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

A R T I C L E   I N F O

Keywords:
Differential microphone arrays
Time-domain processing
Broadband beamforming

A B S T R A C T

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 Nth-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 Nth-order DMAs. We use a similar approach like the one
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 also fit 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., $c = 340$ 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 $\delta$. The direction of the source signal to the array is parameterized by the angle $\theta$, 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 $m$th microphone is given by

$$y_m(n) = s(n) - f_s \tau_m(\theta) + v_m(n)$$

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

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

is the delay time between the first and the $m$th sensor which can be positive or negative, and $v_m(n)$ is the noise picked up by the $m$th 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

$$y_m(n) = \sum_{i=-\infty}^{\infty} s(n - \Delta - \lfloor f_s \tau_m(\theta) \rfloor) + v_m(n)$$

$$\approx \sum_{i=-\infty}^{P-1} s(n - \Delta - \lfloor f_s \tau_m(\theta) \rfloor) + v_m(n),$$

where $\lfloor \cdot \rfloor$ is the floor function;

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

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

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

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

$$[\mathbf{g}_m(\theta)]_{ij} = \text{sinc} [-P - i + j - f_s \tau_m(\theta)],$$

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

$$\mathbf{v}_m(n) = [v_{m1}(n) \cdots v_{mL_h}(n)]^T.$$
It is worth to emphasize that the vector $\mathbf{y}(n)$ 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 $\mathbf{G}(0)$ as a broadband time-domain steering matrix. Moreover, it implies on the next required steps of imposing distortionless and null constraints in the frequency-domain. Therefore, we can treat $\mathbf{G}(0)$ 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 $ML_h$:

$$\mathbf{h} = \left[ \mathbf{h}_1^T \mathbf{h}_2^T \ldots \mathbf{h}_M^T \right]^T,$$

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

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

$$z(n) = \sum_{m=1}^{M} \mathbf{h}_m^T \mathbf{y}_m(n) = \mathbf{h}^T \mathbf{y}(n) = x_{\text{d}}(n) + v_{\text{r}}(n),$$

where

$$x_{\text{d}}(n) = \sum_{m=1}^{M} \mathbf{h}_m^T \mathbf{G}_m(0) \mathbf{s}(n - \Delta)$$

$$= \mathbf{h}^T \mathbf{G}(0) \mathbf{s}(n - \Delta)$$

is the filtered desired signal and

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

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.

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 $\mathbf{y}(n)$ as defined in (13), i.e., $\mathbf{y}_1(n) = s(n - \Delta) + \mathbf{v}_1(n)$. We easily find that

$$\text{iSNR} = \frac{\text{tr}(\mathbf{R}_s)}{\text{tr}(\mathbf{R}_v)} = \frac{\sigma_s^2}{\sigma_v^2},$$

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

$$\text{oSNR}(\mathbf{h}) = \frac{\sigma_s^2}{\sigma_v^2} = \frac{\sigma_s^2}{\sigma_v^2} \left( \frac{\mathbf{h}^T \mathbf{G}(0) \mathbf{G}(0)^T \mathbf{h} \mathbf{h}^T \mathbf{G}(0) \mathbf{G}(0)^T}{\mathbf{h}^T \mathbf{G}(0) \mathbf{G}(0)^T \mathbf{h}} \right),$$

where

$$\Gamma_\Sigma = \mathbf{R}_v / \sigma_v^2$$

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

$$\mathbf{G}(\mathbf{h}) = \text{oSNR}(\mathbf{h}) \text{iSNR} = \frac{\mathbf{h}^T \mathbf{G}(0) \mathbf{G}(0)^T \mathbf{h} \mathbf{h}^T \mathbf{G}(0) \mathbf{G}(0)^T}{\mathbf{h}^T \mathbf{G}(0) \mathbf{G}(0)^T \mathbf{h}}.$$

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

$$\mathbf{W}(\mathbf{h}) = \frac{\mathbf{h}^T \mathbf{G}(0) \mathbf{G}(0)^T \mathbf{h}}{\mathbf{h}^T \mathbf{h}}.$$

