Applied Acoustics 98 (2015) 62–69

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

Study of nonuniform linear differential microphone arrays with the minimum-norm filter

Hao Zhang [a], Jingdong Chen [a,*], Jacob Benesty [b]

[a] Center of Intelligent Acoustics and Immersive Communications, Northwestern Polytechnical University, 127 Youyi West Road, Xi'an, Shaanxi 710072, China
[b] INRS-EMT, University of Quebec, 800 de la Gauchetiere Ouest, Suite 6900, Montreal, QC H5A 1K6, Canada

ARTICLE INFO

Article history:
Received 11 November 2014
Received in revised form 19 March 2015
Accepted 30 April 2015

Keywords:
Differential microphone arrays
Minimum-norm filter
Differential beamforming
Nonuniform linear arrays
Diagonal loading

ABSTRACT

The performance of differential microphone arrays (DMAs) depends on many factors such as the number of sensors and the array geometry. This paper develops an approach that exploits nonuniform linear geometries and the minimum-norm filter to improve the robustness of DMAs against white noise. Unlike the conventional way that forms an Nth-order DMA by subtractively combining the outputs of two DMAs of order $N-1$, this approach works in the short-time Fourier transform (STFT) domain and applies a complex weight to the output of each sensor and then sum the weighted outputs to form the beamforming output in every STFT subband. The minimum-norm filter is obtained by maximizing the white noise gain of the beamformer subject to the so-called fundamental constraints. The nonuniform linear arrays are created by adjusting the interelement spacing according to some rule. We show that the use of nonuniform linear geometries can significantly improve the robustness of DMAs, particularly at low frequencies. We also show that the diagonal loading technique can help improve the robustness of DMA beamformers, though the improvement is not significant.

1. Introduction

Differential microphone arrays (DMAs) refer to the arrays that combine closely spaced sensors to respond to the spatial derivatives of the acoustic pressure field. The basic idea of such arrays can be dated back to the 1930s when the directional ribbon microphones were invented [1,2]. Since then, DMAs have been used in solving many important problems in voice communication such as noise reduction and spatial sound recording. DMAs have many attractive properties in comparison with those arrays that are responsive directly to the acoustic pressure field. First, they can form frequency-invariant beampatterns and, therefore, can be effective in processing signals for both high and low frequencies. As a result, DMAs are suitable for processing broadband signals such as speech. Second, they have the potential to attain maximum directional gains with a given number of sensors [3]. Moreover, DMAs are generally small in size and can be integrated into small communication devices such as smartphones, tablet computers, bluetooth headphones, and hearing aids. With these advantages, the design of DMAs has attracted a significant amount of interest [3–17].

However, the design and implementation of a DMA for real-world systems is not a trivial task. First of all, DMAs suffer from a great drawback of white noise amplification, particularly at low frequencies. This requires either high-quality microphone sensors with low self noise level or robust beamforming algorithms that suffers less white noise amplification. Secondly, the response of an Nth order DMA has an effect of high-pass filtering and, therefore, its frequency response has to be properly compensated to process broadband signals. Although much effort has been devoted to them in the literature [3–17], these issues have not been fully addressed.

Recently, a new approach to the design and implementation of uniform linear DMAs was developed [18,19]. Unlike the conventional way that forms an Nth-order DMA by subtractively combining the outputs of two DMAs of order $N-1$ [3–7], this method works in the short-time Fourier transform (STFT) domain as illustrated in Fig. 1. With this approach, the M noisy signals received at the microphones are partitioned into small overlapping frames of a few milliseconds. Each frame is transformed into the STFT domain. Then, a differential beamformer is designed and applied to the multichannel signals in each STFT subband, thereby producing an estimate of the clean signal in the corresponding subband. Finally, the time-domain clean speech estimate is constructed by using either the overlap-add (or overlap-save) technique with the inverse STFT. With this method, the DMA design problem is converted to one of solving a linear system formed from some fundamental constraints in beampatterns in every STFT subband. This

* Corresponding author.
E-mail address: jingdongchen@ieee.org (J. Chen).

http://dx.doi.org/10.1016/j.apacoust.2015.04.018
0003-682X/

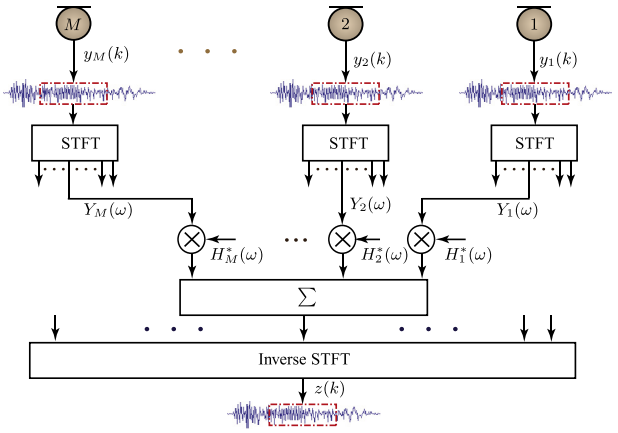

Fig. 1. A schematic diagram of a DMA system in the STFT domain.

approach does not only make the DMA beamforming more flexible as compared to the traditional one, but also makes it possible to derive a minimum-norm filter that can maximize the white noise gain (WNG) given the number of sensors.

This paper is a continuation of the work presented in [18,19]. Our main focus is on improving the white noise amplification problem in DMAs. Based on the minimum-norm filter, we investigate how to use nonuniform geometries and diagonal loading to further improve the WNG.

The remainder of this paper is organized as follows. In Section 2, we briefly describe the signal model and the problem of DMA beamforming. We then list some important measures in Section 3 that can be used to evaluate the performance of DMAs. Section 4 discusses the design of DMA patterns with the minimum-norm method. Then, in Section 5, we present a nonuniform linear array geometry that can be used with the minimum-norm filter to improve the WNG. Section 6 presents some design examples to demonstrate the property of DMAs with a nonuniform linear geometry. Section 7 briefly discusses the diagonal loading technique. Finally, some conclusions are given in Section 8.

2. Signal model and problem formulation

The basic model considered in this work is a nonuniform linear array (NULA) with M omnidirectional microphones as illustrated in Fig. 2. With the assumption that the source is in the far field, the signal received at the mth ($m = 1, 2, \ldots, M$) microphone and at time k can be written as

$$y_m(k) = x(k - \tau_{m,1}\cos\theta) + v_m(k), \tag{1}$$

where $x(k)$ is the source signal, $v_m(k)$ is the additive noise at microphone m, $\tau_{m,1} = \delta_{m,1}/c$ is the relative delay between the mth and the 1st microphones at the angle $\theta = 0^\circ$, $\delta_{m,1}$ is the spacing between the mth and the 1st microphones, θ is the source incidence angle, and $c = 340$ m/s is the speed of sound in the air. In the rest, the first microphone is considered as the reference.

