Efficient Implementations of First-Order Steerable Differential Microphone Arrays with Arbitrary Planar Geometry

Federico Borra, Alberto Bernardini, *Member, IEEE,*
Fabio Antonacci, *Member, IEEE,* and Augusto Sarti, *Senior Member, IEEE*

***Abstract*—We present a spatial filtering approach to first-order steerable Differential Microphone Arrays (DMAs) with arbitrary planar geometry. In particular, the design of the spatial filter is based on a recently proposed frequency-domain design methodology that approximates, in a least-square sense, a target beampattern using the Jacobi-Anger expansion involving Bessel functions. Despite the generality of that approach, however, its computational cost turns out to be excessive when working with limited processing resources. The beamforming technique proposed in this manuscript overcomes this issue by exploiting the fact that in DMAs the spacing between sensors is typically smaller than the smallest wavelength of audio signals of interest. This allows us to substitute zero- and first-order Bessel functions with their Taylor series approximation truncated to the first order. Moreover, we show that this approximation allows us to derive an efficient discrete-time-domain implementation of first-order steerable differential beamformers based on arrays with arbitrary geometries.**

***Index Terms*—Beamforming, Differential Microphones**

I. Introduction

Microphone array signal processing is a constantly growing field of research for both academia and industry [1]–[6]. Spatial and temporal information gathered by microphone arrays can be exploited for a wide variety of applications that range from acoustic source localization [7]–[9], to speech enhancement [10]–[13], and source separation [12]–[16]. Most applications of the sort involve spatial filtering (beamforming) techniques, which allow us to steer beams towards prescribed directions. The simplest and most widespread beamforming method is Delay-And-Sum (DAS) [1], [2], as it operates by individually delaying, weighing; and then finally adding together the sensor signals in order to shape a beampattern that points towards a prescribed direction. The main issue with DAS beamformers is in the fact that their spatial response varies with frequency. In fact, the width of the beampattern's mainlobe increases as the frequency decreases.

One of the solutions adopted in the literature in order to overcome this problem is the adoption of differential microphone arrays (DMAs) [3]–[5], [17]–[29]. Thanks to their small and compact aperture, DMAs exhibit a nearly frequency-invariant beampattern. The literature is rich with DMA methods based on different array geometries. The most widespread ones are the uniform linear DMAs [3], [19], [22], [30] and the uniform circular DMAs [4], [25]. Uniform linear DMAs exhibit limited steering performance as their beampattern strongly depends on the steering direction; in fact, the design of beams with a single mainlobe is possible only when they are steered towards the end-fire directions. Moreover, they suffer from the well-known problem of front-back ambiguity, which often limits their applicability in practical scenarios. Nonuniform linear DMAs have also been proposed in [31], showing that they can significantly improve the robustness against white noise with respect to uniform linear DMAs, especially at low frequencies. Uniform circular DMAs, on the other hand, exhibit much better steering capabilities. The design of circular DMA has been widely addressed in the literature (see [4], [18], [25], [32], [33] and references therein). In particular, the method proposed in [18], [32] enables a continuous steering of first-order beams by linearly combining eigenbeams computed from the microphone signals. In [25], starting from the theory of circular harmonics beamformers [34], [35], the authors propose an optimal method, in a least-squares error sense, for approximating a target beampattern, steered in any direction, using the Jacobi-Anger expansion.

Linear and circular DMAs offer promising features but they have hard constraints on their geometry. In some practical scenarios (e.g., small-size embedded systems), this can be a problem, as it might not be possible to place regularly spaced sensors. In addition, fabrication errors can add uncertainty to sensor positioning. All this raises the issue of how to develop beamforming techniques for DMAs with arbitrary geometry. An approach to the design of DMAs with arbitrary geometry was addressed in [26] as an extension of the results presented in [25]. In [26], for example, using a Jacobi-Anger expansion for the approximation of the beampattern, an algorithm is proposed that can form steerable beampatterns with a microphone array of any planar geometry. Despite the generality of this method, however, a significant computational effort is required. Like most DMA beamformers, in fact, its implementation is based on Short-Time Fourier Transform (STFT) filtering. In addition, STFT filtering involves block processing that can make frequency-domain beamformers unsuitable for audio applications that require small delays (e.g., real-time audio communication). In order to overcome this issue, time-domain beamformers have been discussed in the literature (see [30], [36]–[38] and references therein). In [30], [36], [37], in particular, time-domain differential beamformers with linear and circular geometries are proposed. The design of such beamformers is performed considering a time-domain

Manuscript received ...; revised ...

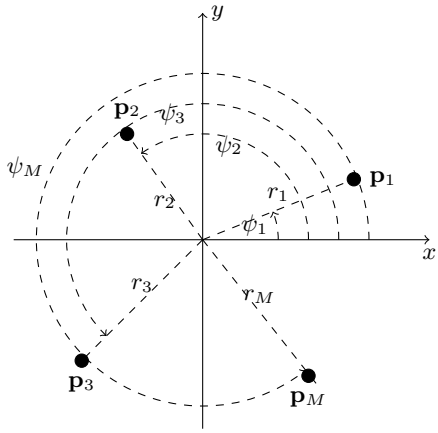


Fig. 1: Illustration of an arbitrary array geometry.

acoustic model and it results in a set of FIR spatial filters (one per sensor) whose derivation involves the solution of a multidimensional linear system of equations; the length of the FIR spatial filter affects the performance of the systems and their computational cost. The author of [38], instead, proposes a robust time-domain modal beamformer for circular arrays with two processing stages: in the former, sensor data are converted to the circular harmonics domain using a real-valued discrete circular Fourier transform and then steered toward a look-direction; the latter is based on a set of FIR spatial filters (one per sensor).

In this manuscript, we develop a computationally inexpensive data-independent differential beamformer for arrays with arbitrary planar geometry, which can be implemented both in the frequency domain and in the time domain. Starting from [26], under the assumption that spacing between sensors is smaller than the smallest wavelength of audio signals of interest, we obtain a simple spatial filtering method that allows us to efficiently steer first-order beams in any direction. We show that the performance of the proposed method is comparable with that of [26] in terms of white noise gain (WNG), directivity factor (DF) and beampattern while the computational cost is significantly reduced.

The rest of the manuscript is organized as follows. Section II presents the signal model and the definition of beampattern. Section III reviews the background on spatial filter design based on the Jacobi-Anger expansion. Section IV and Section V discuss the proposed beamforming method for planar arrays with arbitrary geometry in the frequency-domain and in the time-domain, respectively. Section VI presents some simulation results and comparisons with respect to [26] in terms of WNG, DF and beampattern. Section VII shows a frequency-domain implementation and a time-domain implementation of the proposed system and provides a discussion about the required computational cost. Finally, Section VIII draws conclusions on the proposed solution.

II. Signal Model and Definitions

Let us consider an array of M microphones arbitrarily placed on a two-dimensional plane. Assuming that the center of mass of the array coincides with the origin of the Cartesian coordinate system, we can express the position of the mth microphone as follows

$$\mathbf{p}_m = r_m \left[\cos(\psi_m), \sin(\psi_m)\right], \tag{1}$$

where r_m is the distance from the origin of the axes and ψ_m is the counterclockwise angle with respect to the positive x-axis. A general illustration of the array geometry is shown in Fig. 1. Given a source in the far field, we can express the signal captured by the mth microphone as [25], [26]

$$Y_m(\omega) = D_m(\omega, \theta_s)X(\omega) + V_m(\omega), \tag{2}$$

where $\omega = 2\pi f$ is the angular frequency, f is the temporal frequency, $X(\omega)$ is the source signal, $V_m(\omega)$ models an additive noise and $D_m(\omega, \theta_s)$ models the propagation from the source to the mth microphone with θ_s the direction of arrival of the source. Under the far field assumption, $D_m(\omega, \theta_s)$ can be expressed as [26]

$$D_m(\omega, \theta_s) = e^{j\frac{\omega r_m}{c}\cos(\theta_s - \psi_m)}, \tag{3}$$

where c is the speed of sound and j is the imaginary unit. In vectorial form, (2) can be written as

$$\mathbf{y}(\omega) = \mathbf{d}(\omega, \theta_s)X(\omega) + \mathbf{v}(\omega), \tag{4}$$

where the vectors are defined as

$$\begin{aligned} \mathbf{y}(\omega) &= \left[Y_1(\omega), \ldots, Y_M(\omega)\right]^T \\ \mathbf{d}(\omega, \theta_s) &= \left[D_1(\omega, \theta_s), \ldots, D_M(\omega, \theta_s)\right]^T \\ \mathbf{v}(\omega) &= \left[V_1(\omega), \ldots, V_M(\omega)\right]^T. \end{aligned} \tag{5}$$

In order to perform spatial filtering, microphone signals $Y_m(\omega)$ are combined as follows

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

where H_m^* is the filter coefficient of the mth microphone; the asterisk denotes conjugation; $Z(\omega)$ represents the output of the beamformer; the superscript ^{H} denotes the Hermitian operator; and $\mathbf{h}(\omega) = \left[H_1(\omega), \ldots, H_M(\omega)\right]^T$.

The sensitivity of a beamforming filter $\mathbf{h}(\omega)$ to a plane wave impinging on the array with an angle θ is known as *beampattern* and it is mathematically expressed as

$$\mathcal{B}\left[\mathbf{h}(\omega), \theta\right] = \mathbf{h}^H(\omega)\mathbf{d}(\omega, \theta) = \sum_{m=1}^{M} H_m^*(\omega)e^{j\frac{\omega r_m}{c}\cos(\theta - \psi_m)}. \tag{7}$$

Dealing with data-independent beamformers, when the filter $\mathbf{h}(\omega)$ is designed to steer a beam in the direction θ_s, the function $\left|\mathcal{B}\left[\mathbf{h}(\omega), \theta\right]\right|^2$ exhibits a maximum when evaluated for $\theta = \theta_s$.

