Applied Acoustics 148 (2019) 212–222

Contents lists available at ScienceDirect

Applied Acoustics

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

On the design of time-domain differential microphone arrays

Yaakov Buchris [a,*], Israel Cohen [a], Jacob Benesty [b]

[a] Technion, Israel Institute of Technology, Technion City, Haifa 3200003, Israel
[b] INRS-EMT, University of Quebec, 800 de la Gauchetiere Ouest, Suite 6900, Montreal, QC H5A 1K6, Canada

ARTICLE INFO

Article history:
Received 23 May 2018
Received in revised form 21 October 2018
Accepted 6 December 2018

Keywords:
Differential microphone arrays
Time-domain processing
Broadband beamforming

ABSTRACT

Differential microphone arrays (DMAs) are characterized as compact superdirective beamformers whose beampatterns are almost frequency invariant. In this work, we present a time-domain design of *N*th-order DMAs, which is important in some applications where minimal delay is required, such as real-time audio communications. Moreover, design in the time domain can reduce the computational efforts, compared to the frequency-domain design, especially when short filters are sufficient. We present design examples for DMAs illustrating some of the fundamental properties of the time-domain implementation as well as the equivalence to the frequency-domain design approach.

1. Introduction

Differential microphone arrays (DMAs) constitute a promising solution to some real-world beamforming applications involving speech signals, e.g., hands-free telecommunication, mobile phones, hearing aids, and others. Relative to their physical length, DMAs are characterized by high directivity leading to greatly intelligible speech signals even in heavy reverberant and noisy areas. DMAs first appeared in the literature in the 1930's, designed to respond to the spatial derivatives of an acoustic pressure field [1,2].

In comparison with additive arrays [3], DMAs have several advantages in processing speech signals. First, DMAs may inherently form frequency-invariant beampatterns which are effective for processing both high- and low- frequency signals. In contrast to DMAs, some work has been done on additive arrays to achieve frequency-invariant beampatterns [4–16] by introducing either more hardware and computational complexity, or by allowing model mismatch errors augmentation. Second, DMAs have the potential to attain maximum directional gain with a given number of sensors [17]. Furthermore, DMAs are generally small in size relative to the acoustic wavelength, due to the inherent assumption that the true pressure differentials can be approximated by finite differences of the microphone outputs. Therefore, DMAs can be easily integrated into communication devices. Due to these benefits, DMAs have attracted a significant amount of interest in the field of microphone array processing during the past years (see [18–34] and the references therein).

Broadband array processing algorithms can be implemented both in the time and frequency domains. Design in the time domain is of a great importance for applications requiring small delays such as real-time audio communications [35]. Second, processing in the time domain circumvents the edge effects between successive snapshots of the incoming signals, that are typical to the frequency-domain implementation. Furthermore, in some cases the implementation of time-domain filters is computationally more efficient than the equivalent frequency-domain filters, especially when short filters are sufficient. The benefits of the frequency-domain implementation is mainly due to the ability to implement some widely-used frequency dependent processing algorithms like frequency-selective null-steering and efficient calculation of the sample matrix inversion (SMI) [36]. Although algorithms to reduce the associated delay in the frequency domain exist, they usually introduce non-negligible penalty of high computational complexity [36]; thus, they are rarely used.

Previous works in the design of time-domain beamformers mainly focused on the narrowband case. Recent works have investigated time-domain frequency-invariant broadband beamformers [37–44,19,26,25]. The general approach to design time-domain FIR filters is to solve an optimization problem which incorporates several constraints in the frequency domain. The constraint of a frequency-invariant beampattern is imposed either in all angular directions or only on specific directions such as the mainbeam's direction. Other constraints may be imposed to ensure an adequate performance level such as the white noise gain (WNG) limitation.

In this work, we extend our recent preliminary work [45,46] and present a framework for a broadband time-domain equivalent design of *N*th-order DMAs. We use a similar approach like the one

* Corresponding author.
E-mail address: bucris@technion.ac.il (Y. Buchris).

https://doi.org/10.1016/j.apacoust.2018.12.013
0003-682X/

in [23] which ignores the traditional differential structure of DMAs and develops broadband frequency-domain DMAs up to any order from a signal processing perspective. In contrast to previous approaches to frequency-invariant beampattern design, this framework performs both the design process and the beamforming process in the time domain. It includes formulation of equivalent time-domain expressions for the array incoming signal, the covariance matrices of the noise and signal, and some other useful performance measures like time-domain beampattern, WNG, and directivity factor (DF). We also derive the WNG for the case that there are some model mismatch errors. The array input signal is manipulated to represent it in a separable form, meaning that the input signal is a product of a desired-signal-dependent term and a second term which depends only on the array geometry. This representation enables to apply several array processing algorithms, which were originally developed in the frequency domain, directly into broadband time-domain DMAs. We derive a closed-form solution for time-domain Nth-order DMAs for any given number of sensors. Due to the DMA assumption, the derived solution is much simpler compared to [37–44], since in the latter works several constraints should be imposed in order to ensure the frequency-invariance property. Furthermore, we show that under the DMA's assumption, the derived time-domain beampattern complies with the theoretical beampattern of traditional DMA designs [24]. Finally in the experimental results, we evaluate the performance of the time-domain DMAs in comparison with that of the frequency-domain implementation and demonstrate some of its fundamental properties.

The paper is organized as follows. In Section 2, we formulate the signal model. In Section 3, we define some useful performance measures to evaluate time-domain DMAs. In Section 4, we extend the array gain derived in the previous section, to fit also to the case of model mismatch errors. In Section 5, we develop a general closed-form solution for the Nth-order DMAs. In Section 6, we present design examples for up to third-order DMAs, along with some simulation results confirming the validity of the developed time-domain solution. Section 7 deals with computational complexity and compares time- and frequency- domain designs.

2. Signal model

We consider a broadband source signal, $s(n)$, in the far-field scenario, where n is the discrete-time index, that propagates in an anechoic acoustic environment at the speed of sound, i.e., $c = 340;\mathrm{m/s}$, and impinges on a uniform linear sensor array consisting of M omni-directional microphones, where the distance between two successive sensors is equal to δ. The direction of the source signal to the array is parameterized by the angle θ, where $\theta = 0^\circ$ corresponds to the endfire direction. In the rest, microphone 1 is chosen as the reference sensor. We should say, in passing, that the anechoic far-field model is also used to design conventional DMAs [23]. The proposed approach is more general, though, since the source is considered as broadband instead of narrowband and the design is performed directly in the time domain which has several advantages as discussed before. In this scenario, the signal measured at the mth microphone is given by

$$\begin{aligned} y_m(n) &= s[n-\Delta-f_s\tau_m(\theta)] + v_m(n) \\ &= x_m(n) + v_m(n), \end{aligned} \tag{1}$$

where f_s is the sampling frequency, Δ is the propagation time from the position of the source, $s(n)$, to the reference sensor;

$$\tau_m(\theta) = (m-1)\frac{\delta\cos\theta}{c} \tag{2}$$

is the time delay between the first and the mth sensor which can be positive or negative, and $v_m(n)$ is the noise picked up by the mth sensor. For the general case where $f_s\tau_m(\theta)$ is not an integer, we may apply the Shannon's sampling theorem [47], which implies that

$$\begin{aligned} y_m(n) &= \sum_{l=-\infty}^{\infty} s[n-\Delta-l]\mathrm{sinc}[l-f_s\tau_m(\theta)] + v_m(n) \\ &\approx \sum_{l=-P}^{P+\mu L_h} s[n-\Delta-l]\mathrm{sinc}[l-f_s\tau_m(\theta)] + v_m(n), \end{aligned} \tag{3}$$

where $\mathrm{sinc}(x) = \sin x/x$, $P \gg f_s\tau_m(\theta)$, μ is a fractional number, and L_h is the length of the FIR filter to be defined later. Hence, we can also express (1) as

$$y_m(n) = \mathbf{g}_m^T(\theta)\mathbf{s}(n-\Delta) + v_m(n), \tag{4}$$

where the superscript ^{T} is the transpose operator, the vector $\mathbf{g}_m(\theta)$ is a vector containing the coefficients of the interpolation kernel function, and the vector $\mathbf{s}(n-\Delta)$ contains $L = 2P + \mu L_h$ successive samples of the signal $s(n-\Delta)$:

$$\mathbf{s}(n-\Delta) = [s(n-\Delta+P)\cdots s(n-\Delta)\cdots s(n-\Delta-P-\mu L_h+1)]^T. \tag{5}$$

By considering L_h successive time samples of the mth microphone signal, (4) becomes a vector of length L_h:

$$\mathbf{y}_m(n) = \mathbf{G}_m(\theta)\mathbf{s}(n-\Delta) + \mathbf{v}_m(n), \tag{6}$$

where $\mathbf{G}_m(\theta)$ is a $L_h \times L$ Toeplitz matrix with the elements:

$$[\mathbf{G}_m(\theta)]_{i,j} = \mathrm{sinc}\,[-P-i+j-f_s\tau_m(\theta)], \tag{7}$$

where $i = 1,\ldots,L_h$ and $j = 1,\ldots,L$. The vector $\mathbf{v}_m(n)$ is a vector of length L_h containing the noise samples, i.e.,

$$\mathbf{v}_m(n) = [\,v_m(n) \quad \cdots \quad v_m(n-L_h+1)\,]^T. \tag{8}$$

Now, by concatenating the observations from the M microphones, we get a vector of length ML_h:

$$\begin{aligned} \underline{\mathbf{y}}(n) &= \left[\mathbf{y}_1^T(n) \quad \mathbf{y}_2^T(n) \quad \cdots \quad \mathbf{y}_M^T(n)\right]^T \\ &= \underline{\mathbf{G}}(\theta)\mathbf{s}(n-\Delta) + \underline{\mathbf{v}}(n), \end{aligned} \tag{9}$$

where

$$\underline{\mathbf{G}}(\theta) = \begin{bmatrix} \mathbf{G}_1(\theta) \\ \mathbf{G}_2(\theta) \\ \vdots \\ \mathbf{G}_M(\theta) \end{bmatrix} \tag{10}$$

is a matrix of size $ML_h \times L$ and

$$\underline{\mathbf{v}}(n) = \left[\mathbf{v}_1^T(n) \quad \mathbf{v}_2^T(n) \quad \cdots \quad \mathbf{v}_M^T(n)\right]^T \tag{11}$$

is a vector of length ML_h.