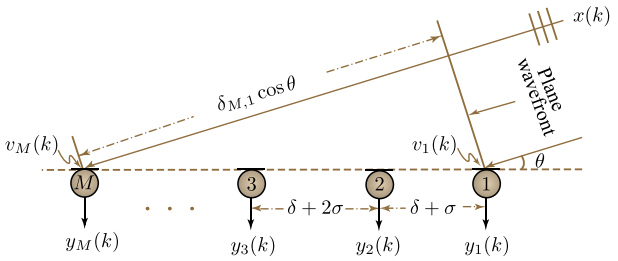

Fig. 2. Illustration of a NULA system.

In the STFT domain, (1) can be expressed as

$$Y_m(\omega) = X(\omega)e^{-\jmath\omega\tau_{m,1}\cos\theta} + V_m(\omega), \tag{2}$$

where $Y_m(\omega)$, $X(\omega)$, and $V_m(\omega)$ are the frequency-domain representations of $y_m(k)$, $x(k)$, and $v_m(k)$, respectively, $\omega = 2\pi f$ is the angular frequency, and $\jmath$ is the imaginary unit with $\jmath^2 = -1$. Putting all these signals in a vector form, we get

$$\mathbf{y}(\omega) \triangleq [\,Y_1(\omega) \quad Y_2(\omega) \quad \cdots \quad Y_M(\omega)\,]^T = \mathbf{d}(\omega, \cos\theta)X(\omega) + \mathbf{v}(\omega), \tag{3}$$

where the superscript ^{T} denotes the transpose operator, the noise signal vector, $\mathbf{v}(\omega)$, is defined similarly to $\mathbf{y}(\omega)$, and

$$\mathbf{d}(\omega, \cos\theta) \triangleq \left[\,1 \quad e^{-\jmath\omega\tau_{2,1}\cos\theta} \quad \cdots \quad e^{-\jmath\omega\tau_{M,1}\cos\theta}\,\right]^T \tag{4}$$

is the phase-delay vector of length M (which is the same as the steering vector used in the literature of traditional beamforming).

In order to recover the desired signal, $X(\omega)$, from $\mathbf{y}(\omega)$, a complex weight $H_m^*(\omega)$ is designed and applied to the output of each microphone, where the superscript * denotes complex conjugation. All the weighted outputs are then summed together to produce an estimate of the clean signal as illustrated in Fig. 1. Mathematically, the beamformer's output is

$$\begin{aligned} Z(\omega) &= \sum_{m=1}^{M} H_m^*(\omega)Y_m(\omega) = \mathbf{h}^H(\omega)\mathbf{y}(\omega) \\ &= \mathbf{h}^H(\omega)\mathbf{d}(\omega, \cos\theta)X(\omega) + \mathbf{h}^H(\omega)\mathbf{v}(\omega), \end{aligned} \tag{5}$$

where $Z(\omega)$ is supposed to be an estimate of $X(\omega)$, the superscript ^{H} is the conjugate–transpose operator, and

$$\mathbf{h}(\omega) \triangleq [\,H_1(\omega) \quad H_2(\omega) \quad \cdots \quad H_M(\omega)\,]^T. \tag{6}$$

With this signal model, the critical issue in DMA beamforming is the design of an optimal filter, $\mathbf{h}(\omega)$, which has been addressed in [18,19]. In this paper, we adopt the minimum-norm method, and explore the use of nonuniform geometries and the diagonal loading technique to further improve the WNG.

To ensure that the designed microphone array is a DMA, the following assumptions are made as in the DMA literature.

- The spacing between any two neighboring sensors is much smaller than the acoustic wavelength. This assumption is required so that the true acoustic pressure differentials can be approximated by finite differences of the microphones' outputs.
- In linear DMAs, the mainlobe of the beampattern is at the angle $\theta = 0^\circ$ (endfire direction) and the desired signal propagates from the same angle.

3. Performance measures

Usually, three important performance measures are used for the evaluation of beamformers. They are the beampattern, the WNG, and the directivity index (DI).

3.1. Beampatterns

The beampattern (or directivity pattern) describes the sensitivity of the beamformer to a plane wave (source signal) impinging on the array from the direction θ. Mathematically, it is written as

$$\mathcal{B}[\mathbf{h}(\omega), \cos\theta] \triangleq \mathbf{d}^H(\omega, \cos\theta)\mathbf{h}(\omega) = \sum_{m=1}^{M} H_m(\omega)e^{\jmath\omega\tau_{m,1}\cos\theta}. \tag{7}$$

It is well known that the frequency-independent beampattern of an Nth-order DMA is [4]

$$\mathcal{B}_N(\theta) = \sum_{n=0}^{N} a_{N,n} \cos^n \theta, \tag{8}$$

where $a_{N,n},\ n = 0, 1, \ldots, N$ are real coefficients. The different values of these coefficients determine the different directivity patterns of the Nth-order DMA. Therefore, the design of an Nth-order DMA boils down to design a beamformer that has a frequency-invariant beampattern similar to (8). This design can be achieved by solving a linear system in the STFT domain with some fundamental constraints as long as the interelement spacing is small.

In the direction of the desired signal, i.e., for $\theta = 0^\circ$, the directivity pattern must be equal to 1, i.e., $\mathcal{B}_N(0^\circ) = 1$. Therefore, we should have

$$\sum_{n=0}^{N} a_{N,n} = 1. \tag{9}$$

As a result, we always choose the first coefficient as

$$a_{N,0} = 1 - \sum_{n=1}^{N} a_{N,n}. \tag{10}$$

The coefficients of some well-known DMA beampatterns are listed in Table 1 and these patterns are plotted in Fig. 3. We can observe that all interesting patterns have at least one null in some direction.

3.2. White noise gain

We compare the input and output signal-to-noise ratios (SNRs) to check whether the beamformer's output is less noisy than its input. According to the signal model given in (2), the input SNR is defined as

$$\text{iSNR}(\omega) \triangleq \frac{\phi_X(\omega)}{\phi_{V_1}(\omega)}, \tag{11}$$

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

The output SNR, according to (5), is defined as

$$\begin{aligned}
\text{oSNR}[\mathbf{h}(\omega)] &= \phi_X(\omega) \frac{|\mathbf{h}^H(\omega)\mathbf{d}(\omega, \cos 0^\circ)|^2}{\mathbf{h}^H(\omega)\mathbf{\Phi}_{\mathbf{v}}(\omega)\mathbf{h}(\omega)} \\
&= \frac{\phi_X(\omega)}{\phi_{V_1}(\omega)} \times \frac{|\mathbf{h}^H(\omega)\mathbf{d}(\omega, \cos 0^\circ)|^2}{\mathbf{h}^H(\omega)\mathbf{\Gamma}_{\mathbf{v}}(\omega)\mathbf{h}(\omega)} \\
&= \text{iSNR}(\omega) \times \frac{|\mathbf{h}^H(\omega)\mathbf{d}(\omega, \cos 0^\circ)|^2}{\mathbf{h}^H(\omega)\mathbf{\Gamma}_{\mathbf{v}}(\omega)\mathbf{h}(\omega)},
\end{aligned} \tag{12}$$

Table 1
Coefficients and DI of some well-known directivity patterns.