III. Background on Beamformer Design

In [26] the filter $\mathbf{h}(\omega)$ is designed in such a way that the resulting beampattern matches a target symmetric frequency-

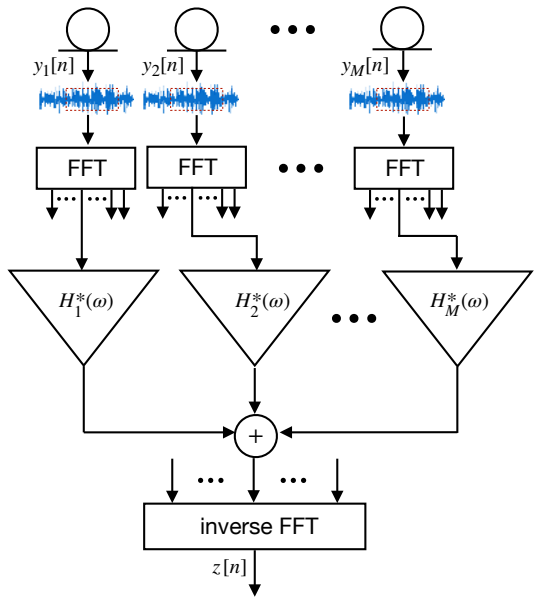


Fig. 2: STFT-based implementation scheme of the beamforming method proposed in [26].

invariant beampattern steered towards θ_s. This target beampattern is given by

$$\begin{aligned}\bar{\mathcal{B}}(\mathbf{b},\theta,\theta_s) &= \sum_{n=0}^{N} a_n \cos(n(\theta-\theta_s)) \\ &= \frac{1}{2}\sum_{n=0}^{N} a_n \left(e^{jn(\theta-\theta_s)} + e^{-jn(\theta-\theta_s)}\right) \\ &= \sum_{n=-N}^{N} b_n e^{jn(\theta-\theta_s)} = \mathbf{b}^T \boldsymbol{\Upsilon}(\theta_s)\mathbf{p}(\theta)\end{aligned} \tag{8}$$

where N is the order of the beampattern,

$$\begin{cases} b_0 = a_0 \\ b_i = b_{-i} = \frac{1}{2}a_i \quad i = 1,2,\ldots,N \end{cases} \tag{9}$$

and

$$\begin{aligned}\mathbf{b} &= [b_{-N},\ldots,b_0,\ldots,b_N]^T \\ \mathbf{p}(\theta) &= \left[e^{-jN\theta},\ldots,1,\ldots,e^{jN\theta}\right]^T \\ \boldsymbol{\Upsilon}(\theta_s) &= \mathrm{diag}\left(e^{jN\theta_s},\ldots,1,\ldots,e^{-jN\theta_s}\right).\end{aligned} \tag{10}$$

The coefficients in the vector $\mathbf{b}$ determine the shape of the beampattern and they can be designed following different optimization criteria [20]. In order to express the beampattern in (7) in a form that is similar to that of the target beampattern in (8), we can use the Jacobi-Anger expansion (or circular harmonics expansion) [4], [39]

$$e^{j\frac{\omega r_m}{c}\cos(\theta-\psi_m)} = \sum_{n=-\infty}^{+\infty} j^n J_n\left(\frac{\omega}{c}r_m\right) e^{jn(\theta-\psi_m)}, \tag{11}$$

where $J_n(\cdot)$ is the nth order Bessel function of the first kind. By truncating the Jacobi-Anger expansion to the Nth order and by substituting (11) in (7), we obtain the following beampattern expression [4], [39]

$$\begin{aligned}\mathcal{B}[\mathbf{h}(\omega),\theta] &= \sum_{m=1}^{M} H_m^*(\omega) \sum_{n=-N}^{N} j^n J_n\left(\frac{\omega}{c}r_m\right) e^{jn(\theta-\psi_m)} \\ &= \sum_{n=-N}^{N} j^n e^{jn\theta} \sum_{m=1}^{M} H_m^*(\omega) J_n\left(\frac{\omega}{c}r_m\right) e^{-jn\psi_m}.\end{aligned} \tag{12}$$

After equating (12) with the target beampattern $\bar{\mathcal{B}}(\mathbf{b},\theta,\theta_s)$ in (8), we obtain

$$\begin{aligned}&\sum_{n=-N}^{N} j^n e^{jn\theta} \sum_{m=1}^{M} H_m^*(\omega) J_n\left(\frac{\omega}{c}r_m\right) e^{-jn\psi_m} \\ &\qquad = \sum_{n=-N}^{N} b_n e^{jn(\theta-\theta_s)}.\end{aligned} \tag{13}$$

By matching mode-by-mode the left-hand side and the right-hand side of (13), we obtain

$$\sum_{m=1}^{M} H_m^*(\omega) j^n J_n\left(\frac{\omega}{c}r_m\right) e^{-jn\psi_m} = b_n e^{-jn\theta_s}. \tag{14}$$

For each $n \in \{-N,\ldots,N\}$ the expression in (14) can be written in matrix form as follows

$$\boldsymbol{\Psi}(\omega)\mathbf{h}(\omega) = \boldsymbol{\Upsilon}^*(\theta_s)\mathbf{b}, \tag{15}$$

where

$$\boldsymbol{\Psi}(\omega) = \begin{bmatrix} (-j)^{-N}\boldsymbol{\psi}_{-N}^H(\omega) \\ \vdots \\ \boldsymbol{\psi}_0^H(\omega) \\ \vdots \\ (-j)^{N}\boldsymbol{\psi}_{N}^H(\omega) \end{bmatrix} \tag{16}$$

is a $(2N+1)\times M$ matrix and

$$\boldsymbol{\psi}_n(\omega) = \left[J_n\left(\frac{\omega r_1}{c}\right)e^{-jn\psi_1},\ldots,J_n\left(\frac{\omega r_M}{c}\right)e^{-jn\psi_M}\right]^T. \tag{17}$$

When $M \geq 2N+1$, the minimum norm solution of (15) can be computed as follows

$$\mathbf{h}(\omega) = \boldsymbol{\Psi}^H(\omega)\left[\boldsymbol{\Psi}(\omega)\boldsymbol{\Psi}^H(\omega)\right]^{-1}\boldsymbol{\Upsilon}^*(\theta_s)\mathbf{b}. \tag{18}$$

The scheme in Fig. 2 shows the implementation of the described beamformer, based on STFT filtering. For each time frame at each sensor a Fast Fourier Transform (FFT) is performed and K multiplications with the filter coefficients are computed, where $2K$ is the FFT length. Lastly, all the filtered signals of the M microphones are summed at each frequency bin and one inverse FFT provides the beamformer output for the considered time frame.

Notice that, in order to derive the filter coefficients $H_m(\omega)$, no particular assumptions have been made on the array geometry. In the following sections, we will show that, assuming the distances between sensors to be negligible with respect to the wavelength of the audio signals of interest (usually true with small-sized arrays used for differential beamforming), the described spatial filtering process in Fig. 2 can be simplified

in such a way that the computational cost will be significantly reduced.

IV. Beamformer Design in the Frequency Domain

In the rest of the manuscript we consider first-order differential microphones ($N = 1$). As DMAs typically use closely spaced omnidirectional sensors, we can assume that [18]

$$\lambda = \frac{2\pi c}{\omega} \gg r_m \Rightarrow \frac{\omega r_m}{c} \to 0 \quad \forall m = 1, \ldots, M, \tag{19}$$

where λ is the wavelength. It is worth noticing that under condition (19) the spatial aliasing problem, which is known for creating undesired grating lobes, is greatly mitigated as discussed in [4], [5], [40].

If (19) is true, zero-order and first-order Bessel functions are well approximated by their Taylor series expansion truncated at the first order [41]

$$\begin{aligned} J_0(x) &\approx 1 + \mathcal{O}\left(x^2\right) \\ J_1(x) &\approx \frac{x}{2} + \mathcal{O}\left(x^3\right) \quad , \end{aligned} \tag{20}$$

where $x = \omega r_m / c$ is close to zero. According to (20), $\mathbf{\Psi}(\omega)$ in eq. (16) can be approximated as

$$\hat{\mathbf{\Psi}}(\omega) = \begin{bmatrix} j\hat{\boldsymbol{\psi}}_{-1}^H(\omega) \\ \hat{\boldsymbol{\psi}}_0^H(\omega) \\ -j\hat{\boldsymbol{\psi}}_1^H(\omega) \end{bmatrix} \tag{21}$$

where

$$\begin{aligned} \hat{\boldsymbol{\psi}}_1(\omega) &= \left[\frac{\omega r_1}{2c} e^{-j\psi_1}, \frac{\omega r_2}{2c} e^{-j\psi_2}, \ldots, \frac{\omega r_M}{2c} e^{-j\psi_M}\right]^T \\ &= \frac{\omega}{2c}\left[r_1 e^{-j\psi_1}, r_2 e^{-j\psi_2}, \ldots, r_M e^{-j\psi_M}\right]^T \\ \hat{\boldsymbol{\psi}}_{-1}(\omega) &= -\hat{\boldsymbol{\psi}}_1^*(\omega) \\ \hat{\boldsymbol{\psi}}_0(\omega) &= [1, 1, \ldots, 1]^T . \end{aligned} \tag{22}$$

