Theoretical Analysis of Differential Microphone Array Beamforming and an Improved Solution

Chao Pan, Jingdong Chen, and Jacob Benesty

***Abstract*—Differential microphone arrays (DMAs), which are responsive to the differential sound pressure field, have attracted much attention due to their properties of frequency-invariant beampatterns, small apertures, and potential of maximum directivity. Traditionally, DMAs are designed and implemented in a multistage (cascade) way, where a proper time delay is used in each stage to form a beampattern of interest. Recently, it was reported that DMAs can be designed by solving a linear system of equations formed from the information about the nulls of the desired beampattern. This paper deals with the problem of beamforming with linear DMAs. Its major contributions are as follows. 1) By using the spatial $\mathcal{Z}$ transform, we present some theoretical analysis of both the traditional cascade and new null-constrained DMA beamforming. It is shown that the cascade and null-constrained DMAs of the same order with the same number of sensors are theoretically identical. 2) We develop a two-stage approach to the study of the robust DMA beamformer, which is based on the principle of maximizing the white noise gain (WNG). The first-stage of this approach is in the structure of the traditional non-robust DMA while the second-stage filter is optimized for improving the WNG. 3) Using the two-stage approach, we show that the robust DMA beamformer may introduce extra nulls in the beampattern at high frequencies; particularly, it introduces $M-N-1$ extra nulls if the interelement spacing is equal to half of the wavelength, where M and N are the number of sensors and the DMA order, respectively. 4) We develop a method that can solve the extra-null problem while maximizing the WNG in robust DMA beamforming, i.e., a robust solution with a frequency-invariant beampattern.**

***Index Terms*—Differential beamforming, differential microphone arrays (DMAs), directivity factor, directivity pattern, frequency-invariant beampattern, microphone arrays, robust DMAs, white noise gain.**

Manuscript received December 15, 2014; revised June 09, 2015; accepted August 08, 2015. Date of publication August 17, 2015; date of current version August 28, 2015. This work was supported in part by the Natural Science Foundation of China (NSFC) "Distinguished Young Scientists Fund" under Grant 61425005. The work of C. Pan was supported in part by the China Scholarship Council. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Nobutaka Ono.

C. Pan is with the Center of Intelligent Acoustics and Immersive Communications, Northwestern Polytechnical University, Xi'an 710072, China, and also with the INRS-EMT, University of Quebec, Montreal, QC H5A 1K6, Canada (e-mail: panchao2nwpu@mail.nwpu.edu.cn).

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

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

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

Digital Object Identifier 10.1109/TASLP.2015.2469142

I. Introduction

MICROPHONE arrays combined with beamforming techniques are often used for extracting a desired source signal from noisy observations. The basic idea of beamforming is to form a spatial filter so that signals coming from the desired direction are preserved as much as possible while signals from other directions are attenuated, where the amount of attenuation depends on the designed beampattern.

Among many microphone array beamforming techniques, differential microphone array (DMA) beamforming has attracted much research interest in recent years because of its properties of frequency-invariant beampatterns, small apertures, and potential of maximum directivity [1], [2]. DMA beamforming is originally inspired from the directional ribbon microphone, which measures the sound pressure gradient instead of the sound pressure [3], [4]. In traditional DMAs, the outputs are formed from a set of omnidirectional sensors in a multistage way as illustrated in Fig. 1 [5]. In such a structure, the output of a first-order DMA is the difference between the outputs of two adjacent omnidirectional microphones, and the output of a second-order DMA is the difference between the outputs of two first-order DMAs. Following the same line of ideas, any order of DMAs can be formed. The time delay in each stage is used for controlling the directions of the nulls and forming a beampattern of interest such as the cardioid, supercardioid, hypercardioid, and dipole [1]–[10]. With this multistage DMA structure, many contributions have been made over the last few decades. For example, an adaptive control of the null of a first-order DMA in the rear-half plane was investigated by combining the forward and backward facing cardioids in [11]. In [12], the performance of a first-order DMA in the presence of sensor mismatch with both additive and multiplicative errors was studied. In [13], the DMA was used for estimating the power spectral density of the noise signal from the microphone observations. The capability of interference rejection of a DMA was investigated in [14]. In [15], a method of beampattern design was proposed by controlling the tradeoff between the front-to-back energy ratio and smoothness in the frontal region of interest.

However, most existing DMAs are developed up to the second order because higher-order DMAs are found sensitive to both the sensors' self noise and the mismatch between sensors. Recently, with the use of the white noise gain (WNG) as the criterion, the robustness of the traditional DMAs was

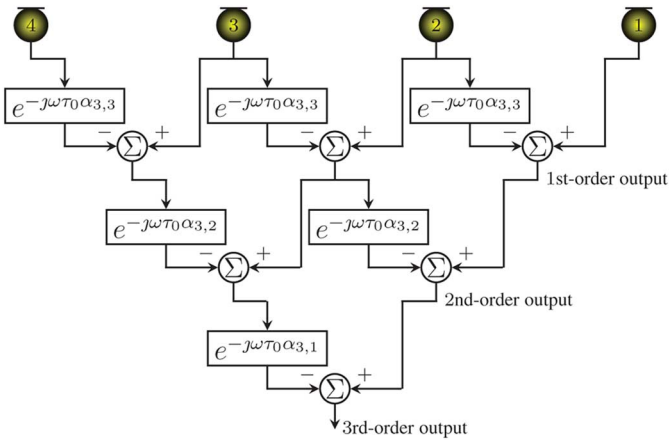

Fig. 1. Schematic diagram of a cascade DMA beamforming in the frequency domain, where $\alpha_{3,1}$, $\alpha_{3,2}$, and $\alpha_{3,3}$ correspond to three nulls in the DMA beampattern.

systematically studied [2]. The most intuitive way to improve the WNG is to increase the number of sensors and exploit the redundancy. A robust DMA design approach based on the maximum WNG (MWNG) principle was then developed in [2], where the filter was derived by minimizing the variance of the residual white noise subject to a group of constraints on the desired and nulls' directions.

This paper deals with the problem of differential beamforming with linear DMAs. The major contributions of this paper are as follows. First of all, it presents some theoretical analysis of both the traditional differential beamforming with the multistage structure and the recently proposed null-constrained DMA approach. By introducing the spatial $\mathcal{Z}$ transform, we show that the traditional and null-constrained DMAs with the same order and same number of sensors are theoretically identical. Second, it studies the limitations of both the traditional (non-robust) and the robust DMAs. We develop a two-stage DMA structure to investigate the performance of the MWNG principle based DMA, where the first-stage filter is equal to the traditional DMA beamformer and the second-stage filter is derived by maximizing the WNG. It is proved that all the zeros of the second-stage filter are located on the unit circle which, on the one hand, introduces extra nulls in the beampattern of the robust DMA and, on the other hand, provides a clue for solving this problem. Third, we develop a more general and flexible approach to the design of robust DMAs, which eliminates the extra nulls of the robust DMA by properly modifying the second-stage filter and pushing its zeros off the unit circle.

The organization of this paper is as follows. In Section II, we discuss the signal model and some important performance measures. In Section III, we introduce the spatial $\mathcal{Z}$ transform and study its properties. Section IV studies the traditional non-robust DMA and its white noise amplification problem. In Section V, we analyze the robust DMA and its extra-null problem. In Section VI, we develop a novel robust differential beamforming approach that circumvents the extra-null problem. Finally, some conclusions are presented in Section VII.

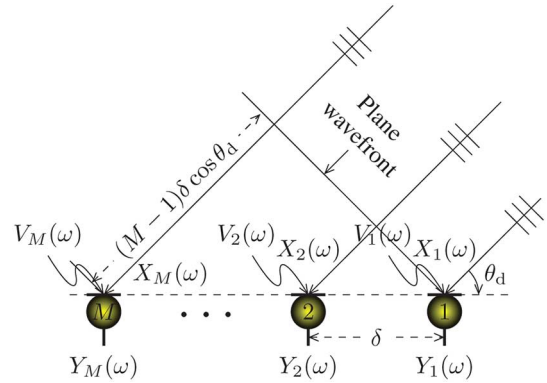

Fig. 2. Illustration of a uniform linear microphone array system.

II. Signal Model and Performance Measures

A. Signal Model

Let us consider the traditional signal model where a farfield source of interest propagates in an anechoic acoustic environment and impinges on a uniform linear array consisting of M omnidirectional microphones as illustrated in Fig. 2. The signal observed at the mth ($m = 1, 2, \ldots, M$) sensor in the frequency domain is

$$\begin{aligned} Y_m(\omega) &= X_m(\omega) + V_m(\omega) \\ &= e^{-j(m-1)\omega\tau_0\cos\theta_{\rm d}} X(\omega) + V_m(\omega), \end{aligned} \tag{1}$$

where ω is the angular frequency, j is the imaginary unit, $\tau_0 = \delta/c$ with δ being the distance between two neighboring sensors and c the sound velocity in the air, $\theta_{\rm d}$ is the incidence angle of the source of interest, $X(\omega)$ is the source signal, and $V_m(\omega)$ is the additive noise at the mth microphone.

If we put all the M observation signals together into a vector, the signal model in (1) can be expressed as

$$\begin{aligned} \mathbf{y}(\omega) &\triangleq [\, Y_1(\omega) \quad Y_2(\omega) \quad \cdots \quad Y_M(\omega) \,]^T \\ &= \mathbf{d}\,(\omega, \cos\theta_{\rm d})\, X(\omega) + \mathbf{v}(\omega), \end{aligned} \tag{2}$$

where the superscript ^{T} is the transpose operator,

$$\begin{aligned} &\mathbf{d}\,(\omega, \cos\theta_{\rm d}) \\ &\triangleq [\, 1 \quad e^{-j\omega\tau_0\cos\theta_{\rm d}} \quad \cdots \quad e^{-j(M-1)\omega\tau_0\cos\theta_{\rm d}} \,]^T \end{aligned} \tag{3}$$

is the phase-delay vector of length M (which is the same as the steering vector used in traditional beamforming), and the noise vector, $\mathbf{v}(\omega)$, is defined in a similar way to $\mathbf{y}(\omega)$.

Generally, the desired source signal coming from the direction $\theta_{\rm d}$ is extracted by applying a complex-valued linear filter:

$$\mathbf{h}(\omega) \triangleq [\, H_1(\omega) \quad H_2(\omega) \quad \cdots \quad H_{M-1}(\omega) \,]^T \tag{4}$$

to the observation signals vector, $\mathbf{y}(\omega)$, i.e.,

$$\begin{aligned} \chi(\omega) &= \mathbf{h}^H \mathbf{y}(\omega) \\ &= \mathbf{h}^H \mathbf{d}\,(\omega, \cos\theta_{\rm d})\, X(\omega) + \mathbf{h}^H \mathbf{v}(\omega), \end{aligned} \tag{5}$$

where the superscript ^{H} is the conjugate-transpose operator, and $\mathbf{h}^H\mathbf{d}(\omega, \cos\theta_{\rm d})X(\omega)$ and $\mathbf{h}^H\mathbf{v}(\omega)$ are the filtered desired signal and residual noise, respectively.