In the model of the DMAs, it is assumed that δ is much smaller than the wavelength of the incoming signal, i.e.,

$$\delta \ll \lambda, \tag{12}$$

and the desired signal propagates at the endfire, i.e., at the angle $\theta = 0$, therefore, the desired signal vector is

$$\underline{\mathbf{y}}(n) = \underline{\mathbf{G}}(0)\mathbf{s}(n-\Delta) + \underline{\mathbf{v}}(n). \tag{13}$$

It is worth to emphasize that the vector $\underline{\mathbf{y}}(n)$ has a separable representation meaning that ignoring the noise term, it is a product between a signal dependent term and a second term which depends only on the array geometry. This structure also exists in the frequency domain. Therefore, we can treat $\underline{\mathbf{G}}(0)$ as a broadband time-domain steering matrix. Moreover, it implies on the next required steps of imposing distortionless and null constraints in order to cancel the influence of the steering matrix $\underline{\mathbf{G}}(0)$ as usually done in the frequency-domain design. Then, we can design all kinds of broadband DMAs, where the main lobe is at the angle $\theta = 0$, with a real-valued spatiotemporal filter of length ML_h:

$$\underline{\mathbf{h}} = \begin{bmatrix} \mathbf{h}_1^T & \mathbf{h}_2^T & \cdots & \mathbf{h}_M^T \end{bmatrix}^T, \tag{14}$$

where $\mathbf{h}_m,\ m = 1, 2, \ldots, M$ are temporal filters of length L_h.

By applying the filter $\underline{\mathbf{h}}$ to the observation vector $\underline{\mathbf{y}}(n)$, we obtain the output of the broadband beamformer:

$$z(n) = \sum_{m=1}^{M} \mathbf{h}_m^T \mathbf{y}_m(n) = \underline{\mathbf{h}}^T \underline{\mathbf{y}}(n) = x_{\text{fd}}(n) + v_{\text{rn}}(n), \tag{15}$$

where

$$\begin{aligned} x_{\text{fd}}(n) &= \sum_{m=1}^{M} \mathbf{h}_m^T \mathbf{G}_m(0)\mathbf{s}(n-\Delta) \\ &= \underline{\mathbf{h}}^T \underline{\mathbf{G}}(0)\mathbf{s}(n-\Delta) \end{aligned} \tag{16}$$

is the filtered desired signal and

$$v_{\text{rn}}(n) = \sum_{m=1}^{M} \mathbf{h}_m^T \mathbf{v}_m(n) = \underline{\mathbf{h}}^T \underline{\mathbf{v}}(n) \tag{17}$$

is the residual noise. Fig. 1 presents a block diagram of the time-domain differential array beamformer.

3. Performance measures

In this section, we introduce several quality measures commonly used in the context of array processing. These measures are usually defined in the frequency domain for the case of broadband beamforming (see e.g. [23]). Herein, we establish their analogous time-domain versions.

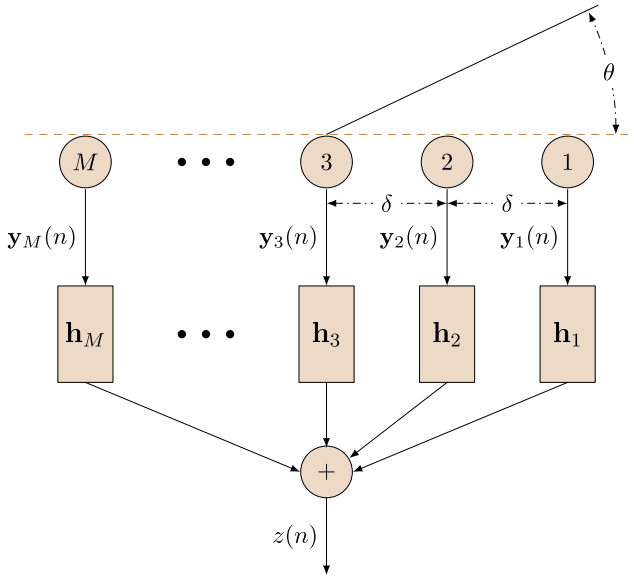

Fig. 1. A block diagram of the time-domain differential array beamformer.

Since our concern is broadband beamforming, we assume for convenience that the source signal, $s(n)$, is white; this way, the whole spectrum is taken into account. Assuming microphone 1 to be the reference sensor, the broadband input signal-to-noise ratio (SNR) is computed from the first L_h components of $\underline{\mathbf{y}}(n)$ as defined in (13), i.e., $\mathbf{y}_1(n) = \mathbf{s}(n-\Delta) + \mathbf{v}_1(n)$. We easily find that

$$\text{iSNR} = \frac{\text{tr}(\mathbf{R}_{\mathbf{s}})}{\text{tr}(\mathbf{R}_{\mathbf{v}_1})} = \frac{\sigma_s^2}{\sigma_{v_1}^2}, \tag{18}$$

where $\text{tr}(\cdot)$ denotes the trace of a square matrix, $\mathbf{R}_{\mathbf{s}}$ and $\mathbf{R}_{\mathbf{v}_1}$ are the correlation matrices of $\mathbf{s}(n-\Delta)$ and $\mathbf{v}_1(n)$, respectively, and $\sigma_{v_1}^2$ and σ_s^2 are the variances of $v_1(n)$ and $s(n)$, respectively. The broadband output SNR is given by

$$\begin{aligned} \text{oSNR}(\underline{\mathbf{h}}) &= \frac{\sigma_{x_{\text{fd}}}^2}{\sigma_{v_{\text{rn}}}^2} \\ &= \frac{\sigma_s^2 \underline{\mathbf{h}}^T \underline{\mathbf{G}}(0)\underline{\mathbf{G}}^T(0)\underline{\mathbf{h}}}{\underline{\mathbf{h}}^T \mathbf{R}_{\underline{\mathbf{v}}} \underline{\mathbf{h}}} \\ &= \frac{\sigma_s^2}{\sigma_{v_1}^2} \times \frac{\underline{\mathbf{h}}^T \underline{\mathbf{G}}(0)\underline{\mathbf{G}}^T(0)\underline{\mathbf{h}}}{\underline{\mathbf{h}}^T \Gamma_{\underline{\mathbf{v}}} \underline{\mathbf{h}}}, \end{aligned} \tag{19}$$

where

$$\Gamma_{\underline{\mathbf{v}}} = \frac{\mathbf{R}_{\underline{\mathbf{v}}}}{\sigma_{v_1}^2} \tag{20}$$

is the pseudo-correlation matrix of $\underline{\mathbf{v}}(t)$. We see from (19) that the gain in SNR is

$$\mathcal{G}(\underline{\mathbf{h}}) = \frac{\text{oSNR}(\underline{\mathbf{h}})}{\text{iSNR}} = \frac{\underline{\mathbf{h}}^T \underline{\mathbf{G}}(0)\underline{\mathbf{G}}^T(0)\underline{\mathbf{h}}}{\underline{\mathbf{h}}^T \Gamma_{\underline{\mathbf{v}}} \underline{\mathbf{h}}}. \tag{21}$$

The WNG is obtained by taking $\Gamma_{\underline{\mathbf{v}}} = \mathbf{I}_{ML_h}$, where $\mathbf{I}_{ML_h}$ is the $ML_h \times ML_h$ identity matrix, i.e.,

$$\mathcal{W}(\underline{\mathbf{h}}) = \frac{\underline{\mathbf{h}}^T \underline{\mathbf{G}}(0)\underline{\mathbf{G}}^T(0)\underline{\mathbf{h}}}{\underline{\mathbf{h}}^H \underline{\mathbf{h}}}. \tag{22}$$

We can also define the broadband beampattern or broadband directivity pattern as

$$|\mathcal{B}(\underline{\mathbf{h}}, \theta)|^2 = \underline{\mathbf{h}}^T \underline{\mathbf{G}}(\theta)\underline{\mathbf{G}}^T(\theta)\underline{\mathbf{h}}. \tag{23}$$

The theoretical frequency-invariant beampattern of an Nth-order DMA is defined as [24]

$$\mathcal{B}_N(\theta) = \sum_{n=0}^{N} a_{N,n} \cos^n \theta, \tag{24}$$

where $a_{N,n}, n = 0, 1, \ldots, N$ are real coefficients that are set according to some desired properties, like null directions, maximal directivity, and more. In the Appendix, we show that for small values of δ we get

$$|\mathcal{B}(\underline{\mathbf{h}}, \theta)|^2 \underset{\delta \ll \lambda}{\rightarrow} \mathcal{B}_N^2(\theta). \tag{25}$$

This result is meaningful since it shows how our general model converges to the theoretical beampattern of DMAs, when the DMA assumptions are satisfied.

We may also define the normalized beampattern error, E_{BP}, to be the integral over θ of the absolute error between the time-domain beampattern (23) and $\mathcal{B}_N^2(\theta)$ (24), divided by the integral over θ of the square of the theoretical beampattern:

$$E_{\text{BP}} = \frac{\int_0^{\pi} \left| |\mathcal{B}(\underline{\mathbf{h}}, \theta)|^2 - \mathcal{B}_N^2(\theta) \right| d\theta}{\int_0^{\pi} \mathcal{B}_N^2(\theta) d\theta}. \tag{26}$$

Finally, we define the DF of the array which is the gain in SNR for the case of spherical isotropic noise. One way to calculate the DF is to use (21) and substitute the time-domain version of $\Gamma_{\underline{\mathbf{v}}}$ for diffuse noise. Yet, an explicit expression for $\Gamma_{\underline{\mathbf{v}}}$ in the time domain is unavailable. Instead, we can use directly the definition of the DF (see for example at [48, ch.2]):

$$\mathcal{D}(\underline{\mathbf{h}}) = \frac{2}{\int_0^{\pi} |\mathcal{B}(\underline{\mathbf{h}}, \theta)|^2 \sin\theta d\theta}, \tag{27}$$

where $\mathcal{B}(\underline{\mathbf{h}}, \theta)$ is defined in (23) and the filter is assumed to be distortionless. These definitions of the SNR, gains, and beampattern, which are extremely useful for the evaluation of any type of beamformer, conclude this section.

