

# A General Approach to the Design and Implementation of Linear Differential Microphone Arrays

Jingdong Chen\* and Jacob Benesty†

\*Northwestern Polytechnical University, 127 Youyi West Road, Xi'an, Shaanxi 710072, China

E-mail: jingdongchen@ieee.org

†INRS-EMT, University of Quebec, 800 de la Gauchetiere Ouest, Suite 6900, Montreal, QC H5A 1K6, Canada

E-mail: benesty@emt.inrs.ca

**Abstract**—The design of differential microphone arrays (DMAs) and the associated beamforming algorithms have become very important problems. Traditionally, an  $N$ th order DMA is formed by subtractively combining the outputs of two DMAs of order  $N - 1$ . This method, though simple and easy to implement, suffers from a number of limitations. For example, it is difficult to design the equalization filter that is needed for compensating the array's non-uniform frequency response, particularly for high-order DMAs. In this paper, we propose a new approach to the design and implementation of linear DMAs for speech enhancement. Unlike the traditional method that works in the time domain, this proposed approach works in the short-time Fourier transform (STFT) domain. The core issue with this framework is how to design the desired differential beamformer in each subband, which is accomplished by solving a linear system consisting of  $N + 1$  fundamental constraints for an  $N$ th-order DMA.

**Index Terms**—Differential microphone arrays, differential beamforming, directivity pattern.

## I. INTRODUCTION

Microphone arrays have attracted a significant amount of interest over the last few decades since they have the potential to solve many important problems such as noise reduction/speech enhancement, source separation, dereverberation, spatial sound recording, and source localization/tracking, to name a few. Many different array systems have been developed since then, which can be categorized into two basic classes based on how they respond to the sound field, i.e., additive and differential arrays.

Additive arrays achieve signal enhancement and noise suppression based on the synchronize-and-add principles; but they have now evolved to include all the arrays with large inter-element spacing (from a couple of centimeters to a couple of decimeters) and optimal beamforming in broadside directions. This kind of arrays have been proven to be useful in dealing with many problems [1], [2], [3]. However, they are also found to suffer from a number of limitations in processing broadband speech signals. First, the beampattern of an additive array is frequency dependent and the beamwidth is inversely proportional to the frequency. As a result, such an array is not

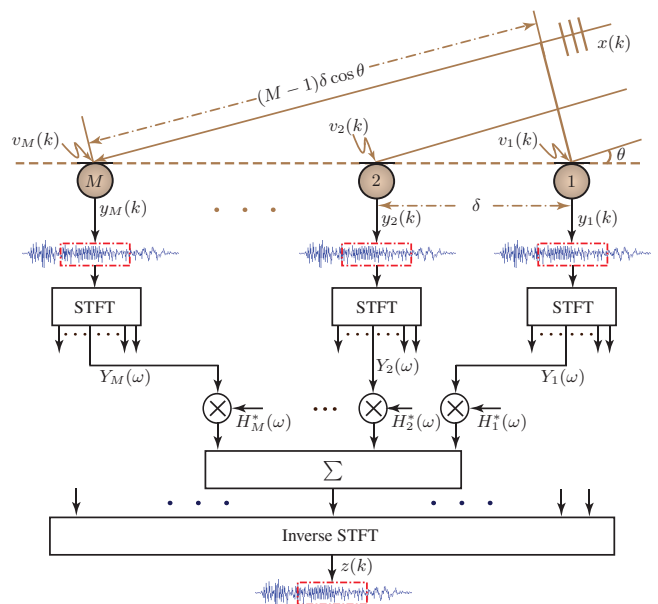


Fig. 1. A schematic diagram of a DMA system in the STFT domain.

effective in dealing with low-frequency noise and interference. Second, noise is generally attenuated in a non-uniform way over its entire spectrum, leading to some disturbing artifacts at the array's output [4]. Furthermore, if the incident angle of the desired speech source is different from the array's look direction, which happens often in practice, the speech signal will be low-pass filtered, leading to speech distortion.

Differential microphone arrays (DMAs) are designed to respond to the spatial derivatives of an acoustic pressure field. In comparison with additive arrays, DMAs can have many advantages in processing speech signals. First, a DMA can form frequency-invariant beampatterns. This makes it effective for processing both high- and low-frequency signals. Second, a DMA has the potential to attain maximum directional gain with a given number of microphone sensors [5]. Furthermore,

the gain of a DMA decreases with the distance between the sound source and the array, and therefore it can inherently suppress noise and interference originating from far-away sources. Additionally, DMAs are generally small in size, and therefore can be easily integrated into communication devices. As a consequence, the design of DMAs and the associated beamforming algorithms have attracted much interest over the past few decades.

Traditionally, an  $N$ th-order DMA is formed by using  $N + 1$  microphones and its output is generated by subtractively combining the outputs of two DMAs of order  $N - 1$  [5]–[18]. This way of DMA design, though simple and easy to implement, has many drawbacks. For example, it is difficult to design the equalization filter that is needed for compensating the array’s non-uniform frequency response, particularly for high-order DMAs. In this paper, we propose a new approach to the design and implementation of linear DMAs for speech enhancement. A schematic diagram of this new approach is illustrated in Fig. 1. The diagram starts with a desired (or clean) speech signal,  $x(k)$ , passing through some acoustic environment and then being corrupted by noise. A DMA consisting of  $M$  omnidirectional microphones is used to pick up the signal and each microphone output consists of both the desired speech signal and some unwanted noise. To estimate the source signal, the  $M$  noisy signals are partitioned into small overlapping frames. Each frame is transformed into the short-time Fourier transform (STFT) domain. In each subband, a differential beamformer is designed and applied to the multichannel STFT coefficients, thereby producing an estimate of the clean signal spectrum in this subband. Finally, the time-domain clean speech estimate is constructed using either the overlap-add or overlap-save technique with the inverse STFT. The most crucial step in this framework is the design of the differential beamformers in the STFT subbands, which is addressed in the rest of this paper.