Given the equations in (22), $\hat{\mathbf{\Psi}}(\omega)$ in (21) can be written as

$$\hat{\mathbf{\Psi}}(\omega) = \mathbf{D}(\omega)\mathbf{\Phi} \tag{23}$$

where

$$\mathbf{D}(\omega) = \operatorname{diag}\left(-\frac{j\omega}{2c}, 1, -\frac{j\omega}{2c}\right) \quad ,$$

$$\mathbf{\Phi} = \begin{bmatrix} r_1 e^{-j\psi_1} & r_2 e^{-j\psi_2} & \ldots & r_M e^{-j\psi_M} \\ 1 & 1 & \ldots & 1 \\ r_1 e^{j\psi_1} & r_2 e^{j\psi_2} & \ldots & r_M e^{j\psi_M} \end{bmatrix}. \tag{24}$$

In equation (23) $\hat{\mathbf{\Psi}}(\omega)$ is expressed as the product of a frequency-dependent diagonal matrix $\mathbf{D}(\omega)$ and a frequency-independent matrix $\mathbf{\Phi}$ that embeds all the information related to the geometry of the array. Substituting $\mathbf{\Psi}(\omega)$ with $\hat{\mathbf{\Psi}}(\omega)$ in (18), and omitting the argument ω for the sake of readability, we get the following approximation of the filter vector $\mathbf{h}(\omega)$

$$\begin{aligned} \hat{\mathbf{h}} &= \hat{\mathbf{\Psi}}^H \left(\hat{\mathbf{\Psi}}\hat{\mathbf{\Psi}}^H\right)^{-1} \mathbf{\Upsilon}^*(\theta_s)\mathbf{b} \\ &= \mathbf{\Phi}^H \mathbf{D}^H \left(\mathbf{D}\mathbf{\Phi}\mathbf{\Phi}^H\mathbf{D}^H\right)^{-1} \mathbf{\Upsilon}^*(\theta_s)\mathbf{b} \\ &= \mathbf{\Phi}^H \mathbf{D}^H \left(\mathbf{D}^H\right)^{-1} \left(\mathbf{\Phi}\mathbf{\Phi}^H\right)^{-1} \mathbf{D}^{-1}\mathbf{\Upsilon}^*(\theta_s)\mathbf{b} \\ &= \mathbf{\Phi}^H \left(\mathbf{\Phi}\mathbf{\Phi}^H\right)^{-1} \mathbf{D}^{-1}\mathbf{\Upsilon}^*(\theta_s)\mathbf{b} \\ &= \mathbf{P}\mathbf{D}^{-1}\mathbf{\Upsilon}^*(\theta_s)\mathbf{b}, \end{aligned} \tag{25}$$

where $\mathbf{P} = \mathbf{\Phi}^H \left(\mathbf{\Phi}\mathbf{\Phi}^H\right)^{-1}$ with

$$\mathbf{\Phi}\mathbf{\Phi}^H = \sum_{m=1}^{M} \begin{bmatrix} r_m^2 & r_m e^{-j\psi_m} & \left(r_m e^{-j\psi_m}\right)^2 \\ r_m e^{j\psi_m} & 1 & r_m e^{-j\psi_m} \\ \left(r_m e^{j\psi_m}\right)^2 & r_m e^{j\psi_m} & r_m^2 \end{bmatrix}. \tag{26}$$

Notice that $\mathbf{\Phi}\mathbf{\Phi}^H$ is a centro-Hermitian matrix. It is known that the inverse of a centro-Hermitian matrix is centro-Hermitian [42], therefore $\left(\mathbf{\Phi}\mathbf{\Phi}^H\right)^{-1}$ is centro-Hermitian. The centro-Hermitian property of $\left(\mathbf{\Phi}\mathbf{\Phi}^H\right)^{-1}$ and eq. (24), allow us to rewrite $\mathbf{P}$ as

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & p_{11}^* \\ p_{21} & p_{22} & p_{21}^* \\ \vdots & \vdots & \vdots \\ p_{M1} & p_{M2} & p_{M1}^* \end{bmatrix}, \tag{27}$$

where $\mathbf{P} \in \mathbb{C}^{M\times 3}$, with $p_{m2} \in \mathbb{R} \quad \forall m \in \{1, \ldots, M\}$.

Applying the Hermitian operator to (25) we obtain

$$\hat{\mathbf{h}}^H(\omega) = \mathbf{b}^H \mathbf{\Upsilon}(\theta_s) \left[(\mathbf{D}(\omega))^{-1}\right]^H \mathbf{P}^H , \tag{28}$$

therefore the mth element of $\hat{\mathbf{h}}^H(\omega)$ can be expressed as

$$\begin{aligned} \hat{H}_m^*(\omega) &= \left[b_{-1}\tfrac{2c}{j\omega}e^{j\theta_s} \quad b_0 \quad b_1\tfrac{2c}{j\omega}e^{-j\theta_s}\right] \begin{bmatrix} p_{m1}^* \\ p_{m2} \\ p_{m1} \end{bmatrix} \\ &= b_0 p_{m2} + \frac{2c}{j\omega}\left(b_{-1}p_{m1}^* e^{j\theta_s} + b_1 p_{m1} e^{-j\theta_s}\right). \end{aligned} \tag{29}$$

In order for the beampattern to be symmetric with respect to the steering angle θ_s, the following condition must be met [4]

$$b_{-1} = b_1. \tag{30}$$

Moreover, in order to prevent any amplitude modification of the signals coming from the direction θ_s, we need $\bar{\mathcal{B}}(\mathbf{b}, \theta_s, \theta_s) = 1$. This is accomplished with the condition [4]

$$b_0 + b_1 + b_{-1} = 1. \tag{31}$$

Given (30) and (31), we can adopt the following parametrization for the coefficients in $\mathbf{b}$

$$b_0 = 1 - q \ , \ b_1 = q/2 \ \text{ with } \ 0 \le q \le 1. \tag{32}$$

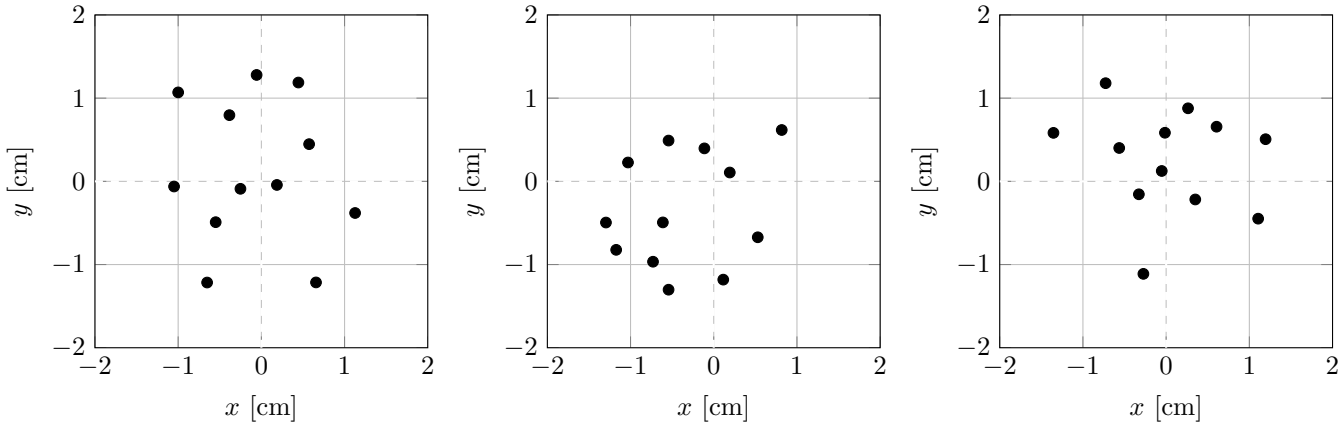
By inserting (32) in (29) we obtain

$$\hat{H}_m^*(\omega) = (1-q)p_{m2} + q2c\frac{1}{j\omega}\left(\Re\left\{p_{m1}^* e^{j\theta_s}\right\}\right), \tag{33}$$

where the operator $\Re\{\cdot\}$ returns the real part of a complex number. It is worth noticing that the only frequency-dependent term in (33) is the factor $1/(j\omega)$.

