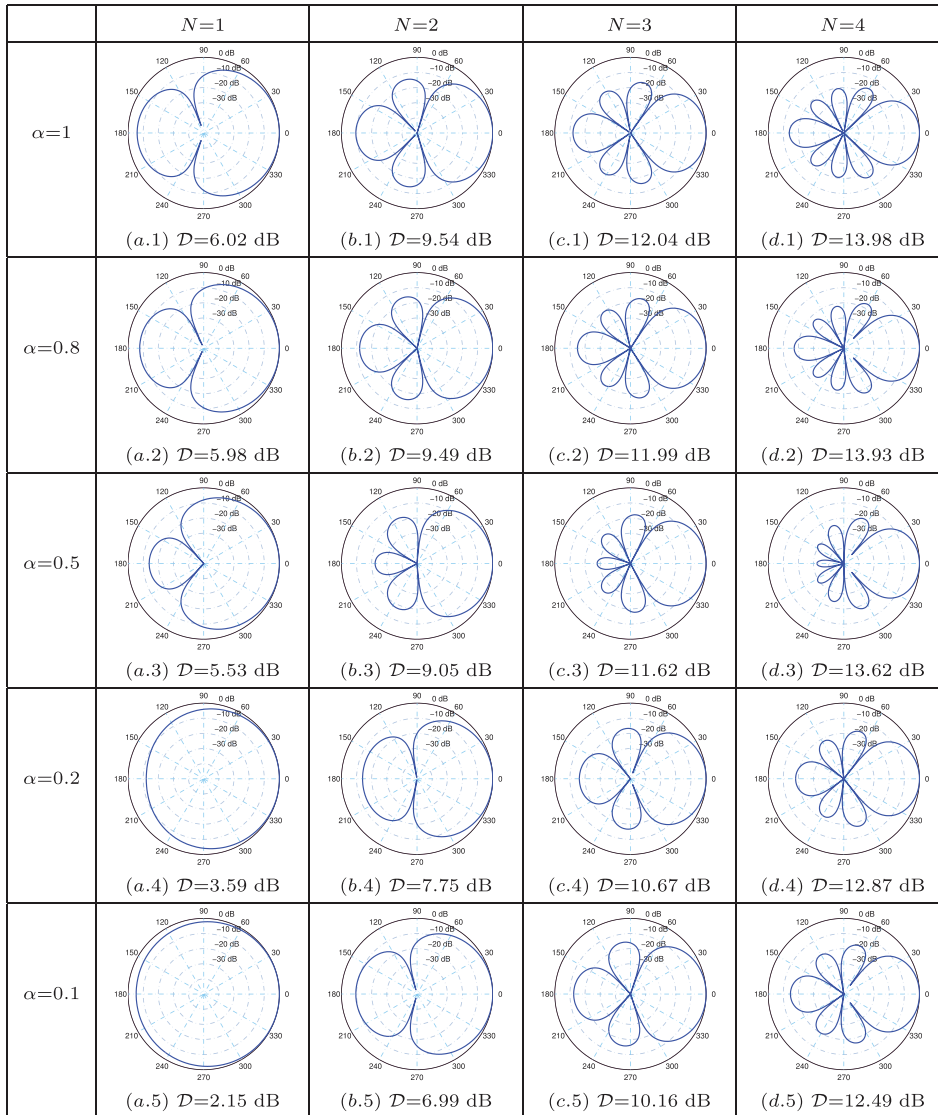
	$N=1$	$N=2$	$N=3$	$N=4$
$\alpha=1$	$(a.1)\ \mathcal{D}=6.02$ dB	$(b.1)\ \mathcal{D}=9.54$ dB	$(c.1)\ \mathcal{D}=12.04$ dB	$(d.1)\ \mathcal{D}=13.98$ dB
$\alpha=0.8$	$(a.2)\ \mathcal{D}=5.98$ dB	$(b.2)\ \mathcal{D}=9.49$ dB	$(c.2)\ \mathcal{D}=11.99$ dB	$(d.2)\ \mathcal{D}=13.93$ dB
$\alpha=0.5$	$(a.3)\ \mathcal{D}=5.53$ dB	$(b.3)\ \mathcal{D}=9.05$ dB	$(c.3)\ \mathcal{D}=11.62$ dB	$(d.3)\ \mathcal{D}=13.62$ dB
$\alpha=0.2$	$(a.4)\ \mathcal{D}=3.59$ dB	$(b.4)\ \mathcal{D}=7.75$ dB	$(c.4)\ \mathcal{D}=10.67$ dB	$(d.4)\ \mathcal{D}=12.87$ dB
$\alpha=0.1$	$(a.5)\ \mathcal{D}=2.15$ dB	$(b.5)\ \mathcal{D}=6.99$ dB	$(c.5)\ \mathcal{D}=10.16$ dB	$(d.5)\ \mathcal{D}=12.49$ dB

FIG. 3. (Color online) Beampatterns of the first-, second-, third-, and fourth-order flexible HD beamformers for different values of α.