## II. SIGNAL MODEL AND PROBLEM FORMULATION

Let us consider a uniform linear array consisting of  $M$  omnidirectional microphones as illustrated in Fig. 1, where the distance between two neighboring sensors is equal to  $\delta$ . Suppose that there is a source signal (plane wave) propagating in an acoustic environment (anechoic farfield model) at the speed of sound, i.e.,  $c = 340$  m/s, and impinging on the array with an incident angle of  $\theta$ . In this scenario, the  $m$ th ( $m = 1, 2, \dots, M$ ) microphone signal in the STFT domain is given by

$$Y_m(\omega) = e^{-j(m-1)\omega\tau_0\alpha} X(\omega) + V_m(\omega), \quad (1)$$

where  $X(\omega)$  and  $V_m(\omega)$  are the STFT of the desired signal,  $x(k)$ , and the noise signal,  $v_m(k)$ , respectively, (note that we neglect the frame indices for the ease of exposition),  $j = \sqrt{-1}$  is the imaginary unit,  $\omega = 2\pi f$  is the angular frequency,  $\tau_0 = \delta/c$  is the delay between two successive sensors at the angle  $\theta = 0^\circ$ , and

$$\alpha = \cos \theta. \quad (2)$$

In a vector form, (1) becomes

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

where the noise signal vector,  $\mathbf{v}(\omega)$ , is defined similarly to  $\mathbf{y}(\omega)$ ,

$$\mathbf{d}(\omega, \alpha) = [ 1 \ e^{-j\omega\tau_0\alpha} \ \dots \ e^{-j(M-1)\omega\tau_0\alpha} ]^T \quad (4)$$

is the steering vector (of length  $M$ ) at the frequency  $\omega$ , and the superscript  $T$  denotes the transpose operator.

The objective of this paper is to design DMAs that can recover the desired signal,  $X(\omega)$ , given  $\mathbf{y}(\omega)$ . For that, a complex weight,  $H_m^*(\omega)$ ,  $m = 1, 2, \dots, M$ , is applied at the output of each microphone, where the superscript  $*$  denotes complex conjugation. The weighted outputs are then summed together to form the beamformer output as shown in Fig. 1. Putting all the gains together in a vector of length  $M$ , we get

$$\mathbf{h}(\omega) = [ H_1(\omega) \ H_2(\omega) \ \dots \ H_M(\omega) ]^T. \quad (5)$$

The beamformer output is then

$$Z(\omega) = \sum_{m=1}^M H_m^*(\omega) Y_m(\omega) = \mathbf{h}^H(\omega) \mathbf{y}(\omega), \quad (6)$$

where  $Z(\omega)$  is supposed to be the estimate of the desired signal,  $X(\omega)$ , and the superscript  $H$  is the transpose-conjugate operator. The problem of beamforming is then to find a weighting vector  $\mathbf{h}(\omega)$  so that  $Z(\omega)$  is a good estimate of  $X(\omega)$ . As pointed out in the previous section, there are two different types of arrays, resulting in two different approaches to beamforming: additive and differential beamforming. In additive beamforming, the filter is optimized in such a way that the microphone signals are aligned in order to steer the main lobe in the direction of the desired signal. In comparison, differential beamforming attempts to optimize the filter to steer a number of nulls in some specific directions [10]. In this paper, our focus is on differential beamforming with small apertures, which is to design beamformers whose beampatterns are very close to the ones obtained with “ideal” DMAs. Since we deal with DMAs, we need to make the following assumptions.

- DMAs are designed to respond to the spatial derivatives of the acoustic pressure field. Therefore, we need to assume that the sensor spacing,  $\delta$ , is much smaller than the acoustic wavelength,  $\lambda = c/f$ , i.e.,  $\delta \ll \lambda$  (this implies that  $\omega\tau_0 \ll 2\pi$ ). 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 main lobe of the beampattern is at the angle  $\theta = 0^\circ$  (endfire direction). We assume that the desired signal propagates at the same angle. As a result, we have

$$\mathbf{y}(\omega) = \mathbf{d}(\omega, 1) X(\omega) + \mathbf{v}(\omega). \quad (7)$$

The approach to DMA beamforming taken here is based on the fundamental observation that for all beampatterns of interest, some constraints must be fulfilled at all frequencies [11]. In fact, the number of the fundamental constraints is equal to the number microphones, i.e.,  $M$ . In other words, we select  $M$  fundamental constraints from a well-defined beampattern of a DMA [11]. For example, in the first-order dipole with two microphones, the two fundamental constraints are a one at the angle  $0^\circ$  and a null at the angle  $90^\circ$ . Since we have two microphones and two constraints, we have a simple linear system of two equations to solve. As a result, the obtained solution is optimal from a mathematical point of view and the derived dipole is the best we can get.