Pattern	N	Coefficients $a_{N,n}$	$\mathcal{D}$
Dipole	1	$[0, 1]$	3
	2	$[0, 0, 1]$	4.3
	3	$[0, 0, 0, 1]$	5.1
Cardioid	1	$[\frac{1}{2}, \frac{1}{2}]$	4.3
	2	$[0, \frac{1}{2}, \frac{1}{2}]$	6.6
	3	$[0, 0, \frac{1}{2}, \frac{1}{2}]$	7.6
Hypercardioid	1	$[\frac{1}{3}, \frac{2}{3}]$	4.8
	2	$[-\frac{1}{5}, \frac{2}{5}, \frac{4}{5}]$	7.0
	3	$[-\frac{1}{7}, -\frac{4}{7}, \frac{4}{7}, \frac{8}{7}]$	8.4
Supercardioid	1	$[0.414, 0.586]$	4.6
	2	$[0.103, 0.484, 0.413]$	6.3
	3	$[0.022, 0.217, 0.475, 0.286]$	7.2

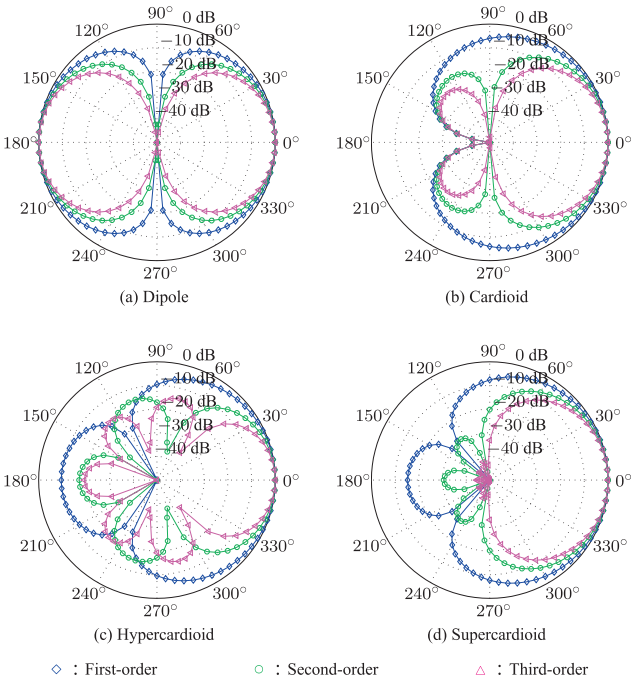

Fig. 3. Some well-known directivity patterns: first-, second-, and third-orders.

where

$$\mathbf{\Phi}_{\mathbf{v}}(\omega) \triangleq E\left[\mathbf{v}(\omega)\mathbf{v}^H(\omega)\right] \tag{13}$$

and

$$\mathbf{\Gamma}_{\mathbf{v}}(\omega) \triangleq \frac{\mathbf{\Phi}_{\mathbf{v}}(\omega)}{\phi_{V_1}(\omega)} \tag{14}$$

are the correlation and pseudo-coherence matrices of $\mathbf{v}(\omega)$, respectively.

The definition of the gain in SNR is then

$$\mathcal{G}[\mathbf{h}(\omega)] \triangleq \frac{\text{oSNR}[\mathbf{h}(\omega)]}{\text{iSNR}[\mathbf{h}(\omega)]} = \frac{|\mathbf{h}^H(\omega)\mathbf{d}(\omega, \cos 0^\circ)|^2}{\mathbf{h}^H(\omega)\mathbf{\Gamma}_{\mathbf{v}}(\omega)\mathbf{h}(\omega)}. \tag{15}$$

Now, let us assume that the matrix $\mathbf{\Gamma}_{\mathbf{v}}(\omega)$ is nonsingular. In this case, using the Cauchy–Schwarz inequality [20], we have

$$\begin{aligned}
\left|\mathbf{h}^H(\omega)\mathbf{d}(\omega, \cos 0^\circ)\right|^2 &= \left|\mathbf{h}^H(\omega)\mathbf{\Gamma}_{\mathbf{v}}^{\frac{1}{2}}(\omega)\mathbf{\Gamma}_{\mathbf{v}}^{-\frac{1}{2}}(\omega)\mathbf{d}(\omega, \cos 0^\circ)\right|^2 \\
&\leqslant \left\{\mathbf{h}^H(\omega)\mathbf{\Gamma}_{\mathbf{v}}^{\frac{1}{2}}(\omega)\left[\mathbf{\Gamma}_{\mathbf{v}}^{\frac{1}{2}}(\omega)\right]^H\mathbf{h}(\omega)\right\} \\
&\quad \times \left\{\mathbf{d}^H(\omega, \cos 0^\circ)\left[\mathbf{\Gamma}_{\mathbf{v}}^{-\frac{1}{2}}(\omega)\right]^H\mathbf{\Gamma}_{\mathbf{v}}^{-\frac{1}{2}}(\omega)\mathbf{d}(\omega, \cos 0^\circ)\right\} \\
&= \left[\mathbf{h}^H(\omega)\mathbf{\Gamma}_{\mathbf{v}}(\omega)\mathbf{h}(\omega)\right] \\
&\quad \times \left[\mathbf{d}^H(\omega, \cos 0^\circ)\mathbf{\Gamma}_{\mathbf{v}}^{-1}(\omega)\mathbf{d}(\omega, \cos 0^\circ)\right],
\end{aligned} \tag{16}$$

with equally if and only if $\mathbf{h}(\omega) \propto \mathbf{\Gamma}_{\mathbf{v}}^{-1}(\omega)\mathbf{d}(\omega, \cos 0^\circ)$. Using the inequality (16) in (15), we deduce an upper bound for the gain:

$$\begin{aligned}
\mathcal{G}[\mathbf{h}(\omega)] &\leqslant \mathbf{d}^H(\omega, \cos 0^\circ)\mathbf{\Gamma}_{\mathbf{v}}^{-1}(\omega)\mathbf{d}(\omega, \cos 0^\circ) \\
&\leqslant \text{tr}\left[\mathbf{\Gamma}_{\mathbf{v}}^{-1}(\omega)\right]\text{tr}\left[\mathbf{d}(\omega, \cos 0^\circ)\mathbf{d}^H(\omega, \cos 0^\circ)\right] \\
&\leqslant M\text{tr}\left[\mathbf{\Gamma}_{\mathbf{v}}^{-1}(\omega)\right],
\end{aligned} \tag{17}$$

where $\text{tr}[\cdot]$ is the trace of a square matrix. We observe how the gain is upper bounded [as long as $\mathbf{\Gamma}_{\mathbf{v}}(\omega)$ is nonsingular] and depends on the number of microphones as well as on the nature of the noise.

If the noise is temporally and spatially white with the same variance at all microphones, we have $\mathbf{\Gamma}_{\mathbf{v}}(\omega)=\mathbf{I}_M$, where $\mathbf{I}_M$ is the $M\times M$ identity matrix. In this case, the gain in SNR is called WNG, which is

$$\mathcal{G}_{\text{wn}}[\mathbf{h}(\omega)]=\frac{|\mathbf{h}^H(\omega)\mathbf{d}(\omega,\cos 0^\circ)|^2}{\mathbf{h}^H(\omega)\mathbf{h}(\omega)}=\frac{1}{\mathbf{h}^H(\omega)\mathbf{h}(\omega)}, \tag{18}$$

where we have used the premise that $\mathcal{B}_N(0^\circ)=1$, i.e., the distortionless constraint $\mathbf{h}^H(\omega)\mathbf{d}(\omega,\cos 0^\circ)=1$. The WNG generally tells us how robust a DMA beamformer is with respect to sensor self noise.