As discussed previously, one way to improve the WNG of the MD beamformer is by reducing the order. For example, as shown in Fig. 4, the 3rd-order MD beamformer can achieve a DF of approximately 12 dB. However, the WNG at low frequencies is very small, which makes it difficult to use in practice. To avoid serious white noise amplification, one can use the same SMA to design a 2nd-order MD beamformer. Now, the WNG is significantly improved, but the corresponding DF drops down to approximately 9 dB.

The flexible HD beamformer offers a convenient way to control the tradeoff between a high DF and a reasonable level of the WNG. Figure 5 plots the DF and the WNG of the 3rd-order flexible HD beamformer as a function of frequency for $\alpha \in \{0, 0.1, 0.2, 0.5, 1\}$. It is clearly seen that this beamformer can achieve a good compromise between the performance of the 3rd-order MD beamformer and that of the 2nd-order MD beamformer.

The value of the tuning parameter, α, plays an important role on the beamforming performance. Figure 6 plots the DF of the flexible HD beamformer as a function of α for different orders. It is seen that the DF increases with α. As the value of α increases from 0 to 1, the DF of the Nth-order flexible HD beamformer increases from $20\log_{10}N$ dB to $20\log_{10}(N+1)$ dB. Figure 7 plots the WNG of the flexible HD beamformer as a function of α for different orders at, respectively, $f=500$ and 1000 Hz. It is seen that the WNG decreases with α. Clearly, we can choose a proper value of α to achieve a good compromise between the WNG of the Nth-order MD beamformer and that of the $(N-1)$th-order MD beamformer.

As discussed in Sec. VI, one can determine the value of the tuning parameter analytically given the value of either the DF or the WNG. Of course, we assumed that the distortionless constraint holds in the analysis, i.e., $|\mathbf{h}_N^H(\omega)\mathbf{d}(\omega, \theta_s, \phi_s)| \approx 1$. Figure 8 plots the value of $|\mathbf{h}_N^H(\omega)\mathbf{d}(\omega, \theta_s, \phi_s)|$ as a function of frequency for $N \in \{1, 2, 3, 4\}$ with the uniform sampling. It is clearly seen that the approximation error on the distortionless constraint is negligible.

Figure 9 plots the computed tuning parameter α with a given value of the DF, where for the Nth-order flexible HD beamformer, the DF changes from $20\log_{10}N$ dB to $20\log_{10}(N+1)$ dB. As shown in Eq. (A5), α has two

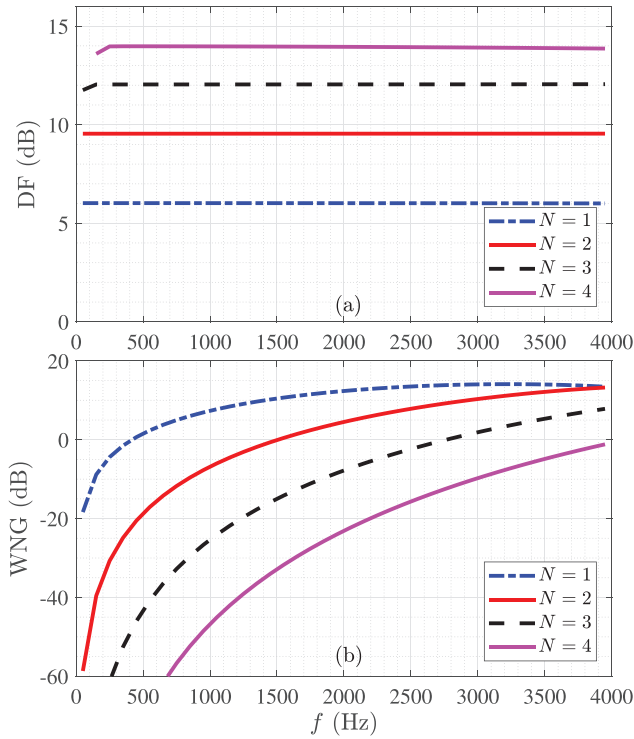J. Acoust. Soc. Am. **143** (5), May 2018