4. Performance in the presence of array perturbations

In this section, we extend the analytical expression of the array gain (21) also to the practical case where the actual array has some perturbations from the nominal model.

Array perturbations usually arise from some errors in the gain, phase, as well as the locations of the sensors. Returning to our proposed model of the received signal (9), we may model the array perturbations by adding an error matrix to the steering matrix $\underline{\mathbf{G}}(\theta)$:

$$\underline{\mathbf{G}}_{\text{pert}}(\theta) = \underline{\mathbf{G}}(\theta) + \Delta\underline{\mathbf{G}}, \tag{28}$$

where $\Delta\underline{\mathbf{G}}$ is the error matrix whose entries are given by

$$[\Delta\underline{\mathbf{G}}]_{i,j} = \epsilon_{i,j}, \quad \epsilon_{i,j} \approx N(0, \sigma^2). \tag{29}$$

This model is reasonable because the matrix $\underline{\mathbf{G}}(\theta)$ is considered as the time-domain steering matrix which depends only on the array geometry.

Substituting the matrix $\underline{\mathbf{G}}_{\text{pert}}(\theta)$ into (9), we get that

$$\underline{\mathbf{y}}(n) = \underline{\mathbf{G}}(\theta)\mathbf{s}(n-\Delta) + \underline{\tilde{\mathbf{v}}}(n), \tag{30}$$

where

$$\underline{\tilde{\mathbf{v}}}(n) = \Delta\underline{\mathbf{G}}\mathbf{s}(n-\Delta) + \underline{\mathbf{v}}(n) \tag{31}$$

is a vector containing the additive background noise and the perturbation noise.

The correlation matrix of the noise vector $\underline{\tilde{\mathbf{v}}}(n)$ is given by

$$\mathbf{R}_{\underline{\tilde{\mathbf{v}}}} = \sigma_s^2\mathbf{D} + \sigma_{v_1}^2\Gamma_{\underline{\mathbf{v}}}, \tag{32}$$

where $\mathbf{D} = \text{diag}\left([\sigma^2, \sigma^2, \ldots, \sigma^2]^T\right)$ is an $ML_h \times ML_h$ diagonal matrix. The WNG for that case is given by

$$\mathcal{W}_{\text{pert}}(\underline{\mathbf{h}}) = \frac{\underline{\mathbf{h}}^T\underline{\mathbf{G}}(0)\underline{\mathbf{G}}^T(0)\underline{\mathbf{h}}}{\underline{\mathbf{h}}^T(\text{iSNR}\mathbf{D} + \mathbf{I})\underline{\mathbf{h}}}. \tag{33}$$

Note that for the ideal case when there are no model mismatch errors (i.e., $\sigma = 0$), (33) reduced to (22). We now move to present the design of time-domain DMAs up to any desired order.

5. Design of *N*th-order DMAs

It is well known that the design of Nth-order DMAs requires at least $N+1$ microphones [23,24]. We present here the general derivation for time-domain design of DMAs for any order $N \geqslant 1$ and any number of sensors $M \geqslant N+1$. The number of constraints is exactly equal to $N+1$. The first one is a distortionless constraint which can be formulated as follows. We see from (16) that the distortionless constraint is

$$\underline{\mathbf{h}}^T\underline{\mathbf{G}}(0) = \mathbf{i}^T(D), \tag{34}$$

where $\mathbf{i}(D)$ is a vector of length L with all its elements equal to zero except for its Dth element. The remaining N constraints are of the form:

$$\underline{\mathbf{h}}^T\underline{\mathbf{G}}(\theta_n) = \alpha_n\mathbf{i}^T(D), \ n = 1, 2, \ldots, N, \tag{35}$$

where α_n, $n = 1, 2, \ldots, N$, are the attenuation parameters, with $0 \leqslant \alpha_n \leqslant 1$. Combining these $N+1$ constraints together, we get the following linear system to solve

$$\mathbf{C}_{N,M}(\theta)\underline{\mathbf{h}} = \underline{\mathbf{i}}_N(\boldsymbol{\alpha}, D), \tag{36}$$

where

$$\mathbf{C}_{N,M}(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{G}_1^T(0) & \mathbf{G}_2^T(0) & \cdots & \mathbf{G}_M^T(0) \\ \mathbf{G}_1^T(\theta_1) & \mathbf{G}_2^T(\theta_1) & \cdots & \mathbf{G}_M^T(\theta_1) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{G}_1^T(\theta_N) & \mathbf{G}_2^T(\theta_N) & \cdots & \mathbf{G}_M^T(\theta_N) \end{bmatrix} \tag{37}$$

is the $(N+1)L \times ML_h$ constraints matrix, $\theta_n \in (0, \pi]$, $n = 1, 2, \ldots, N$, with $\theta_1 \neq \theta_2 \neq \cdots \neq \theta_N$, are the corresponding directions where the attenuations are desired. The vector:

$$\underline{\mathbf{i}}_N(\boldsymbol{\alpha}, D) = \begin{bmatrix} \mathbf{i}^T(D) & \alpha_1\mathbf{i}^T(D) & \cdots & \alpha_N\mathbf{i}^T(D) \end{bmatrix}^T \tag{38}$$

is of length NL. We can derive the beamformer using the pseudo-inverse solution:

$$\underline{\mathbf{h}}_{N,M;\text{Pinv}} = \mathbf{P}^{\dagger}_{\mathbf{C}_{N,M}}(\boldsymbol{\theta})\underline{\mathbf{i}}_N(\boldsymbol{\alpha}, D), \tag{39}$$

where

$$\mathbf{P}^{\dagger}_{\mathbf{C}_{N,M}}(\boldsymbol{\theta}) = \left[\mathbf{C}_{N,M}^T(\boldsymbol{\theta})\mathbf{C}_{N,M}(\boldsymbol{\theta}) + \eta\mathbf{I}\right]^{-1}\mathbf{C}_{N,M}^T(\boldsymbol{\theta}) \tag{40}$$

is the pseudo-inverse of the matrix $\mathbf{C}_{N,M}(\boldsymbol{\theta})$, and the scalar η is a regularization parameter.

In order to set the appropriate delay, D, we suggest to choose the value that minimizes the error introduced by the solution of (39), i.e.

$$D^{\text{opt}} = \min_D \|\mathbf{C}_{N,M}(\boldsymbol{\theta})\underline{\mathbf{h}}_{N,M;\text{Pinv}} - \underline{\mathbf{i}}_N(\boldsymbol{\alpha}, D)\|^2. \tag{41}$$

Fig. 2 shows the error defined by (41) as a function of the delay, D. One can see that the optimal delay is not necessary the theoretical value expected from the model of (5). This means that we have a degree of freedom by allowing an additional delay to the DMA's output signals and improving their performance. Using this value of delay provides the best results with respect to other values. Finding improved solutions that both achieve a reduced delay and sufficient level of performance is a future research topic.

Note also that in order to solve (36) in the direct way, one have to invert the matrix $\mathbf{C}_{N,M}(\boldsymbol{\theta})$. This operation is computationally expensive as it will be discussed later. Instead, we may exploit

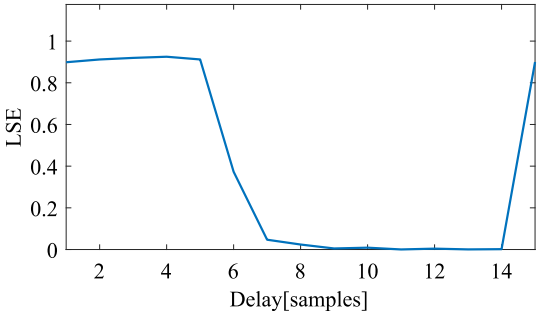

Fig. 2. The least-squares error (LSE) (41) vs. D.

the structure of this matrix. This matrix is composed of several sub-matrices, each contains the sinc function in a circulant structure. This implies that $\mathbf{C}_{N,M}(\boldsymbol{\theta})$ is a rank deficient matrix. Therefore, we can calculate the singular value decomposition (SVD) of $\mathbf{C}_{N,M}(\boldsymbol{\theta})$ and invert it using only the most dominant singular values. Fig. 3 shows an image of a typical matrix, $\mathbf{C}_{N,M}(\boldsymbol{\theta})$ for the case of second order DMAs with $M = 3$ sensors (a), and its singular values (b). One can see that this matrix is indeed a rank-deficient matrix and its effective rank is approximately equal to L, therefore, we can exploit it for more efficient calculation of its pseudo-inverse. Using the SVD decomposition, we can derive the r - rank approximation of the matrix $\mathbf{C}_{N,M}(\boldsymbol{\theta})$:

$$\mathbf{C}^{(r)}_{N,M}(\boldsymbol{\theta}) \approx \sum_{l=1}^{r} \sigma_l \mathbf{u}_l \mathbf{v}_l^T, \tag{42}$$

where $\sigma_l, \mathbf{v}_l$, and $\mathbf{u}_l$ are the lth singular value, lth right-singular vector, and lth left-singular vector of the matrix $\mathbf{C}_{N,M}(\boldsymbol{\theta})$, respectively. The parameter r may be chosen according to

$$||\mathbf{C}_{N,M}(\boldsymbol{\theta}) - \mathbf{C}^{(r)}_{N,M}(\boldsymbol{\theta})||^2 \leqslant \epsilon, \tag{43}$$

where ϵ is a small positive integer that control the accuracy of the r-rank approximation of the matrix $\mathbf{C}_{N,M}(\theta)$. We can invert the matrix using the following formula:

$$\mathbf{P}^{\dagger,(r)}_{\mathbf{C}_{N,M}}(\boldsymbol{\theta}) \approx \sum_{l=1}^{r} \frac{1}{\sigma_l} \mathbf{v}_l \mathbf{u}_l^T. \tag{44}$$

Note that by using (44) and setting an appropriate value for ϵ, we can also achieve a stable solution.