3.3. Directivity index

In diffuse (spherically isotropic) noise, we have

$$[\mathbf{\Gamma}_{\mathbf{v}}(\omega)]_{ij}=[\mathbf{\Gamma}_{\text{dn}}(\omega)]_{ij}=\frac{\sin(\omega\tau_{i,j})}{\omega\tau_{i,j}}=\text{sinc}(\omega\tau_{i,j}), \tag{19}$$

where $\tau_{i,j}=\delta_{i,j}/c$. Substituting (19) into (15), we obtain the SNR gain in diffuse noise, which is also called the directivity factor:

$$\mathcal{G}_{\text{dn}}[\mathbf{h}(\omega)]=\frac{|\mathbf{h}^H(\omega)\mathbf{d}(\omega,\cos 0^\circ)|^2}{\mathbf{h}^H(\omega)\mathbf{\Gamma}_{\text{dn}}(\omega)\mathbf{h}(\omega)}. \tag{20}$$

With the premise that $\mathcal{B}_N(0^\circ)=1$, we have

$$\mathcal{G}_{\text{dn}}[\mathbf{h}(\omega)]=\frac{1}{\mathbf{h}^H(\omega)\mathbf{\Gamma}_{\text{dn}}(\omega)\mathbf{h}(\omega)}. \tag{21}$$

The DI is simply defined as [3,4]

$$\mathcal{D}[\mathbf{h}(\omega)]=10\log_{10}\mathcal{G}_{\text{dn}}[\mathbf{h}(\omega)]. \tag{22}$$

The DI values of some well-known directivity patterns are listed in Table 1. With the same pattern, the DI value increases with the order of the DMA.

4. The minimum-norm filter

Suppose that we have M microphones with $M\geqslant N+1$ and we want to design an Nth-order DMA. The objective is then to find a filter so that its corresponding pattern as given in (7) approaches the ideal Nth-order DMA pattern in (8). It is shown in [18] that the DMA filter with a uniform linear array (ULA) can be obtained by solving the following linear systems of $N+1$ equations:

$$\mathbf{D}(\omega,\boldsymbol{\alpha})\mathbf{h}(\omega)=\boldsymbol{\beta}, \tag{23}$$

where

$$\mathbf{D}(\omega,\boldsymbol{\alpha})=\begin{bmatrix}\mathbf{d}^H(\omega,1)\\ \mathbf{d}^H(\omega,\alpha_{N,1})\\ \vdots\\ \mathbf{d}^H(\omega,\alpha_{N,N})\end{bmatrix} \tag{24}$$

is the constraint matrix of size $(N+1)\times M$, $\mathbf{d}(\omega,\alpha_{N,n})$ is the phase-delay vector of length M as defined in (4), $\mathbf{h}(\omega)$ is a filter of length M defined in (6), and

$$\boldsymbol{\alpha}=\begin{bmatrix}1 & \alpha_{N,1} & \cdots & \alpha_{N,N}\end{bmatrix}^T, \tag{25}$$

$$\boldsymbol{\beta}=\begin{bmatrix}1 & \beta_{N,1} & \cdots & \beta_{N,N}\end{bmatrix}^T, \tag{26}$$

are vectors of length $N+1$ containing the design coefficients of the desired pattern. The value of α equals $\cos\theta$, where θ is the desired null direction, and β is the corresponding value of that direction in the ideal beampattern.

The choice of the coefficients $\alpha_{N,n}$ and $\beta_{N,n}$, $n=1,2,\ldots,N$ is critical for the proper design of a desired DMA beamformer. The rules of thumb are as follows.

- The N coefficients $\alpha_{N,n}$ should be chosen in such a way that $\mathbf{D}(\omega,\boldsymbol{\alpha})$ is full row rank, i.e., its rank is equal to $N+1$.
- The N pairs of coefficients $(\alpha_{N,n},\beta_{N,n})$ should take values from a desired ideal DMA pattern. In general, they should correspond to the nulls of the desired DMA pattern; but they can take other values as well.

If they satisfy the above two conditions, the constraints are called fundamental constraints.

Now, if $M=N+1$, $\mathbf{D}(\omega,\boldsymbol{\alpha})$ is a square matrix. The beamforming filter is then obtained as

$$\mathbf{h}(\omega)=\mathbf{D}^{-1}(\omega,\boldsymbol{\alpha})\boldsymbol{\beta}, \tag{27}$$

If $M>N+1$, the beamforming filter can be determined by the minimum-norm approach, i.e.,

$$\mathbf{h}_{\text{MN}}(\omega)=\mathbf{D}^H(\omega,\boldsymbol{\alpha})\left[\mathbf{D}(\omega,\boldsymbol{\alpha})\mathbf{D}^H(\omega,\boldsymbol{\alpha})\right]^{-1}\boldsymbol{\beta}. \tag{28}$$

It has been shown that this minimum-norm filter maximizes the WNG. As a matter of fact, considering both (18) and (23), one can check that the beamforming filter that maximizes the WNG is the solution of the following optimization problem:

$$\begin{aligned}\mathbf{h}_{\text{MaxWNG}}(\omega)=\arg\max_{\mathbf{h}(\omega)}&\frac{1}{\mathbf{h}^H(\omega)\mathbf{h}(\omega)}\\ \text{subject to}\quad&\mathbf{D}(\omega,\boldsymbol{\alpha})\mathbf{h}(\omega)=\boldsymbol{\beta},\end{aligned} \tag{29}$$

which is equivalent to

$$\begin{aligned}\mathbf{h}_{\text{MaxWNG}}(\omega)=\arg\min_{\mathbf{h}(\omega)}&\ \mathbf{h}^H(\omega)\mathbf{h}(\omega)\\ \text{subject to}\quad&\mathbf{D}(\omega,\boldsymbol{\alpha})\mathbf{h}(\omega)=\boldsymbol{\beta}.\end{aligned} \tag{30}$$

It can be checked that the solution of (29) or (30) is (28), i.e., $\mathbf{h}_{\text{MN}}(\omega)=\mathbf{h}_{\text{MaxWNG}}(\omega)$. So, the minimum-norm filter maximizes the WNG.

It can be observed from (17) that the WNG should approach M with a large number of microphones. If the value of M is much larger than $N+1$, the order of the DMA may not be equal to N anymore but the Nth order DMA fundamental constraints will always be fulfilled. As a result, the resulting shape of the directivity pattern may be slightly different from the one obtained with $M=N+1$. This approach is the best we can do to solve the conflicting requirement of a high-order DMA that does not amplify much the white noise. Meanwhile, it has to pay some cost, which includes (1) the increased number of sensors and A/D channels, and (2) the beampattern may vary slightly from the specified pattern particularly at high frequencies. However, slight dependency of the beampattern on frequency can be avoided by posting more constraints. Since it maximizes the WNG, this method is more robust than its traditional counterpart. With this approach, we can easily design DMAs of different orders.