Huang *et al.* 3031

FIG. 4. (Color online) DF and WNG of the MD beamformer as a function of the frequency, f, for four different orders: (a) DF and (b) WNG. Conditions: $M = 32$ and $r = 3$ cm.

solutions. But α_2 is the correct solution because its value is in the range between 0 and 1 as clearly seen in Fig. 9 (pink shadowed area). This result corroborates with the theoretical analysis.
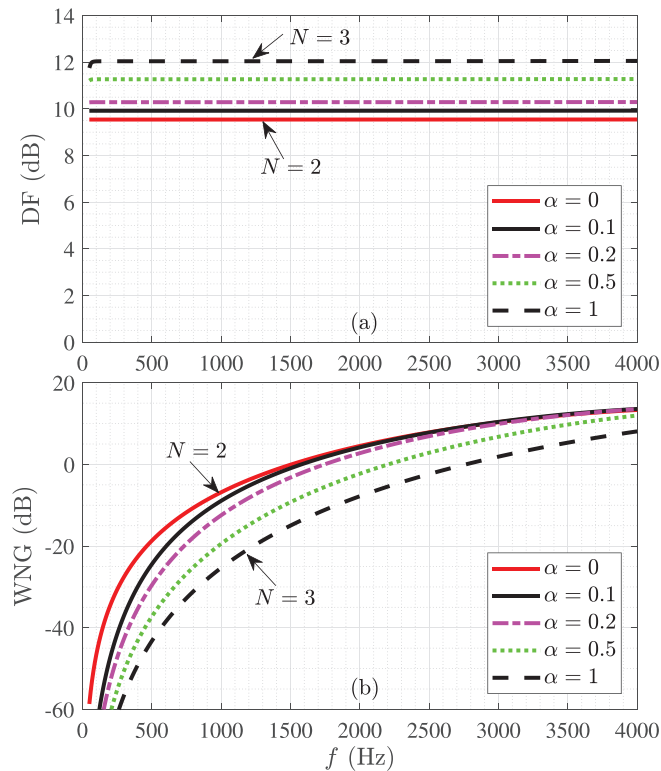

FIG. 5. (Color online) DF and WNG of the third-order flexible HD beamformer as a function of the frequency, f, for different values of α: (a) DF and (b) WNG. Conditions: $M = 32$ and $r = 3$ cm.
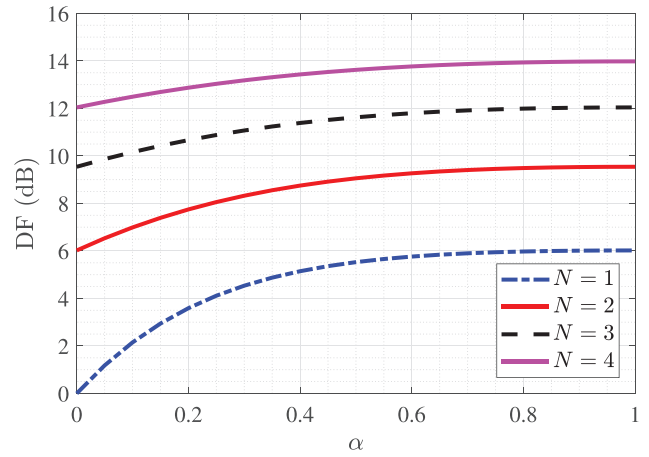

FIG. 6. (Color online) DF of the flexible HD beamformer as a function of α, for different orders, N. Conditions: $M = 32$ and $r = 3$ cm.

Similarly, Fig. 10 plots the computed tuning parameter α with a given value of the WNG. The result confirms that α_1 is the correct solution because its value is in the range between 0 and 1 (shadowed area in Fig. 9). This, again, corroborates the theoretical analysis.