B. Beampattern

The most popular performance measure in beamforming is the beampattern (or directivity pattern), which describes the sensitivity of a beamformer to a plane wave impinging on the array from the direction θ. Mathematically, it is defined as

$$\begin{aligned}\mathcal{B}\left[\mathbf{h}\left(\omega\right),\theta\right] &\triangleq \mathbf{d}^H\left(\omega,\cos\theta\right)\mathbf{h}\left(\omega\right)\\ &= \sum_{m=1}^{M} H_m\left(\omega\right)e^{j(m-1)\omega\tau_0\cos\theta}.\end{aligned} \tag{6}$$

Ideally, an Nth-order DMA has a beampattern of the following form:

$$\mathcal{B}_N\left(\theta\right) = \sum_{n=0}^{N} \varrho_{N,n}\cos^n\theta, \tag{7}$$

where $\varrho_{N,n}, n = 0, 1, \ldots, N$ are real coefficients.

In differential beamforming with linear microphone arrays, the mainlobe of the beampattern is generally at the endfire direction, i.e., at 0°, which indicates that the linear DMAs do not have much flexibility in terms of beam steering. However, this does not reduce the wide usefulness of linear DMAs. If beam steering is needed in some applications, one may consider to use multiple linear DMAs or two-dimensional or three-dimensional array geometries with a better electronically steering capability. Since the focus here is on linear DMAs, we assume that $\theta_{\mathrm{d}} = 0^\circ$ in the rest of this paper.

C. SNR Gains

The signal-to-noise ratio (SNR), which quantifies the level of the signal of interest relative to the level of the unwanted noise, is one of the most important measures for evaluating the quality of the observations and beamforming.

By taking the first microphone as the reference point, the input SNR (iSNR) is defined, according to (1), as

$$\mathrm{iSNR}\left(\omega\right) \triangleq \frac{\phi_X\left(\omega\right)}{\phi_{V_1}\left(\omega\right)}, \tag{8}$$

where $\phi_X(\omega) \triangleq E[|X(\omega)|^2]$ and $\phi_{V_1}(\omega) \triangleq E[|V_1(\omega)|^2]$ are the variances of $X(\omega)$ and $V_1(\omega)$, respectively, with $E[\cdot]$ denoting mathematical expectation.

The output SNR (oSNR) is obtained from (5), i.e.,

$$\begin{aligned}\mathrm{oSNR}\left[\mathbf{h}\left(\omega\right)\right] &\triangleq \frac{\phi_X\left(\omega\right)\left|\mathbf{h}^H\left(\omega\right)\mathbf{d}\left(\omega,\cos 0^\circ\right)\right|^2}{\mathbf{h}^H\left(\omega\right)\mathbf{\Phi}_{\mathbf{v}}\left(\omega\right)\mathbf{h}\left(\omega\right)}\\ &= \frac{\phi_X\left(\omega\right)}{\phi_{V_1}\left(\omega\right)}\times\frac{\left|\mathbf{h}^H\left(\omega\right)\mathbf{d}\left(\omega,\cos 0^\circ\right)\right|^2}{\mathbf{h}^H\left(\omega\right)\mathbf{\Gamma}_{\mathbf{v}}\left(\omega\right)\mathbf{h}\left(\omega\right)},\end{aligned} \tag{9}$$

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

From (8) and (9), we deduce that the SNR gain is

$$\mathcal{G}\left[\mathbf{h}\left(\omega\right)\right] \triangleq \frac{\mathrm{oSNR}\left[\mathbf{h}\left(\omega\right)\right]}{\mathrm{iSNR}\left(\omega\right)} = \frac{\left|\mathbf{h}^H\left(\omega\right)\mathbf{d}\left(\omega,\cos 0^\circ\right)\right|^2}{\mathbf{h}^H\left(\omega\right)\mathbf{\Gamma}_{\mathbf{v}}\left(\omega\right)\mathbf{h}\left(\omega\right)}. \tag{10}$$

From (10), if the SNR gain is smaller than 1 (or 0 dB), the output SNR will be smaller than the input SNR. In this case, the noise is amplified relative to the desired signal.

D. White Noise Gain

The WNG, as its name indicates, is the beamformer's SNR gain in spatially white noise. In this case, the noise pseudo-coherence matrix becomes $\mathbf{\Gamma}_{\mathbf{v}}(\omega) = \mathbf{I}_M$, where $\mathbf{I}_M$ is the $M \times M$ identity matrix. Substituting $\mathbf{\Gamma}_{\mathbf{v}}(\omega) = \mathbf{I}_M$ into (10), we get the WNG due to the beamforming filter, $\mathbf{h}(\omega)$:

$$\mathcal{G}_{\mathrm{wn}}\left[\mathbf{h}\left(\omega\right)\right] \triangleq \frac{\left|\mathbf{h}^H\left(\omega\right)\mathbf{d}\left(\omega,\cos 0^\circ\right)\right|^2}{\mathbf{h}^H\left(\omega\right)\mathbf{h}\left(\omega\right)}, \tag{11}$$

which is generally used to measure the robustness of a beamformer against the sensors' self noise. It can be checked that

$$\mathcal{G}_{\mathrm{wn}}\left[\mathbf{h}\left(\omega\right)\right] \leq \mathbf{d}^H\left(\omega,\cos 0^\circ\right)\mathbf{d}\left(\omega,\cos 0^\circ\right) = M, \tag{12}$$

where the equality holds if and only if $\mathbf{h}(\omega) \propto \mathbf{d}(\omega, \cos 0^\circ)$, which corresponds to the traditional delay-and-sum beamformer. The WNG of a DMA beamformer is usually much smaller than M.

E. Directivity Factor

In a spherically isotropic (diffuse) noise field, the noise has an equal energy flow in all directions. In this case, the (i,j)th element of the noise pseudo-coherence matrix with a uniform linear array is

$$\left[\mathbf{\Gamma}_{\mathrm{dn}}(\omega)\right]_{i,j} = \mathrm{sinc}\left[\omega\tau_0(i-j)\right] = \frac{\sin\left[\omega\tau_0(i-j)\right]}{\omega\tau_0(i-j)}. \tag{13}$$

Substituting (13) into (10), we get the directivity factor:

$$\mathcal{G}_{\mathrm{dn}}\left[\mathbf{h}\left(\omega\right)\right] \triangleq \frac{\left|\mathbf{h}^H\left(\omega\right)\mathbf{d}(\omega,\cos 0^0)\right|^2}{\mathbf{h}^H\left(\omega\right)\mathbf{\Gamma}_{\mathrm{dn}}\left(\omega\right)\mathbf{h}\left(\omega\right)}. \tag{14}$$

It can be checked that

$$\mathcal{G}_{\mathrm{dn}}\left[\mathbf{h}\left(\omega\right)\right] \leq \mathbf{d}^H\left(\omega,\cos 0^\circ\right)\mathbf{\Gamma}_{\mathrm{dn}}^{-1}\left(\omega\right)\mathbf{d}\left(\omega,\cos 0^\circ\right). \tag{15}$$

This upper bound corresponds to the directivity factor of the superdirective beamformer [20], [21].

III. Spatial $\mathcal{Z}$ Transform

A. Definition

Similar to the definition of the $\mathcal{Z}$ transform of a discrete linear system, the spatial $\mathcal{Z}$ transform of the filter, $\mathbf{h}(\omega)$, at the output of the uniform linear array system, is defined as

$$\mathcal{H}\left(\mathcal{Z}\right) \triangleq \sum_{m=1}^{M} H_m\left(\omega\right)\mathcal{Z}^{-(m-1)}, \tag{16}$$

where $\mathcal{Z}$ is a complex number.

B. Some Properties and Relations

Property 3.1: When a beamformer is implemented in a multistage (cascade) way, its spatial $\mathcal{Z}$ transform is equal to the product of the spatial $\mathcal{Z}$ transforms of the beamforming filters at all the different stages, i.e.,

$$\mathcal{H}(\mathcal{Z}) = \prod_{\ell=1}^{L}\mathcal{H}^{(L,\ell)}(\mathcal{Z}), \tag{17}$$

where L is the total number of stages,

$$\mathcal{H}^{(L,\ell)}(\mathcal{Z}) = \sum_{i=1}^{K_\ell} H_i^{(L,\ell)}(\omega)\mathcal{Z}^{-(i-1)}, \tag{18}$$

$H_i^{(L,\ell)}(\omega), i = 1, 2, \ldots, K_\ell$, are the elements of the ℓth stage filter,

$$\mathbf{h}^{(L,\ell)}(\omega) \triangleq \left[H_1^{(L,\ell)}(\omega) \quad H_2^{(L,\ell)}(\omega) \quad \cdots \quad H_{K_\ell}^{(L,\ell)}(\omega) \right]^T,$$

and K_ℓ is the length of the filter at the ℓth stage.

The proof of this property is equivalent to the proof that the global filter, $\mathbf{h}(\omega)$, is equal to the convolutions of filters from the first stage to the last one, which is presented in Appendix A.

For example, the spatial $\mathcal{Z}$ transform of the multistage beamformer shown in Fig. 1 can be written as

$$\mathcal{H}(\mathcal{Z}) = \prod_{\ell=1}^{3} \left(1 - e^{-j\omega\tau_0\alpha_{3,\ell}}\mathcal{Z}^{-1}\right), \tag{19}$$

and the corresponding filter at each stage is

$$\begin{aligned}\mathbf{h}^{(3,\ell)}(\omega) &= \left[H_1^{(3,\ell)}(\omega) \quad H_2^{(3,\ell)}(\omega) \right]^T \\ &= \left[1 \quad -e^{-j\omega\tau_0\alpha_{3,\ell}} \right]^T,\end{aligned}$$

which is a vector of length 2.

From the definitions of the beampattern and spatial $\mathcal{Z}$ transform, it is clear that

$$\mathcal{B}\left[\mathbf{h}\left(\omega\right), \theta\right] = \mathcal{H}\left(e^{-j\omega\tau_0\cos\theta}\right). \tag{20}$$

Since the distortionless constraint at the direction 0° is required in a DMA, i.e., $\mathcal{B}[\mathbf{h}(\omega), 0^\circ] = 1$, the equivalent form in the spatial $\mathcal{Z}$ domain is

$$\mathcal{H}\left(e^{-j\omega\tau_0}\right) = 1. \tag{21}$$

Property 3.2: The nulls in $\mathcal{B}[\mathbf{h}(\omega), \theta]$ correspond to the zeros of $\mathcal{H}(\mathcal{Z})$ on the unit circle.