In the simulation section, we study the design of time-domain DMAs up to the third order. We also evaluate the performance of the various orders and patterns by means of the beampattern, WNG, DF, and E_{BP} as defined in Section 3. We start with the design of first-order time-domain DMAs, then we move to higher orders.

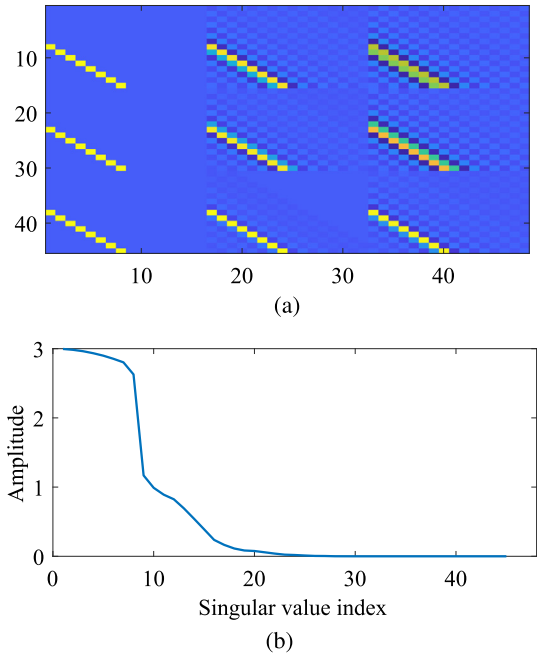

Fig. 3. An example of a typical image of the matrix $\mathbf{C}_{N,M}(\boldsymbol{\theta})$ (a), and its singular values (b).

6. Simulations

6.1. First-order DMAs

The first-order standard DMAs directivity patterns are dipole, cardioid, hypercardioid, and supercardioid, each with one distinct null in the following directions: $\theta_{\mathrm{Dp}} = \pi/2, \theta_{\mathrm{Cd}} = \pi, \theta_{\mathrm{Hc}} = 2\pi/3$, and $\theta_{\mathrm{Sc}} = 3\pi/4$. We choose the sensor spacing $\delta = 1$ cm as this value is much smaller with respect to the maximal wavelength of the signal, and first examine the case of $M = N + 1 = 2$ sensors, where $N = 1$ stands for the order. We choose the fractional delay length $P = 6$ taps which is a sufficient value for practical implementations and $\mu = 0.5$ and get $L = 24$ taps. We choose the filter length $L_h = 16$ taps and the sampling frequency to be $f_s = 8000$ Hz. The parameter ϵ is set to be 10^{-3}.

Fig. 4 shows a comparison between the broadband beampattern of the time-domain implementation (39) (black dashed line), and the theoretical beampattern [23, ch.2] (blue circles line). Using (24), it is easy to show that the theoretical beampatterns of first-order DMAs have the following forms:

$$\mathcal{B}_1(\theta) = (1 - a_{1,1}) + a_{1,1} \cos\theta, \tag{45}$$

where the parameter $a_{1,1}$ is obtained by imposing the null constraint on (45). These patterns were also achieved by the frequency-domain implementation in [23, ch.3]. The comparison illustrates the equivalence between the time- and frequency-domain implementations.

Table 1 summarizes the results for the four basic shapes by means of $\mathcal{W}(\underline{\mathbf{h}}), \mathcal{D}(\underline{\mathbf{h}})$, and E_{BP}. The main difference between the WNG and the DF of the frequency-domain implementation presented at [23, ch.3] to the values in Table 1 is that for the frequency-domain case, these quantities are frequency-dependent while for the time-domain case each of these quantities is a scalar. Yet, the mean values of the theoretical WNG and the DF over the frequency obtained in [23, ch.3] are very close to the

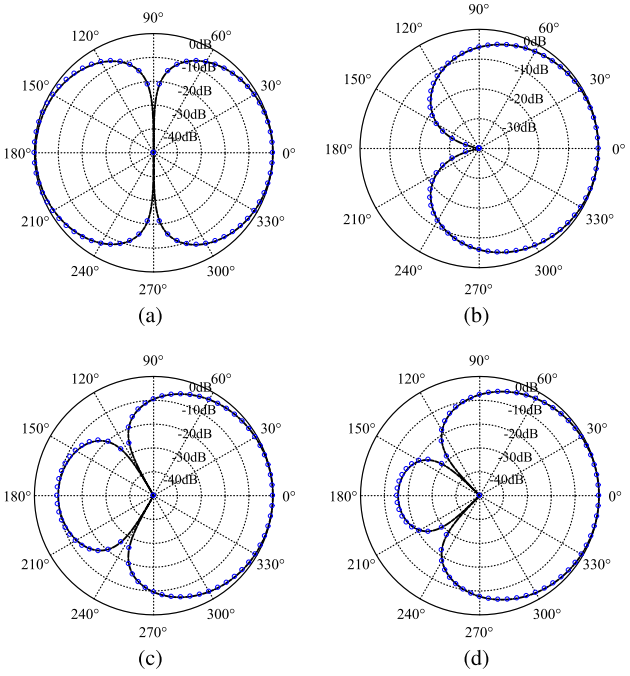

Fig. 4. Beampatterns for the four basic shapes of first-order DMAs produced by the time-domain implementation (dark dashed line): (a) dipole, (b) cardioid, (c) hypercardioid, and (d) supercardioid. The theoretical patterns are also presented (blue circles line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Results of $\mathcal{W}(\underline{\mathbf{h}})$, $\mathcal{D}(\underline{\mathbf{h}})$, and E_{BP} for the four basic shapes of first-order DMAs.

Pattern	$\mathcal{W}(\underline{\mathbf{h}})$ [dB]	$\mathcal{D}(\underline{\mathbf{h}})$ [dB]	E_{BP}
Dipole	−14.4	4.90	0.040
Cardioid	−8.68	4.75	0.007
Hypercardioid	−11.00	5.84	0.010
Supercardioid	−10.00	5.42	0.007

corresponding values in Table 1. This fact shows also the equivalence between the time-domain approach and the frequency-domain approach.

The obtained filters were tested by simulating a speech signal, whose spectrogram is presented in Fig. 5, impinging on the DMA using the model of (6). The received vector was fed into the temporal filters (39). Fig. 6 shows the time-domain waveform (dark blue line) of the signals arrived from the endfire direction, the null direction, and an arbitrary direction of $\theta = 88^\circ$ for the supercardioid beamformer. It also shows the waveforms of the recovered signals in the output of the DMA (light red line). One can see that the produced filters provide perfect recovery of the desired endfire signal while completely suppressing the signal arriving from the null direction. For signals impinging from an arbitrary direction which the DMA was not designed to suppress at all, the output signal is reasonably suppressed as compared to the input signal. Table 2 shows the normalized energy of the output signal for all the three signals:

$$\mathbf{E}_{\mathrm{norm}}(\theta) = \frac{\mathbf{E}_{\mathrm{out}}(\theta)}{\mathbf{E}_{\mathrm{i}}(\theta)}, \tag{46}$$

where $\mathbf{E}_{\mathrm{i}}(\theta)$ is the energy of a signal arrived from direction θ measured at the reference sensor (dark blue line), and $\mathbf{E}_{\mathrm{out}}(\theta)$ is the energy of the output array signal due to the input signal from direction θ (light red line). There is a reduction of about 6 dB for the signal arriving from $\theta = 88^\circ$ which corresponds to the gain of the beampattern of a supercardioid plotted in Fig. 4(d) in the direction of $\theta = 88^\circ$. These results demonstrate the practical ability of first-order DMAs to spatially filter undesired signals while perfectly recovering the endfire desired signal.

Note that unlike the traditional [24] and the frequency-domain [23] implementations of DMAs where the beamformers coefficient vectors can be explicitly expressed and the differential structure is presented, in the time-domain design each sensor is filtered by an FIR filter vector and the differential structure cannot be presented explicitly. Yet, both implementations obtain almost identical frequency-invariant beampattern as reflected from Figs. 4 and 6, meaning that the proposed design indeed implements a differential beamformer.

Fig. 7 shows the time-domain WNG and the time-domain DF as a function of the number of sensors, M, for the case of a first-order hypercardioid. One can see that the WNG increases with the number of sensors while the DF is slightly above the value of 5 dB and does not change at all. From this result, we can infer that one of the

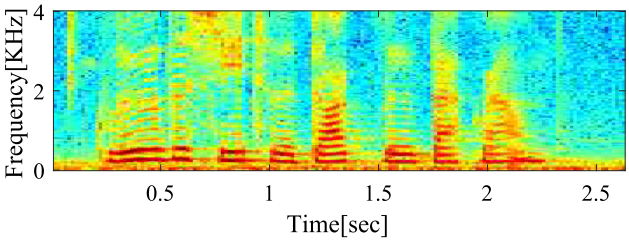

Fig. 5. Spectrogram of the speech signal.