### III. DMA BEAMPATTERNS

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

$$\begin{aligned} \mathcal{B}[\mathbf{h}(\omega), \alpha] &= \mathbf{d}^H(\omega, \alpha) \mathbf{h}(\omega) \\ &= \sum_{m=1}^M H_m(\omega) e^{j(m-1)\omega\tau_0\alpha}. \end{aligned} \quad (8)$$

The frequency-independent beampattern of an  $N$ th-order DMA is well known. It is defined as [6]

$$\mathcal{B}_N(\alpha) = \sum_{n=0}^N a_{N,n} \alpha^n, \quad (9)$$

where  $a_{N,n}$ ,  $n = 0, 1, \dots, N$ , are real coefficients. The different values of these coefficients determine the different directivity patterns of the  $N$ th-order DMA. In the direction of the desired signal, i.e., for  $\theta = 0^\circ$  (or  $\alpha = 1$ ), the beampattern must be equal to 1, i.e.,  $\mathcal{B}_N(1) = 1$ . Therefore, we have

$$\sum_{n=0}^N a_{N,n} = 1. \quad (10)$$

As a result, we always choose the first coefficient as

$$a_{N,0} = 1 - \sum_{n=1}^N a_{N,n}. \quad (11)$$

All interesting patterns have at least one null in some direction. Since  $\cos\theta$  is an even function, so is  $\mathcal{B}_N(\alpha)$ . Therefore, on a polar plot,  $\mathcal{B}_N(\alpha)$  is symmetric about the axis  $0^\circ - 180^\circ$  and any DMA design can be restricted to this range. It follows from (9) that an  $N$ th-order DMA has at most  $N$  nulls in this range.

According to (9), first-order directivity patterns have the form:

$$\mathcal{B}_1(\alpha) = (1 - a_{1,1}) + a_{1,1}\alpha. \quad (12)$$

For the most important patterns, the values  $a_{1,1}$  are as follows:

- dipole:  $a_{1,1} = 1$ , null at  $\alpha = 0$ ;
- cardioid:  $a_{1,1} = \frac{1}{2}$ , null at  $\alpha = -1$ ;

- hypercardioid:  $a_{1,1} = \frac{2}{3}$ , null at  $\alpha = -1/2$ ; and
- supercardioid:  $a_{1,1} = 2 - \sqrt{2}$ , null at  $\alpha = \frac{(1-\sqrt{2})}{(2-\sqrt{2})}$ .

Second-order beampatterns are described by the equation:

$$\mathcal{B}_2(\alpha) = (1 - a_{2,1} - a_{2,2}) + a_{2,1}\alpha + a_{2,2}\alpha^2. \quad (13)$$

The second-order dipole has a null at  $\alpha = 0$  and a one (maximum) at  $\alpha = -1$ . Replacing these values in (13), we find that  $a_{2,1} = 0$  and  $a_{2,2} = 1$ . By analogy with the first-order and second-order dipoles, we define the  $N$ th-order dipole as

$$\mathcal{B}_{D,N}(\alpha) = \alpha^N, \quad (14)$$

implying that  $a_{N,N} = 1$  and  $a_{N,N-1} = a_{N,N-2} = \dots = a_{N,0} = 0$ . The  $N$ th-order dipole has only one (distinct) null (in the range  $0^\circ - 180^\circ$ ) at  $\theta = 90^\circ$ .

The most well-known second-order cardioid has two nulls; one at  $\alpha = -1$  and the other one at  $\alpha = 0$ . From these values, we easily deduce from (13) that  $a_{2,1} = a_{2,2} = \frac{1}{2}$ . By analogy with the first-order and second-order cardioids, we define the  $N$ th-order cardioid as

$$\mathcal{B}_{C,N}(\alpha) = \left(\frac{1}{2} + \frac{1}{2}\alpha\right) \alpha^{N-1}, \quad (15)$$

implying that  $a_{N,N} = a_{N,N-1} = \frac{1}{2}$  and  $a_{N,N-2} = a_{N,N-3} = \dots = a_{N,0} = 0$ . This  $N$ th-order cardioid has only two distinct nulls (in the range  $0^\circ - 180^\circ$ ): one at  $\theta = 90^\circ$  and the other one at  $\theta = 180^\circ$ .

The  $N$ th-order hypercardioid and supercardioid are characterized by the fact that they have  $N$  distinct nulls in the interval  $0^\circ < \theta < 180^\circ$ . Hence, their general beampattern is

$$\mathcal{B}_{HS,N}(\alpha) = \prod_{n=1}^N [\zeta_{N,n} + (1 - \zeta_{N,n})\alpha]. \quad (16)$$

Third-order beampatterns have the form:

$$\mathcal{B}_3(\alpha) = (1 - a_{3,1} - a_{3,2} - a_{3,3}) + a_{3,1}\alpha + a_{3,2}\alpha^2 + \alpha^3. \quad (17)$$

The values of  $a_{N,n}$  for some examples of hypercardioid and supercardioid [6], [12] are as follows:

- second-order hypercardioid,  $a_{2,1} = \frac{2}{5}$ ,  $a_{2,2} = \frac{4}{5}$ ;
- second-order supercardioid,  $a_{2,1} \approx 0.484$ ,  $a_{2,2} \approx 0.413$ ;
- third-order hypercardioid,  $a_{3,1} = -\frac{4}{7}$ ,  $a_{3,2} = \frac{4}{7}$ ,  $a_{3,3} = \frac{8}{7}$ ; and
- third-order supercardioid,  $a_{3,1} \approx 0.217$ ,  $a_{3,2} \approx 0.475$ ,  $a_{3,3} \approx 0.286$ .