Property 3.3: The proof of this property is straightforward from (20). From (17) with (20), we see that the beampattern of a beamformer is equal to the product of the beampatterns from all stages, i.e.,

$$\begin{aligned}\mathcal{B}\left[\mathbf{h}\left(\omega\right), \theta\right] &= \prod_{\ell=1}^{L} \mathcal{B}\left[\mathbf{h}^{(L,\ell)}\left(\omega\right), \theta\right] \\ &= \prod_{\ell=1}^{L} \left[\sum_{i=1}^{K_\ell} H_i^{(L,\ell)}\left(\omega\right) e^{j(i-1)\omega\tau_0\cos\theta}\right]. \end{aligned}\tag{22}$$

This property shows that putting a null in any stage causes a null in the global beampattern, which is how the traditional cascade DMA forms a desired beampattern.

IV. Traditional DMA Beamforming

The traditional Nth-order DMA is constructed to respond to the Nth-order derivative of a sound pressure field along the axis of the linear array by utilizing the finite difference between the sensors' outputs to approximate the corresponding derivative [1]. The schematic diagram of the first-order, second-order, and third-order DMAs with this approach is shown in Fig. 1, where the time delays are determined by the nulls' directions. Also, in the traditional DMA, we always have $N + 1 = M$, which is assumed in this section.

According to *Property* 3.1, for a traditional Nth-order DMA with N nulls in the directions $\alpha_{N,n} = \cos\theta_{N,n}, n = 1, 2, \ldots, N$, its spatial $\mathcal{Z}$ transform can generally be expressed as

$$\mathcal{H}_{\mathrm{DMA}}(\mathcal{Z}) = \zeta(\omega) \prod_{n=1}^{N} \left(1 - e^{-j\omega\tau_0\alpha_{N,n}}\mathcal{Z}^{-1}\right), \tag{23}$$

where

$$\zeta(\omega) = \frac{1}{\prod_{n=1}^{N}\left[1 - e^{j\omega\tau_0(1-\alpha_{N,n})}\right]} \tag{24}$$

is a normalization factor to satisfy the distortionless constraint. The coefficients of this $\mathbf{h}_{\mathrm{DMA}}(\omega)$ filter are derived in Appendix B.

Replacing $\mathcal{Z}$ with $e^{-j\omega\tau_0\cos\theta}$ in (23), we deduce the beampattern of the traditional DMA beamformer:

$$\mathcal{B}\left[\mathbf{h}_{\mathrm{DMA}}\left(\omega\right), \theta\right] = \zeta(\omega) \prod_{n=1}^{N} \left[1 - e^{j\omega\tau_0(\cos\theta - \alpha_{N,n})}\right]. \tag{25}$$

It can be observed that, indeed, the beampattern has nulls in the directions $\theta_{N,n}, n = 1, 2, \ldots, N$.

Using the first-order approximation of the exponential function, i.e., $e^x \approx 1 + x$, the beampattern can be expressed as

$$\mathcal{B}\left[\mathbf{h}_{\mathrm{DMA}}\left(\omega\right), \theta\right] \approx \frac{\prod_{n=1}^{N}\left(\cos\theta - \alpha_{N,n}\right)}{\prod_{n=1}^{N}\left(1 - \alpha_{N,n}\right)}, \tag{26}$$

which is an Nth-order polynomial with respect to $\cos\theta$. Since the beampattern of an nth-order directional derivative of a plane wave along the array axis is $\cos^n\theta$, the DMA beampattern can be viewed as a linear combination of derivative patterns from the 0th order to the Nth order.

A. Equivalent Form

Recently, a new approach to the design of DMA beamformers was proposed based on the use of the nulls' information [2], which converts the differential beamforming problem to one of solving a simple linear system of equations. In the case of an Nth-order DMA with N distinct nulls, where $0^\circ < \theta_{N,1} < \theta_{N,2} < \cdots < \theta_{N,N} \leq 180^\circ$, the DMA beamformer is deduced by solving the following linear system of $N + 1$ equations:

$$\mathbf{D}\left(\omega, \alpha\right)\mathbf{h}\left(\omega\right) = \mathbf{i}, \tag{27}$$

where $\boldsymbol{\alpha} \triangleq \left[1 \quad \alpha_{N,1} \quad \cdots \quad \alpha_{N,N} \right]^T$ is a vector of length $N+1$,

$$\mathbf{D}\left(\omega, \alpha\right) \triangleq \begin{bmatrix} \mathbf{d}^H\left(\omega, 1\right) \\ \mathbf{d}^H\left(\omega, \alpha_{N,1}\right) \\ \vdots \\ \mathbf{d}^H\left(\omega, \alpha_{N,N}\right) \end{bmatrix} \tag{28}$$

is the constraint matrix of size $(N + 1) \times (N + 1)$, $\mathbf{d}(\omega, \alpha_{N,n})$ is defined in (3), and

$$\mathbf{i} \triangleq \left[1 \quad 0 \quad \cdots \quad 0 \right]^T \tag{29}$$

is a vector of length $N + 1$.

In Appendix C, the solution of (27) is deduced; it is observed that the resulting filter coefficients are equal to those of the traditional cascade DMA beamformer with the same order. As a matter of fact, according to *Property* 3.2, it is straightforward to see that this approach is equivalent to designing a beamformer with the zeros of $\mathcal{H}(\mathcal{Z})$ on the unit circle.

B. White Noise Amplification Problem

The filter coefficients of the traditional Nth-order DMA beamformer are given in Appendix B. Substituting (64) into (11), we deduce that the corresponding WNG is

$$\mathcal{G}_{\text{wn}}\left[\mathbf{h}_{\text{DMA}}(\omega)\right]=\frac{\prod_{n=1}^{N}\left|1-e^{j\omega\tau_0(1-\alpha_{N,n})}\right|^2}{\sum_{n=0}^{N}\left|\mathcal{F}_N^{(n)}\right|^2}. \tag{30}$$

Since $\mathcal{F}_N^{(N)}=1$, we have

$$0\le\mathcal{G}_{\text{wn}}\left[\mathbf{h}_{\text{DMA}}(\omega)\right]\le\Pi_{n=1}^{N}\left|1-e^{j\omega\tau_0(1-\alpha_{N,n})}\right|^2, \tag{31}$$

which means that the WNG approaches zero as the frequency decreases to zero. Moreover, the right-hand side of (31) shows that the WNG is also equal to zero when

$$\omega\tau_0(1-\alpha_{N,n})=2k\pi, k=0,\pm1,\pm2,\ldots. \tag{32}$$

For example, for the first-order cardioid, $\alpha_{1,1}=-1$ and $\omega\tau_0=\pi$ (see Fig. 3). Therefore, in the DMA design, the interelement spacing should always satisfy

$$\omega_{\max}\tau_0<\pi, \tag{33}$$

where $\omega_{\max}$ is the maximal angular frequency of interest.

Since the denominator on the right-hand side of (30) is both lower and upper bounded, replacing the exponential function of the numerator with its first-order approximation, we can deduce the behavior of the WNG at low frequencies:

$$\mathcal{G}_{\text{wn}}\left[\mathbf{h}_{\text{DMA}}(\omega)\right]\propto(\omega\tau_0)^{2N}, \tag{34}$$

which is equivalent to saying that the WNG approaches minus infinity at a speed of $6N$ dB/oct as the frequency approaches zero, where N is the order of the DMA. This well-known phenomenon can be seen in Fig. 3. One can conclude that the traditional DMA suffers badly from white noise amplification.

V. MWNG DMA Beamforming

As discussed above, traditional DMAs suffer from white noise amplification, particularly at low frequencies. To deal with this issue, a maximum WNG (MWNG) principle based DMA was developed in [2], [7], in which the filter is obtained from the optimization problem:

$$\min_{\mathbf{h}(\omega)}\mathbf{h}^H(\omega)\mathbf{h}(\omega)\quad\text{subject to}\quad\mathbf{D}(\omega,\alpha)\,\mathbf{h}(\omega)=\mathbf{i}, \tag{35}$$

where now $M\ge N+1$ and $\mathbf{D}(\omega,\alpha)$ is a matrix of size $(N+1)\times M$, which is no longer a square matrix. The solution is

$$\mathbf{h}_{\text{MWNG}}(\omega)=\mathbf{D}^H(\omega,\alpha)\left[\mathbf{D}(\omega,\alpha)\,\mathbf{D}^H(\omega,\alpha)\right]^{-1}\mathbf{i}. \tag{36}$$

This beamformer is also called minimum-norm filter since it is the minimum-norm solution of (27) with $M\ge N+1$. This filter maximizes the WNG by fully exploiting the fact that we have more microphones than required for a given order DMA. However, the corresponding beampattern introduces extra nulls at high frequencies, which will be discussed in the next subsection.

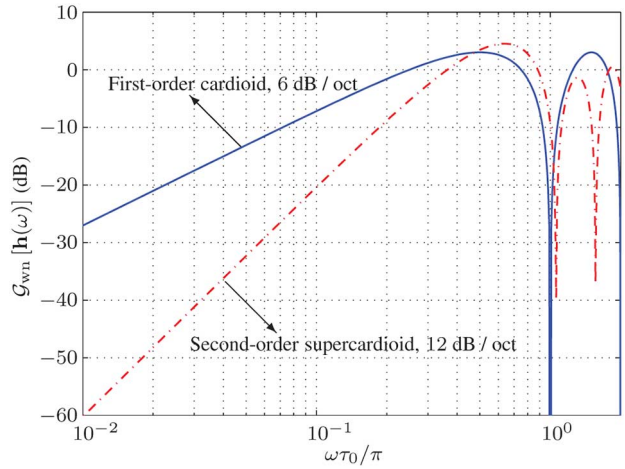

Fig. 3. WNG of the traditional DMA beamformer with different beampatterns, where the first-order cardioid is formed with 2 omnidirectional sensors and the second-order supercardioid is formed with 3 omnidirectional sensors.

A. Two-Stage Robust Filter

Since the nulls in the beampattern correspond to the zeros of the filter on the unit circle, the spatial $\mathcal{Z}$ transform of the filter that satisfies the constraints in (35) can be expressed as

$$\mathcal{H}(\mathcal{Z})=\mathcal{H}_{\text{DMA}}^{(2,1)}(\mathcal{Z})\mathcal{H}^{(2,2)}(\mathcal{Z}), \tag{37}$$

where

$$\mathcal{H}_{\text{DMA}}^{(2,1)}(\mathcal{Z})=\mathcal{H}_{\text{DMA}}(\mathcal{Z}) \tag{38}$$

is the spatial $\mathcal{Z}$ transform of the traditional Nth-order DMA beamformer, superscripts $(2,1)$ and $(2,2)$ denote the first- and second-stage filters in the two-stage structure, respectively, $\mathcal{H}^{(2,2)}(\mathcal{Z})$ is an $(M-N-1)$th-order polynomial with respect to $\mathcal{Z}^{-1}$ which corresponds to $M-N$ filter coefficients to be optimized.