VIII. CONCLUSIONS

In this paper, we studied the problem of frequency-invariant beamforming to achieve maximum directivity factor while maintaining a reasonable level of white noise
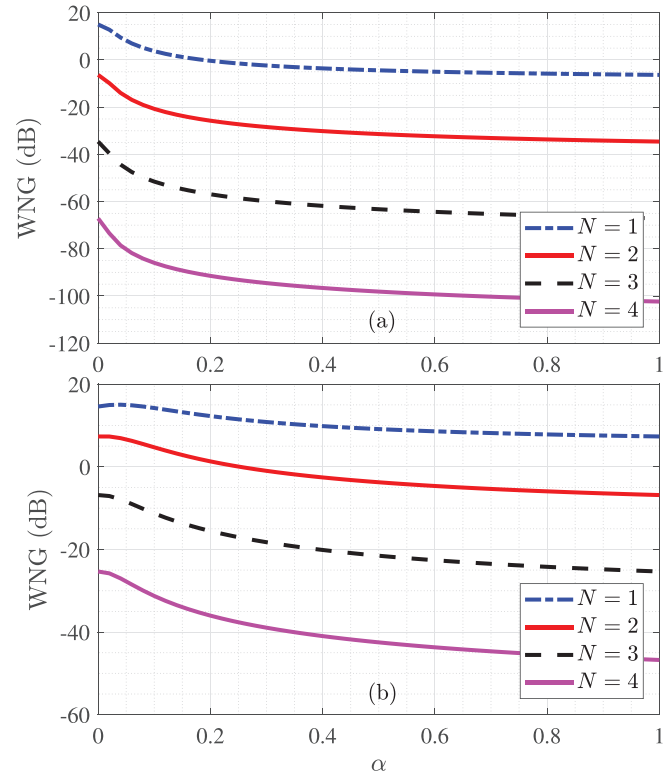

FIG. 7. (Color online) WNG of the flexible HD beamformer as a function of α for different orders, N: (a) $f = 500$ Hz and (b) $f = 1000$ Hz. Conditions: $M = 32$ and $r = 3$ cm.
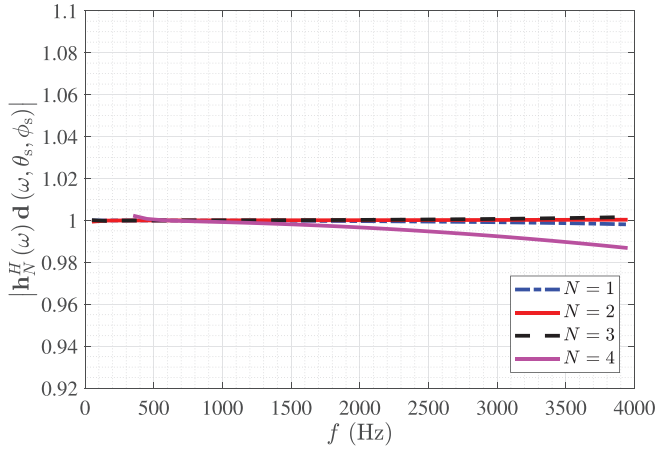
FIG. 8. (Color online) The value of $|\mathbf{h}_N^H(\omega)\mathbf{d}(\omega, \theta_s, \phi_s)|$ as a function of the frequency, f, for different orders, N. Conditions: $M = 32$ and $r = 3\,\mathrm{cm}$.

FIG. 10. (Color online) Solutions of the tuning factor, α, as a function of the desired WNG for different orders, N. Conditions: $M = 32$ and $r = 3\,\mathrm{cm}$.