### IV. DMA BEAMFORMING

As pointed out earlier, the underlying principle of differential beamforming is to optimize the beamforming filter by posting one constraint at  $\theta = 0^\circ$  so that the resulting gain is equal to 1 at this direction, and then steering a number of nulls in some specific directions. A DMA of order  $N$  can have a maximum of  $N$  nulls. Therefore, any DMA of order  $N$  can

be designed by solving the following linear system of  $N + 1$  equations [11]:

$$\mathbf{D}(\omega, \alpha) \mathbf{h}(\omega) = \beta, \quad (18)$$

where

$$\mathbf{D}(\omega, \alpha) = \begin{bmatrix} \mathbf{d}^H(\omega, 1) \\ \mathbf{d}^H(\omega, \alpha_{N,1}) \\ \vdots \\ \mathbf{d}^H(\omega, \alpha_{N,N}) \end{bmatrix} \quad (19)$$

is the constraint matrix of size  $(N + 1) \times M$ ,  $M$  is the number of microphones,  $\mathbf{d}(\omega, \alpha_{N,n})$  is a steering vector of length  $M$  as defined in (4),  $\mathbf{h}(\omega)$  is a filter of length  $M$  defined in (5), and

$$\alpha = [1 \ \alpha_{N,1} \ \cdots \ \alpha_{N,N}]^T, \quad (20)$$

$$\beta = [1 \ \beta_{N,1} \ \cdots \ \beta_{N,N}]^T, \quad (21)$$

are vectors of length  $N + 1$  containing the design coefficients of the directivity pattern. In this paper, we only consider the case  $M = N + 1$  so that  $\mathbf{D}(\omega, \alpha)$  is a square matrix. This is also the case in all known approaches in the literature [6].

The choice of the coefficients  $\alpha_{N,n}$  and  $\beta_{N,n}$ ,  $n = 1, 2, \dots, N$  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  $\mathbf{D}(\omega, \alpha)$  is well conditioned so that its inverse can be computed without any numerical problem; as a result, white noise amplification can be better.
- The  $N$  pairs of coefficients  $(\alpha_{N,n}, \beta_{N,n})$  should take values from a desired “ideal” DMA beampattern. In general, they should correspond to the nulls of the desired “ideal” DMA beampattern; but they can take other values as well. For example, if the “ideal” beampattern has multiple nulls at the same angle, we should choose a different angle. Then  $\beta_{N,n}$  will no longer be 0.

If they satisfy the above two conditions, the constraints are called fundamental constraints. Then, the beamforming filter can be easily solved from (18), i.e.,

$$\mathbf{h}(\omega) = \mathbf{D}^{-1}(\omega, \alpha) \beta. \quad (22)$$

In what follows, we give a number of examples on how to design the first-, second-, and third-order DMAs.

#### A. First-Order DMAs

For a general first-order DMA, we have the following linear system of two equations:

$$\begin{bmatrix} 1 & e^{j\omega\tau_0} \\ 1 & e^{j\omega\tau_0}\alpha_{1,1} \end{bmatrix} \mathbf{h}(\omega) = \begin{bmatrix} 1 \\ \beta_{1,1} \end{bmatrix}, \quad (23)$$

where  $-1 \leq \alpha_{1,1} < 1$  and  $0 \leq \beta_{1,1} \leq 1$ . We immediately find that the solution to (23) is

$$\mathbf{h}(\omega) = \frac{1}{1 - e^{j\omega\tau_0}(1 - \alpha_{1,1})} \begin{bmatrix} 1 - \beta_{1,1}e^{j\omega\tau_1} \\ -(1 - \beta_{1,1})e^{-j\omega\tau_2} \end{bmatrix}, \quad (24)$$