According to (37), any filter satisfying the constraints in (35) can be expressed as

$$\mathbf{h}(\omega)=\mathbf{H}_{\text{DMA}}^{(2,1)}(\omega)\mathbf{h}^{(2,2)}(\omega), \tag{39}$$

where the (i,j)th element of $\mathbf{H}_{\text{DMA}}^{(2,1)}(\omega)$ is

$$\left[\mathbf{H}_{\text{DMA}}^{(2,1)}(\omega)\right]_{i,j}=\begin{cases}\left[\mathbf{h}_{\text{DMA}}^{(2,1)}(\omega)\right]_{i-j+1}, & j\le i\le j+N\\ 0, & \text{else}\end{cases}, \tag{40}$$

$\mathbf{h}_{\text{DMA}}^{(2,1)}(\omega)=\mathbf{h}_{\text{DMA}}(\omega)$ is the first-stage filter of length $N+1$, and $\mathbf{h}^{(2,2)}(\omega)$ is the second-stage filter of length $M-N$.

The first-stage filter is determined by the positions of the nulls in the desired beampattern. The second-stage filter corresponds to the redundancy. In the MWNG principle, the second-stage filter is deduced by maximizing the WNG.

Using the distortionless constraint of the filter $\mathbf{h}(\omega)$, we get

$$\mathbf{d}^H(\omega,1)\,\mathbf{H}_{\text{DMA}}^{(2,1)}(\omega)\mathbf{h}^{(2,2)}(\omega)=1, \tag{41}$$

which is equivalent to

$$\mathbf{h}^{(2,2)H}(\omega)\mathbf{d}_{\mathbf{H}}(\omega,1)=1, \tag{42}$$

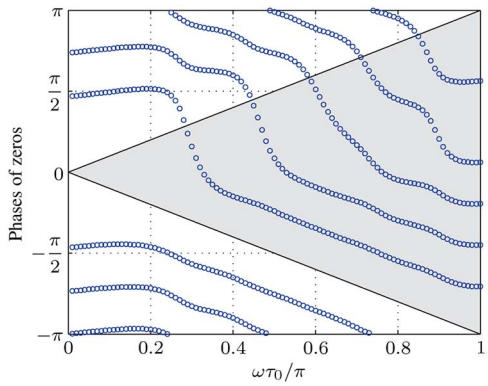

Fig. 4. Distribution of the zeros of $\mathcal{H}_{\mathrm{MWNG}}^{(2,2)}(\mathcal{Z})$ on the unit circle, where the desired beampattern is the second-order supercardioid and $M = 8$.

where

$$\mathbf{d}_{\mathbf{H}}(\omega,1) = \mathbf{H}_{\mathrm{DMA}}^{(2,1)H}(\omega)\mathbf{d}(\omega,1) = \left[1 \quad e^{-j\omega\tau_0} \quad \cdots \quad e^{-j(M-N-1)\omega\tau_0}\right]^T \tag{43}$$

is a vector of length $M - N$. The proof of (43) is given in Appendix D.

Therefore, the second-stage filter that maximizes the WNG is deduced by solving the following problem:

$$\begin{aligned} \min_{\mathbf{h}^{(2,2)}(\omega)} \quad & \mathbf{h}^{(2,2)H}(\omega)\mathbf{\Psi}(\omega)\mathbf{h}^{(2,2)}(\omega) \\ \text{subject to} \quad & \mathbf{h}^{(2,2)H}(\omega)\,\mathbf{d}_{\mathbf{H}}(\omega,1) = 1, \end{aligned} \tag{44}$$

where

$$\mathbf{\Psi}(\omega) = \mathbf{H}_{\mathrm{DMA}}^{(2,1)H}(\omega)\mathbf{H}_{\mathrm{DMA}}^{(2,1)}(\omega) \tag{45}$$

is a Hermitian matrix of size $(M - N) \times (M - N)$ whose elements are determined by the coefficients of the traditional DMA. From (44), the second-stage filter with maximum WNG principle is

$$\mathbf{h}_{\mathrm{MWNG}}^{(2,2)}(\omega) = \frac{\mathbf{\Psi}^{-1}(\omega)\mathbf{d}_{\mathbf{H}}(\omega,1)}{\mathbf{d}_{\mathbf{H}}^{H}(\omega,1)\mathbf{\Psi}^{-1}(\omega)\mathbf{d}_{\mathbf{H}}(\omega,1)}. \tag{46}$$

Property 5.1: All the $M - N - 1$ zeros of $\mathcal{H}_{\mathrm{MWNG}}^{(2,2)}(\mathcal{Z})$ lie on the unit circle.

Proof: See the proof in Appendix E.

Fig. 4 shows the phases of the zeros of $\mathbf{h}_{\mathrm{MWNG}}^{(2,2)}(\omega)$ in the $\mathcal{Z}$ plane, where the gray region is the visible region (once a zero appears in the visible region, there will be a null in the corresponding beampattern). As we can see, more and more zeros appear in the visible region as $\omega\tau_0$ increases.

From (37) and (20), the beampattern of the MWNG DMA can be written as

$$\mathcal{B}\left[\mathbf{h}_{\mathrm{MWNG}}(\omega),\theta\right] = \mathcal{B}\left[\mathbf{h}_{\mathrm{DMA}}^{(2,1)}(\omega),\theta\right] \times \mathcal{B}\left[\mathbf{h}_{\mathrm{MWNG}}^{(2,2)}(\omega),\theta\right]. \tag{47}$$

Since the beampattern of the MWNG filter is equal to the beampattern of the traditional DMA filter multiplied by the beampattern of the second-stage MWNG filter, the extra nulls are introduced by the second-stage filter.

Property 5.2: As $\omega\tau_0$ approaches π, the beampattern of the MWNG DMA beamformer introduces $M - N - 1$ extra nulls.

Proof: The proof is directly obtained from *Property* 5.1 because the visible region occupies the entire unit circle when $\omega\tau_0 = \pi$.

B. Performance Evaluation

In this subsection, we design the second-order supercardioid, which has two nulls in the directions 106° and 153°, with the MWNG principle to illustrate how it works, where $M = 8$, $\delta = 1$ cm, and the maximum frequency is $f_{\max} = 8$ kHz.

The beampatterns corresponding to the MWNG filter at four different frequencies are shown in the right column of Fig. 5. It is observed that these beampatterns are almost frequency-invariant at low frequencies [Fig. 5(c.2) and (d.2)]. The underlying reason is that the second-stage filter of the MWNG filter does not form nulls at low frequencies [see the dashdot-red lines in Fig. 5(c.1) and Fig. 5(d.1)]. However, at high frequencies, the second-stage filter forms nulls as expected [see dashdot-red lines in Fig. 5(a1) and (a.2)]; and as a result, the beampatterns of the MWNG filter introduces extra nulls, as shown in Fig. 5(a.2) and Fig. 5(b.2). Basically, the number of extra nulls increases as the frequency increases.

The SNR gains of the traditional DMA and MWNG filters are shown in Fig. 6. As one can see, the WNG of the MWNG filter is much higher than that of the traditional DMA filter, especially at low frequencies where the improvement is more than 20 dB. We observe that the directivity factor of the MWNG filter is smaller than that of the traditional DMA filter at frequencies lower than 5.6 kHz. The reason is that the second-stage filter does not have the maximum array response in the endfire direction [see the dashdot-red line in Fig. 5(c.1)]; at the same time, it amplifies the noise from all other directions. It is interesting to see that the MWNG filter achieves a higher directivity factor at high frequencies. The underlying reason is that the amount of spatial noise rejected by the extra nulls is more than the amount of amplified noise in other directions.

In conclusion, the MWNG filter can achieve a much higher WNG than the traditional DMA filter; however, it may sacrifice the directivity factor at low frequencies and the frequency invariance of the beampattern at high frequencies.

VI. ZOU DMA Beamforming

The nulls in the beampattern correspond to the zeros of the beamformer on the unit circle. So, if we push the zeros off the unit circle, the corresponding nulls are eliminated. Using this principle, we develop a new beamformer in this section to solve the extra-null problem of the MWNG filter by pushing the zeros of the second-stage filter off the unit circle. This new beamformer is named as ZOU DMA beamformer where "ZOU" stands for "zeros off unit circle."

A. Derivation

All the zeros of $\mathcal{H}_{\mathrm{MWNG}}^{(2,2)}(\mathcal{Z})$ are located on the unit circle in the $\mathcal{Z}$ plane. The second-stage filter can then be modified as follows:

$$\mathbf{h}_{\mathrm{ZOU}}^{(2,2)}(\omega) \propto \mathbf{\Sigma}(\omega)\,\mathbf{h}_{\mathrm{MWNG}}^{(2,2)}(\omega), \tag{48}$$

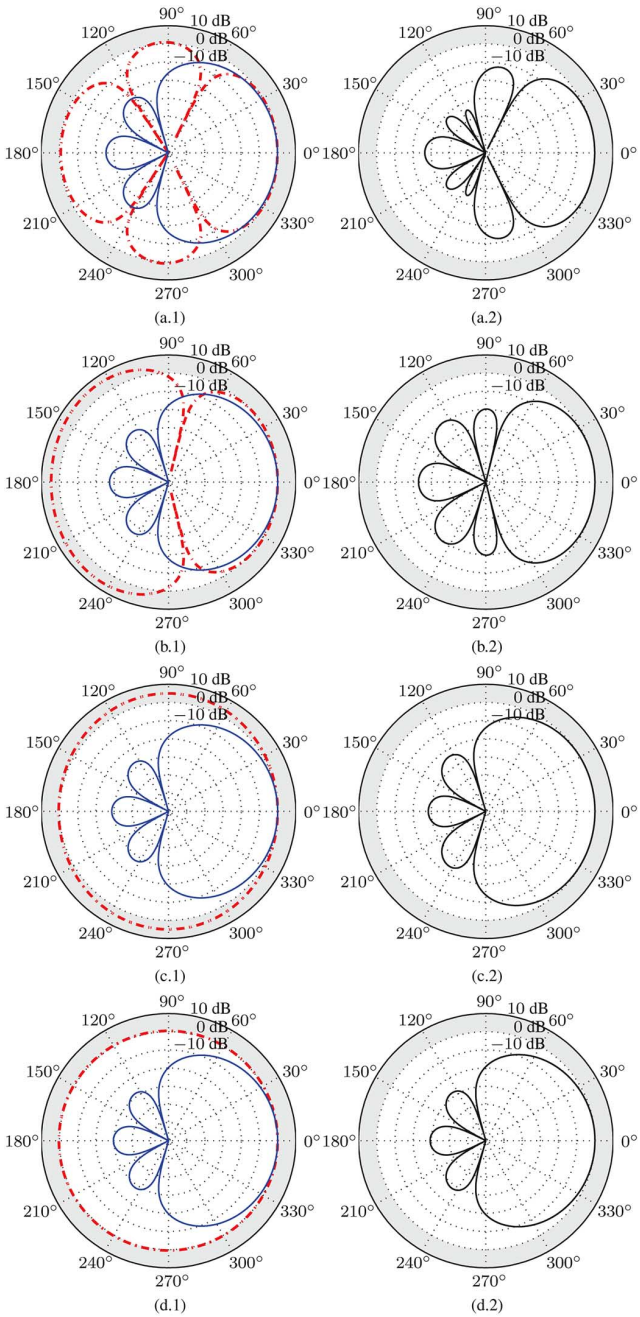

Fig. 5. Beampatterns of the MWNG beamformer and its two cascade filters at different frequencies (the solid-blue and dashdot-red lines in the left column correspond to the first- and second- stage filters, respectively, and the solid-black line in the right column corresponds to the MWNG filter). The desired beampattern is the second-order supercardioid (with two nulls in the directions 106° and 153°) with $M = 8$ and $\delta = 1$ cm. (a.1) Two cascade filters with $f = 8$ kHz (a.2) MWNG filter with $f = 8$ kHz (b.1) Two cascade filters with $f = 6$ kHz (b.2) MWNG filter with $f = 6$ kHz (c.1) Two cascade filters with $f = 4$ kHz (c.2) MWNG filter with $f = 4$ kHz (d.1) Two cascade filters with $f = 1$ kHz (d.2) MWNG filter with $f = 1$ kHz.

where

$$\mathbf{\Sigma}(\omega) = \mathrm{diag}\left[1, \sigma(\omega), \sigma^2(\omega), \ldots, \sigma^{M-N-1}(\omega)\right]. \quad (49)$$

One can verify that: if $\sigma(\omega) = 1$, the zeros of the modified filter are not changed; if $0 < \sigma(\omega) < 1$, all the zeros are inside the unit circle; and if $\sigma(\omega) > 1$, all the zeros are outside the unit circle.