5. Nonuniform linear arrays

ULAs are most often studied and used in the design of DMAs while nonuniform linear ones have not been investigated for DMAs. In this section, we investigate nonuniform linear DMAs. It is shown in the next section that the nonuniform geometry, if properly designed, can significantly improve the WNG, particularly at low frequencies.

In this paper, we focus only on a particular nonuniform linear geometry based on the ULA by adding an arithmetic sequence of distance with a common difference σ to the basic spacing δ. In this case, we have

$$\delta_{i,i-1}=\delta+(i-1)\sigma,\quad i=2,3,\ldots,M \tag{31}$$

and

$$\tau_{i,i-1} = \tau_0 + (i-1)\tau, \quad i = 2,3,\ldots,M, \tag{32}$$

where $\tau_0 = \delta/c$ and $\tau = \sigma/c$. The time delay between the mth microphone and the reference one becomes

$$\tau_{m,1} = (m-1)\tau_0 + \frac{(m-1)m}{2}\tau, \quad i = 2,3,\ldots,M. \tag{33}$$

When σ is 0, the array degenerates to a ULA.

Note that the interelement spacing should be limited according to the assumption made in Section 2. For a DMA with a given order and a given number of microphones, if we assume that the DMA works in the frequency band ranging from 500 Hz to 8 kHz $= f_{\max}$, the maximum value of $\delta_{i,i-1}$ should be smaller than $c/f_{\max}$ (i.e., 4.25 cm).

6. Design examples and performance of nonuniform linear DMAs

Any pattern can be designed with the minimum-norm filter. In this section, we show how to design three types of patterns with ULAs and NULAs: the first-order cardioid, the second-order cardioid, and a third-order pattern.

6.1. First-order cardioid

For the first-order cardioid, there is a one at 0° and a null at 180°. So, the fundamental constraints with this patten can be written as

$$\boldsymbol{\alpha} = [1 \quad -1]^T, \tag{34a}$$

$$\boldsymbol{\beta} = [1 \quad 0]^T. \tag{34b}$$

Let us set $\delta = 1$ cm and $\sigma = 0.1$ cm to form the nonuniform linear DMA. With a given number of microphones and substituting the fundamental constraints (34) into (28), we easily obtain the minimum-norm filter.

We compare the performance of the minimum-norm filter ($M > N+1$) with that of the traditional method with $M = N+1$ in Fig. 4, where the case with $M = 2$ corresponds to the traditional DMA ($M = N+1$) and the $M = 6$ case shows the results of the minimum-norm filter DMA.

Fig. 4(a) and (b) plot the patterns of the traditional method and the minimum-norm filter for three different frequencies, i.e., 1 kHz, 3 kHz, and 5 kHz. It is seen that the designed patterns for $M = 2$ are almost the same as the desired one. With 6 microphones, the designed pattern at 3 kHz is more directional than the ideal one, while at 5 kHz, it tends to become the second-order cardioid.

Fig. 4(c) plots the WNG of the designed pattern as a function of frequency for both the ULA and NULA for $M = 2$ and $M = 6$, where "u" and "n" stand for uniform and nonuniform, respectively. It can be seen that the WNG of the minimum-norm filter is higher than that of the traditional method for both arrays, which validates that the minimum-norm filter is more robust than its traditional counterpart. When $M = 2$ and with the ULA, the WNG is less than 0 dB for $f < 4.25$ kHz, indicating that the white noise is amplified for frequencies less than 4.25 kHz. But with the NULA, white noise amplification only happens for $f < 3.91$ Hz. When the number of microphones increases to 6, white noise amplification happens for frequencies below 0.67 kHz with the ULA while that happens for frequencies below 0.52 kHz with the NULA.

To further visualize the influence of σ on the performance of the DMAs, we plot in Fig. 5 the WNG of the designed pattern with $\delta = 1$ cm and $M = 3$ as a function of f for various values of σ. It is again clearly seen that the nonuniform geometry can help deal with the white noise amplification problem in DMA beamforming.

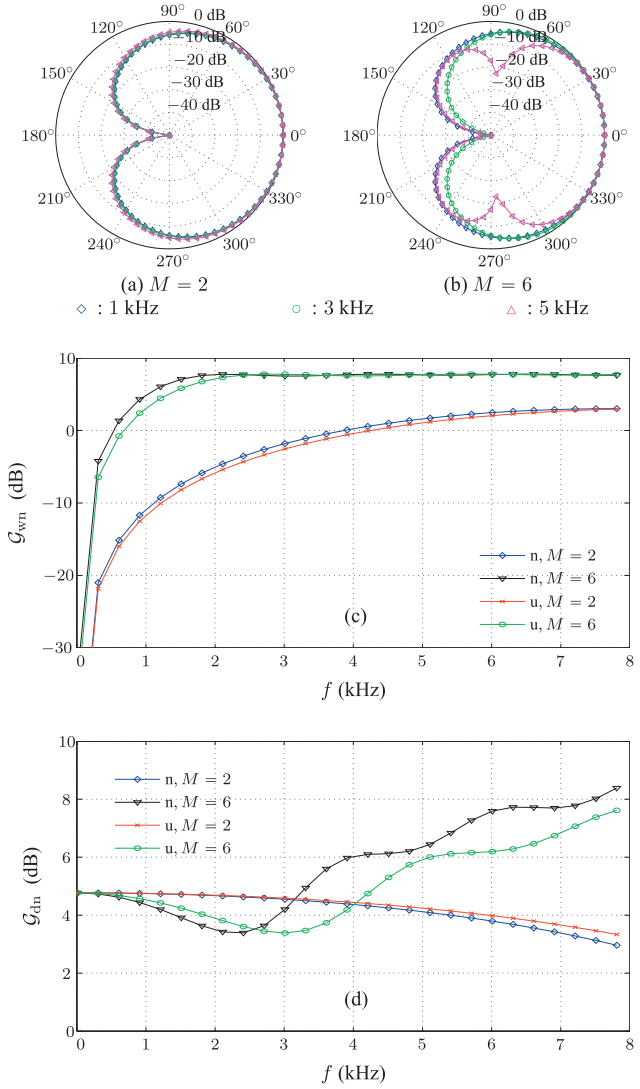

Fig. 4. Patterns, WNG, and DI of a first-order DMA (cardioid) with a NULA: (a), (b) patterns (◇: 1 kHz, ○: 3 kHz, △: 5 kHz), (c) WNG, and (d) DI. $\delta = 1$ cm and $\sigma = 0.1$ cm.

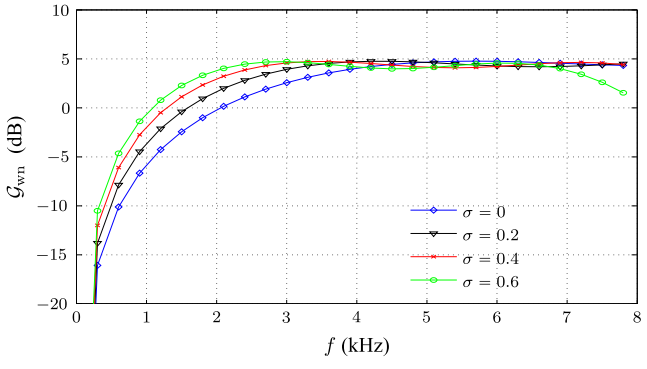

Fig. 5. WNG of a first-order DMA (cardioid) with a NULA as a function of f for various values of σ. $\delta = 1$ cm and $M = 3$.