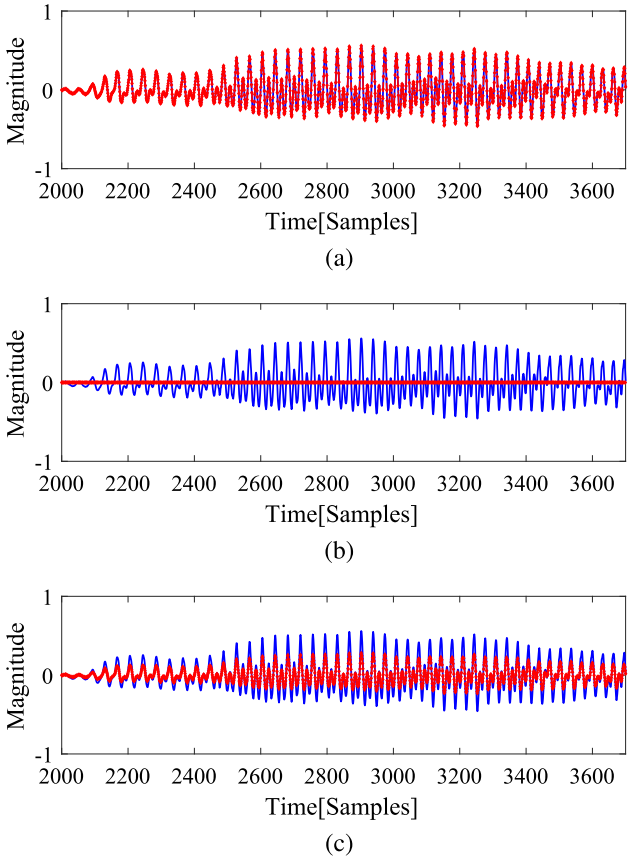

Fig. 6. Time-domain waveforms of the original signals (dark blue line) and the output recovered signals (light red line) for the supercardioid: (a) source in the endfire direction, (b) source in the null direction, and (c) source in an arbitrary direction of 88°.

Table 2
Normalized energy of the output signals for the case of a first-order supercardioid DMA.

	Endfire signal	Null signal	Arbitrary signal
$\mathbf{E}_{\mathrm{norm}}$[dB]	−0.009	−46.810	−6.569

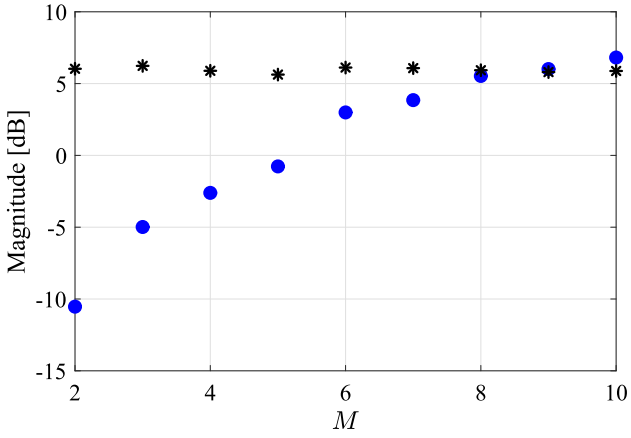

Fig. 7. WNG (circles) and DF (stars) vs. M, for the case of a first-order hypercardioid.

effective ways to increase the robustness of DMAs is to add more sensors. Increasing the number of sensors is an effective way to improve the WNG also for the case of additive arrays. For example,

the delay and sum (DAS) beamformer obtains the largest WNG for a given number of sensors which is [48]

$$\mathcal{W}_{\mathrm{DAS}} = M. \tag{47}$$

While the WNG of the DMA beamformer is about -10 dB for two microphones, the WNG of a DAS beamformer for two microphones is 3 dB.

Fig. 8 shows the time-domain WNG as a function of the number of sensors, M, for the case that some model mismatch errors exist as derived in Section 4. It compares between the following cases: $\sigma^2 = 0$ (blue circles), $\sigma^2 = 0.05$ and $\mathrm{iSNR} = 2$ (black stars), $\sigma^2 = 0.25$ and $\mathrm{iSNR} = 2$ (red diamonds), $\sigma^2 = 0.05$ and $\mathrm{iSNR} = 4$ (magenta squares), and $\sigma^2 = 0.25$ and $\mathrm{iSNR} = 4$ (green triangles). As expected, the WNG is influenced by the variance of the mismatch error. It is also influenced by the input SNR since as the SNR increased, the mismatch noise is more dominant than the background noise. For all these cases, the WNG is still an increasing function of the number of sensors, M.

Before we move to higher orders of DMAs, we present two more design aspects of DMAs. The first aspect is related to the minimal required filter length, L_h, and the second aspect is related to the element spacing parameter, δ.

6.1.1. Minimal required filter length

We examine the issue of the minimal required length of the filter vector, $\underline{\mathbf{h}}_{N,M;\mathrm{Pinv}}$, as given in (39). In Fig. 9, the time-domain waveforms of the filters for the case of a first-order hypercardioid with $M = 2$ sensors are plotted as a function of the number of taps. In the bottom plot, the filters cumulative energy as a function of the number of taps is plotted. The dashed line in the bottom plot is the total energy of the filters, justifying the assumption that in practice, a small number of coefficients is sufficient. In this case, most of the filters energy is concentrated in the 15 first coefficients.

Fig. 10 shows the performance by means of WNG and DF as a function of the filter length, L_h, for the case of a first-order hypercardioid with $M = 2$ sensors. It shows that up to $L_h = 15$ taps there are fluctuations and the performance is unstable, but starting from the length of $L_h = 15$ taps, the performance is almost constant. Similar results were obtained for the other checked patterns of first order DMAs. From this example we can see that the required length of the filter should be on the order of 10 to 20 taps which

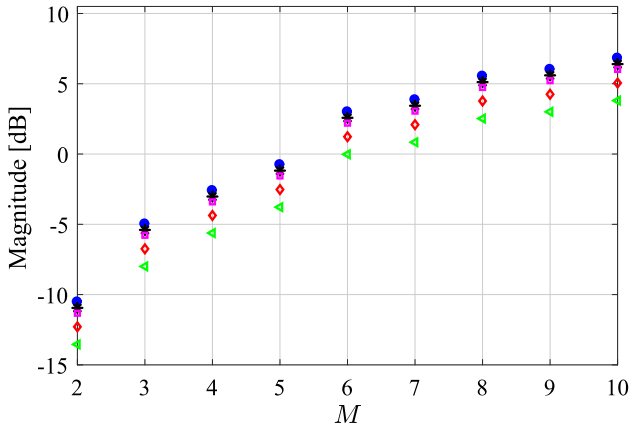

Fig. 8. WNG (blue circles) vs. M, for the case of a first-order hypercardioid. Five cases are compared: $\sigma^2 = 0$ (blue circles), $\sigma^2 = 0.05$ and $\mathrm{iSNR} = 2$ (black stars), $\sigma^2 = 0.25$ and $\mathrm{iSNR} = 2$ (red diamonds), $\sigma^2 = 0.05$ and $\mathrm{iSNR} = 4$ (magenta squares), and $\sigma^2 = 0.25$ and $\mathrm{iSNR} = 4$ (green triangles). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

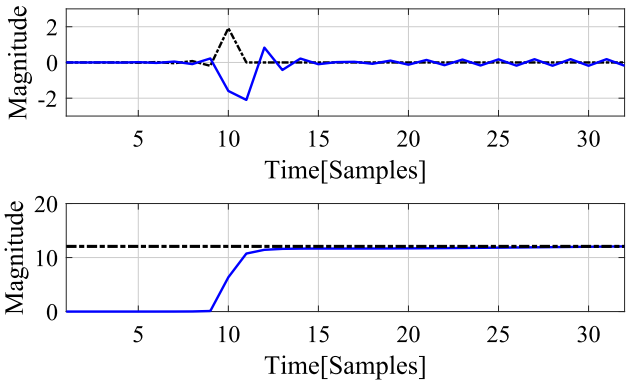

Fig. 9. Waveform of the filters for the case of a first-order hypercardioid with $M = 2$ sensors as a function of time (top). The cumulative energy vs. time (solid line) and the total filters energy (dashed line) (bottom).

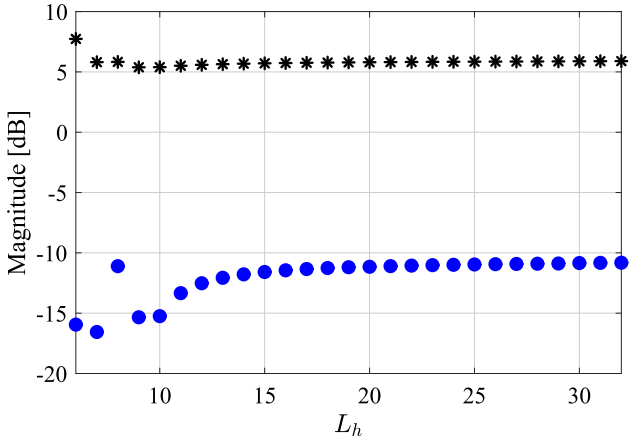

Fig. 10. WNG (circles) and DF (stars) vs. L_h for the case of a first-order hypercardioid with $M = 2$.

makes the time-domain design computationally more efficient in some cases as will be discussed later.

6.1.2. Spacing between two sensors

As discussed before, DMAs are very small and can provide high directivity relative to their small physical length. Moreover, when broadband signals are employed, they can provide nearly constant directivity patterns. In the case of a time-domain implementation, there is no clear definition of a frequency dependent WNG and DF. So, we try to exhibit the influence of the sensor spacing, δ, by simulation results. Fig. 11 shows the beampattern of a first-order hypercardioid with element spacing $\delta = 15$ cm. We can see that both the mainbeam and the sidelobes are saturated which are expected to behave like that because for such a choice of δ, the mainbeam cannot be constant over the entire bandwidth. We expect good spatial properties in the low frequencies and poor results in the high frequency range. To prove this fact, Fig. 12 shows the output spectrum (light red line) and the original spectrum (dark blue line) for the case of $\delta = 15$ cm (top) and for the case of $\delta = 1$ cm (bottom) for a white signal impinging on the array from direction of $\theta = 145°$. While for the case of a small value of δ, the entire output spectrum is flat implying a frequency-invariant beampattern, it is not the case when δ is large, resulting in a very good attenuation in low frequencies while insufficient attenuation in high frequencies. This result demonstrates the claim that under the DMAs assumption, the proposed solution is sufficient to provide frequency-invariant beampattern. However, when the

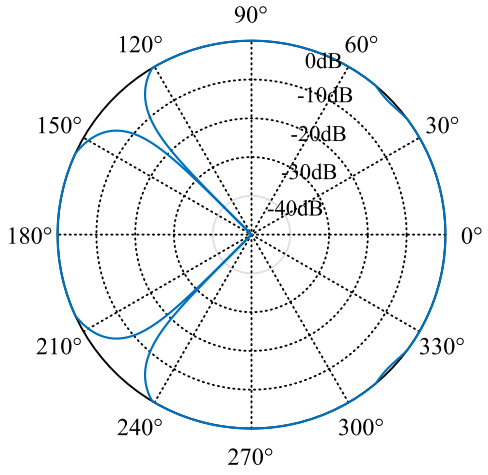

Fig. 11. Beampattern of a first-order hypercardioid ($M = 2$) for the case of $\delta = 15$ cm.