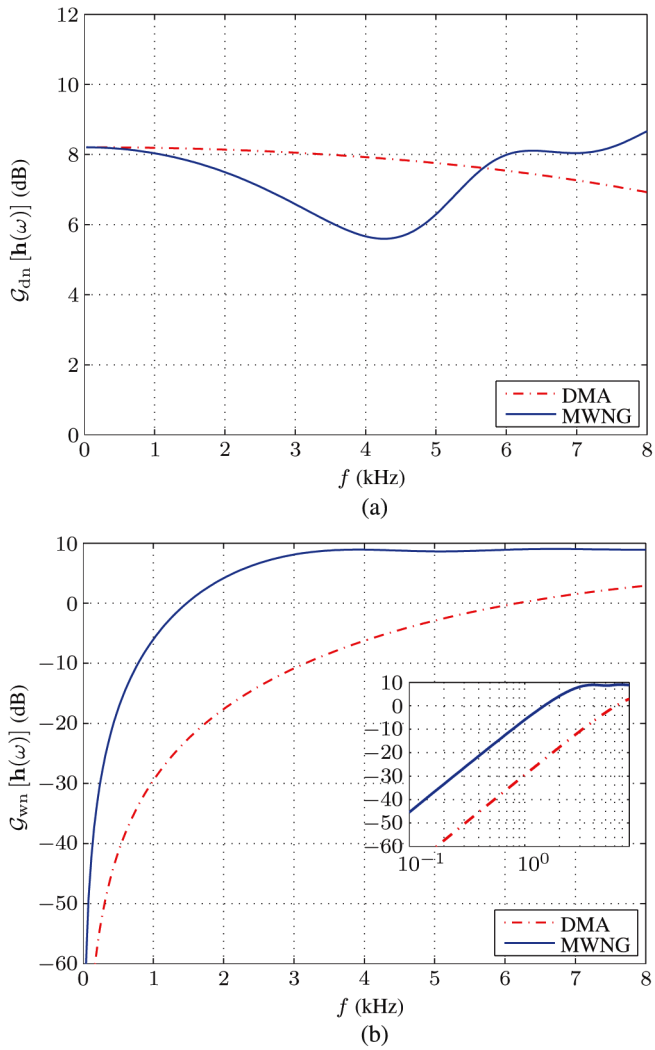

Fig. 6. SNR gains of the traditional DMA and MWNG filters to form the second-order supercardioid with $M = 8$ and $\delta = 1$ cm: (a) directivity factor and (b) WNG.

Taking the distortionless constraint into account, we deduce the second-stage ZOU filter:

$$\mathbf{h}_{\mathrm{ZOU}}^{(2,2)}(\omega) = \frac{\mathbf{\Sigma}(\omega)\,\mathbf{\Psi}^{-1}(\omega)\mathbf{d}_{\mathbf{H}}(\omega,1)}{\mathbf{d}_{\mathbf{H}}^{H}(\omega,1)\mathbf{\Sigma}(\omega)\,\mathbf{\Psi}^{-1}(\omega)\mathbf{d}_{\mathbf{H}}(\omega,1)}. \quad (50)$$

Substituting (50) into (39), we deduce the ZOU filter:

$$\mathbf{h}_{\mathrm{ZOU}}(\omega) = \mathbf{H}_{\mathrm{DMA}}^{(2,1)}(\omega)\mathbf{h}_{\mathrm{ZOU}}^{(2,2)}(\omega). \quad (51)$$

In particular, for $\sigma(\omega) = 0$ and $\sigma(\omega) = 1$, the ZOU filter degenerates, respectively, to the traditional DMA and MWNG beamformers, respectively. Therefore, the ZOU beamformer is in fact a tradeoff filter between the traditional DMA and MWNG filters, which handles the compromise between the WNG and the frequency invariance of the beampattern.

B. Performance Evaluation

Same as the evaluation of the MWNG filter, we consider to design the second-order supercardioid with the ZOU filter and compare its performance with the traditional DMA and MWNG

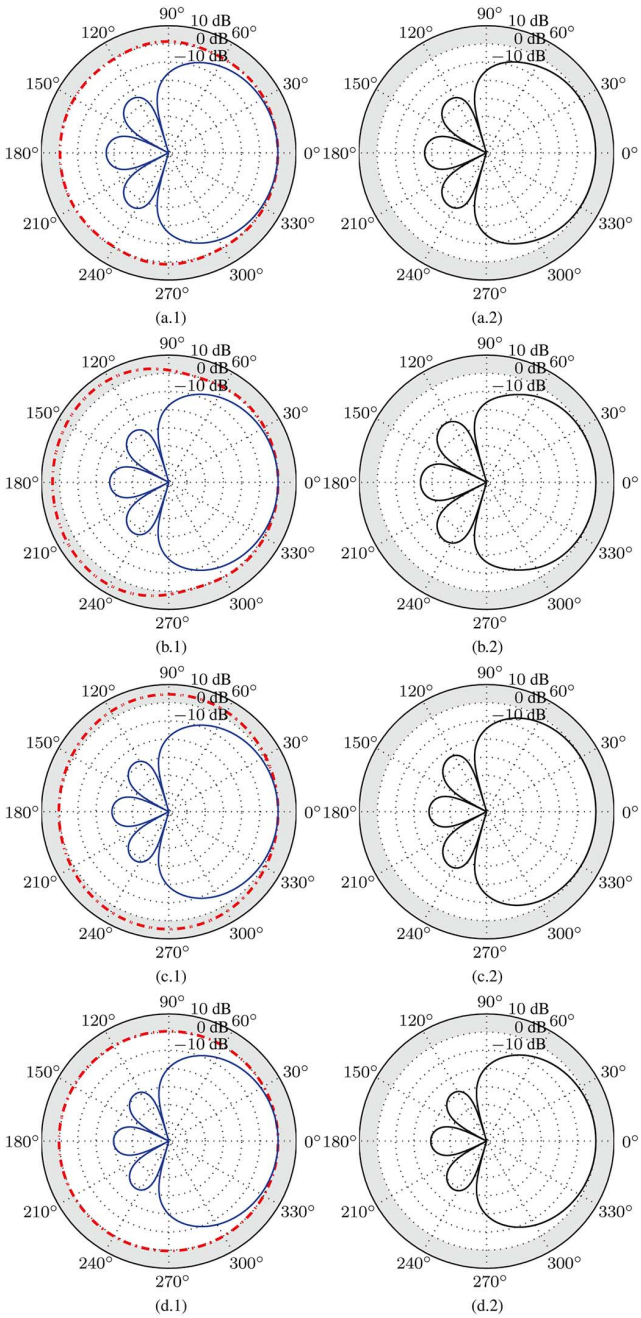

Fig. 7. Beampatterns of the ZOU beamformer and its two cascade filters at four different frequencies (the solid-blue and dashdot-red lines in the left column correspond to the first- and second- stage filters, respectively, and the solid-black line in the right column corresponds to the ZOU filter). The desired beampattern is the second-order supercardioid (with two nulls in the directions 106° and 153°) with $M = 8$ and $\delta = 1$ cm. (a.1) Two cascade filters with $f = 8$ kHz (a.2) ZOU filter with $f = 8$ kHz (b.1) Two cascade filters with $f = 6$ kHz (b.2) ZOU filter with $f = 6$ kHz (c.1) Two cascade filters with $f = 4$ kHz (c.2) ZOU filter with $f = 4$ kHz (d.1) Two cascade filters with $f = 1$ kHz (d.2) ZOU filter with $f = 1$ kHz.

filters. The simulation conditions are the same as those for the MWNG filter design. The parameter $\sigma(\omega)$ is chosen as

$$\sigma(\omega) = 1 - \frac{\omega\tau_0}{\pi}. \tag{52}$$

There are two reasons for this. 1) When $\omega\tau_0 = \pi$, $\sigma(\omega)$ should be equal to 0 because all the $M - N - 1$ extra nulls will occur in this case. 2) When $\omega\tau_0 = 0$, $\sigma(\omega)$ should be equal to 1 since no extra null will be introduced by the second-stage filter in this case.

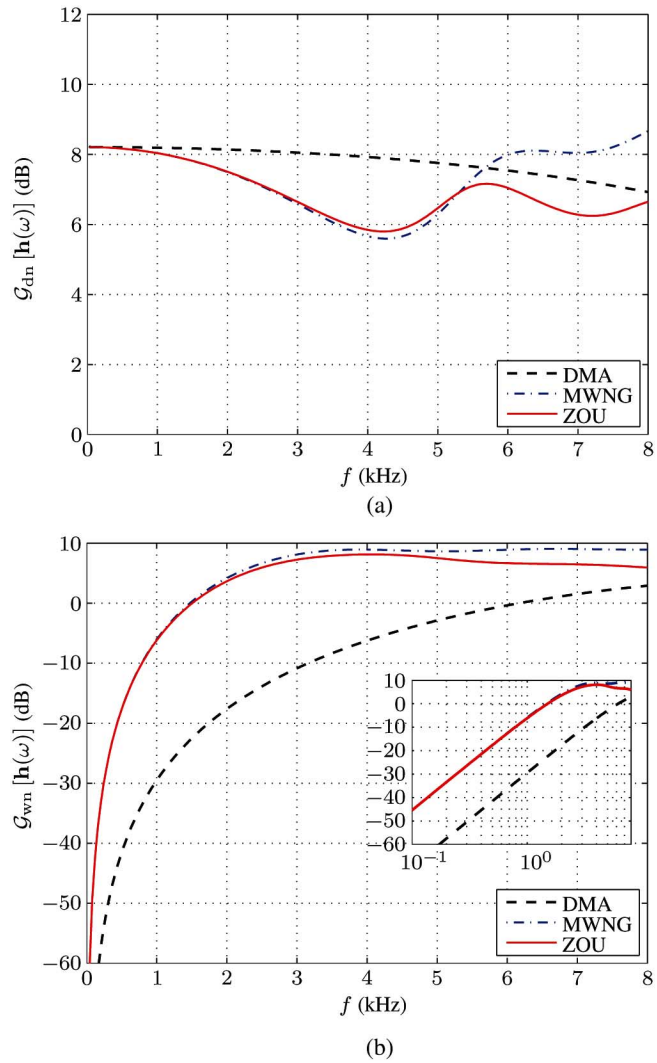

Fig. 8. SNR gains of the traditional DMA, MWNG, and ZOU filters to form the second-order supercardioid with $M = 8$ and $\delta = 1$ cm: (a) directivity factor and (b) WNG.