gain with SMAs. We discussed the problem formulation of beamforming and the ideal directivity pattern, i.e., a maximum response at the desired direction, and zero response in all other directions. Then, by using the spherical harmonics expansion, we obtained the relationship between the ideal directivity pattern and the beamformer's beampattern. A flexible HD beamformer is subsequently deduced by solving a linear system of equations, which can achieve a good compromise between the performances of the MD beamformers of order N and order $N - 1$. An important parameter of the flexible HD beamformer is the tuning coefficient, whose value plays an important role on the beamformer's performance. The proper value of this parameter can be determined by experiments. But we derived analytical links between this parameter and the DF and the WNG. Based on these links, we showed how to determine the optimal value of the tuning parameter once the value of the DF or the WNG is specified. Simulation results validated the theoretical analysis and demonstrated the useful properties of the flexible HD beamformer.
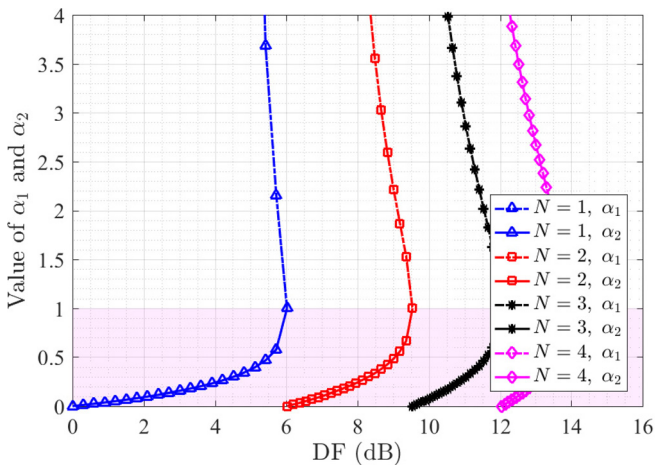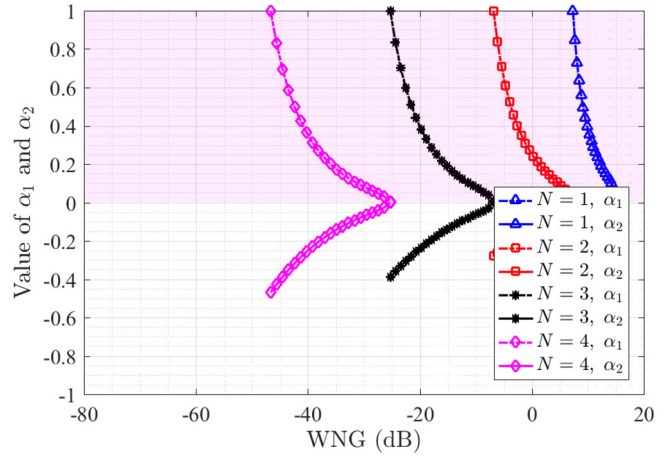
APPENDIX A

In this appendix, we discuss how to determine the value of the tuning parameter, α, given the value of the DF, i.e., $\mathcal{D}$. Given $\mathcal{D}$, one can easily find the order, N, of the flexible HD beamformer according to Eq. (48). We then have

$$\mathcal{D} = \frac{\left[N^2 + \alpha(2N + 1)\right]^2}{N^2 + \alpha^2(2N + 1)}. \tag{A1}$$

From Eq. (A1), we easily deduced that

$$(2N + 1)(\mathcal{D} - 2N - 1)\alpha^2 - 2N^2(2N + 1)\alpha \\ + N^2(\mathcal{D} - N^2) = 0. \tag{A2}$$

If $\mathcal{D} = 2N + 1$, Eq. (A2) is a linear equation, and the solution is easily obtained as

$$\alpha = \frac{\mathcal{D} - N^2}{2(2N + 1)}. \tag{A3}$$

If $\mathcal{D} \neq 2N - 1$, Eq. (A2) becomes a quadratic equation and its discriminant is

$$\Delta = 4N^2\mathcal{D}(2N + 1)\left[(N + 1)^2 - \mathcal{D}\right] \geq 0, \tag{A4}$$

where the equality holds if and only if $\mathcal{D} = (N + 1)^2$. So, the quadratic equation of Eq. (A2) has two solutions,

$$\alpha_{1,2} = \frac{2N^2(2N + 1) \pm \sqrt{\Delta}}{2(2N + 1)(\mathcal{D} - 2N - 1)}. \tag{A5}$$

FIG. 9. (Color online) Solutions of the tuning factor, α, as a function of the desired DF for different orders, N.

J. Acoust. Soc. Am. **143** (5), May 2018

Huang *et al.*   3033

Now we have the following two cases.

- $\mathcal{D} < 2N + 1$ (this only happens for $N = 1$ and $N = 2$). In this case, it can be verified that $\sqrt{\Delta} \geq 2N^2(2N + 1)$. As a result, we have $\alpha_1 \leq 0$ and $\alpha_2 \geq 0$, where the equality holds if and only if $\mathcal{D} = N^2$. Since the value of α should be greater than 0, the correct solution must be α_2.
- $\mathcal{D} > 2N + 1$. In this case, we have $0 < \mathcal{D} - 2N - 1 \leq N^2$, so

$$\frac{2N^2(2N + 1)}{2(2N + 1)(\mathcal{D} - 2N - 1)} = \frac{N^2}{(\mathcal{D} - 2N - 1)} \geq 1. \quad \text{(A6)}$$

From Eq. (A6), it is seen that $\alpha_1 > 1$. Since the range of the tuning parameter is $[0, 1]$, it is safe to say that the correct solution must be α_2.