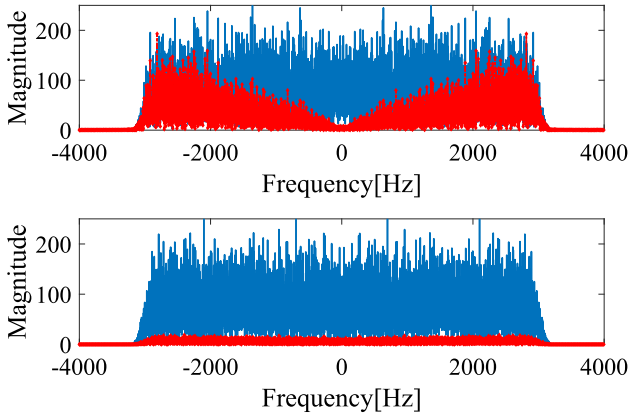

Fig. 12. Waveform for the case of $\delta = 15$ cm (top) and for the case of $\delta = 1$ cm (bottom) for a signal arrived from 145°. The dark blue spectrum is the original spectrum and the light red one is the output signal spectrum. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

assumption is not valid anymore, solutions with additional constraints should be applied in order to achieve frequency-invariant beampattern.

The results presented in this section show equivalence between time- and frequency- domain implementations of first-order DMAs. Moreover, testing the time-domain filters with actual broadband signals confirms that the desired endfire signal is undistorted while suppressing the undesired signals, even those not arriving exactly from a null direction. We now move to design examples of higher orders of DMAs.

6.2. Second-order DMAs

In this section, we present the design of the second-order DMAs. The second-order standard DMAs directivity patterns are dipole, cardioid, hypercardioid, and supercardioid. The second-order dipole has one null in $\theta_{\mathrm{Dp},1} = \pi/2$ and a second distortionless constraint at $\theta_{\mathrm{Dp},2} = \pi$. The three other shapes have two distinct nulls in the following directions: $\theta_{\mathrm{Cd},1} = \pi/2, \theta_{\mathrm{Cd},2} = \pi, \theta_{\mathrm{Hc},1} = 0.8005\pi, \theta_{\mathrm{Hc},2} = 0.3997\pi$, and $\theta_{\mathrm{Sc},1} = 0.8493\pi, \theta_{\mathrm{Sc},2} = 0.5903\pi$. We choose $L_h = 20$ taps. All the rest design parameters are identical to those of the first-order example.

Fig. 13 shows the time-domain implementation (dark dashed line) and the theoretical (blue circles line) beampatterns of the four basic shapes of second-order DMAs. Table 3 summarizes the performance measures by means of $\mathcal{W}(\underline{\mathbf{h}}), \mathcal{D}(\underline{\mathbf{h}})$, and E_{BP}. Using (24), it is easy to show that the theoretical beampatterns of second-order DMAs have the following form:

$$\mathcal{B}_2(\theta) = (1 - a_{2,1} - a_{2,2}) + a_{2,1}\cos\theta + a_{2,2}\cos^2\theta, \tag{48}$$

where the parameters $a_{2,1}$ and $a_{2,2}$ are obtained by imposing the null constraints on (48).

Fig. 14 shows the time-domain WNG and DF as a function of the number of sensors, M, for the case of a second-order cardioid. In Fig. 15 the corresponding beampatterns are plotted for (a) $M = 5$ and (b) $M = 8$ sensors, respectively. One can see that the directivity patterns is close to the theoretical value also for the robust case of increased number of sensors.

From the above examples of the second-order DMAs, one can inspect similar trends like those presented in the first-order case. These trends are also compatible to the theoretical results and to the frequency-domain results (see [23, ch.4]).

6.3. Third-order DMAs

In the case of third-order DMAs, we choose to present two out of the three patterns presented at [23, ch.5], because these patterns contain three distinct nulls in the following directions: $\theta_{\mathrm{case1},1} = \frac{\pi}{2}, \theta_{\mathrm{case1},2} = \frac{2\pi}{3}, \theta_{\mathrm{case1},3} = \pi$, and $\theta_{\mathrm{case2},1} = \frac{\pi}{3}, \theta_{\mathrm{case2},2} = \frac{2\pi}{3}$, $\theta_{\mathrm{case2},3} = \pi$. We choose $L_h = 24$ taps and $M = 4$ sensors. All the rest design parameters are like those in the previous examples. In Fig. 16 the beampatterns of the two basic shapes are plotted.

Table 4 summarizes the results for the two basic shapes by means of $\mathcal{W}(\underline{\mathbf{h}}), \mathcal{D}(\underline{\mathbf{h}})$, and E_{BP}, where the last measure is calculated using the following theoretical beampattern for third-order DMAs:

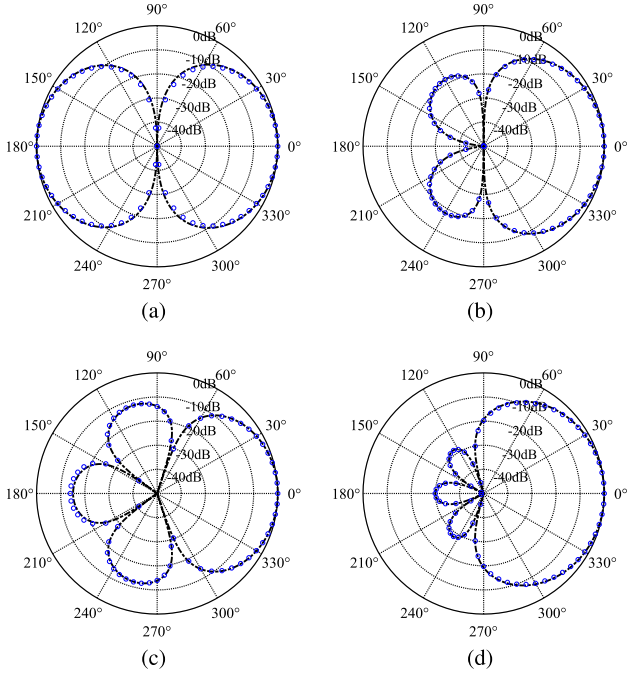

Fig. 13. Beampatterns for the four basic shapes of second-order DMAs produced by the time-domain implementation (dark dashed line): (a) dipole, (b) cardioid, (c) hypercardioid, and (d) supercardioid. The theoretical patterns are also presented (blue circles line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Results of $\mathcal{W}(\underline{\mathbf{h}}), \mathcal{D}(\underline{\mathbf{h}})$, and E_{BP} for the four basic shapes of second-order DMAs.

Pattern	$\mathcal{W}(\underline{\mathbf{h}})$ [dB]	$\mathcal{D}(\underline{\mathbf{h}})$ [dB]	E_{BP}
Dipole	−29.18	6.70	0.03
Cardioid	−25.48	8.60	0.02
Hypercardioid	−29.30	9.36	0.01
Supercardioid	−23.91	8.01	0.03

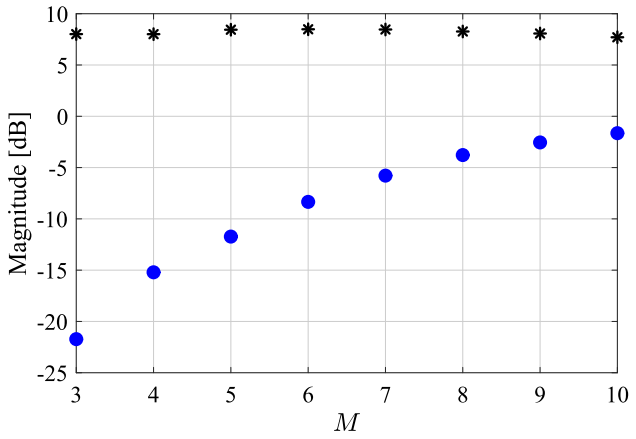

Fig. 14. WNG (circles) and DF (stars) vs. M, for the case of a second-order cardioid.

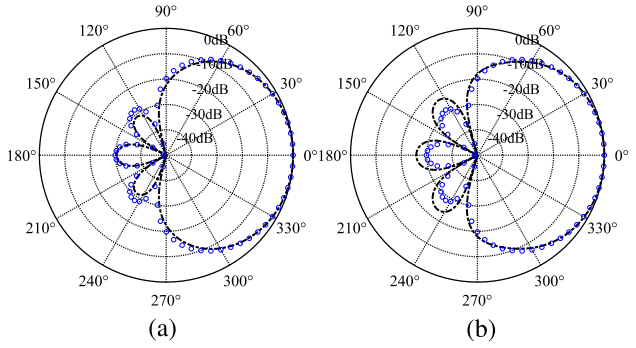

Fig. 15. Beampatterns for the case of a second-order cardioid for the case of (a) $M = 5$ sensors, and (b) $M = 8$ sensors.

$$\mathcal{B}_3(\theta) = (1 - a_{3,1} - a_{3,2} - a_{3,3}) + a_{3,1}\cos\theta + a_{3,2}\cos^2\theta + a_{3,3}\cos^3\theta, \tag{49}$$

where the parameters $a_{3,1}, a_{3,2}$, and $a_{3,3}$ are obtained by imposing the three null constraints on (49).

We finally plot in Fig. 17 the time-domain WNG and DF for various values of the number of sensors, M, for a third-order DMA (case 1). Comparing between the WNG and DF of the first-, second-, and third-order DMAs one can see that as the DMA's order increases proportionally, the WNG decreases but the DF increases proportionally to $(N+1)^2$ with accordance to the case of superdirective beamformers [17].