Substituting $H_m^*(\omega)$ with $\hat{H}_m^*(\omega)$ in (6) we obtain

$$\begin{aligned} \hat{Z}(\omega) &= \sum_{m=1}^{M} \hat{H}_m^*(\omega) Y_m(\omega) \\ &= (1-q)\sum_{m=1}^{M} p_{m2} Y_m(\omega) + \\ &\quad + q\frac{2c}{j\omega}\sum_{m=1}^{M} \Re\left\{p_{m1}^* e^{j\theta_s}\right\} Y_m(\omega) \ , \end{aligned} \tag{34}$$



(a) (b) (c)

Fig. 3: Examples of random array geometries obtained using (57).

where the term $1/(j\omega)$ was collected, as it is shared by all filter coefficients $\hat{H}_m^*(\omega)$. It follows that the output of the beamformer can be compactly expressed as

$$\hat{Z}(\omega) = (1-q)\hat{Z}_o(\omega) + q\hat{Z}_d(\omega). \tag{35}$$

where $\hat{Z}_o(\omega)$ is the omnidirectional component defined as

$$\hat{Z}_o(\omega) = \sum_{m=1}^{M} p_{m2} Y_m(\omega) \ , \tag{36}$$

and $\hat{Z}_d(\omega)$ is the steered dipole component defined as

$$\hat{Z}_d(\omega) = \frac{2c}{j\omega} Y_d(\omega), \tag{37}$$

with

$$Y_d(\omega) = \sum_{m=1}^{M} \Re\left\{p_{m1}^* e^{j\theta_s}\right\} Y_m(\omega) \ . \tag{38}$$

The Special Case of Uniform Circular Arrays

Though the filter definition in (33) is valid for DMAs of arbitrary geometry, in the following we show that its expression further simplifies when working with uniform circular arrays, which are widely used in a variety of spatial filtering applications [4]. In a uniform circular array we have

$$\begin{aligned} r_m &= r \quad \forall m \in \{1,\dots,M\} \\ \psi_m &= \frac{2\pi}{M}(m-1) \quad \forall m \in \{1,\dots,M\}. \end{aligned} \tag{39}$$

Given the conditions in (39), the matrix $\mathbf{\Phi}\mathbf{\Phi}^H$ in (26) becomes a diagonal matrix in the form

$$\mathbf{\Phi}\mathbf{\Phi}^H = M \mathrm{diag}\left(r^2, 1, r^2\right). \tag{40}$$

In fact, as shown in the Appendix A, the non-diagonal elements of $\mathbf{\Phi}\mathbf{\Phi}^H$ become zero. In this case, the mth row of matrix $\mathbf{P}$ becomes

$$[p_{m1}, p_{m2}, p_{m1}^*] = \frac{1}{M}\left[\frac{1}{r}e^{j\frac{2\pi}{M}(m-1)}, 1, \frac{1}{r}e^{-j\frac{2\pi}{M}(m-1)}\right] \tag{41}$$

As a consequence, the mth element of the filter vector $\hat{\mathbf{h}}^H(\omega)$ in (33) can be expressed analytically as

$$\begin{aligned} \hat{H}_m^*(\omega) &= \frac{1}{M}\left((1-q) + q\frac{2c}{j\omega}\Re\left\{\frac{e^{-j\frac{2\pi}{M}(m-1)+\theta_s}}{r}\right\}\right) \\ &= \frac{1}{M}\left((1-q) + q\frac{2c}{j\omega}\frac{\cos(\theta_s-\psi_m)}{r}\right) \ . \end{aligned} \tag{42}$$

It follows that the output of the beamformer can be expressed in the form

$$\begin{aligned} \hat{Z}(\omega) =& (1-q)\frac{1}{M}\sum_{m=1}^{M} Y_m(\omega) + \\ &+ \frac{q}{M}\frac{2c}{j\omega r}\sum_{m=1}^{M}\cos(\theta_s-\psi_m)Y_m(\omega) \\ =& (1-q)\frac{1}{M}\sum_{m=1}^{M} Y_m(\omega) + q\frac{2c}{j\omega}Y_d(\omega) \\ =& (1-q)\hat{Z}_o(\omega) + q\hat{Z}_d(\omega). \end{aligned} \tag{43}$$

V. Beamformer Design in the Time Domain

In order to derive a continuous-time expression of the output signal of the beamformer, we apply a inverse Fourier transform to (35), obtaining

$$\hat{z}(t) = (1-q)\hat{z}_o(t) + q\hat{z}_d(t), \tag{44}$$

where, according to (37), $\hat{z}_d(t)$ satisfies

$$y_d(t) = \frac{1}{2c}\frac{d\hat{z}_d(t)}{dt}. \tag{45}$$

Starting from (44) and (45) there are many possible discrete-time realizations of the beamformer, which depend on the discretization method (e.g., finite difference methods) used for approximating time derivatives in (45). For example, adopting the trapezoidal rule with a fixed sampling step (known as bilinear transform [43]), we obtain

$$\hat{z}[n] = (1-q)\hat{z}_o[n] + q\hat{z}_d[n] \tag{46}$$

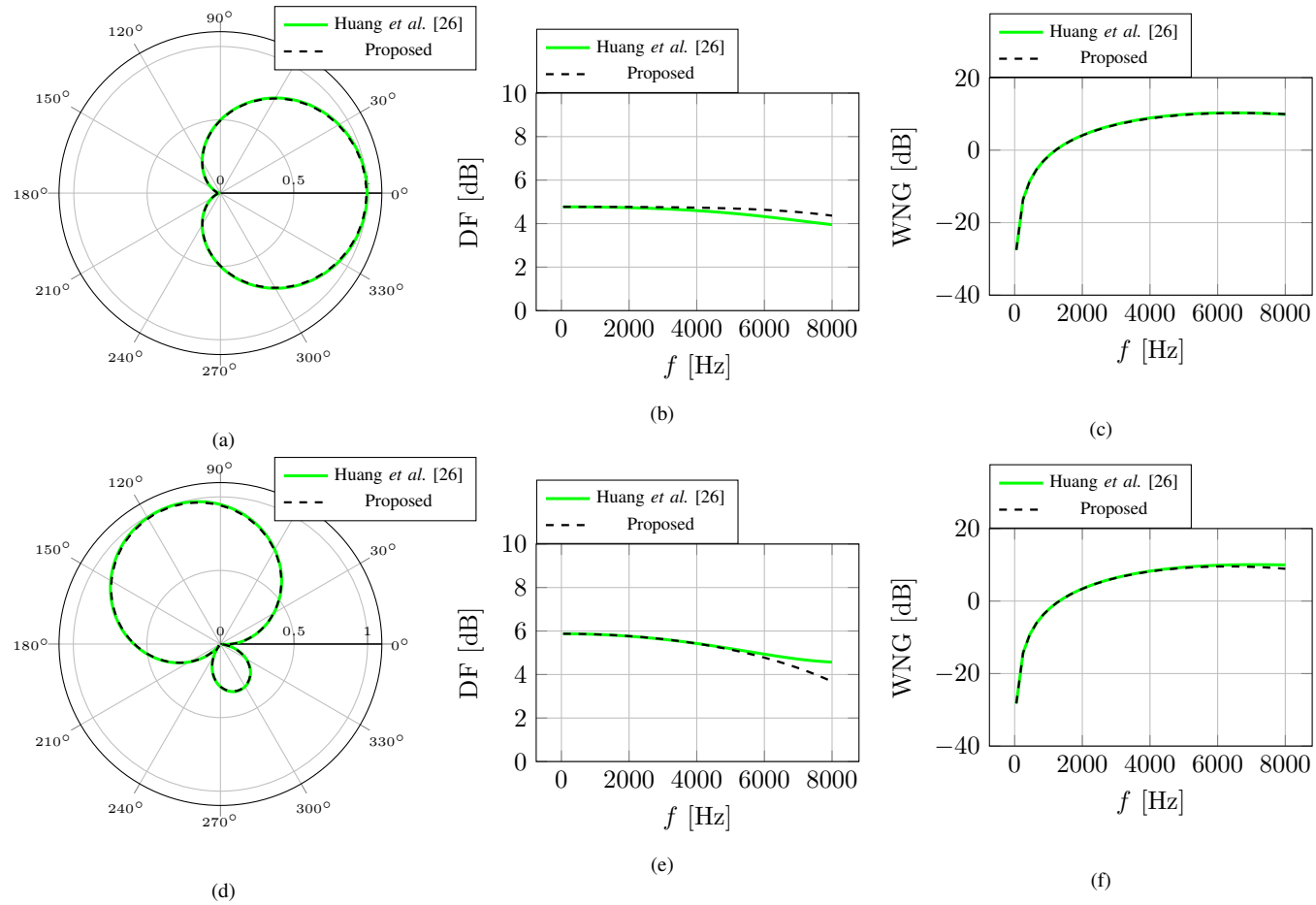


Fig. 4: Comparison, in terms of spatial response for $f = 1\,\text{kHz}$, for the DF and WNG, between the proposed method (dashed line) and [26] (solid line) for a first-order differential beamformer based on the geometry in Fig. 3a. (a), (b) and (c) refer to a first-order beamformer with $q = 0.5$ and steered toward direction $\theta_s = 0°$. (d), (e) and (f), instead, refer to a first-order beamformer with $q = 0.67$ and steered toward direction $\theta_s = 120°$.

where n is the sample index and

$$\begin{aligned} \hat{z}_o[n] &= \sum_{m=1}^{M} p_{m2} y_m[n] \\ \hat{z}_d[n] &= R_C \left(y_d[n] + y_d[n-1] \right) + \hat{z}_d[n-1], \end{aligned} \tag{47}$$

with

$$R_C = cT_s, \quad y_d[n] = \sum_{m=1}^{M} \Re \left\{ p_{m1}^* e^{j\theta_s} \right\} y_m[n]. \tag{48}$$

$T_s = 1/F_s$ is the sampling period defined as the inverse of the sampling frequency F_s. According to (46) the output of the system $\hat{z}[n]$ is computed as a linear combination of the omnidirectional component $\hat{z}_o[n]$ and the steered dipole component $\hat{z}_d[n]$.