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

$$|B(\mathbf{h}, 0)|^2 = \mathbf{h}^T \mathbf{G}(0) \mathbf{G}(0)^T \mathbf{h}.$$

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

$$B_N(\theta) = \sum_{n=0}^{N} a_n \cos^n \theta,$$

where $a_{n+1}, n = 0, 1, \ldots, N$ 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 $\delta$ we get

$$|B(\mathbf{h}, 0)|^2 \rightarrow B_N^2(\theta).$$

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, $E_{\text{BP}}$, to be the integral over $\theta$ of the absolute error between the time-domain beampattern (23) and $B_N^2(\theta)$ (24), divided by the integral over $\theta$ of the square of the theoretical beampattern:

$$E_{\text{BP}} = \frac{\int_0^{2\pi} |B(\mathbf{h}, 0)|^2 - B_N^2(\theta) \, d\theta}{\int_0^{2\pi} B_N^2(\theta) \, d\theta}.$$
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 $\Gamma_x$ for diffuse noise, yet, an explicit expression for $\Gamma_x$ in the time domain is unavailable. Instead, we can use directly the definition of the DF (see for example at [48, ch.2]):

$$D(h) = \frac{2}{\int_0^\infty |B(h, \theta)|^2 \sin(\theta) d\theta},$$  

(27)

where $B(h, \theta)$ is defined in (23) and the filter is assumed to be distortionless. These definitions of the SNR, gains, and beampattern, 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 $G(\theta)$:

$$G_{pert}(\theta) = G(\theta) + \Delta G,$$

(28)

where $\Delta G$ is the error matrix whose entries are given by

$$[\Delta G]_{ij} = \epsilon_{ij}, \quad \epsilon_{ij} \approx N(0, \sigma^2).$$

(29)

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

Substituting the matrix $G_{pert}(\theta)$ into (9), we get that

$$y(n) = G(\theta)s(n - \Delta) + \bar{y}(n),$$

(30)

where

$$\bar{y}(n) = \Delta Gs(n - \Delta) + \bar{y}(n)$$

(31)

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

The correlation matrix of the noise vector $\bar{y}(n)$ is given by

$$R_y = \sigma^2 D + \sigma^2 \Gamma_y,$$

(32)

where $D = \text{diag}\{\sigma^2, \sigma^2, \ldots, \sigma^2\}^T$ is an $M_{\text{LA}} \times M_{\text{LA}}$ diagonal matrix. The WNG for that case is given by

$$\mathcal{W}_{pert}(h) = \frac{h^T G(0) G^T(0) h}{h^T (\text{SNRD} + \Gamma) h},$$

(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 $N$th-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 $N \geq 1$ and any number of sensors $M \geq N + 1$. 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

$$h^T G(0) = i^T(D),$$

(34)

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

$$h^T G(\theta_n) = x_n i^T(D), \quad n = 1, 2, \ldots, N,$$

(35)

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

$$C_{N,M}(\theta) b_N = i_N(x, D),$$

(36)

where

$$C_{N,M}(\theta) = \left[ \begin{array}{cccc} G_1^T(0) & G_1^T(0) & \cdots & G_1^T(0) \\ \vdots & \vdots & \ddots & \vdots \\ G_N^T(0) & G_N^T(0) & \cdots & G_N^T(0) \end{array} \right],$$

(37)

is the $(N + 1) \times ML_h$ matrix which has $n$ constraints, $\theta_n \in [0, \pi], \quad n = 1, 2, \ldots, N$, are the corresponding directions where the attenuations are desired. The vector:

$$i_N(x, D) = \left[ \begin{array}{c} i^T(D) \\ x_1 i^T(D) \\ \vdots \\ x_N i^T(D) \end{array} \right]^T$$

(38)