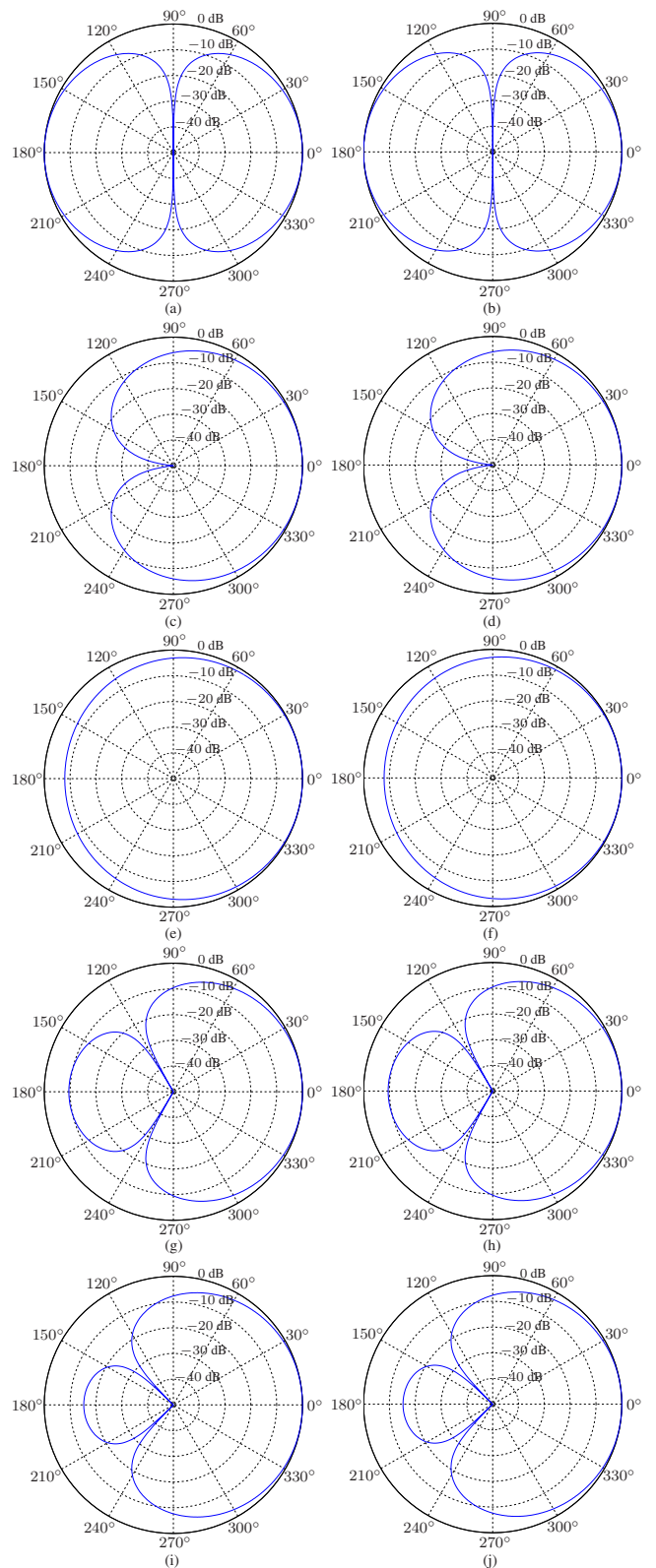


Fig. 2. Beampatterns of a first-order DMA for  $f = 0.5$  kHz: (a) dipole with  $\delta = 1$  cm, (b) dipole with  $\delta = 2$  cm, (c) cardioid with  $\delta = 1$  cm, (d) cardioid with  $\delta = 2$  cm, (e) subcardioid with  $\delta = 1$  cm, (f) subcardioid with  $\delta = 2$  cm, (g) hypercardioid with  $\delta = 1$  cm, (h) hypercardioid with  $\delta = 2$  cm, (i) supercardioid with  $\delta = 1$  cm, and (j) supercardioid with  $\delta = 2$  cm.



where  $\tau_1 = \tau_0(1 - \alpha_{1,1})$  and  $\tau_2 = \tau_0\alpha_{1,1}$ . Depending on how the two coefficients  $\alpha_{1,1}$  and  $\beta_{1,1}$  are chosen, we can have different beampatterns. The most popularly used beampatterns and the corresponding values of  $\alpha_{1,1}$  and  $\beta_{1,1}$  are as follows.

- Dipole:  $\alpha_{1,1} = \beta_{1,1} = 0$ .
- Cardioid:  $\alpha_{1,1} = -1, \beta_{1,1} = 0$ .
- Subcardioid [13], [12]:  $\alpha_{1,1} = -1, \beta_{1,1} = 0.4$ .
- Hypercardioid:  $\alpha_{1,1} = -\frac{1}{2}, \beta_{1,1} = 0$ .
- Supercardioid:  $\alpha_{1,1} = \frac{1-\sqrt{2}}{2-\sqrt{2}}, \beta_{1,1} = 0$ .

Figure 2 displays the dipole, cardioid, subcardioid, hypercardioid, supercardioid beampatterns of a first-order DMA according to (24) for  $f = 0.5$  kHz and two values of  $\delta$ , i.e.,  $\delta = 1$  cm and  $\delta = 2$  cm.

### B. Second-Order DMAs

The general linear system of equations to design any second-order differential array is

$$\begin{bmatrix} \mathbf{d}^H(\omega, 1) \\ \mathbf{d}^H(\omega, \alpha_{2,1}) \\ \mathbf{d}^H(\omega, \alpha_{2,2}) \end{bmatrix} \mathbf{h}(\omega) = \begin{bmatrix} 1 \\ \beta_{2,1} \\ \beta_{2,2} \end{bmatrix}, \quad (25)$$

where  $-1 \leq \alpha_{2,1} < 1$ ,  $-1 \leq \alpha_{2,2} < 1$ ,  $\alpha_{2,1} \neq \alpha_{2,2}$ ,  $0 \leq \beta_{2,1} \leq 1$ , and  $0 \leq \beta_{2,2} \leq 1$ .

The values of these parameters for the different patterns of interest are as follows.

- Dipole:  $\alpha_{2,1} = 0, \alpha_{2,2} = -1, \beta_{2,1} = 0$ , and  $\beta_{2,2} = 1$ . In this case, the DMA filter is

$$\mathbf{h}(\omega) = \frac{1}{(1 - e^{j\omega\tau_0})(1 - e^{j2\omega\tau_0})} \begin{bmatrix} 1 \\ -1 - e^{j\omega\tau_0} \\ e^{j\omega\tau_0} \end{bmatrix} + \frac{1}{(1 - e^{-j\omega\tau_0})(1 - e^{-j2\omega\tau_0})} \begin{bmatrix} 1 \\ -1 - e^{-j\omega\tau_0} \\ e^{-j\omega\tau_0} \end{bmatrix}. \quad (26)$$

- Cardioid:  $\alpha_{2,1} = -1, \alpha_{2,2} = 0, \beta_{2,1} = \beta_{2,2} = 0$ .
- Hypercardioid:  $\alpha_{2,1} = -0.81, \alpha_{2,2} = 0.31, \beta_{2,1} = \beta_{2,2} = 0$ .
- Supercardioid:  $\alpha_{2,1} = -0.89, \alpha_{2,2} = -0.28, \beta_{2,1} = \beta_{2,2} = 0$ .
- Quadrupole:  $\alpha_{2,1} = -\frac{1}{\sqrt{2}}, \alpha_{2,2} = \frac{1}{\sqrt{2}}, \beta_{2,1} = \beta_{2,2} = 0$ .