A. The Special Case of Uniform Circular Arrays

As a special case of the general model in (46), (47) and (48), we can derive the discrete-time implementation of a beamformer based on a uniform circular array as

$$\hat{z}[n] = (1-q)\hat{z}_o[n] + q\hat{z}_d[n] \tag{49}$$

where

$$\begin{aligned} \hat{z}_o[n] &= \frac{1}{M} \sum_{m=1}^{M} y_m[n] \\ \hat{z}_d[n] &= R_C \left(y_d[n] + y_d[n-1] \right) + \hat{z}_d[n-1], \end{aligned} \tag{50}$$

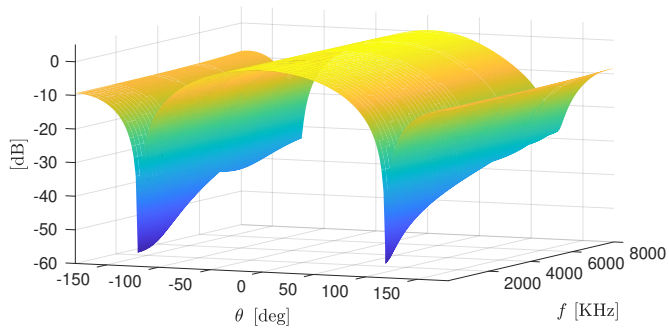
with

$$R_C = cT_s, \quad y_d[n] = \frac{1}{Mr} \sum_{m=1}^{M} \cos(\theta_s - \psi_m) y_m[n]. \tag{51}$$

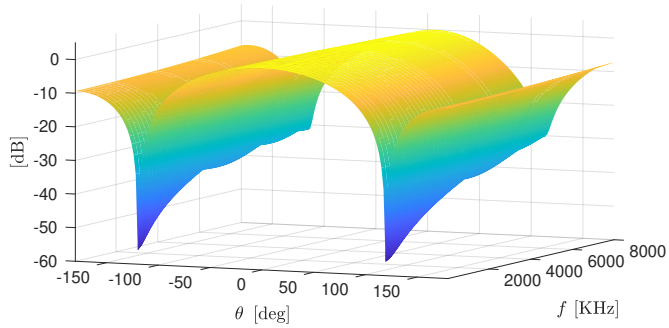
B. Frequency Analysis of the Discrete-Time Beamformer

The proposed discrete-time model is derived from the frequency domain model in the previous section by applying the inverse Fourier transform and approximating the time derivative with the bilinear transform [43]. Such an approximation results in a frequency warping, since points on the $j\omega$ axis of the complex s-plane of the Laplace transform are mapped to the unit circle in the z-plane of the Z transform, according to the substitution

$$j\omega \leftarrow \frac{2}{T_s} \frac{e^{j\tilde{\omega}T_s} - 1}{e^{j\tilde{\omega}T_s} + 1} \tag{52}$$



(a) Huang *et al.* [26]



(b) Proposed method

Fig. 5: Spatial-frequency beampattern plots of both Huang *et al.* [26] (a) and the proposed method (b) for the geometry in Fig. 3a with $q = 0.67$.

where $\widetilde{\omega}$ is here referred to as the *digital frequency*. It is possible to express in closed-form the reference "analog" frequency ω as a function of the digital frequency $\widetilde{\omega}$ by [43]

$$\omega = \frac{2}{T_s}\tan\left(\widetilde{\omega}\frac{T_s}{2}\right) \quad . \tag{53}$$

It follows that ω is really close to $\widetilde{\omega}$ at low frequencies, while they differ more and more at high frequencies. It is evident that the higher the sampling frequency $F_s = 1/T_s$, the more the difference between ω and $\widetilde{\omega}$ becomes negligible in the whole frequency range of interest. Eq. (53) allows us to perform a frequency analysis of the discrete-time realization of the spatial filter $\hat{\mathbf{h}}(\omega)$ as the sampling frequency F_s is changed. An analysis of the sort has already been performed in [44] where a discrete-time implementation of differential beamformers characterized by small-size uniform linear arrays is discussed.

VI. Simulations

In this section we study the behavior of the proposed beamforming method in terms of spatial response (beampattern), WNG and DF. WNG and DF both quantify the improvement in Signal-to-Noise Ratio (SNR), assuming that we are dealing with spatially white noise and spatially isotropic diffuse noise, respectively [4]. Given a spatial filter $\mathbf{h}(\omega)$, the WNG is mathematically defined as

$$\mathrm{WNG}\left[\mathbf{h}(\omega)\right] = \frac{|\mathbf{h}^H(\omega)\mathbf{d}(\omega,\theta_s)|^2}{\mathbf{h}^H(\omega)\mathbf{h}(\omega)} \quad , \tag{54}$$

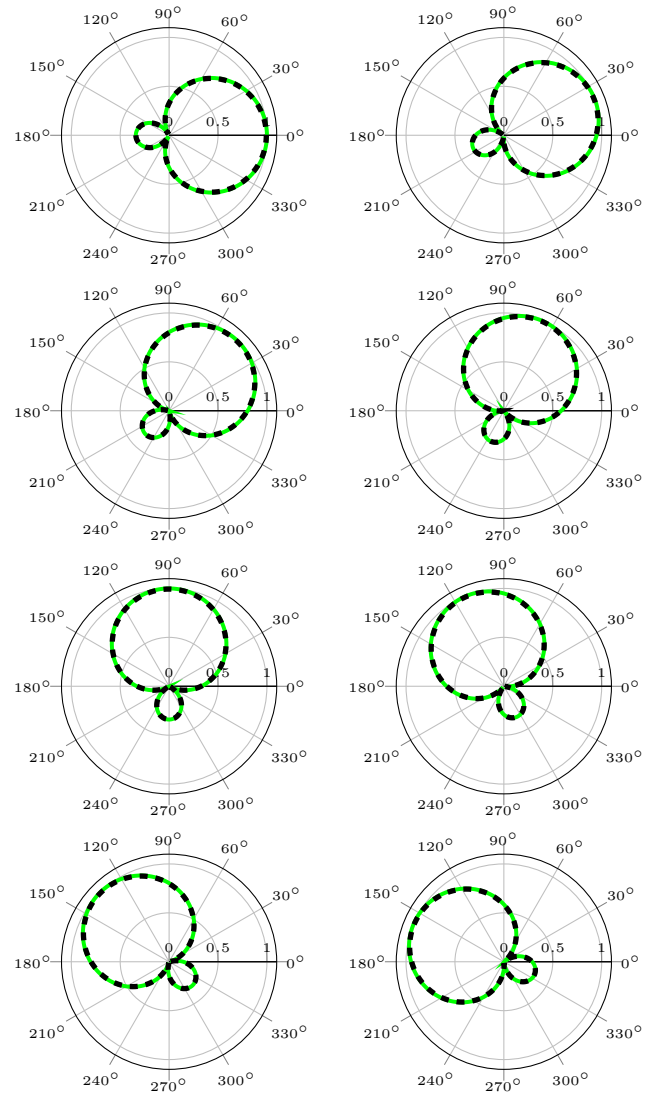


Fig. 6: Beampattern of both Huang *et al.* [26] (green solid line) and the proposed method (black dashed line) for different steering angle $\theta_s = \{0, 22.5°, 45°, 67.5°, 90°, 112.5°, 135°, 157,5°\}$ for $f = 1\,\mathrm{kHz}$.

while the DF as

$$\mathrm{DF}\left[\mathbf{h}(\omega)\right] = \frac{|\mathbf{h}^H(\omega)\mathbf{d}(\omega,\theta_s)|^2}{\mathbf{h}^H(\omega)\Gamma_{\mathrm{dn}}(\omega)\mathbf{h}(\omega)}, \tag{55}$$

where

$$\left[\Gamma_{\mathrm{dn}}(\omega)\right]_{ij} = \frac{\sin\left(\omega\delta_{ij}/c\right)}{\omega\delta_{ij}/c}, \tag{56}$$

with δ_{ij} the distance between the ith microphone and the jth microphone. A comparison with the beamforming method discussed in [26] by G. Huang *et al.* is also provided. The purpose of this comparison is to show that, despite the approximation introduced in (20), the behavior of the proposed first-order differential spatial filters is in line with that of [26]. However, in Section VII we will show that our method is less demanding than [26] in terms of computational load.

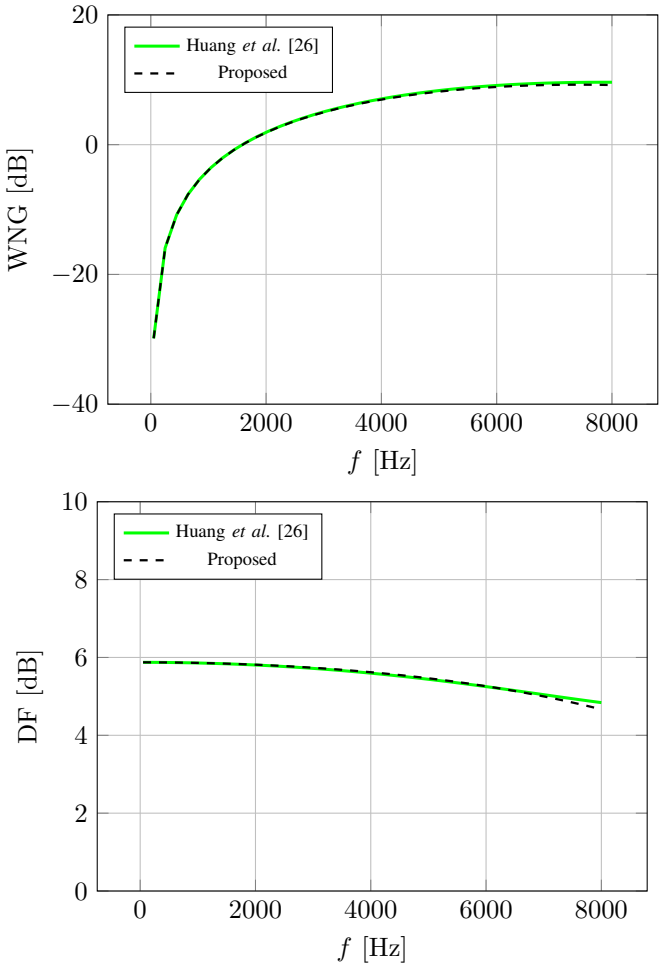


Fig. 7: Average WNG and average DF resulting from 300 realizations of beamformers characterized by $M = 12$ sensors and different array geometries obtained using (57).