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

$$b_{N,M,\text{Pinv}} = P_{C_{N,M}}^{-1}(\theta) b_N(x, D),$$

(39)

where

$$P_{C_{N,M}}^{-1}(\theta) = \left[ C_{N,M}(\theta) C_{N,M}(\theta) + \eta I \right]^{-1} C_{N,M}(\theta)$$

(40)

is the pseudo-inverse of the matrix $C_{N,M}(\theta)$, and the scalar $\eta$ 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.

$$D_{\text{opt}} = \min_{D} \| C_{N,M}(\theta) b_{N,M,\text{Pinv}} - i_N(x, D) \|^2.$$

(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 $C_{N,M}(\theta)$. This operation is computationally expensive as it will be discussed later. Instead, we may exploit
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 $C_{N,M}(\theta)$ is a rank deficient matrix. Therefore, we can calculate the singular value decomposition (SVD) of $C_{N,M}(\theta)$ and invert it using only the most dominant singular values. Fig. 3 shows an image of a typical matrix, $C_{N,M}(\theta)$ 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 $C_{N,M}(\theta)$:

$$C_{N,M}^{(r)}(\theta) \approx \sum_{i=1}^{r} \sigma_i u_i v_i^T,$$

(42)

where $\sigma_i, v_i,$ and $u_i$ are the $i$th singular value, $i$th right-singular vector, and $i$th left-singular vector of the matrix $C_{N,M}(\theta)$, respectively. The parameter $r$ may be chosen according to

$$||C_{N,M}(\theta) - C_{N,M}^{(r)}(\theta)||^2 \leq \epsilon,$$

(43)

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

$$P_{C_{N,M}}^{(r)}(\theta) \approx \sum_{i=1}^{r} \frac{1}{\sigma_i} v_i u_i^T.$$

(44)

Note that by using (44) and setting an appropriate value for $\epsilon$, 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.

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: $\theta_{DP} = \pi/2$, $\theta_{WC} = \pi$, $\theta_{HC} = 2\pi/3$, and $\theta_{SC} = 3\pi/4$. 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 $L_b = 16$ taps and the sampling frequency to be $f_s = 8000$ Hz. The parameter $\epsilon$ 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:

$$B_1(\theta) = (1 - a_{1,1}) + a_{1,1} \cos \theta,$$

(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 $W(\mathbf{h}), D(\mathbf{h}),$ 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
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:

\[
E_{\text{norm}}(\theta) = \frac{E_{\text{out}}(\theta)}{E_i(\theta)},
\]

where \( E_i(\theta) \) is the energy of a signal arrived from direction \( \theta \) measured at the reference sensor (dark blue line), and \( E_{\text{out}}(\theta) \) is the energy of the output array signal due to the input signal from direction \( \theta \) (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 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,

<table>
<thead>
<tr>
<th>Pattern</th>
<th>( W(h) ) [dB]</th>
<th>( D(h) ) [dB]</th>
<th>( E_{BP} ) [dB]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dipole</td>
<td>-14.4</td>
<td>4.90</td>
<td>0.040</td>
</tr>
<tr>
<td>Cardioid</td>
<td>-8.68</td>
<td>4.75</td>
<td>0.007</td>
</tr>
<tr>
<td>Hypercardioid</td>
<td>-11.00</td>
<td>5.84</td>
<td>0.010</td>
</tr>
<tr>
<td>Supercardioid</td>
<td>-10.00</td>
<td>5.42</td>
<td>0.007</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>System</th>
<th>Endfire signal</th>
<th>Null signal</th>
<th>Arbitrary signal</th>
</tr>
</thead>
<tbody>
<tr>
<td>( E_{\text{norm}}[\text{dB}] )</td>
<td>-0.009</td>
<td>-46.810</td>
<td>-6.569</td>
</tr>
</tbody>
</table>

Table 2 Normalized energy of the output signals for the case of a first-order supercardioid DMA.
the delay and sum (DAS) beamformer obtains the largest WNG for a given number of sensors which is [48]

\[ W_{\text{DAS}} = M. \] (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 iSNR = 2 (black stars), \(\sigma^2 = 0.25\) and iSNR = 2 (red diamonds), \(\sigma^2 = 0.05\) and iSNR = 4 (magenta squares), and \(\sigma^2 = 0.25\) and 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, \(d\).