In the last four cases, the DMA filter is

$$\mathbf{h}(\omega) = \frac{1}{\begin{bmatrix} 1 - e^{j\omega\tau_0}(1 - \alpha_{2,1}) \\ 1 - e^{j\omega\tau_0}(1 - \alpha_{2,2}) \end{bmatrix}} \times \begin{bmatrix} 1 \\ -e^{-j\omega\tau_0}\alpha_{2,1} - e^{-j\omega\tau_0}\alpha_{2,2} \\ e^{-j\omega\tau_0}(\alpha_{2,1} + \alpha_{2,2}) \end{bmatrix}. \quad (27)$$

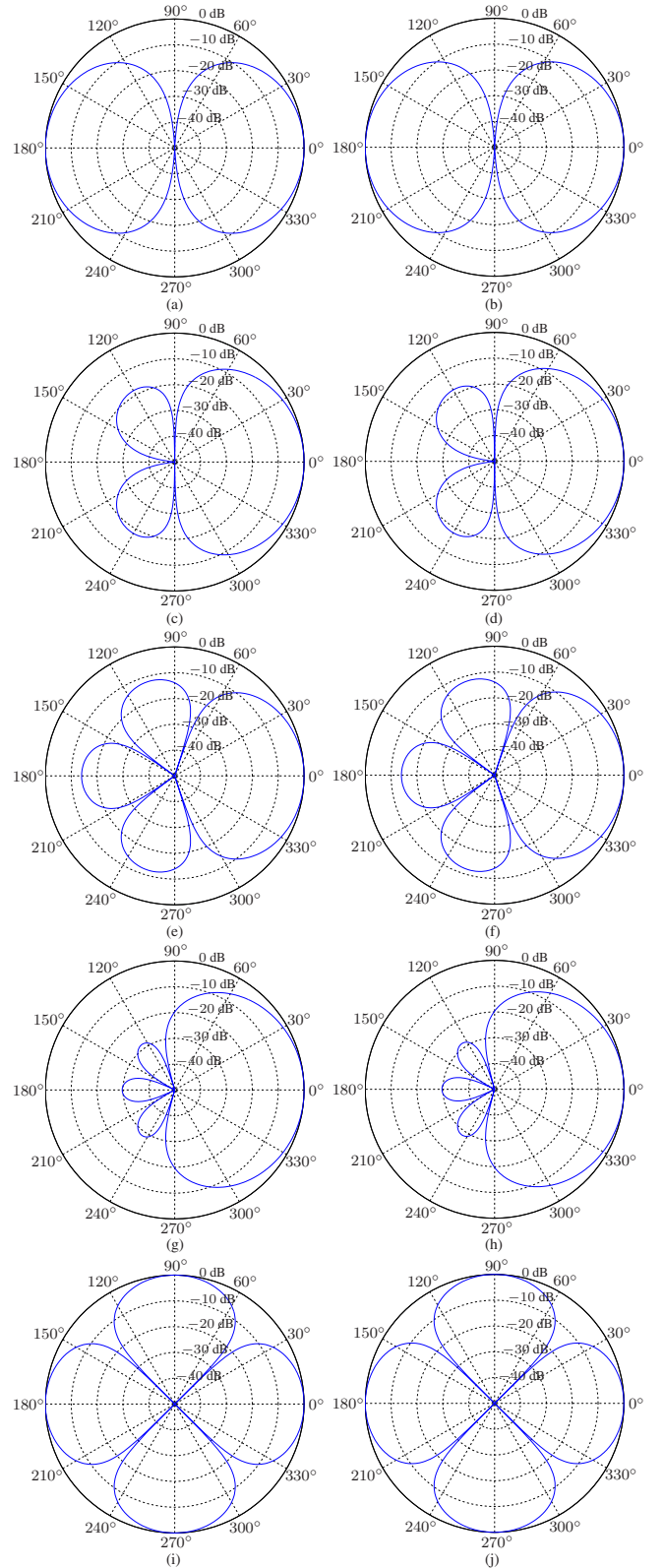


Fig. 3. Beampatterns of a second-order DMA for  $f = 0.5$  kHz: (a) dipole with  $\delta = 1$  cm, (b) dipole with  $\delta = 2$  cm, (c) cardioid with  $\delta = 1$  cm, (d) cardioid with  $\delta = 2$  cm, (e) hypercardioid with  $\delta = 1$  cm, (f) hypercardioid with  $\delta = 2$  cm, (g) supercardioid with  $\delta = 1$  cm, (h) supercardioid with  $\delta = 2$  cm, (i) quadrupole with  $\delta = 1$  cm, and (j) quadrupole with  $\delta = 2$  cm.

Figure 3 displays the dipole, cardioid, hypercardioid, supercardioid, and quadrupole beampatterns of a second-order DMA according to (26) and (27) for  $f = 0.5$  kHz and two values of  $\delta$ , i.e.,  $\delta = 1$  cm and  $\delta = 2$  cm.