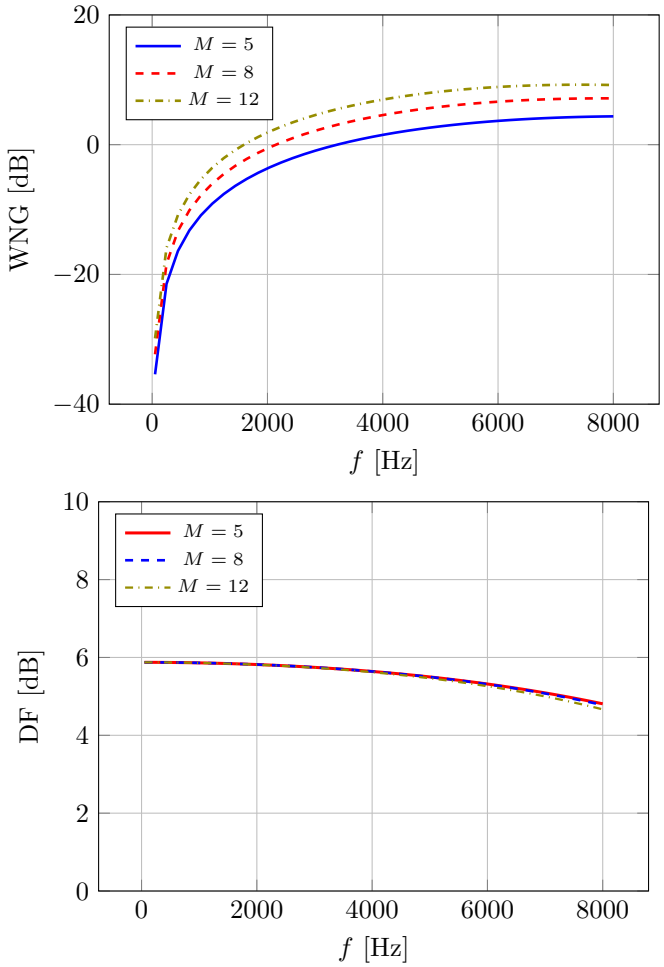


Fig. 8: Average WNG and average DF resulting from 300 realizations of beamformers with different array geometries obtained using (57). The number of sensors M is varied at each experiment.

In all the simulations presented in this section, the positions $\mathbf{p}_m$ of the M omnidirectional sensors are obtained by modeling r_m and ψ_m as random variables. More precisely,

$$\begin{aligned} r_m &\sim \mathcal{U}(0.1\,\text{cm}, 1.5\,\text{cm}) \ , \\ \psi_m &\sim \mathcal{U}(0, 2\pi) \ , \end{aligned} \tag{57}$$

where $\sim \mathcal{U}(a, b)$ indicates a uniform distribution with boundaries a and b. As further constraints over the array geometry, we set the minimum distance between each pair of sensors to $0.35\,\text{cm}$, i.e., $\delta_{ij} \geq 0.35\,\text{cm}\ \forall \{i, j\} \,|\, i \neq j$. Fig. 3 shows three array geometries of the sort.

For example, let us now design two first-order differential beamformers, the first with $q = 0.5$, steered toward $\theta_s = 0°$, and the second with $q = 0.67$, steered toward $\theta_s = 120°$, both based on the array geometry with $M = 12$ sensors in Fig. 3a. Fig. 4 shows the obtained results in terms of spatial response for $f = 1\,\text{kHz}$, WNG and DF, and it provides a comparison with respect to [26]. Notice that the differences of behavior between the proposed method and [26] are generally negligible, except for small discrepancies of DF at high frequencies. Moreover, Fig. 5 shows that the proposed beamformer and the one in [26] are both characterized by an almost frequency-invariant beampattern.

As a further analysis, we average the spatial responses, the WNG and the DF computed for 300 different array geometries with $M = 12$ obtained using (57), as described above. Fig. 6 shows the averaged spatial responses for $f = 1\,\text{kHz}$, $q = 0.67$ and 8 different steering angles θ_s. We also computed the standard deviations of the spatial responses that are close to zero at all angles and they exhibit a maximum of 0.324 in correspondence to the zeros of the beampatterns. The results of the proposed method are basically identical, both in terms of average and standard deviation, to those obtained with [26]. Fig. 7 shows that the average DF and the average WNG obtained with the two methods are matching, except for small differences at high frequencies. In particular, since it is known that the WNG is a measure of robustness of a beamformer against the deviations of the assumed sensor characteristics, like gain, phase and position, [45], we deduce that the proposed spatial method is comparable to the method in [26] in terms of sensitivity to microphone mismatches or misplacements, which is a crucial aspect in differential beamformers using small-sized arrays.

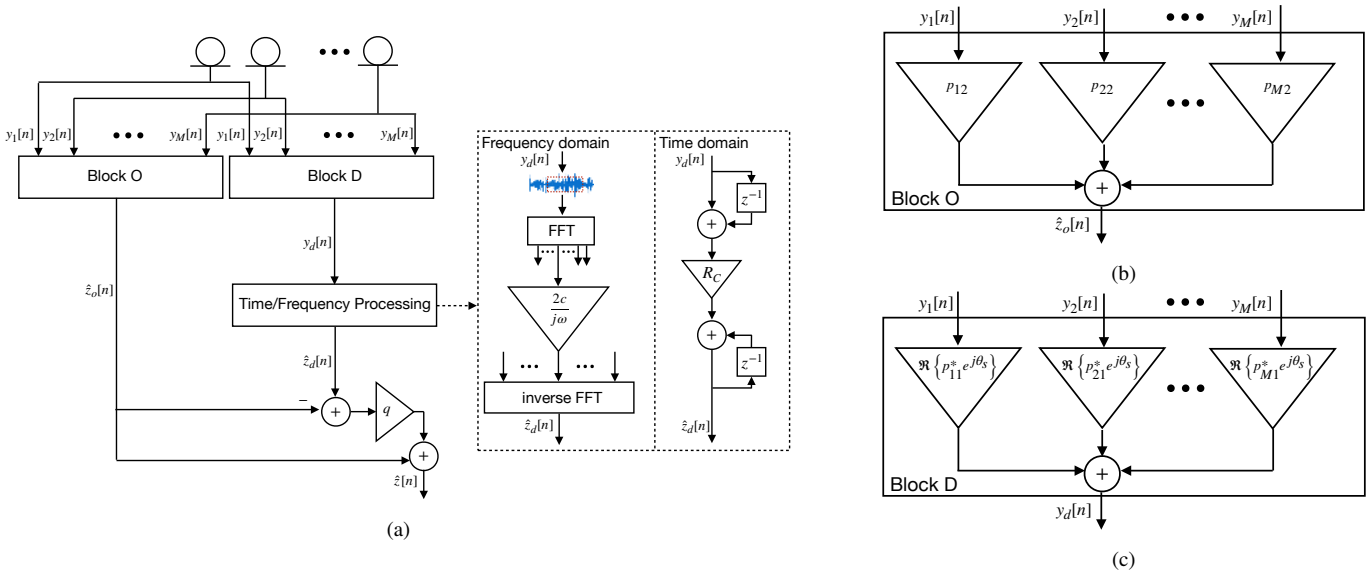


Fig. 9: Implementation scheme of the proposed beamforming system (a). The *Time/Frequency Processing* block varies depending on the chosen implementation strategy. A detail of *Block O* is given in (b), while a detail of *Block D* is given in (c).

As a last experiment, we vary the number of sensors M with the purpose of investigating how the DF and the WNG are affected. In particular, we test three different array configurations with $M = 5$, $M = 8$ and $M = 12$ sensors, respectively. For each value of M, 300 realizations are tested. The average DF and the average WNG are reported in Fig. 8. As expected, as the number of sensors M increases, the WNG increases as well. The DF, instead, is substantially unaltered.

VII. System Implementation

In this section we discuss how the proposed system can be implemented both in the frequency domain and in the time domain. As far as the frequency-domain is concerned we also compare its computational cost with the one required by the method in [26].

A. Implementation Schemes

The signal flow that describes both the STFT-based implementation and the discrete-time implementation of the proposed system is shown in Fig. 9. The signals $y_m[n]$ with $m \in 1, \ldots, M$ captured by the M sensors feed two blocks, namely *Block O* and *Block D*. The output of *Block O* is the signal $\hat{z}_o[n]$ in (47) that is the discrete-time counterpart of $\hat{Z}_o(\omega)$ in (37). Instead, the output of *Block D* is the signal $y_d[n]$ in (48) that is the discrete-time counterpart of $Y_d(\omega)$ in (37). The signal $y_d[n]$ is then fed to the block *Time/Frequency Processing*. The output of such block is the signal $\hat{z}_d[n]$ that can be obtained using either the frequency domain approach, according to (37), or the discrete-time domain approach, according to (47). Notice that, with reference to Fig. 9, despite the output of the *Time/Frequency Processing* block being denoted as $\hat{z}_d[n]$ in both time and frequency domains, the two outputs can be different from each other. This is due to the fact that the time derivative in (45) is approximated by a finite difference. However, it has already been shown that, using a sufficiently high sampling frequency this difference becomes negligible [44]. Finally, according to (46), the signals $\hat{z}_o[n]$ and $\hat{z}_d[n]$ are linearly combined in order to obtain the beamformer output $\hat{z}[n]$.