Other results presented in previous sections are valid also for the third-order case and are not presented here. Comparing between the results of all orders one can see similar trends of the WNG and the DF also in the frequency-domain implementation presented in [23]. This fact is important because a time-domain framework for the design of DMAs up to any order is useful in a large variety of applications with some hard design considerations like small-delay, memory, computational complexity and more. In the next section, we evaluate the computational complexity of the time-domain implementation and compare it to that of the frequency-domain implementation.

Table 4
Results of $\mathcal{W}(\underline{\mathbf{h}}), \mathcal{D}(\underline{\mathbf{h}})$, and E_{BP} for the two basic shapes of third-order DMAs.

Pattern	$\mathcal{W}(\underline{\mathbf{h}})$ [dB]	$\mathcal{D}(\underline{\mathbf{h}})$ [dB]	E_{BP}
Case 1	−29.32	10.16	0.007
Case 2	−36.78	11.78	0.030

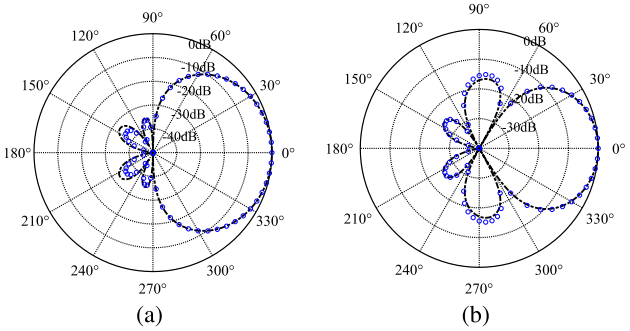

Fig. 16. Beampatterns for two cases of third-order DMAs with 3 distinct nulls, produced by the time-domain implementation (dark dashed line): (a) case 1, and (b) case 2. The theoretical patterns are also presented (blue circles line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

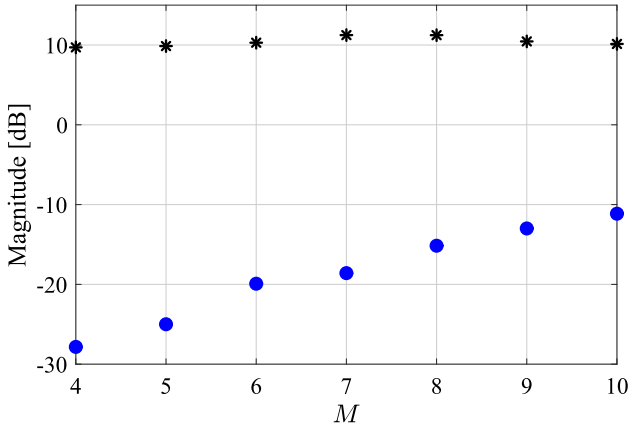

Fig. 17. WNG (circles) and DF (stars) vs. M, for the case of a third-order DMA (case 1).

7. Computational complexity

As discussed in the introduction, there are some benefits to implementing DMAs in the time domain rather than in the frequency domain. One of the most important benefits is reduced computational complexity in some design cases. Let us consider separately the filter design stage and the convolution stage in which the incoming data is filtered.

In the filter design stage, the most expensive operation in the time-domain implementation is the pseudo-inverse operation (39). Direct implementation according to (39) involves multiplication of the transpose of $\mathbf{C}_{N,M}(\boldsymbol{\theta})$ by itself, its inverse and another multiplication by the transpose of $\mathbf{C}_{N,M}(\boldsymbol{\theta})$. These operations require approximately $3M^3L_h^3$ multiplications and a similar number of additions. Lower complexity may be achieved by using the SVD method in (44) since we can use a reduced number of singular values.

If we implement the filters in the frequency domain and return to the time domain, this requires us to design the filter in each frequency bin, and then use inverse fast Fourier transform (FFT). Assuming that the number of frequency bins is also L_h, so it requires approximately $4M^3L_h^2\log_2 L_h$ multiplications and a similar number of additions, where the factor four is due to the complex operations. Moreover, the evaluated number of operations for the frequency-domain implementation assumes that L_h is an integer power of two. If this is not satisfied, a zero padding operation should be performed which increases the computational complexity even further.

Comparing the last two results, one can easily be convinced that there are non-negligible number of cases in which there will be fewer operations in direct design in time, especially for small values of the filter length, L_h. Indeed, the experimental results of the previous sections claim for a typical filter length in the range between 10 to 30 taps.

In the convolution stage, one may refer to [49, ch.5] where a detailed comparison between the computational complexities of the time and frequency convolutions is presented. The bottom line is that the choice between time or frequency implementation is related to the length of the filter and to the length of the incoming signal.

8. Conclusions

We have presented a framework for time-domain implementation of DMAs. A closed-form solution for temporal DMAs filters of any order and any number of sensors was provided. We have also formulated time-domain performance measures for microphone arrays. The proposed solution is very simple with respect to other methods due to the DMA assumption which enables to get frequency-invariant beampatterns without additional constraints in the frequency domain.

We showed that the proposed model converges to the classical DMAs model since the time-domain beampattern converges to the theoretical beampattern. Several design examples were presented in the simulation part demonstrating the equivalence of the time- and frequency- domain implementations. Each of the possible implementations involves some benefits and shortcomings. The time-domain implementation has some benefits like a short delay, which is essential in some applications such as real-time communications. On the other hand, a frequency-domain implementation enables frequency-dependent null steering, which can be useful in frequency-selective fading channels.

From a computational point of view, depending on the length of the filters, the time-domain design can be advantageous in some cases, especially when short filters are sufficient, while the design in the frequency domain may be preferred in other cases. The possibility to choose time- or frequency- domain implementation for a specific system, is of a great importance because it provides flexibility in the design considerations of several real-world applications involving DMAs.

Future research directions may focus on time-domain implementation for different array geometries, robustness issues, extension of the proposed time-domain design to support also reverberant environments, and comparisons to more advanced frequency-implementations like subband filtering implementation.

Funding

This research was supported by the Israel Science Foundation (Grant No. 576/16), the ISF-NSFC joint research program (Grant No. 2514/17) and MAFAAT-Israel Ministry of Defense.

Appendix A. Proof of the equivalence between (23) and (24) for $\delta \ll \lambda$

The analytic term for the time-domain beampattern (23) contains the product of the matrix $\underline{\mathbf{G}}(\theta)$ with its transpose:

$$\underline{\mathbf{G}}(\theta)\underline{\mathbf{G}}^T(\theta) = \begin{bmatrix} \mathbf{G}_1(\theta)\mathbf{G}_1^T(\theta) & \cdots & \mathbf{G}_1(\theta)\mathbf{G}_M^T(\theta) \\ \vdots & \vdots & \vdots \\ \mathbf{G}_M(\theta)\mathbf{G}_1^T(\theta) & \cdots & \mathbf{G}_M(\theta)\mathbf{G}_M^T(\theta) \end{bmatrix}. \tag{50}$$

According to (7), the matrix $\mathbf{G}_i(\theta)$ has the following structure:

$$\mathbf{G}_i(\theta) = \begin{bmatrix} \mathbf{t}^T(l-\tau_i(\theta)f_s) \\ \mathbf{t}^T(l-1-\tau_i(\theta)f_s) \\ \vdots \\ \mathbf{t}^T(l-L_h+1-\tau_i(\theta)f_s) \end{bmatrix}, \tag{51}$$

where $\mathbf{t}(l-\tau_i(\theta)f_s) = \mathrm{sinc}[-P-1+l-f_s\tau_i(\theta)]$, $l=1,\ldots,L$ is a column vector. From now on, we omit the dependency of $\tau_i(\theta)$ on θ for notation simplicity. In order to simplify (50), we may exploit the following property of the scalar product between two time-shifted sinc functions:

$$\mathbf{t}^T(\tau_1)\mathbf{t}(\tau_2) \approx \begin{cases} \xi_{\tau_1,\tau_2} & , \quad |\tau_1-\tau_2|<1 \\ 0 & , \quad |\tau_1-\tau_2|>1 \end{cases}, \tag{52}$$

where

$$\xi_{\tau_1,\tau_2} = \mathrm{sinc}^2\left(\frac{\tau_1+\tau_2}{2}f_s\right) \tag{53}$$

is a scalar which expresses the value of the shifted version of the sinc function at the best matching point between $\mathbf{t}(\tau_1)$ and $\mathbf{t}(\tau_2)$. Therefore, each block of (50) can be simplified to

$$\mathbf{G}_i(\theta)\mathbf{G}_j^T(\theta) \approx \mathrm{sinc}^2\left(\frac{\tau_i+\tau_j}{2}f_s\right)\mathbf{I}, \tag{54}$$

where $\mathbf{I}$ is an $L_h \times L_h$ identity matrix. We may conclude that $\underline{\mathbf{G}}(\theta)\underline{\mathbf{G}}^T(\theta)$ is a matrix which is composed of diagonal sub-matrices of the form (54).