The beampatterns of the ZOU filter and its two cascade filters at four different frequencies are shown in the left column of Fig. 7. It can be seen that the second-stage filter forms no nulls and that the beampatterns are almost frequency-invariant, as shown in Fig. 7(a.1), (b.1), (c.1), and (d.1).

Fig. 8(a) and (b) shows the directivity factors and white noise gains of the traditional DMA, MWNG, and ZOU filters. As one can see, the ZOU filter achieves a WNG similar to the one with the MWNG filter at low frequencies. At the same time, the degradation of WNG at high frequencies is acceptable. Notice that the directivity factor of the ZOU filter at high frequencies is much smaller than that of the MWNG filter. The underlying reason is that the extra nulls and their capabilities of spatial noise rejection no longer exist in the ZOU filter.

VII. Conclusions

In this paper, we studied both the traditional cascade DMA and the null-constrained DMA using the spatial $\mathcal{Z}$ transform.

We proved that the traditional cascade DMA and the null-constrained DMA with the same order and the same number of microphones are theoretically identical. A two-stage DMA structure was then introduced to analyze the null-constrained robust DMA beamformer that maximizes the WNG given the order of the DMA and the number of sensors. In such a structure, the first-stage filter is the traditional DMA beamformer that forms the beampattern of interest and the second-stage filter attempts to maximize the WNG. Using this two-stage DMA structure, we showed that the robust DMA beamformer may introduce extra nulls to the beampattern and the extra nulls are from the second-stage filter. Meanwhile, we proved that all the zeros of the second-stage filter are located on the unit circle. By pushing the zeros of the second-stage filter off the unit circle, we developed a ZOU DMA beamformer that can eliminate the extra nulls in the robust DMA beamformer.

Appendix
Proof of Property 3.1

In this appendix, for the ease of exposition, we drop the dependency on the angular frequency in all variables. Let us further denote the output of the ℓth stage as $\chi_m^{(L,\ell)}, m = 1, 2, \ldots, M - \sum_{i=1}^{\ell}(K_i - 1)$. From the multistage structure, the mth output in the first stage can be expressed as

$$\chi_m^{(L,1)} = \sum_{i=1}^{K_1} Y_{m+i-1} H_i^{(L,1)}, \tag{53}$$

where $m = 1, 2, \ldots, M - K_1 + 1$ in this stage. In a similar way, we deduce the mth output in the second stage as

$$\chi_m^{(L,2)} = \sum_{j=1}^{K_2} \chi_{m+j-1}^{(L,1)} H_j^{(L,2)}, \tag{54}$$

where $m = 1, 2, \ldots, M - \sum_{i=1}^{2}(K_i - 1)$ in this stage. Substituting (53) into (54), we deduce that

$$\begin{aligned}\chi_m^{(L,2)} &= \sum_{j=1}^{K_2}\sum_{i=1}^{K_1} Y_{m+i+j-2} H_i^{(L,1)} H_j^{(L,2)} \\ &= \sum_{j=1}^{\sum_{k=1}^{2}(K_k-1)+1} Y_{m+j-1} \sum_{i=1}^{K_1} H_i^{(L,1)} H_{j-i+1}^{(L,2)}.\end{aligned} \tag{55}$$

Since the jth $[j = 1, 2, \ldots, \sum_{k=1}^{2}(K_k - 1) + 1]$ element of the convolution between the first- and second-stage filters is

$$\left[H^{(L,1)} * H^{(L,2)}\right]_j = \sum_{i=1}^{K_1} H_i^{(L,1)} H_{j-i+1}^{(L,2)}, \tag{56}$$

where $*$ denotes convolution and $[\cdot]_j$ is the jth element of the sequence, we can rewrite (55) as

$$\chi_m^{(L,2)} = \sum_{j=1}^{\sum_{k=1}^{2}(K_k-1)+1} Y_{m+j-1} \left[H^{(L,1)} * H^{(L,2)}\right]_j. \tag{57}$$

Continuing this process, we finally deduce the output of the last stage:

$$\begin{aligned}\chi_m^{(L,L)} = &\sum_{j=1}^{\sum_{k=1}^{L}(K_k-1)+1} Y_{m+j-1} \\ &\times \left[H^{(L,1)} * H^{(L,2)} * \cdots * H^{(L,L)}\right]_j,\end{aligned} \tag{58}$$

which means that 1) to implement this L-stage beamformer we need $\sum_{k=1}^{L}(K_k - 1) + 1$ sensors and 2) the global filter is equal to the convolutions of the sub-filters from the first stage to the last one. This completes the proof of *Property* 3.1.

Appendix
Filter Coefficients of the Cascade DMA Beamformer

Let us define

$$v_0 = e^{j\omega\tau_0}, \tag{59}$$

$$v_n = e^{j\omega\tau_0\alpha_{N,n}}, n = 1, 2, \ldots, N. \tag{60}$$

We can rewrite the spatial $\mathcal{Z}$ transform of the cascade DMA as

$$\begin{aligned}\mathcal{H}_{\mathrm{DMA}}(\mathcal{Z}) &= \frac{1}{\Pi_{n=1}^{N}(v_n - v_0)} \prod_{n=1}^{N}\left(v_n - \mathcal{Z}^{-1}\right) \\ &= \frac{1}{\Pi_{n=1}^{N}(v_n - v_0)} \sum_{n=0}^{N} (-1)^n \mathcal{F}_N^{(n)} \mathcal{Z}^{-n},\end{aligned} \tag{61}$$

where $\mathcal{F}_N^{(n)}$ $(n = 0, 1, \ldots, N)$ is an $(N-n)$th-order combination factor of $\{v_1, v_2, \ldots, v_N\}$ and has $C_N^n = \frac{N!}{n!(N-n)!}$ terms. Specifically, we have

$$\begin{aligned}\mathcal{F}_N^{(0)} &= \Pi_{n=1}^{N} v_n, \\ \mathcal{F}_N^{(1)} &= (v_2 v_3 v_4 \cdots v_N) + (v_1 v_3 v_4 \cdots v_N) \\ &\quad + (v_1 v_2 v_4 \cdots v_N) + \cdots + (v_1 v_2 v_3 \cdots v_{N-1}), \\ \mathcal{F}_N^{(2)} &= (v_3 v_4 v_5 \cdots v_N) + (v_1 v_4 v_5 \cdots v_N) \\ &\quad + (v_1 v_2 v_5 \cdots v_N) + \cdots + (v_1 v_2 v_3 \cdots v_{N-2}), \\ &\vdots \\ \mathcal{F}_N^{(N-1)} &= (v_N) + (v_{N-1}) + (v_{N-2}) + \cdots + (v_1), \\ \mathcal{F}_N^{(N)} &= 1.\end{aligned}$$

Since

$$\begin{aligned}\prod_{n=1}^{N}\left(v_n - \mathcal{Z}^{-1}\right) &= v_N \prod_{n=1}^{N-1}\left(v_n - \mathcal{Z}^{-1}\right) \\ &\quad - \mathcal{Z}^{-1} \prod_{n=1}^{N-1}\left(v_n - \mathcal{Z}^{-1}\right) \\ &= v_N \sum_{n=0}^{N-1} (-1)^n \mathcal{F}_{N-1}^{(n)} \mathcal{Z}^{-n} \\ &\quad - \mathcal{Z}^{-1} \sum_{n=0}^{N-1} (-1)^n \mathcal{F}_{N-1}^{(n)} \mathcal{Z}^{-n} \\ &= \sum_{n=0}^{N} (-1)^n \mathcal{F}_N^{(n)} \mathcal{Z}^{-n},\end{aligned} \tag{62}$$

it is not difficult to deduce the relationship of these combination factors:

$$\mathcal{F}_N^{(n)} = \mathcal{F}_{N-1}^{(n-1)} + v_N \mathcal{F}_{N-1}^{(n)}, n = 1, \ldots, N-1. \quad (63)$$

Finally, the traditional cascade DMA filter coefficients can be expressed as

$$H_{n+1}(\omega) = (-1)^n \frac{\mathcal{F}_N^{(n)}}{\Pi_{i=1}^N (v_i - v_0)}, n = 0, 1, \ldots, N. \quad (64)$$

Appendix
Filter Coefficients of the Null-Constrained DMA Beamformer

For the null-constrained DMA with distinct nulls, the filter is the solution of the following equation:

$$\mathbf{V}_{N+1}\mathbf{h}(\omega) = \mathbf{i}, \quad (65)$$

where

$$\mathbf{V}_{N+1} = \begin{bmatrix} 1 & v_0 & \cdots & v_0^N \\ 1 & v_1 & \cdots & v_1^N \\ 1 & v_2 & \cdots & v_2^N \\ \vdots & \vdots & \ddots & \vdots \\ 1 & v_{N-1} & \cdots & v_{N-1}^N \\ 1 & v_N & \cdots & v_N^N \end{bmatrix}, \quad (66)$$

$$\mathbf{h}(\omega) = [\, H_1(\omega) \quad H_2(\omega) \quad \cdots \quad H_{N+1}(\omega) \,]^T, \quad (67)$$

$$\mathbf{i} = [\, 1 \quad 0 \quad \cdots \quad 0 \,]^T. \quad (68)$$

Eliminating the first row and $(n+1)$th $(n = 0, 1, \ldots, N)$ column of $\mathbf{V}_{N+1}$, we get

$$\mathbf{V}_{N+1}^{(n)} = \begin{bmatrix} 1 & \cdots & v_1^{n-1} & v_1^{n+1} & \cdots & v_1^N \\ 1 & \cdots & v_2^{n-1} & v_2^{n+1} & \cdots & v_2^N \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & \cdots & v_{N-1}^{n-1} & v_{N-1}^{n+1} & \cdots & v_{N-1}^N \\ 1 & \cdots & v_N^{n-1} & v_N^{n+1} & \cdots & v_N^N \end{bmatrix}. \quad (69)$$

According to (65), the filter coefficients can be rewritten as

$$H_{n+1}(\omega) = (-1)^n \frac{\det\left[\mathbf{V}_{N+1}^{(n)}\right]}{\det\left[\mathbf{V}_{N+1}\right]}, n = 0, 1, \ldots, N, \quad (70)$$

where $\det[\cdot]$ denotes the determinant of a matrix.