Table 2 lists the frequencies above which the WNG of the designed pattern is greater than 0 dB for different values of M.

Fig. 4(d) plots the DI of the designed pattern as a function of frequency for both the ULA and NULA for $M = 3$ and $M = 6$. It is seen that the DI of the NULA with 6 microphones is slightly lower than

Table 2
Frequencies below which the WNG of a first-order DMA (cardioid) with a ULA is less than 0.

f(kHz)	M						
	2	3	4	5	6	7	8
$N=1$	4.25	2.06	1.28	0.89	0.67	0.53	0.43

that of the ULA in frequencies below 2.59 kHz, but slightly higher at high frequencies. In the 3-microphone case, the DI for the ULA and NULA does not differ much.

Fig. 6 plots the DI of the designed pattern with $\delta = 1$ cm and $M = 3$ as a function of f for various values of σ. It is observed that the DI does not change much with σ below 5 kHz, but it may vary slightly for frequencies larger than 5 kHz in the studied case.

6.2. Second-order cardioid

The second-order cardioid pattern has a one at 0° and two nulls at 90° and 180°. The fundamental constraints with this patten can be written as

$$\boldsymbol{\alpha} = [1 \quad 0 \quad -1]^T, \tag{35a}$$

$$\boldsymbol{\beta} = [1 \quad 0 \quad 0]^T. \tag{35b}$$

Again, we choose $\delta = 1$ cm and $\sigma = 0.1$ cm to form the NULA. Given the number of microphones and substituting the fundamental constraints (35) into (28), we easily obtain the minimum-norm filter.

Fig. 7(a) and (b) plot the patterns of the minimum-norm filter with $M = 4$ and $M = 7$, respectively, for three different frequencies, i.e., 1 kHz, 3 kHz, and 5 kHz. It is seen that all the patterns are close to the desired one.

It can be seen from Fig. 7(c) that the WNG increases with M for both arrays. This validates the property of the minimum-norm method, which takes advantage of multiple microphones to improve the WNG. Furthermore, the nonuniform geometry helps improve the WNG. As for the DI, there is not much difference between the two arrays with 4 microphones; but for 7 microphones, the DI is smaller below 4.39 kHz while larger above 4.39 kHz with the NULA.

6.3. A third-order pattern

Now, let us consider a third-order pattern that has a one at 0° and three distinct nulls at 90°, 120°, and 180°. We have

$$\boldsymbol{\alpha} = [1 \quad 0 \quad -0.5 \quad -1]^T, \tag{36a}$$

$$\boldsymbol{\beta} = [1 \quad 0 \quad 0 \quad 0]^T. \tag{36b}$$

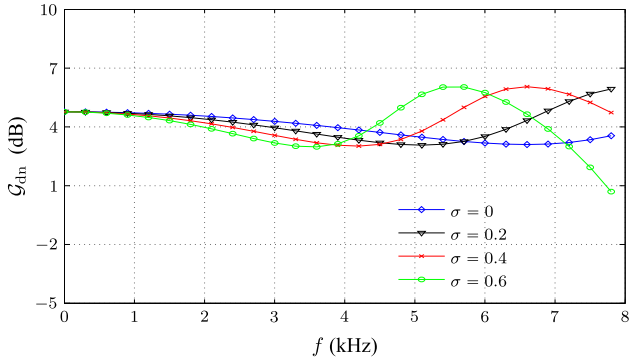

Fig. 6. DI of a first-order DMA (cardioid) with a NULA as a function of f for various values of σ. $\delta = 1$ cm and $M = 3$.

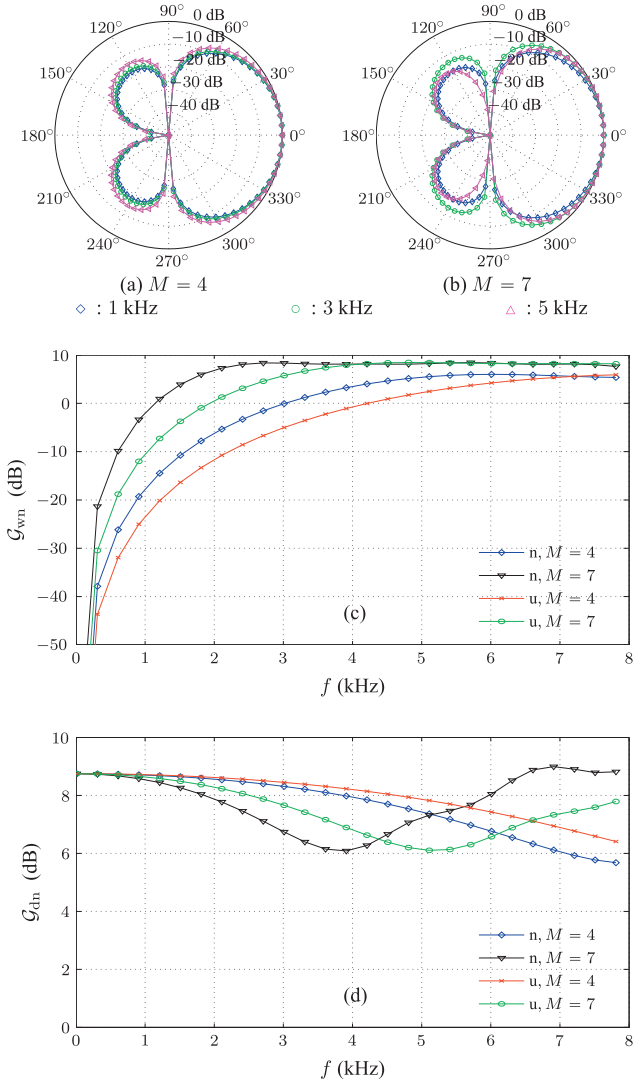

Fig. 7. Patterns, WNG, and DI of a second-order DMA (cardioid) with a NULA: (a), (b) patterns (◇: 1 kHz, ○: 3 kHz, △: 5 kHz), (c) WNG, and (d) DI. $\delta = 1$ cm and $\sigma = 0.1$ cm.

Similar to the previous simulations, we set $\delta = 1$ cm and $\sigma = 0.1$ cm to form the NULA. Substituting the fundamental constraints (36) into (28), we obtain the minimum-norm filter.

Fig. 8 plots the patterns, WNG, and DI of the minimum-norm filter for $M = 5$ and $M = 8$. It is observed that all the designed patterns are close to the ideal one. The WNG of this third-order DMA increases with M for both the ULA and NULA. Also, the nonuniform geometry helps improve the WNG.