### 6.1.1. Minimal required filter length

We examine the issue of the minimal required length of the filter vector, \(D_{M\text{PMV}}\), 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 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, \(d\), by simulation results. Fig. 11 shows the beampattern of a first-order hypercardioid with element spacing \(d = 15\) 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 \(d\), 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 \(d = 15\) cm (top) and for the case of \(d = 1\) cm (bottom) for a white signal impinging on the array from direction of \(\theta = 145^\circ\). While for the case of a small value of \(d\), the entire output spectrum is flat implying a frequency-invariant beampattern, it is not the case when \(d\) 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
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 \( \theta_{D} = \pi/2 \) and a second distortionless constraint at \( \theta_{D} = \pi \). The three other shapes have two distinct nulls in the following directions: \( \theta_{C} = \pi/2, \theta_{O} = \pi, \theta_{H} = 0.8005\pi, \theta_{C} = 0.3997\pi, \) and \( \theta_{S} = 0.8493\pi, \theta_{s} = 0.5903\pi \). We choose \( L_0 = 20 \) 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 \( W(h), \), \( D(h) \), and \( E_{BP} \). Using (24), it is easy to show that the theoretical beampatterns of second-order DMAs have the following form:

\[
B_2(\theta) = (1 - a_{2,1} - a_{2,2}) + a_{2,1} \cos \theta + a_{2,2} \cos^2 \theta,
\]

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

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: \( \theta_{case1,1} = \frac{\pi}{3}, \theta_{case1,2} = \frac{2\pi}{3}, \theta_{case1,3} = \pi, \) and \( \theta_{case2,1} = \frac{\pi}{3}, \theta_{case2,2} = \frac{2\pi}{3}, \theta_{case2,3} = \pi. \) We choose \( L_0 = 24 \) 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 \( W(h), D(h), \) and \( E_{BP} \), where the last measure is calculated using the following theoretical beampattern for third-order DMAs:

![Image](image-url)

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.)

![Image](image-url)

Fig. 11. Beampattern of a first-order hypercardoid (\( M = 2 \)) for the case of \( \delta = 15 \) cm.
\[ B_3(\theta) = (1 - a_{3,1} - a_{3,2} - a_{3,3}) + a_{3,1} \cos \theta + a_{3,2} \cos^2 \theta + a_{3,3} \cos^3 \theta, \]

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.

### 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 \(C_M(\theta)\) by itself, its inverse and another multiplication by the transpose of \(C_M(\theta)\). These operations require approximately \(3M^2L^3\) 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.

---

**Table 3**

<table>
<thead>
<tr>
<th>Pattern</th>
<th>(\Psi(h)) [dB]</th>
<th>(\Delta(h)) [dB]</th>
<th>(E_{BP}) [dB]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dipole</td>
<td>-29.18</td>
<td>6.70</td>
<td>0.03</td>
</tr>
<tr>
<td>Cardioid</td>
<td>-25.48</td>
<td>8.60</td>
<td>0.02</td>
</tr>
<tr>
<td>Hypercardioid</td>
<td>-29.30</td>
<td>9.36</td>
<td>0.01</td>
</tr>
<tr>
<td>Supercardioid</td>
<td>-23.91</td>
<td>8.01</td>
<td>0.03</td>
</tr>
</tbody>
</table>

**Table 4**

<table>
<thead>
<tr>
<th>Pattern</th>
<th>(\Psi(h)) [dB]</th>
<th>(\Delta(h)) [dB]</th>
<th>(E_{BP}) [dB]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>-29.32</td>
<td>10.16</td>
<td>0.007</td>
</tr>
<tr>
<td>Case 2</td>
<td>-36.78</td>
<td>11.78</td>
<td>0.030</td>
</tr>
</tbody>
</table>