Substituting the last result into (23) we get the following:

$$\begin{aligned} \underline{\mathbf{h}}^T\underline{\mathbf{G}}(\theta)\underline{\mathbf{G}}^T(\theta)\underline{\mathbf{h}} &\approx \sum_{i=1}^{M}\sum_{j=1}^{M}\mathbf{h}_i^T\mathbf{G}_i(\theta)\mathbf{G}_j^T(\theta)\mathbf{h}_j \\ &= \sum_{i=1}^{M}\sum_{j=1}^{M}\mathrm{sinc}^2\left(\tfrac{\tau_i+\tau_j}{2}f_s\right)\mathbf{h}_i^T\mathbf{h}_j \\ &= \sum_{i=1}^{M}\sum_{j=1}^{M}\mathrm{sinc}^2\left(\tfrac{\delta f_s\mu_{i,j}}{c}\cos\theta\right)\mathbf{h}_i^T\mathbf{h}_j, \end{aligned} \tag{55}$$

where $\mu_{i,j} = i+j-2$. According to the DMA assumption [23], $\frac{2\pi f\delta}{c} \ll 2\pi$, therefore, $\frac{f_s\delta}{c} \approx \frac{f_{\max}\delta}{c} \ll 1$, where $f_{\max}$ is the highest frequency bin of the signal. Let us assign $x_{i,j} \equiv \mu_{i,j}\frac{f_s\delta}{c}\cos\theta$ and approximate $f(x_{i,j}) = \mathrm{sinc}^2(x_{i,j})$ using Taylor series around the point x_0:

$$f(x) = \sum_{n=0}^{\infty}\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n. \tag{56}$$

We can calculate (56) by using the well-known Taylor series $\mathrm{sinc}(x) = \sum_{n=0}^{\infty}\beta_n(-1)^n x^{2n}$ and its derivatives which lead to the following general form:

$$\mathrm{sinc}^2(x_{i,j}) \approx \sum_{n=0}^{2N} f_n(x_{i,j})\frac{\delta f_s\mu_{i,j}}{c}(\cos\theta-1)^n, \tag{57}$$

where we choose x_0 as the corresponding to the case of $\theta = 0^\circ$ and cut the infinite series after $2N$ terms. The functions f_n contain the derivatives of the sinc function. Opening the brackets in the last expression and substituting into (55) we get the general structure:

$$\underline{\mathbf{h}}^T\underline{\mathbf{G}}(\theta)\underline{\mathbf{G}}^T(\theta)\underline{\mathbf{h}} \approx \sum_{n=0}^{2N} \varrho_{N,n}(\Theta)\cos^n\theta, \tag{58}$$

where Θ is the parameters vector containing all the above parameters like $\delta, c, \underline{\mathbf{h}}$, etc.

On the other hand, the square of (24) is

$$\mathcal{B}_N^2(\theta) = \left(\sum_{n=0}^{N} a_{N,n}\cos^n\theta\right)^2 = \sum_{n=0}^{2N} \tilde{a}_{N,n}\cos^n\theta. \tag{59}$$

Comparing between (58) and (59), it is obvious that both are equivalent.

Q.E.D

References

[1] Olson HF. A uni-directional ribbon microphone. J Acoust Soc Am 1932;3:139–47.
[2] Olson HF. Gradient microphones. J Acoust Soc Am 1946;17:192–8.
[3] Benesty J, Chen J, Huang Y. Microphone Array Signal Processing, Berlin. Germany: Springer-Verlag; 2008.
[4] Crocco M, Trucco A. Design of robust superdirective arrays with a tunable tradeoff between directivity and frequency-invariance. IEEE Trans Signal Process 2011;59(5):2169–81.
[5] Crocco M, Trucco A. Stochastic and analytic optimization of sparse aperiodic arrays and broadband beamformers with robust superdirective patterns. IEEE Trans Audio, Speech, Language Process 2012;20(9):2433–47.
[6] Buchris Y, Cohen I, Doron M. Bayesian focusing for coherent wideband beamforming. IEEE Trans Audio, Speech, Language Process 2012;20 (4):1282–95.
[7] Flangan J, Berkeley D, Elko G, West J, Sondhi M. Autodirective microphone systems. Acustica 1991;73:58–71.
[8] Doles IJH, Benedict FD. Broad-band array design using the asymptotic theory of unequally spaced arrays. IEEE Trans Antennas Propagat 1991;36:27–33.
[9] Ward DB, Kennedy RA, Williamson RC. Theory and design of broadband sensor arrays with frequency invariant farfield beam-patterns. J Acoust Soc Amer 1995;97(2):1023–34.
[10] Chan SC, Chen HH. Uniform concentric circular arrays with frequency-invariant characteristics - theory, design, adaptive beamforming and doa estimation. IEEE Trans Signal Process 2007;55:165–77.
[11] Liu W, Weiss S. Design of frequency invariant beamformers for broadband arrays. IEEE Trans Signal Process 2008;56(2):855–60.
[12] Doclo S, Moonen M. Design of broadband beamformers robust against gain and phase errors in the microphone array characteristics. IEEE Trans Signal Process 2003;51(10):2511–26.
[13] Van Veen BD, Buckley KV. Beamforming: a versatile approach to spatial filtering. IEEE Mag Acoust, Speech, Signal Process 1988;5:4–24.
[14] E. Mabande, A. Schad, W. Kellermann, Design of robust superdirective beamformers as a convex optimization problem, Int. Conf. Acoust., Speech, Signal Process., Taipei, Taiwan.
[15] Benesty J, Chen J, Huang Y, Dmochowski J. On microphone-array beamforming from a MIMO acoustic signal processing perspective. IEEE Trans Audio, Speech, Language Process 2007;15(3):1053–65.
[16] Do-Hong T, Russer P. Spatial signal processing for wideband beamforming. p. 73–6.
[17] G.W. Elko, J. Meyer, Microphone arrays, in: J. Benesty, M.M. Sondhi, Y. Huang (Eds.), Springer Handbook of Speech Processing, Berlin, Germany: Springer-Verlag, 2008, Ch. 50, Part I, pp. 1021–1041.
[18] Benesty J, Souden M, Huang Y. A perspective on differential microphone arrays in the context of noise reduction. IEEE Trans Audio, Speech, Language Process 2012;20(2):699–704.
[19] De Sena E, Hacihabibğlu H, Cavetković Z. On the design and implementation of higher order differential microphones. IEEE Trans Audio, Speech, Language Process 2012;20(1):162–74.
[20] Abhayapala TD, Gupta A. Higher order differential-integral microphone arrays. J Acoust Soc Am 2010;127:227–33.
[21] Buck M. Aspects of first-order differential microphone arrays in the presence of sensor imperfections. Eur Trans Telecommun 2002;13:115–22.
[22] Kolundija M, Faller C, Vetterli M. Spatiotemporal gradient analysis of differential microphone arrays. J Audio Eng Soc 2011;59:20–8.
[23] Benesty J, Chen J. Study and Design of Differential Microphone Arrays, Berlin. Germany: Springer-Verlag; 2012.
[24] Elko GW. Superdirectional microphone arrays. In: Gay SL, Benesty J, editors. Acoustical Signal Processing for Telecommunication. Boston, MA: Kluwer Academic Publishers; 2000. p. 181–237. Ch. 10.
[25] Elko GW. Differential microphone arrays. In: Huang Y, Benesty J, editors. Audio Signal Processing for Next-Generation Multimedia Communication Systems, 2. Norwell, MA: Kluwer; 2004. p. 11–65.
[26] Teutsch H, Elko GW. First- and second-order adaptive differential microphone arrays. In: 7th International Workshop on Acoustic Echo and Noise Control, Darmstadt, Germany. p. 57–60.
[27] Elko GW, Pong ATN. A simple first-order differential microphone. p. 169–72.
[28] Elko GW, Pong ATN. A steerable and variable firstorder differential microphone array. Proc ICASSP 1997:223–6.
[29] Derkx RMM, Janse K. Theoretical analysis of a first-order azimuth-steerable superdirective microphone array. IEEE Trans Audio, Speech, Language Process 2009;17:150–62.
[30] Chen J, Benesty J. A general approach to the design and implementation of linear differential microphone arrays. p. 1–7.
[31] Zhao L, Benesty J, Chen J. Design of robust differential microphone arrays. IEEE Trans Audio, Speech, Language Process 2014;22(10):1455–64.
[32] Chao P, Benesty J, Chen J. Design of robust differential microphone arrays with orthogonal polynomials. J Acous Soc Am 2015;138(2):1079–89.
[33] Zhang H, Chen J, Benesty J. Study of nonuniform linear differential microphone arrays with the minimum-norm filter. Appl Acoust 2015;98:62–9.
[34] Chao P, Chen J, Benesty J. Theoretical analysis of differential microphone array beamforming and an improved solution. IEEE Trans Audio, Speech, Language Process 2015;23(11):2093–105.
[35] Godara LC. Application of the fast fourier transform to broadband beamforming. J Acoust Soc Am 1995;98(1):230–40.
[36] Compton RT. The relationship between tapped delay-line and FFT processing in adaptive arrays. IEEE Trans Antennas Propag 1988;36(1):15–26.
[37] Nongpiur RC, Shpak DJ. L-infinity norm design of linear-phase robust broadband beamformers using constrained optimization. IEEE Trans Signal Process 2013;61(23):6034–46.
[38] Mabande E, Schad A, Kellermann W. A time-domain implementation of data-independent robust broadband beamformers with low filter order. In: Proc. Joint Workshop on Hands-free Speech Commun. Microphone Arrays, Edinburgh, U.K.
[39] Godara LC, Jahromi MRS. Convolution constraints for broadband antenna arrays. IEEE Trans Anthennas Propag 2007;55(11):3146–54.
[40] Yan S, Ma YL, Hou CH. Optimal array pattern synthesis for broadband arrays. J Acoust Soc Am 2007;122(5):2686–96.
[41] Yan S. Optimal design of fir beamformer with frequency invariant patterns. Appl Acoust 2006;67:511–28.
[42] Chen HH, Chan SC, Ho KL. Adaptive beamforming using frequency invariant uniform concentric circular arrays. IEEE Trans Circuits Syst I, Reg. Papers 2007;54(9):1938–49.
[43] Ward DB, Kennedy RA, Williamson RC. FIR filter design for frequency invariant beamformers. IEEE Signal Process Lett 1996;3:69–71.
[44] Yan S, Sun H, Ma X, Svensson PU, Hou C. Time-domain implementation of broadband beamformer in spherical harmonics domain. IEEE Trans Audio, Speech, Language Process 2011;19(5):1221–30.
[45] Y. Buchris, I. Cohen, J. Benesty, First-order differential microphone arrays from a time-domain broadband perspective, in: International Workshop on Acoustic Echo and Noise Control (IWAENC) 2016 conf.
[46] Buchris Y, Cohen I, Benesty J. Analysis and design of time-domain first-order circular differential microphone arrays.
[47] Shannon CE. Communications in the presence of noise. Proc IRE 1949;37:10–21.
[48] Van-trees HL. Detection, Estimation and Modulation Theory, Part IV – Optimum Array Processing. New York: Wiley Interscience; 2002.
[49] Porat B. A Course in Digital Signal Processing. John Wiley & Sons, INC.; 1997.