### C. Third-Order DMAs

For the third-order DMAs, we only consider the case where the beamformer has three distinct nulls. In this situation, the linear system of four equations becomes

$$\begin{bmatrix} \mathbf{d}^H(\omega, 1) \\ \mathbf{d}^H(\omega, \alpha_{3,1}) \\ \mathbf{d}^H(\omega, \alpha_{3,2}) \\ \mathbf{d}^H(\omega, \alpha_{3,3}) \end{bmatrix} \mathbf{h}(\omega) = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad (28)$$

where  $-1 \leq \alpha_{3,1} < 1$ ,  $-1 \leq \alpha_{3,2} < 1$ ,  $-1 \leq \alpha_{3,3} < 1$ , and  $\alpha_{3,1} \neq \alpha_{3,2} \neq \alpha_{3,3}$ . The solution is then

$$\mathbf{h}(\omega) = \frac{1}{\begin{bmatrix} 1 - e^{j\omega\tau_0}(1 - \alpha_{3,1}) \\ 1 - e^{j\omega\tau_0}(1 - \alpha_{3,2}) \end{bmatrix}} \times \frac{1}{\begin{bmatrix} 1 - e^{j\omega\tau_0}(1 - \alpha_{3,3}) \end{bmatrix}} \begin{bmatrix} 1 \\ \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix}, \quad (29)$$

where

$$\gamma_1 = -e^{-j\omega\tau_0\alpha_{3,1}} - e^{-j\omega\tau_0\alpha_{3,2}} - e^{-j\omega\tau_0\alpha_{3,3}}, \quad (30)$$

$$\gamma_2 = e^{-j\omega\tau_0(\alpha_{3,1} + \alpha_{3,2})} + e^{-j\omega\tau_0(\alpha_{3,2} + \alpha_{3,3})} + e^{-j\omega\tau_0(\alpha_{3,1} + \alpha_{3,3})}, \quad (31)$$

$$\gamma_3 = -e^{-j\omega\tau_0(\alpha_{3,1} + \alpha_{3,2} + \alpha_{3,3})}. \quad (32)$$

Figure 4 displays the patterns of the third-order DMAs according to (29) with three interesting cases:

- Case 1:  $\alpha_{3,1} = 0$ ,  $\alpha_{3,2} = -\frac{1}{2}$ , and  $\alpha_{3,3} = -1$ ;
- Case 2:  $\alpha_{3,1} = \frac{1}{2}$ ,  $\alpha_{3,2} = -\frac{1}{2}$ , and  $\alpha_{3,3} = -1$ ;
- Case 3:  $\alpha_{3,1} = \frac{\sqrt{2}}{2}$ ,  $\alpha_{3,2} = \frac{\sqrt{2}-2}{2}$ , and  $\alpha_{3,3} = \frac{\sqrt{2}-\sqrt{10}}{2}$ .

## V. CONCLUSIONS

In this paper, we developed an approach to the design and implementation of linear DMAs for speech enhancement. Unlike the traditional method that works in the time domain, this proposed approach works in the STFT domain. Specifically, we first partition the input microphone signals into short-time frames. Each frame is then transformed into the STFT domain. In each STFT subband, a differential beamformer is designed and applied to the multichannel spectra, thereby producing an estimate of the desired signal spectrum in this subband. Finally, the time-domain desired speech estimate is constructed using the overlap-add (or overlap-save) technique with the inverse STFT. With this framework, the core issue of DMA is how to design the desired differential beamformer in each subband. To accomplish this, we form a linear system with  $N + 1$  fundamental constraints for an  $N$ th-order DMA. The DMA filter coefficients can then be obtained by solving a simple linear system.