---

**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.

**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).
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 \( \mathbf{G}(\theta) \) with its transpose:

\[
\mathbf{G}(\theta)\mathbf{G}^\top(\theta) = \begin{bmatrix}
\mathbf{G}_1(\theta)\mathbf{G}_1^\top(\theta) & \cdots & \mathbf{G}_i(\theta)\mathbf{G}_i^\top(\theta) \\
\vdots & \ddots & \vdots \\
\mathbf{G}_L(\theta)\mathbf{G}_L^\top(\theta) & \cdots & \mathbf{G}_L(\theta)\mathbf{G}_L^\top(\theta)
\end{bmatrix},
\]

(50)

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

\[
\mathbf{G}_i(\theta) = \begin{bmatrix}
\mathbf{t}^\top(l - \tau_i(0) f_s) \\
\mathbf{t}^\top(l - 1 - \tau_i(0) f_s) \\
\vdots \\
\mathbf{t}^\top(l - L_h + 1 - \tau_i(0) f_s)
\end{bmatrix},
\]

(51)

where \( \mathbf{t}(l - \tau_i(0) f_s) = \text{sinc}(l - \tau_i(0)) \), \( l = 1, \ldots, L \) is a column vector. From now on, we omit the dependency of \( \tau_i(0) \) on \( \theta \) for notation simplicity. In order to simplify (50), we may exploit the following property of the scalar product between two time-shifted sinc functions:

\[
\mathbf{t}^\top(\tau_1)\mathbf{t}(\tau_2) \approx \begin{cases} 1, & |\tau_1 - \tau_2| < 1 \\ 0, & |\tau_1 - \tau_2| > 1 \end{cases},
\]

(52)

where

\[
\tau_{t_1, t_2} = \text{sinc}^2\left(\frac{\tau_1 + \tau_2}{2} f_s\right)
\]

(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

\[
\mathbf{G}_i(\theta)\mathbf{G}_i^\top(\theta) \approx \text{sinc}^2\left(\frac{\tau_1 + \tau_2}{2} f_s\right)\mathbf{I},
\]

(54)

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

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

\[
\mathbf{h}^\top\mathbf{G}(\theta)\mathbf{G}^\top(\theta)\mathbf{h} \approx \sum_{i=1}^{M} \sum_{j=1}^{M} \text{sinc}^2\left(\frac{i+j}{2} f_s\right) \mathbf{h}_i^\top\mathbf{h}_j,
\]

(55)

where \( \mu_{ij} = i + j - 2 \). According to the DMA assumption \[23\], \( \frac{2\pi f_s}{\lambda} \ll 2\pi \). Therefore, \( \frac{2\pi f_s}{\lambda} \approx \frac{2\pi f_{\text{max}}}{\lambda} \ll 1 \), where \( f_{\text{max}} \) is the highest frequency bin of the signal. Let us assign \( x_{0,i} = \mu_{ij} f_s \cos \theta \) and approximate \( f(x_{0,i}) = \text{sinc}^2(x_{0,i}) \) using Taylor series around the point \( x_0 \):

\[
f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n.
\]

(56)

We can calculate (56) by using the well-known Taylor series

\[
\text{sinc}^2(x_{0,i}) \approx \sum_{n=0}^{2N} \delta f_{x_{0,i}}^n (\cos \theta - 1)^n.
\]

(57)
where we choose \( x_0 \) as the corresponding to the case of \( \theta = 0 \) 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:

\[
b^h \Theta(\theta) = \sum_{n=0}^{2N} b_{n} \cos n \theta.
\]

(58)

where \( \Theta \) is the parameters vector containing all the above parameters of \( \Delta, C, b, \) etc.

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

\[
B^2_n(\theta) = \left( \sum_{n=0}^{N} a_{n} \cos \theta \right)^2 = \sum_{n=0}^{2N} b_{n} \cos n \theta.
\]

(59)

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

Q.E.D.

References