Combining (64) with (70), the two DMA approaches are equivalent if, $\forall n = 0, 1, \ldots, N$,

$$\det\left[\mathbf{V}_{N+1}^{(n)}\right] = \frac{\mathcal{F}_N^{(n)}}{\Pi_{p=1}^N (v_p - v_0)} \det\left[\mathbf{V}_{N+1}\right]. \quad (71)$$

Let us evaluate the determinant of $\mathbf{V}_{N+1}$. $\forall i = 2, 3, \ldots, N+1$, we replace the ith column with

$$[\mathbf{V}_{N+1}]_{:,i} - v_N [\mathbf{V}_{N+1}]_{:,i-1},$$

where $[\cdot]_{:,i}$ denotes the ith column of a matrix. The determinant of the resulting matrix is the same as the original one, i.e.,

$$\begin{aligned} &\det\left[\mathbf{V}_{N+1}\right] \\ &= \begin{vmatrix} 1 & (v_0 - v_N) & \cdots & v_0^{N-1}(v_0 - v_N) \\ 1 & (v_1 - v_N) & \cdots & v_2^{N-2}(v_1 - v_N) \\ 1 & (v_2 - v_N) & \cdots & v_3^{N-2}(v_2 - v_N) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & (v_{N-1} - v_N) & \cdots & v_{N-1}^{M-2}(v_{N-1} - v_N) \\ 1 & 0 & \cdots & 0 \end{vmatrix} \\ &= \det\left[\mathbf{V}_N\right](-1)^{N+2} \prod_{p=0}^{N-1} (v_p - v_N) \\ &= \det\left[\mathbf{V}_N\right] \prod_{p=0}^{N-1} (v_N - v_p) \\ &= \det\left[\mathbf{V}_{N-1}\right] \left[\prod_{p=0}^{N-2} (v_{N-1} - v_p)\right] \left[\prod_{p=0}^{N-1} (v_N - v_p)\right] \\ &= \prod_{0 \le i < N, i < j \le N} (v_j - v_i). \end{aligned} \quad (72)$$

It is then easy to deduce that

$$\det\left[\mathbf{V}_{N+1}\right] = \frac{\det\left[\mathbf{V}_N\right]}{\prod_{p=0}^{N-1} (v_N - v_p)}. \quad (73)$$

For $n = 0$, $\mathbf{V}_{N+1}^{(n)}$ is a Vandermonde matrix. The corresponding determinant is

$$\begin{aligned} \det\left[\mathbf{V}_{N+1}^{(0)}\right] &= \prod_{p=1}^N v_p \prod_{1 \le i < N, i < j \le N} (v_j - v_i) \\ &= \frac{\prod_{p=1}^N v_p}{\prod_{p=1}^N (v_p - v_0)} \prod_{0 \le i < N, i < j \le N} (v_j - v_i) \\ &= \frac{\prod_{p=1}^N v_p}{\prod_{p=1}^N (v_p - v_0)} \det\left[\mathbf{V}_{N+1}\right] \\ &= \frac{\mathcal{F}_N^{(0)}}{\prod_{p=1}^N (v_p - v_0)} \det\left[\mathbf{V}_{N+1}\right]. \end{aligned} \quad (74)$$

Therefore, the first filter coefficients of the two DMA approaches are equal.

For $n = N$, $\mathbf{V}_{N+1}^{(N)}$ is still a Vandermonde matrix. The corresponding determinant is

$$\begin{aligned} \det\left[\mathbf{V}_{N+1}^{(N)}\right] &= \prod_{1 \le i < N, i < j \le N} (v_j - v_i) \\ &= \frac{1}{\prod_{p=1}^N (v_p - v_0)} \prod_{0 \le i < N, i < j \le N} (v_j - v_i) \\ &= \frac{\mathcal{F}_N^{(N)}}{\prod_{p=1}^N (v_p - v_0)} \det\left[\mathbf{V}_{N+1}\right]. \end{aligned} \quad (75)$$

Therefore, the last filter coefficients of the two DMA approaches are also equal.

For $0 < n < N$, replacing the $(n+1)$th column of $\mathbf{V}_{N+1}^{(n)}$ with

$$\left[\mathbf{V}_{N+1}^{(n)}\right]_{:,i} - v_N^2 \left[\mathbf{V}_{N+1}^{(n)}\right]_{:,i-1}$$

and the other columns with

$$\left[\mathbf{V}_{N+1}^{(n)}\right]_{:,i} - v_N \left[\mathbf{V}_{N+1}^{(n)}\right]_{:,i-1},$$

the determinant of $\mathbf{V}_{N+1}^{(n)}$ can be deduced as in (76). Since $\forall i = 1, 2, \ldots, N-1$ (see equation at the bottom of the page),

$$v_i^{n-1}(v_i - v_N)^2 = \left(v_i^n + v_N v_i^{n-1}\right)(v_i - v_N), \tag{77}$$

we can see that

$$\begin{aligned}&\det\left[\mathbf{V}_{N+1}^{(n)}\right]\\ &= \left\{\det\left[\mathbf{V}_N^{(n-1)}\right] + v_N \det\left[\mathbf{V}_N^{(n)}\right]\right\} \times \Pi_{p=1}^{N-1}(v_N - v_p).\end{aligned} \tag{78}$$

Using the mathematical induction method, we can prove that (71) holds $\forall N \geq 1$.

Basic Step: For $N = 1$, (71) obviously holds according to (74) and (75).

Inductive Step: Assuming that $\forall n = 0, 1, \ldots, N-1$,

$$\det\left[\mathbf{V}_N^{(n)}\right] = \frac{\mathcal{F}_{N-1}^{(n)}}{\Pi_{p=1}^{N-1}(v_p - v_0)} \det\left[\mathbf{V}_N\right]. \tag{79}$$

Substituting (79) into (78), we can deduce that

$$\begin{aligned}&\det\left[\mathbf{V}_{N+1}^{(n)}\right]\\ &= \left\{\det\left[\mathbf{V}_N^{(n-1)}\right] + v_N \det\left[\mathbf{V}_N^{(n)}\right]\right\} \times \Pi_{p=1}^{N-1}(v_N - v_p)\\ &= \frac{\Pi_{p=1}^{N-1}(v_N - v_p)}{\Pi_{p=1}^{N-1}(v_p - v_0)} \times \left[\mathcal{F}_{N-1}^{(n-1)} + v_N \mathcal{F}_{N-1}^{(n)}\right] \times \det\left[\mathbf{V}_N\right].\end{aligned} \tag{80}$$

Substituting (73) into (80), we get

$$\begin{aligned}&\det\left[\mathbf{V}_{N+1}^{(n)}\right]\\ &= \frac{1}{\Pi_{p=1}^{N}(v_p - v_0)} \times \left[\mathcal{F}_{N-1}^{(n-1)} + v_N \mathcal{F}_{N-1}^{(n)}\right]\\ &\quad \times \det\left[\mathbf{V}_{N+1}\right].\end{aligned} \tag{81}$$

Substituting (63) into (81), we finally find that

$$\det\left[\mathbf{V}_{N+1}^{(n)}\right] = \frac{\mathcal{F}_N^{(n)}}{\Pi_{p=1}^{N}(v_p - v_0)} \det\left[\mathbf{V}_{N+1}\right], \tag{82}$$

which completes the proof for $0 < n < N$.

Appendix
Proof of (43)

From (43), the nth ($n = 1, 2, \ldots, M-N$) element of the vector $\mathbf{d}_{\mathbf{H}}(\omega, 1)$ can be written as

$$\begin{aligned}\left[\mathbf{d}_{\mathbf{H}}(\omega,1)\right]_n &= \sum_{m=1}^{M} \left[\mathbf{H}_{\mathrm{DMA}}^{(2,1)}(\omega)\right]_{m,n}^{*} \times \left[\mathbf{d}(\omega,1)\right]_m\\ &= \sum_{m=1}^{M} \left[\mathbf{H}_{\mathrm{DMA}}^{(2,1)}(\omega)\right]_{m,n}^{*} \times e^{-j(m-1)\omega\tau_0}.\end{aligned} \tag{83}$$

where the superscript * denotes the complex-conjugate operation. Substituting (40) into (83), we deduce that

$$\begin{aligned}\left[\mathbf{d}_{\mathbf{H}}(\omega,1)\right]_n &= \sum_{m=n}^{n+N} \left[\mathbf{h}_{\mathrm{DMA}}^{(2,1)}(\omega)\right]_{m-n+1}^{*} \times e^{-j(m-1)\omega\tau_0}\\ &= e^{-j(n-1)\omega\tau_0} \sum_{i=1}^{N+1} \left[\mathbf{h}_{\mathrm{DMA}}^{(2,1)}(\omega)\right]_i^{*}\\ &\quad \times e^{-j(i-1)\omega\tau_0}.\end{aligned} \tag{84}$$

According to the distortionless constraint in (27), we have

$$\sum_{i=1}^{N+1} \left[\mathbf{h}_{\mathrm{DMA}}^{(2,1)}(\omega)\right]_i^{*} \times e^{-j(i-1)\omega\tau_0} = 1. \tag{85}$$

Substituting (85) into (84), we finally deduce that

$$\left[\mathbf{d}_{\mathbf{H}}(\omega,1)\right]_n = e^{-j(n-1)\omega\tau_0}, \tag{86}$$

which completes the proof.