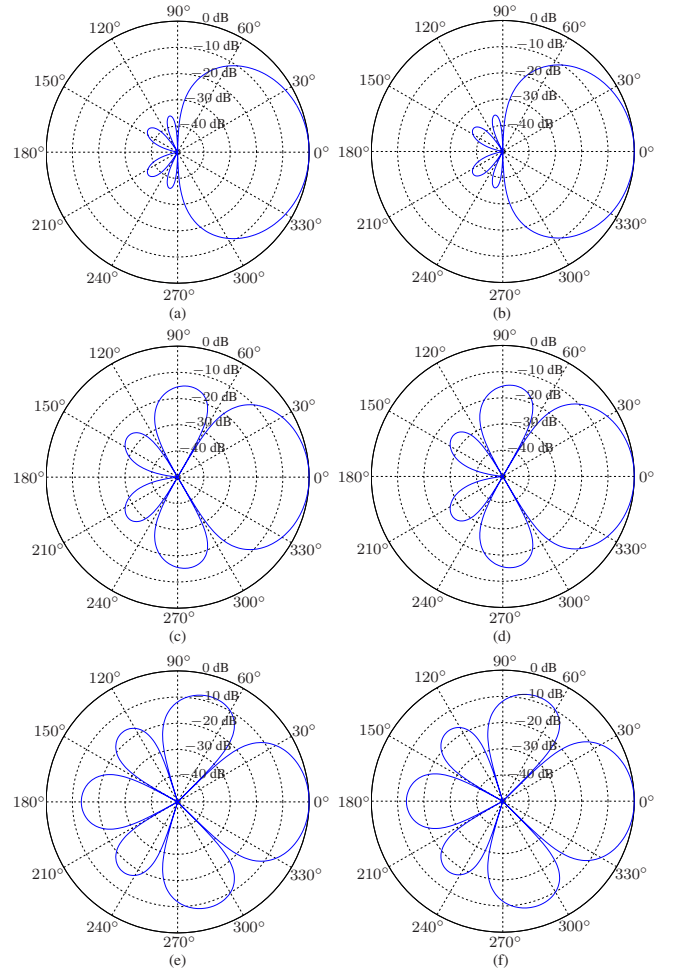


Fig. 4. Beampatterns of a third-order DMA for  $f = 0.5$  kHz: (a)  $\alpha_{3,1} = 0$ ,  $\alpha_{3,2} = -\frac{1}{2}$ ,  $\alpha_{3,3} = -1$ , and  $\delta = 1$  cm, (b)  $\alpha_{3,1} = 0$ ,  $\alpha_{3,2} = -\frac{1}{2}$ ,  $\alpha_{3,3} = -1$ , and  $\delta = 2$  cm, (c)  $\alpha_{3,1} = \frac{1}{2}$ ,  $\alpha_{3,2} = -\frac{1}{2}$ ,  $\alpha_{3,3} = -1$ ,  $\delta = 1$  cm, (d)  $\alpha_{3,1} = \frac{1}{2}$ ,  $\alpha_{3,2} = -\frac{1}{2}$ ,  $\alpha_{3,3} = -1$ , and  $\delta = 2$  cm, (e)  $\alpha_{3,1} = \frac{\sqrt{2}}{2}$ ,  $\alpha_{3,2} = \frac{\sqrt{2}-2}{2}$ ,  $\alpha_{3,3} = \frac{\sqrt{2}-\sqrt{10}}{2}$ , and  $\delta = 1$  cm, (f)  $\alpha_{3,1} = \frac{\sqrt{2}}{2}$ ,  $\alpha_{3,2} = \frac{\sqrt{2}-2}{2}$ ,  $\alpha_{3,3} = \frac{\sqrt{2}-\sqrt{10}}{2}$ , and  $\delta = 2$  cm.

## REFERENCES

- [1] M. Brandstein and D. B. Ward, Eds., *Microphone Arrays: Signal Processing Techniques and Applications*. Berlin, Germany: Springer-Verlag, 2001.
- [2] Y. Huang, J. Benesty, and J. Chen, *Acoustic MIMO Signal Processing*. Berlin, Germany: Springer-Verlag, 2006.
- [3] J. Benesty, J. Chen, and Y. Huang, *Microphone Array Signal Processing*. Berlin, Germany: Springer-Verlag, 2008.
- [4] D. B. Ward, R. C. Williamson, and R. A. Kennedy, "Broadband microphone arrays for speech acquisition," *Acoust. Australia*, vol. 26, pp. 17-20, Apr. 1998.
- [5] 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, Chapter 50, Part I, pp. 1021-1041.
- [6] G. W. Elko, "Superdirectional microphone arrays," in *Acoustic Signal Processing for Telecommunication*, S. L. Gay and J. Benesty, Eds. Boston, MA: Kluwer Academic Publishers, 2000, Chapter 10, pp. 181-237.
- [7] G. M. Sessler and J. E. West, "Directional transducers," *IEEE Trans. Audio Electroacoustic.*, vol. 19, pp. 19-23, Mar. 1971.

- [8] H. F. Olson, "Gradient microphones," *J. Acoust. Soc. Am.*, vol. 17, pp. 192–198, Jan. 1946.
- [9] G. W. Elko, J. E. West, and S. Thompson, "Differential and gradient microphone arrays," *J. Acoust. Soc. Am.*, vol. 114, pp. 2426–2426, Oct. 2003.
- [10] M. Buck, "Aspects of first-order differential microphone arrays in the presence of sensor imperfections," *European Trans. Telecommunications*, vol. 13, pp. 115–122, Mar.-Apr. 2002.
- [11] J. Benesty and J. Chen, *Study and Design of Differential Microphone Arrays*. Berlin: Springer-Verlag, 2012.
- [12] E. De Sena, H. Hacıhabiboğlu, and Z. Cvetković, "On the design and implementation of higher-order differential microphones," *IEEE Trans. Audio, Speech, Language Process.*, vol. 20, pp. 162–174, Jan. 2012.
- [13] J. Merimaa, "Applications of 3-D microphone array," in *Proc. AES 112th Convention*, 2002, pp. 1–11.
- [14] G. W. Elko and A.-T. Nguyen Pong, "A steerable and variable first-order differential microphone array," in *Proc. IEEE ICASSP*, 1997, pp. 223–226.
- [15] R. M. M. Derkx and K. Janse, "Theoretical analysis of a first-order azimuth-steerable superdirective microphone array," *IEEE Trans. Audio, Speech, Language Process.*, vol. 17, pp. 150–162, Jan. 2011.
- [16] R. Derkx, "Optimal azimuthal steering of a first-order superdirectional microphone response," in *Proc. IWAENC*, 2008.
- [17] R. N. Miles, Y. Liu, Q. Su, and E. Cui, "A silicon directional microphone with second-order directivity," in *Proc. Inter. Congress Acoust.*, Sept. 2007.
- [18] M. Kolundžija, C. Faller, and M. Vetterli, "Spatio-temporal gradient analysis of differential microphone arrays," in *Proc. Audio Engineering Society Convention*, 2009.