To summarize, the correct solution of Eq. (A3) is

$$\alpha = \alpha_2 = \frac{2N^2(2N + 1) - \sqrt{\Delta}}{2(2N + 1)(\mathcal{D} - 2N - 1)}. \quad \text{(A7)}$$

Substituting Eq. (A4) into Eq. (A7), we finally obtain the correct solution of the tunable parameter, α, with a specified value of the DF, $\mathcal{D}$, i.e.,

$$\alpha = \frac{N^2(2N + 1) - N\sqrt{(2N + 1)\mathcal{D}\left[(N + 1)^2 - \mathcal{D}\right]}}{(2N + 1)(\mathcal{D} - 2N - 1)}. \quad \text{(A8)}$$

APPENDIX B

In this appendix, we discuss how to determine the value of α with a given value of the WNG, i.e., $\mathcal{W}$. Given $\mathcal{W}$, one can easily find the order, N, of the flexible HD beamformer according to Eq. (55). Following a similar way as in Appendix A, we have

$$\mathcal{W} = \frac{M\left[N^2 + (2N + 1)\alpha\right]^2}{\dfrac{(2N + 1)}{\mathcal{J}_N^2(\varpi)}\alpha^2 + \displaystyle\sum_{n=0}^{N-1}\frac{(2n + 1)}{\mathcal{J}_n^2(\varpi)}}. \quad \text{(B1)}$$

From Eq. (B1), we easily get

$$\mathcal{A}\alpha^2 + \mathcal{B}\alpha + \mathcal{C} = 0, \quad \text{(B2)}$$

where

$$\mathcal{A} = \mathcal{W}\frac{(2N + 1)}{\mathcal{J}_N^2(\varpi)} - M(2N + 1)^2,$$
$$\mathcal{B} = -2MN^2(2N + 1),$$
$$\mathcal{C} = \mathcal{W}\sum_{n=0}^{N-1}\frac{(2n + 1)}{\mathcal{J}_n^2(\varpi)} - MN^4. \quad \text{(B3)}$$

It can be verified from Eq. (55) that

$$\mathcal{W}\frac{(2N + 1)}{\mathcal{J}_N^2(\varpi)} \geq \frac{M(N + 1)^4}{\displaystyle\sum_{n=0}^{N}\frac{(2n + 1)}{\mathcal{J}_n^2(\varpi)}} \times \frac{(2N + 1)}{\mathcal{J}_N^2(\varpi)}$$
$$= \frac{M(N + 1)^4(2N + 1)}{\displaystyle\sum_{n=0}^{N}(2n + 1)\frac{\mathcal{J}_N^2(\varpi)}{\mathcal{J}_n^2(\varpi)}}. \quad \text{(B4)}$$

Generally, white noise amplification occurs at low frequencies. So, we set our focal point of designing the flexible HD beamformer with a specified value of the WNG to low frequencies, i.e., the values of ϖ are small. In this case, for $n \leq N$, we have $\mathcal{J}_N^2(\varpi) \leq \mathcal{J}_n^2(\varpi)$, i.e., $[\mathcal{J}_N^2(\varpi)/\mathcal{J}_n^2(\varpi)] \leq 1$. Then, it can be verified that

$$0 < \sum_{n=0}^{N}(2n + 1)\frac{\mathcal{J}_N^2(\varpi)}{\mathcal{J}_n^2(\varpi)} < \sum_{n=0}^{N}(2n + 1) = (N + 1)^2. \quad \text{(B5)}$$

From Eqs. (B4) and (B5), we derive

$$\mathcal{A} = \mathcal{W}\frac{(2N + 1)}{\mathcal{J}_N^2(\varpi)} - M(2N + 1)^2$$
$$> \frac{M(N + 1)^4(2N + 1)}{(N + 1)^2} - M(2N + 1)^2$$
$$= MN^2(2N + 1) > 0. \quad \text{(B6)}$$

From Eq. (55), we have

$$\mathcal{C} = \mathcal{W}\sum_{n=0}^{N-1}\frac{(2n + 1)}{\mathcal{J}_n^2(\varpi)} - MN^4$$
$$\leq \frac{MN^4}{\displaystyle\sum_{n=0}^{N-1}\frac{(2n + 1)}{\mathcal{J}_n^2(\varpi)}}\sum_{n=0}^{N-1}\frac{(2n + 1)}{\mathcal{J}_n^2(\varpi)} - MN^4 = 0. \quad \text{(B7)}$$