Fig. 10 shows the implementation of *Block O* and *Block D* for a uniform circular array. In this case the output $\hat{z}_o[n]$ of the *Block O* is obtained using (50), while the output $\hat{z}_d[n]$ of *Block D* is obtained according to (51).

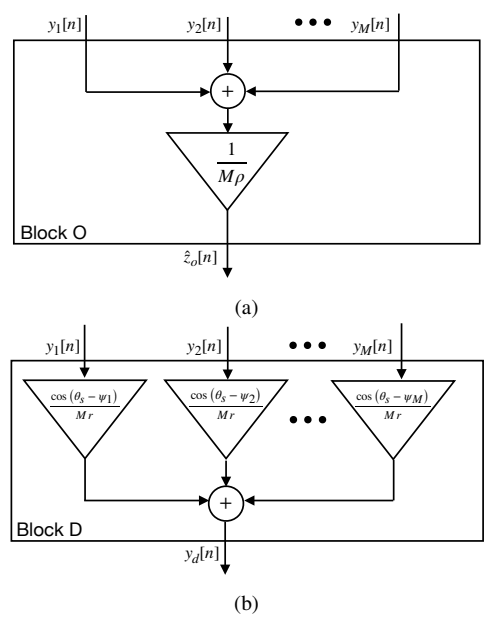


Fig. 10: Block O and Block D for the UCA case

B. A Discussion on Computational Cost

We compare the computational cost of the proposed frequency domain beamformer (see Fig. 9) to that of [26] (see Fig. 2). We consider both the general scenario of an arbitrary

TABLE I: Frequency domain implementation cost: comparison between the proposed method and [26] in terms of number of FFTs, additions (ADD) and multiplications (MUL). M is the number of sensors, while K indicates the number FFT bins.

	FFT	ADD	MUL
Proposed Arbitrary	2	$2M$	$2M+1+K$
Proposed Uniform Circular	2	$2M$	$M+1+K$
Huang *et al.* [26]	$M+1$	$K(M-1)$	MK

array geometry and the special case of a uniform circular array. Table I shows the comparison in terms of number of FFTs, additions and multiplications. The number of frequency bins is indicated by K. Considering the number of FFTs, one can see that only two FFTs are needed for the proposed approach, while the number of FFTs for [26] linearly increases with the number of microphones M. As far as the number of additions is concerned, it linearly increases with M for the proposed method but, unlike [26], it does not depend on K. More precisely, by solving the inequality $K(M-1) > 2M$ we obtain

$$M > \frac{K}{K-2} \approx 1. \tag{58}$$

Equation (58) shows that, by employing more than one microphone, the number of additions needed for the proposed method is less than that needed for [26]. The same holds for the number of multiplications. In fact, considering the general case of an arbitrary array geometry, and solving the inequality $MK > 2M+1+K$ we obtain

$$M > \frac{K+1}{K-2} \approx 1. \tag{59}$$

The computational cost of the proposed discrete-time domain beamformer, instead, is hereafter discussed in terms of number of required operations per sampling step. In particular, Table II reports the number of additions and multiplications to perform, along with the number of scalar values to store (i.e., memory requirements), at each sampling step. Also in this case, we consider both the general scenario of an arbitrary array geometry and the special case of a uniform circular array.

TABLE II: Discrete-time domain implementation cost in terms of number of adds (ADD), multiplications (MUL) and samples to store (MEM) per sampling step.

	ADD	MUL	MEM
Arbitrary Geometry	$2M+2$	$2M+2$	2
Uniform Circular Geometry	$2M+2$	$M+3$	2

A comparison between the proposed frequency domain beamformer and the proposed discrete-time domain beamformer in terms of computational cost could be developed by considering K samples of the output signal $\hat{z}[n]$, where K is the length of the FFT frames. In this case, the cost to output K samples required by the frequency domain beamformer is already reported in Table I, while the cost of the discrete-time domain beamformer can be obtained by multiplying by K the values in Table II that refer to a single sampling step. A more precise cost comparison should take into account that frames in STFT-filtering usually overlap with the amount of overlap depending on the used window functions [43].

As a further remark, we note that the proposed discrete-time domain beamformer exhibits an higher time resolution than the proposed frequency domain beamformer since the steering direction θ_s can be varied at each sampling step.

VIII. Conclusions and Future Works

In this manuscript we presented a method that allows us to implement steerable first-order DMAs with arbitrary planar geometry, under the assumption of small-size arrays (spacing between sensors smaller than the audio wavelength of interest). This method's performance is comparable with that of differential spatial filtering solutions available in the literature in terms of WNG, DF and beampattern, but it requires a significantly lower computational cost. Both frequency-domain and discrete-time domain implementations have been discussed. Future works will be devoted to the extension of the proposed approach to higher order DMAs. Such an extension would require us to find approximations of higher order Bessel functions, thus making possible to move frequency dependent terms to the last stages of the beamformer signal flow. However, a thorough analysis on the limitations of such an approximation would be necessary to study the trade-off between computational complexity and accuracy of the resulting beamformer. Another appealing development will be the integration of the proposed DMA models into two-stage spatial filtering methods like those discussed in [46].

Appendix A
Derivation of $\mathbf{\Phi}\mathbf{\Phi}^H$ in the Uniform Circular Case

We show that, in the case of a uniform circular array geometry, the matrix multiplication $\mathbf{\Phi}\mathbf{\Phi}^H$ reduces to a diagonal matrix as stated in (40). Looking at the expression in (26), under the assumption in (39), we show that the following two equations are verified

$$\sum_{m=1}^{M} r_m e^{-j\psi_m} = 0 \ , \tag{60}$$

$$\sum_{m=1}^{M} \left(r_m e^{-j\psi_m}\right)^2 = 0 \ . \tag{61}$$

The first equation (60) can be derived as follows

$$\begin{aligned}
\sum_{m=1}^{M} r_m e^{-j\psi_m} &= r \sum_{m=1}^{M} e^{-j\frac{2\pi}{M}(m-1)} \\
&= r e^{j\frac{2\pi}{M}} \sum_{m=1}^{M} e^{-j\frac{2\pi}{M}(m)} \quad (n = m-1) \\
&= r e^{j\frac{2\pi}{M}} \sum_{n=0}^{M-1} e^{-j\frac{2\pi}{M}(n)} \\
&= r e^{j\frac{2\pi}{M}} \left(\frac{1 - e^{-jM\frac{2\pi}{M}}}{1 - e^{-j\frac{2\pi}{M}}} \right) = 0.
\end{aligned} \tag{62}$$

The second equation (61), instead, can be derived as

$$\begin{aligned}\sum_{m=1}^{M}\left(r_m e^{-j\psi_m}\right)^2 &= \sum_{m=1}^{M} r^2 e^{-j2\frac{2\pi}{M}(m-1)} \\ &= r^2 e^{j\frac{4\pi}{M}} \sum_{m=1}^{M} e^{-j\frac{4\pi}{M}(m)} \quad (n=m-1) \\ &= r^2 e^{j\frac{4\pi}{M}} \sum_{n=0}^{M-1} e^{-j\frac{4\pi}{M}(n)} \\ &= r^2 e^{j\frac{4\pi}{M}} \left(\frac{1-e^{-jM\frac{4\pi}{M}}}{1-e^{-j\frac{4\pi}{M}}}\right) = 0.\end{aligned} \tag{63}$$

REFERENCES

[1] M. Brandstein and D. Ward, Eds., *Microphone Arrays: Signal Processing, Techniques and Applications*, ser. Digital Signal Processing. Berlin: Springer, 2001.

[2] J. Benesty, J. Chen, and Y. Huang, *Microphone Array Signal Processing*, ser. Springer Topics in Signal Processing. Berlin: Springer, 2008, no. v. 1, oCLC: ocn233534508.

[3] J. Benesty and J. Chen, *Study and Design of Differential Microphone Arrays*, ser. Springer Topics in Signal Processing. Heidelberg; New York: Springer, 2013, no. Volume 6.

[4] J. Benesty, J. Chen, and I. Cohen, *Design of Circular Differential Microphone Arrays*, ser. Springer Topics in Signal Processing. Cham: Springer International Publishing, 2015, vol. 12.

[5] J. Benesty, J. Chen, and C. Pan, *Fundamentals of Differential Beamforming*, ser. SpringerBriefs in Electrical and Computer Engineering. Singapore: Springer Singapore, 2016.

[6] D. P. Jarrett, E. A. P. Habets, and P. A. Naylor, *Theory and applications of spherical microphone array processing*. Springer, 2017, vol. 9.

[7] M. S. Brandstein and H. F. Silverman, "A Practical Methodology for Speech Source Localization with Microphone Arrays," *Computer Speech & Language*, vol. 11, no. 2, pp. 91–126, Apr. 1997.

[8] J. H. DiBiase, H. F. Silverman, and M. S. Brandstein, "Robust Localization in Reverberant Rooms," in *Microphone Arrays*. Berlin: Springer, 2001, pp. 157–180.

[9] J. Valin, F. Michaud, J. Rouat, and D. Létourneau, "Robust Sound Source Localization Using a Microphone Array on a Mobile Robot," in *International Conference on Intelligent Robots and Systems (IROS)*, vol. 2. IEEE, 2003, pp. 1228–1233.