Appendix
Proof of Property 5.1

In this appendix, we prove that the zeros of $\mathbf{h}_{\mathrm{MWNG}}^{(2,2)}(\omega)$ are all located on the unit circle in the $\mathcal{Z}$ plane. This proof is similar to the one given in [22] showing that all the zeros of the maximum eigenvector of a symmetric Toeplitz matrix are located on the unit circle.

$$\det\left[\mathbf{V}_{N+1}^{(n)}\right] = \begin{vmatrix} 1 & (v_1 - v_N) & \cdots & v_1^{n-2}(v_1 - v_N) & v_1^{n-1}\left(v_1^2 - v_N^2\right) & \cdots & v_1^{N-1}(v_1 - v_N) \\ 1 & (v_2 - v_N) & \cdots & v_2^{n-2}(v_2 - v_N) & v_2^{n-1}\left(v_2^2 - v_N^2\right) & \cdots & v_2^{N-1}(v_2 - v_N) \\ \vdots & & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & (v_{N-1} - v_N) & \cdots & v_{N-1}^{n-2}(v_{N-1} - v_N) & v_{N-1}^{n-1}\left(v_{N-1}^2 - v_N^2\right) & \cdots & v_{N-1}^{N-1}(v_{N-1} - v_N) \\ 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{vmatrix}. \tag{76}$$

One can easily verify that the normalized $\mathbf{h}_{\mathrm{MWNG}}^{(2,2)}(\omega)$ is also the unique solution of

$$\max_{\mathbf{h}} \frac{\mathbf{h}^H\mathbf{A}\mathbf{h}}{\mathbf{h}^H\mathbf{B}\mathbf{h}}, \tag{87}$$

where

$$\mathbf{h} = [\, h(0) \quad h(1) \quad \cdots \quad h(L-1) \,]^T, \tag{88}$$
$$\mathbf{A} = \mathbf{d}_{\mathbf{H}}(\omega)\mathbf{d}_{\mathbf{H}}^H(\omega), \tag{89}$$
$$\mathbf{B} = \mathbf{H}_{\mathrm{DMA}}^{(2,1)H}(\omega)\mathbf{H}_{\mathrm{DMA}}^{(2,1)}(\omega). \tag{90}$$

It can be checked that both $\mathbf{A}$ and $\mathbf{B}$ are Hermitian Toeplitz matrices of size $L \times L$ with $L = M - N$. By defining

$$a(i-j) = [\mathbf{A}]_{i,j}, i, j = 1, 2, \ldots, L, \tag{91}$$
$$b(i-j) = [\mathbf{B}]_{i,j}, i, j = 1, 2, \ldots, L, \tag{92}$$

the optimization problem in (87) can be expressed as

$$\max_{\{g(k)\}} \frac{\sum_{k=-(L-1)}^{L-1} a(k)g(k)}{\sum_{k=-(L-1)}^{L-1} b(k)g(k)}, \tag{93}$$

where

$$g(k) = \sum_{i=0}^{L-1} h^*(i)h(i-k). \tag{94}$$

It can be deduced that

$$\begin{aligned}
\mathcal{G}(\mathcal{Z}) &= \sum_{k=-(L-1)}^{L-1} g(k)\mathcal{Z}^{-k} \\
&= \sum_{k=-(L-1)}^{L-1} \sum_{i=0}^{L-1} h^*(i)h(i-k)\mathcal{Z}^{-k} \\
&= \sum_{i=0}^{L-1} h^*(i) \sum_{k=-(L-1)}^{L-1} h(i-k)\mathcal{Z}^{-k} \\
&= \sum_{i=0}^{L-1} h^*(i)\mathcal{Z}^{-i} \sum_{j=0}^{L-1} h(j)\mathcal{Z}^{j} \\
&= \mathcal{H}^*(\mathcal{Z}^*)\mathcal{H}(\mathcal{Z}^{-1}) \\
&= \prod_{i=1}^{L-1} (1 - \mathcal{Z}_i^*\mathcal{Z}^{-1})(1 - \mathcal{Z}_i\mathcal{Z})
\end{aligned} \tag{95}$$

where $\mathcal{Z}_i, i = 1, 2, \ldots, L-1$ are the roots of $\mathcal{H}(\mathcal{Z})$. Therefore, we have at most 2^{L-1} different $\{h(k)\}$ with the same $\{g(k)\}$. If $\{h(k)\}$ is the unique solution to (86), the roots should satisfy

$$\mathcal{Z}_i^* = \mathcal{Z}_i^{-1}, \forall i,$$

which means that all the roots are on the unit circle in the $\mathcal{Z}$ plane.

REFERENCES

[1] G. W. Elko and J. Meyer, "Microphone arrays," in *Springer Handbook of Speech Processing*, J. Benesty, M. M. Sondhi, and Y. Huang, Eds. Berlin, Germany: Springer-Verlag, 2008, ch. 50, pp. 1021–1041.

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

[3] H. F. Olson, "A uni-directional ribbon microphone," *J. Acoust. Soc. Amer.*, vol. 5, pp. 139–147, Jun. 1932.

[4] H. F. Olson, "Gradient microphones," *J. Acoust. Soc. Amer.*, vol. 17, pp. 192–198, Jan. 1946.

[5] G. M. Sessler and J. E. West, "Directional transducers," *IEEE Trans. Audio Electroacoust.*, vol. AE-19, no. 1, pp. 19–23, Mar. 1971.

[6] J. Chen and J. Benesty, "A general approach to the design and implementation of linear differential microphone arrays," in *Proc. Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA)*, 2013.

[7] J. Chen, J. Benesty, and C. Pan, "On the design and implementation of linear differential microphone arrays," *J. Acoust. Soc. Amer.*, vol. 136, pp. 3097–3113, Dec. 2014.

[8] J. Eargle, *The Microphone Book*. Waltham, MA, USA: Focal Press, 2004.

[9] T. D. Abhayapala and A. Gupta, "Higher order differential-integral microphone arrays," *J. Acoust. Soc. Amer.*, vol. 127, pp. EL227–EL233, May 2010.

[10] R. N. Marshall and W. R. Harry, "A new microphone providing uniform directivity over an extended frequency range," *J. Acoust. Soc. Amer.*, vol. 12, pp. 481–497, Apr. 1941.

[11] G. W. Elko and A.-T. N. Pong, "A simple adaptive first-order differential microphone," in *Proc. IEEE Workshop Appl. Signal Process. Audio Acoust. (WASPAA)*, 1995, pp. 169–172.

[12] M. Buck, "Aspects of first-order differential microphone arrays in the presence of sensor imperfections," *Eur. Trans. Telecomm.*, vol. 13, pp. 115–122, Mar.–Apr. 2002.

[13] M. Ihle, "Differential microphone arrays for spectral subtraction," in *Proc. Int. Workshop Acoust. Signal Enhanc. (IWAENC)*, 2003.

[14] J. Benesty, M. Souden, and Y. Huang, "A perspective on differential microphone arrays in the context of noise reduction," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 20, no. 2, pp. 699–704, Feb. 2012.

[15] E. De Sena, H. Hacihabiboglu, and Z. Cvetkovic, "On the design and implementation of higher-order differential microphones," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 20, no. 1, pp. 162–174, Jan. 2012.

[16] C. Pan, J. Chen, and J. Benesty, "Performance study of the MVDR beamformer as a function of the source incidence angle," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 22, no. 1, pp. 67–79, Jan. 2014.

[17] E. N. Gilbert and S. P. Morgan, "Optimum design of directive antenna arrays subject to random variations," *Bell Syst. Tech. J.*, vol. 34, no. 3, pp. 637–663, May 1955.

[18] M. Crocco and A. Trucco, "Design of robust superdirective arrays with a tunable tradeoff between directivity and frequency-invariance," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2169–2181, May 2011.

[19] S. Yan and Y. Ma, "Robust supergain beamforming for circular array via second-order cone programming," *Appl. Acoust.*, vol. 66, no. 9, pp. 1018–1032, Sep. 2005.

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

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

[22] J. Makhoul, "On the eigenvectors of symmetric Toeplitz matrices," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-29, no. 4, pp. 868–872, Aug. 1981.

Chao Pan (SM'13) was born in 1989. He received the Bachelor degree in electronics and information engineering from the Northwestern Polytechnical University (NPU) in 2011. He is currently a Ph.D. student in information and communication engineering at NPU and also a visiting Ph.D. student at INRS-EMT, University of Quebec. His research interests are in speech enhancement, noise reduction, and microphone array signal processing for hands-free speech communications.

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

From 1998 to 1999, he was with ATR Interpreting Telecommunications Research Laboratories, Kyoto, Japan, where he conducted research on speech synthesis, speech analysis, as well as objective measurements for evaluating speech synthesis. He then joined the Griffith University, Brisbane, Australia, where he engaged in research on robust speech recognition and signal processing. From 2000 to 2001, he worked at ATR Spoken Language Translation Research Laboratories on robust speech recognition and speech enhancement. From 2001 to 2009, he was a Member of Technical Staff at Bell Laboratories, Murray Hill, New Jersey, working on acoustic signal processing for telecommunications. He subsequently joined WeVoice Inc. in New Jersey, serving as the Chief Scientist. He is currently a Professor at the Northwestern Polytechnical University in Xi'an, China. His research interests include acoustic signal processing, adaptive signal processing, speech enhancement, adaptive noise/echo control, microphone array signal processing, signal separation, and speech communication. Dr. Chen served as an Associate Editor of the IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING from 2008 to 2014 and is currently a technical committee (TC) member of the IEEE Signal Processing Society (SPS) TC on Audio and Electroacoustics, and a member of the editorial advisory board of the *Open Signal Processing Journal*. He was the Technical Program Co-Chair of the 2009 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA) and the Technical Program Chair of IEEE TENCON 2013, and helped organize many other conferences.

Dr. Chen received the 2008 Best Paper Award from the IEEE Signal Processing Society (with Benesty, Huang, and Doclo), the best paper award from the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA) in 2011 (with Benesty), the Bell Labs Role Model Teamwork Award twice, respectively, in 2009 and 2007, the NASA Tech Brief Award twice, respectively, in 2010 and 2009, the Japan Trust International Research Grant from the Japan Key Technology Center in 1998, and the Young Author Best Paper Award from the 5th National Conference on Man–Machine Speech Communications in 1998.

Jacob Benesty was born in 1963. He received a Masters degree in microwaves from Pierre & Marie Curie University, France, in 1987, and a Ph.D. degree in control and signal processing from Orsay University, France, in April 1991. During his Ph.D. (from Nov. 1989 to Apr. 1991), he worked on adaptive filters and fast algorithms at the Centre National d'Etudes des Telecommunications (CNET), Paris, France. From January 1994 to July 1995, he worked at Telecom Paris University on multichannel adaptive filters and acoustic echo cancellation. From October 1995 to May 2003, he was first a Consultant and then a Member of the Technical Staff at Bell Laboratories, Murray Hill, NJ, USA. In May 2003, he joined the University of Quebec, INRS-EMT, in Montreal, Quebec, Canada, as a Professor. He is also a Visiting Professor at the Technion, Haifa, in Israel, an Adjunct Professor at Aalborg University, in Denmark, and a Guest Professor at Northwestern Polytechnical University, Xi'an, Shaanxi, in China.

His research interests are in signal processing, acoustic signal processing, and multimedia communications. He is the inventor of many important technologies. In particular, he was the lead researcher at Bell Labs who conceived and designed the world-first real-time hands-free full-duplex stereophonic teleconferencing system. Also, he conceived and designed the world-first PC-based multi-party hands-free full-duplex stereo conferencing system over IP networks.

He was the co-chair of the 1999 International Workshop on Acoustic Echo and Noise Control and the general co-chair of the 2009 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. He is the recipient, with Morgan and Sondhi, of the IEEE Signal Processing Society 2001 Best Paper Award. He is the recipient, with Chen, Huang, and Doclo, of the IEEE Signal Processing Society 2008 Best Paper Award. He is also the co-author of a paper for which Huang received the IEEE Signal Processing Society 2002 Young Author Best Paper Award. In 2010, he received the "Gheorghe Cartianu Award" from the Romanian Academy. In 2011, he received the Best Paper Award from the IEEE WASPAA for a paper that he co-authored with Chen.