From Eqs. (B6) and (B7), it can be checked that $\mathcal{AC} \leq 0$. The discriminant of the quadratic equation (B2) is

$$\Delta = \mathcal{B}^2 - 4\mathcal{AC} \geq \mathcal{B}^2 = 2M^2N^4(2N + 1)^2 > 0. \quad \text{(B8)}$$

Therefore, the quadratic Eq. (B2) have two distinct roots,

$$\alpha_{1,2} = \frac{-\mathcal{B} \pm \sqrt{\Delta}}{2\mathcal{A}}. \quad \text{(B9)}$$

Since $\mathcal{A} > 0$, $\mathcal{B} < 0$, and $\Delta \geq \mathcal{B}^2$, it can be checked that $\alpha_1 = -(\mathcal{B} + \sqrt{\Delta})/2\mathcal{A} \geq 0$, $\alpha_2 = -(\mathcal{B} - \sqrt{\Delta})/2\mathcal{A} \leq 0$. We recall that the value of α is in the range of $[0, 1]$. Consequently, the correct solution for α is

$$\alpha = \alpha_1 = \frac{2MN^2(2N + 1) + \sqrt{\Delta}}{2\mathcal{W}\dfrac{(2N + 1)}{\mathcal{J}_N^2(\varpi)} - 2M(2N + 1)^2}. \quad \text{(B10)}$$

Abhayapala, T. D. (**2008**). "Theory and design of high order sound field microphones using spherical microphone array," in *Proceedings of IEEE ICASSP*, p. 5268–5271.

Abhayapala, T. D., and Ward, D. B. (**2002**). "Theory and design of high order sound field microphones using spherical microphone array," in *Proceedings of IEEE ICASSP*.

Benesty, J., and Chen, J. (**2012**). *Study and Design of Differential Microphone Arrays* (Springer-Verlag, Berlin).

Benesty, J., Chen, J., and Cohen, I. (**2015**). *Design of Circular Differential Microphone Arrays* (Springer-Verlag, Berlin).

Benesty, J., Chen, J., and Huang, Y. (**2008**). *Microphone Array Signal Processing* (Springer-Verlag, Berlin).

Beranek, L. L. (**1986**). *Acoustics* (AIP, Woodbury, NY), pp. 1–491.

Brandstein, M., and Ward, D. (**2001**). *Microphone Arrays: Signal Processing Techniques and Applications* (Springer-Verlag, Berlin).

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.

Cox, H., Zeskind, R. M., and Kooij, T. (**1986**). "Practical supergain," *IEEE Trans. Acoust. Speech Sign. Process.* **34**, 393–398.

Crocco, M., and Trucco, A. (**2011**). "Design of robust superdirective arrays with a tunable tradeoff between directivity and frequency-invariance," *IEEE Trans. Sign. Process.* **59**, 2169–2181.

Doclo, S., and Moonen, M. (**2007**). "Superdirective beamforming robust against microphone mismatch," *IEEE Trans. Audio, Speech Lang. Process.* **15**, 617–631.

Elko, G. W. (**1996**). "Microphone array systems for hands-free telecommunication," *Speech Commun.* **20**, 229–240.

Elko, G. W. (**2000**). "Superdirectional microphone arrays," in *Acoustic Signal Processing for Telecommunication* (Springer-Verlag, Berlin), 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-Verlag, Berlin, Germany), Chap. 48, pp. 1021–1041.

Gannot, S., Burshtein, D., and Weinstein, E. (**2004**). "Analysis of the power spectral deviation of the general transfer function GSC," *IEEE Trans. Sign. Process.* **52**, 1115–1120.

Huang, G., Benesty, J., and Chen, J. (**2016a**). "Subspace superdirective beamforming with uniform circular microphone arrays," in *Proceedings of IEEE IWAENC*, pp. 1–5.

Huang, G., Benesty, J., and Chen, J. (**2016b**). "Superdirective beamforming based on the Krylov matrix," *IEEE/ACM Trans. Audio Speech Lang. Process.* **24**, 2531–2543.

Huang, G., Benesty, J., and Chen, J. (**2017a**). "Design of robust concentric circular differential microphone arrays," *J. Acoust. Soc. Am.* **141**, 3236–3249.