To clearly see the improvement in WNG of a NULA as compared to a ULA, we define $\Delta\mathcal{G}_{\rm wn} = \mathcal{G}_{\rm NULAwn} - \mathcal{G}_{\rm ULAwn}$ and plot it in Fig. 9 as a function of f for the third-order DMA (with different values of M) and different order DMAs (with $M = 4$). It is seen that non-uniform microphone distributions can improve the WNG of a DMA, especially at low frequencies. The value of $\Delta\mathcal{G}_{\rm wn}$ is above 0 for frequencies from 0 Hz to 8 kHz for most of the cases plotted in the figure, in other words, the WNG of a NULA is higher than that of a ULA at these frequencies. Furthermore, the improvement in WNG increases with M for DMAs of the same order at low frequencies. However, the differences in WNG between NULAs and ULAs are small at high frequencies. It can also be seen that once M is given, the higher the order of the DMA, the larger the improvement in WNG. These results validate the fact that non-uniform DMAs can reduce the white noise amplification at low frequencies.

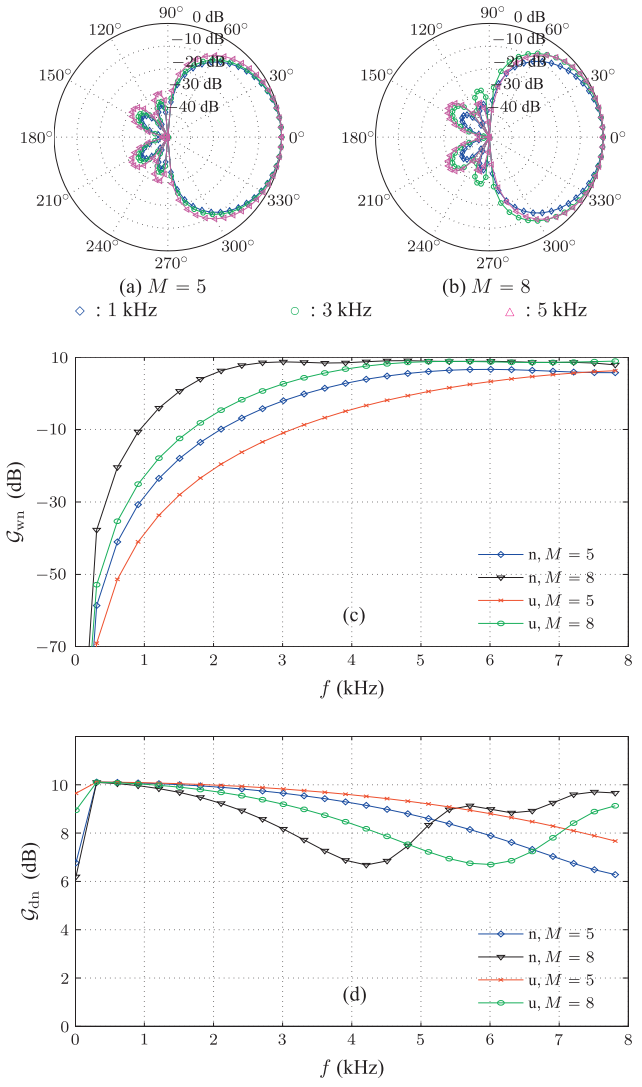

Fig. 8. Patterns, WNG, and DI of a third-order DMA with a NULA: (a), (b) patterns (◇: 1 kHz, ○: 3 kHz, △: 5 kHz), (c) WNG, and (d) DI. $\delta = 1$ cm and $\sigma = 0.1$ cm.

7. Diagonal loading technique

The diagonal loading method is a simple yet effective approach to improve the robustness of many beamformers. In this section, we briefly discuss the possibility of using this technique to improve the WNG of DMA beamformers. Unlike most beamformers that add a scaled identity matrix to the correlation matrix, we consider here to add diagonal loading to the product matrix $\mathbf{D}(\omega,\boldsymbol{\alpha})\mathbf{D}^H(\omega,\boldsymbol{\alpha})$ before computing its inverse, which can reduce the norm of the filter and consequently increase the WNG at low frequencies.

The modified minimum-norm filter can then be written as

$$\mathbf{h}_\mu(\omega) = \mathbf{D}^H(\omega,\boldsymbol{\alpha})[\mathbf{D}(\omega,\boldsymbol{\alpha})\mathbf{D}^H(\omega,\boldsymbol{\alpha}) + \mu\mathbf{I}]^{-1}\boldsymbol{\beta}, \tag{37}$$

where $\mathbf{I}$ is the identity matrix of size $(N+1)\times(N+1)$ and μ is a positive parameter controlling the diagonal loading level.

The matrix $\boldsymbol{\Upsilon}(\omega) = \mathbf{D}(\omega,\boldsymbol{\alpha})\mathbf{D}^H(\omega,\boldsymbol{\alpha})$ can be decomposed as

$$\boldsymbol{\Upsilon}(\omega) = \mathbf{D}(\omega,\boldsymbol{\alpha})\mathbf{D}^H(\omega,\boldsymbol{\alpha}) = \mathbf{Q}(\omega)\boldsymbol{\Lambda}(\omega)\mathbf{Q}^H(\omega), \tag{38}$$

where $\boldsymbol{\Lambda}(\omega) = \mathrm{diag}[\lambda_1(\omega),\lambda_2(\omega),\ldots,\lambda_{N+1}(\omega)]$ is a diagonal matrix consisting of all the eigenvalues of $\boldsymbol{\Upsilon}(\omega)$ and $\mathbf{Q}(\omega) = [\mathbf{q}_1(\omega)\mathbf{q}_2(\omega)\cdots\mathbf{q}_{N+1}(\omega)]$ is the eigenvector matrix. It follows immediately that

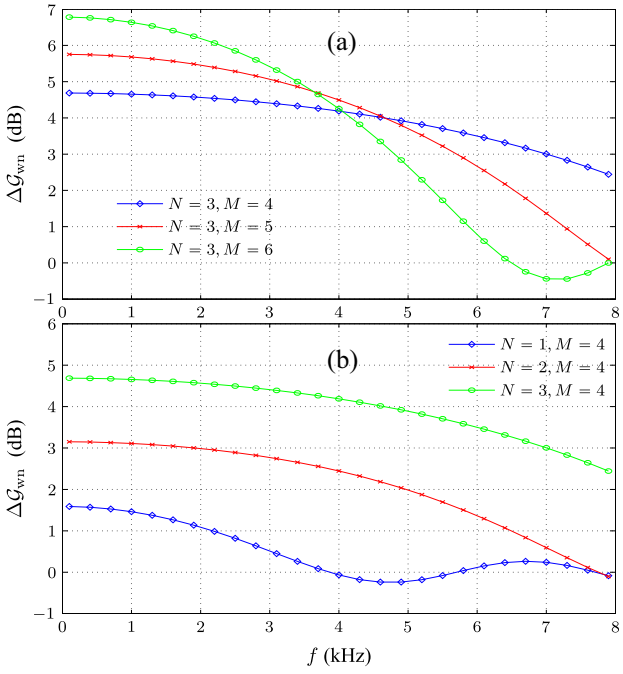

Fig. 9. The improvement in WNG of a NULA as compared to a ULA with $\delta = 1$ cm, $\sigma = 0.1$ cm: (a) third-order DMA with different values of M and (b) first-, second-, and third-order DMA with $M = 4$.

$$\mathbf{h}_\mu^H(\omega)\mathbf{h}_\mu(\omega) = \mathbf{b}^H(\omega)[\boldsymbol{\Lambda}(\omega)+\mu\mathbf{I}]^{-1}\boldsymbol{\Lambda}(\omega)[\boldsymbol{\Lambda}(\omega)+\mu\mathbf{I}]^{-1}\mathbf{b}(\omega) = \sum_{i=1}^{N+1}\frac{\lambda_i(\omega)}{[\lambda_i(\omega)+\mu]^2}|b_i(\omega)|^2, \tag{39}$$

where $\mathbf{b}(\omega) = \mathbf{Q}^H(\omega)\boldsymbol{\beta}$ and $b_i(\omega)$, $i = 1,2,\ldots,N+1$, is the ith element of $\mathbf{b}(\omega)$. It is clearly seen from (39) that the norm of the filter $\mathbf{h}_\mu$ decreases as the parameter μ increases.