Huang, G., Benesty, J., and Chen, J. (**2017b**). "On the design of frequency-invariant beampatterns with uniform circular microphone arrays," *IEEE/ACM Trans. Audio Speech Lang. Process.* **25**, 1140–1153.

Li, C., Benesty, J., Huang, G., and Chen, J. (**2016**). "Subspace superdirective beamformers based on joint diagonalization," in *Proceedings of IEEE ICASSP*, pp. 400–404.

Li, Z., and Duraiswami, R. (**2007**). "Flexible and optimal design of spherical microphone arrays for beamforming," *IEEE Trans. Audio Speech Lang. Process.* **15**, 702–714.

Meyer, J. (**2001**). "Beamforming for a circular microphone array mounted on spherically shaped objects," *J. Acoust. Soc. Am.* **109**, 185–193.

Meyer, J., and Elko, G. (**2002**). "A highly scalable spherical microphone array based on a orthonormal decomposition of the soundfield," in *Proceedings of IEEE ICASSP*, pp. II178–II1784.

Meyer, J., and Elko, G. (**2008**). "Spherical harmonic modal beamforming for an augmented circular microphone array," in *Proceedings of IEEE ICASSP*, pp. 5280–5283.

Monzingo, R. A., and Miller, T. W. (**1980**). *Introduction to Adaptive Arrays* (SciTech, Raleigh, NC).

Rafaely, B. (**2005**). "Phase-mode versus delay-and-sum spherical microphone array processing," *IEEE Sign. Process. Lett.* **12**, 713–716.

Rafaely, B. (**2015**). *Fundamentals of Spherical Array Processing* (Springer-Verlag, Berlin).

Rafaely, B., Balmages, I., and Eger, L. (**2007a**). "High-resolution plane-wave decomposition in an auditorium using a dual-radius scanning spherical microphone array," *J. Acoust. Soc. Am.* **122**, 2661–2668.

Rafaely, B., and Khaykin, D. (**2011**). "Optimal model-based beamforming and independent steering for spherical loudspeaker arrays," *IEEE Trans. Audio Speech Lang. Process.* **19**, 2234–2238.

Rafaely, B., Weiss, B., and Bachmat, E. (**2007b**). "Spatial aliasing in spherical microphone arrays," *IEEE Trans. Sign. Process.* **55**, 1003–1010.

Shabtai, N. R., and Rafaely, B. (**2014**). "Generalized spherical array beamforming for binaural speech reproduction," *IEEE/ACM. Trans. Audio Speech Lang. Process.* **22**, 238–247.

Sun, H., Kellermann, W., Mabande, E., and Kowalczyk, K. (**2012**). "Localization of distinct reflections in rooms using spherical microphone array eigenbeam processing," *J. Acoust. Soc. Am.* **131**, 2828–2840.

Teal, P. D., Abhayapala, T. D., and Kennedy, R. A. (**2002**). "Spatial correlation for general distributions of scatterers," *IEEE Signal Process. Lett.* **9**, 305–308.

Teutsch, H. (**2007**). *Modal Array Signal Processing: Principles and Applications of Acoustic Wavefield Decomposition* (Springer-Verlag, Berlin).

Teutsch, H., and Kellermann, W. (**2006**). "Acoustic source detection and localization based on wavefield decomposition using circular microphone arrays," *J. Acoust. Soc. Am.* **120**, 2724–2736.

Wang, Y., Yang, Y., Ma, Y., and He, Z. (**2014**). "Robust high-order superdirectivity of circular sensor arrays," *J. Acoust. Soc. Am.* **136**, 1712–1724.

Williams, E. G. (**1999**). *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography* (Academic, New York).

Yan, S., Ma, Y., and Hou, C. (**2007**). "Optimal array pattern synthesis for broadband arrays," *J. Acoust. Soc. Am.* **122**, 2686–2696.

Yan, S., Sun, H., Svensson, U. P., Ma, X., and Hovem, J. M. (**2011**). "Optimal modal beamforming for spherical microphone arrays," *IEEE Trans. Audio, Speech, Lang. Process.* **19**, 361–371.

Zotkin, D. N., Duraiswami, R., and Gumerov, N. A. (**2008**). "Sound field decomposition using spherical microphone arrays," in *Proceedings of IEEE ICASSP*, pp. 277–280.

J. Acoust. Soc. Am. **143** (5), May 2018

Huang *et al.*   3035