The critical issue with the diagonal loading is to determine the value of μ. This parameter should satisfy $0 < \mu \ll M$ and also

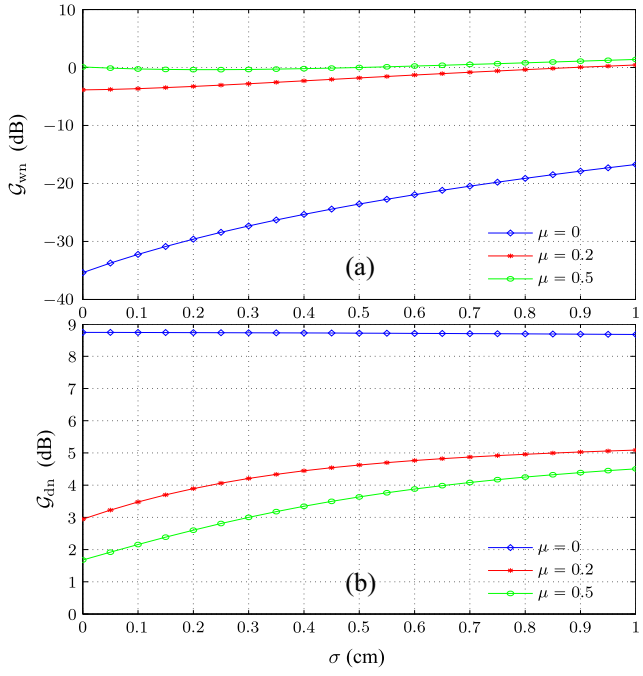

Fig. 10. WNG and DI for a second-order cardioid with $M = 4$ with and without diagonal loading at $f = 500$ Hz: (a) WNG and (b) DI.

should be much smaller than the average of the diagonal elements of the product matrix $\mathbf{D}(\omega,\boldsymbol{\alpha})\mathbf{D}^H(\omega,\boldsymbol{\alpha})$. While the upper bound is given, the proper value of the loading parameter has to be determined through simulations.

To examine whether the diagonal loading can help improve the WNG, we designed a second-order cardioid pattern using a NULA with 4 microphones. Fig. 10 plots the WNG and DI of the designed pattern as a function of σ for $f = 500$ Hz and $\mu = 0,\ 0.2$, and 0.5. It is seen that diagonal loading can help improve the WNG, and the larger the loading parameter, the larger the WNG. However, diagonal loading also causes significant degradation of the DI. In other words, diagonal loading improves WNG by paying a price of sacrificing the DI of the array, leading to less suppression of directional noise.

8. Conclusions

Differential microphone arrays (DMAs) have the great potential to be used in devices such as smartphones, tablets, hearing aids, and smart televisions for reducing noise and interference, thereby enhancing the voice signal of interest. However, the white noise amplification problem, which is inherent to DMA beamforming, particularly at low frequencies, has to be addressed before the potentiality of DMAs can be fully exploited. This paper was devoted to the white noise amplification problem in DMAs. Three approaches were investigated: the minimum-norm filter, nonuniform linear arrays, and the diagonal loading. They can be used either separately or in a combined manner. Several design examples were provided. The results showed that each method can help improve the white noise gain (WNG) if used separately, and greater WNG was achieved by combining those methods. It should be noted, however, that diagonal loading improves the WNG at the cost of worsening the directivity index (DI) of the array.

Acknowledgment

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

References

[1] Olson HF. A uni-directional ribbon microphone. J Acoust Soc Am 1932;3:315.
[2] Olson HF. Gradient microphones. J Acoust Soc Am 1946;17:192–8.
[3] Elko GW, Meyer J. Microphone arrays. In: Benesty J, Sondhi MM, Huang Y, editors. Springer handbook of speech processing. Berlin, Germany: Springer-Verlag; 2008. p. 1021–41 [chapter 50], Part I.
[4] Elko GW. Superdirectional microphone arrays. In: Gay SL, Benesty J, editors. Acoustic signal processing for telecommunication. Boston (MA): Kluwer Academic Publishers; 2000. p. 181–237 [chapter 10].
[5] Sessler GM, West JE. Directional transducers. IEEE Trans Audio Electroacoustic 1971;19:19–23.
[6] Olson HF. Gradient microphones. J Acoust Soc Am 1946;17:192–8.
[7] Elko GW, West JE, Thompson S. Differential and gradient microphone arrays. J Acoust Soc Am 2003;114. p. 2426–2426.
[8] Buck M. Aspects of first-order differential microphone arrays in the presence of sensor imperfections. Eur Trans Telecommun 2002;13:115–22.
[9] Abhayapala TD, Gupta A. Higher order differential–integral microphone arrays. J Acoust Soc Am 2010;127:EL227–33.
[10] De Sena E, Hacihabiboğlu H, Cvetković Z. On the design and implementation of higher-order differential microphones. IEEE Trans Audio, Speech, Language Process 2012;20:162–74.
[11] Buck M, Rößler M. First order differential microphone arrays for automotive applications. In: Proc IWAENC; 2001.
[12] Merimaa J. Applications of 3-D microphone array. In: Proc AES 112th convention; 2002. p. 1–11.
[13] Elko GW, Nguyen Pong A-T. A steerable and variable first-order differential microphone array. In: Proc IEEE ICASSP; 1997. p. 223–6.
[14] Derkx RMM, Janse K. Theoretical analysis of a first-order azimuth-steerable superdirective microphone array. IEEE Trans Audio, Speech, Language Process 2011;17:150–62.
[15] Derkx R. Optimal azimuthal steering of a first-order superdirectional microphone response. In: Proc IWAENC; 2008.
[16] Miles RN, Liu Y, Su Q, Cui E. A silicon directional microphone with second-order directivity. In: Proc inter congress acoust; September 2007.
[17] Kolundžija M, Faller C, Vetterli M. Spatio-temporal gradient analysis of differential microphone arrays. In: Proc audio engineering society convention; 2009.
[18] Benesty J, Chen J. Study and design of differential microphone arrays. Berlin: Springer-Verlag; 2012.
[19] Chen J, Benesty J, Pan C. On the design and implementation of linear differential microphone arrays. J Acoust Soc Am 2014;136(6):3097–113.
[20] Horn RA, Johnson CR. Matrix analysis. Cambridge University Press; 1985.