[10] S. Gannot, D. Burshtein, and E. Weinstein, "Signal Enhancement Using Beamforming and Nonstationarity with Applications to Speech," *IEEE Transactions on Signal Processing*, vol. 49, no. 8, pp. 1614–1626, Aug. 2001.

[11] J. Benesty, S. Makino, and J. Chen, *Speech Enhancement*, ser. Signals and Communication Technology. Berlin: Springer, 2005.

[12] S. Gannot, E. Vincent, S. Markovich-Golan, and A. Ozerov, "A consolidated perspective on multimicrophone speech enhancement and source separation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 4, pp. 692–730, 2017.

[13] E. Vincent, T. Virtanen, and S. Gannot, *Audio source separation and speech enhancement*. John Wiley & Sons, 2018.

[14] L. C. Parra and C. V. Alvino, "Geometric Source Separation: Merging Convolutive Source Separation with Geometric Beamforming," *IEEE Transactions on Speech and Audio Processing*, vol. 10, no. 6, pp. 352–362, Sep. 2002.

[15] F. Asano, S. Ikeda, M. Ogawa, H. Asoh, and N. Kitawaki, "Combined Approach of Array Processing and Independent Component Analysis for Blind Separation of Acoustic Signals," *IEEE Transactions on Speech and Audio Processing*, vol. 11, no. 3, pp. 204–215, May 2003.

[16] S. M., *Audio Source Separation*. Berlin: Springer, 2018.

[17] G. W. Elko and Anh-Tho N. P., "A Simple Adaptive First-Order Differential Microphone," in *Workshop on Applications of Signal Processing to Audio and Accoustics (WASPAA)*. New Paltz, NY, USA: IEEE, 1995, pp. 169–172.

[18] ——, "A Steerable and Variable First-Order Differential Microphone Array," in *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 1, Apr. 1997, pp. 223–226 vol.1.

[19] G. W. Elko, "Differential Microphone Arrays," in *Audio Signal Processing for Next-Generation Multimedia Communication Systems*, Y. Huang and J. Benesty, Eds. Boston: Kluwer Academic Publishers, 2004, pp. 11–65.

[20] E. D. Sena, H. Hacihabiboglu, and Z. Cvetkovic, "On the Design and Implementation of Higher Order Differential Microphones," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 1, pp. 162–174, Jan. 2012.

[21] L. Zhao, J. Benesty, and J. Chen, "Design of Robust Differential Microphone Arrays," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 10, pp. 1455–1466, Oct. 2014.

[22] C. Pan, J. Chen, and J. Benesty, "Theoretical Analysis of Differential Microphone Array Beamforming and an Improved Solution," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 11, pp. 2093–2105, Nov. 2015.

[23] X. Wu and H. Chen, "Directivity Factors of the First-Order Steerable Differential Array With Microphone Mismatches: Deterministic and Worst-Case Analysis," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 2, pp. 300–315, Feb. 2016.

[24] A. Bernardini, M. D'Aria, R. Sannino, and A. Sarti, "Efficient Continuous Beam Steering for Planar Arrays of Differential Microphones," *IEEE Signal Processing Letters*, vol. 24, no. 6, pp. 794–798, Jun. 2017.

[25] G. Huang, J. Benesty, and J. Chen, "On the Design of Frequency-Invariant Beampatterns With Uniform Circular Microphone Arrays," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 5, pp. 1140–1153, May 2017.

[26] G. Huang, J. Chen, and J. Benesty, "On the Design of Differential Beamformers with Arbitrary Planar Microphone Array Geometry," *The Journal of the Acoustical Society of America*, vol. 144, no. 1, pp. EL66–EL70, Jul. 2018.

[27] G. Huang, J. Benesty, I. Cohen, and J. Chen, "A simple theory and new method of differential beamforming with uniform linear microphone arrays," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2020.

[28] ——, "Differential beamforming on graphs," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 901–913, 2020.

[29] X. Wang, J. Benesty, J. Chen, and I. Cohen, "Beamforming with small-spacing microphone arrays using constrained/generalized lasso," *IEEE Signal Processing Letters*, vol. 27, pp. 356–360, 2020.

[30] Y. Buchris, I. Cohen, and J. Benesty, "On the Design of Time-Domain Differential Microphone Arrays," *Applied Acoustics*, vol. 148, pp. 212–222, may 2019.

[31] H. Zhang, J. Chen, and J. Benesty, "Study of nonuniform linear differential microphone arrays with the minimum-norm filter," *Applied Acoustics*, vol. 98, pp. 62–69, 2015.

[32] G. W. Elko, "Steerable and Variable First-Order Differential Microphone Array," Patent 6,041,127, Mar., 2000.

[33] J. Lovatello, A. Bernardini, and A. Sarti, "Steerable Circular Differential Microphone Arrays," in *European Signal Processing Conference (EUSIPCO)*. IEEE, Sep. 2018, pp. 11–15.

[34] A. M. Torres, M. Cobos, B. Pueo, and J. J. Lopez, "Robust Acoustic Source Localization Based on Modal Beamforming and Time–Frequency Processing Using Circular Microphone Arrays," *The Journal of the Acoustical Society of America*, vol. 132, no. 3, pp. 1511–1520, Sep. 2012.

[35] S. Yan, "Optimal Design of Modal Beamformers for Circular Arrays," *The Journal of the Acoustical Society of America*, vol. 138, no. 4, pp. 2140–2151, Oct. 2015.

[36] Y. Buchris, I. Cohen, and J. Benesty, "First-order differential microphone arrays from a time-domain broadband perspective," in *International Workshop on Acoustic Signal Enhancement (IWAENC)*. IEEE, Sep. 2016, pp. 1–5.

[37] Y. Buchris, I. Cohen, and J. Benesty, "Analysis and design of time-domain first-order circular differential microphone arrays," in *International Congress on Acoustics (ICA)*, Sep. 2016.

[38] S. Yan, "Robust Time-domain Wideband Modal Beamforming with Circular Arrays," *IEEE Transactions on Aerospace and Electronic Systems*, pp. 1–1, 2019.

[39] L. Zhao, J. Benesty, and J. Chen, "Design of Robust Differential Microphone Arrays with the Jacobi–Anger Expansion," *Applied Acoustics*, vol. 110, pp. 194–206, Sep. 2016.

[40] G. W. Elko and J. Meyer, *Microphone arrays*. Springer, 2008.

[41] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*. Courier Corporation, 1965, vol. 55.

[42] I. J. Good, "The Inverse of a Centrosymmetric Matrix," *Technometrics*, vol. 12, no. 4, pp. 925–928, 1970.

[43] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*, 2nd ed. Upper Saddle River, N.J: Prentice Hall, 1999.
[44] A. Bernardini, F. Antonacci, and A. Sarti, "Wave Digital Implementation of Robust First-Order Differential Microphone Arrays," *IEEE Signal Processing Letters*, vol. 25, no. 2, pp. 253–257, Feb 2018.
[45] S. Doclo and M. Moonen, "Superdirective beamforming robust against microphone mismatch," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 2, pp. 617–631, 2007.
[46] F. Borra, A. Bernardini, F. Antonacci, and A. Sarti, "Uniform Linear Arrays of First-Order Steerable Differential Microphones," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 12, pp. 1906–1918, Dec. 2019.

Federico Borra (S'17) received the B.S. degree, in 2014, and the M.S. degree (cum laude), in 2016, in computer engineering from the Politecnico di Milano, Italy. In 2020 he received his Ph.D. degree in information engineering from the Politecnico di Milano, Italy, where he is currently a postdoctoral researcher. His main research interests concern space-time audio signal processing.

Alberto Bernardini (S'16–M'19) received his B.S. degree from University of Bologna, Italy, in 2012 and his M.S. degree (cum laude) from Politecnico di Milano, Italy, in 2015, both in Computer Engineering. In 2019, he received his Ph.D. degree (cum laude) in Information Engineering from Politecnico di Milano, Italy, where he is currently a postdoctoral researcher. His main research interests are audio signal processing and modeling of nonlinear systems. He authored about 20 publications in international journals and proceedings of international conferences. He is also the first author of an international patent.

Fabio Antonacci (M'14) was born in Bari, Italy, on July 26, 1979. He received the Laurea degree in 2004 in telecommunication engineering and the Ph.D. degree in information engineering in 2008, both from the Politecnico di Milano, Milan, Italy. He is currently an Assistant Professor at the Politecnico di Milano. His research focuses on space-time processing of audio signals, for both speaker and microphone arrays (source localization, acoustic scene analysis, rendering of spatial sound) and on modeling of acoustic propagation. He is a member of the IEEE Audio and Acoustic Signal Processing Technical Committee and of the EURASIP SAT on Audio, Speech and Music Signal Processing.

Augusto Sarti (M'04–SM'13) received his Ph.D. Information Engineering from the University of Padova, Italy, in 1993, with a joint graduate program with University of California, Berkeley. In 1993, he joined the Faculty of the Politecnico di Milano, Italy, where he is currently a Full Professor. In 2013, he also joined the University of California, Davis. He coordinates the activities of the Musical Acoustics Laboratory and the Sound and Music Computing Laboratory of the Politecnico di Milano. He promoted/coordinated and/or contributed to numerous European projects in the area of multimedia signal processing. He has coauthored over 300 scientific publications on international journals and congresses and numerous patents in the multimedia signal processing area. His main research interests are in the area of audio and acoustic signal processing, with particular focus on sound analysis, synthesis, and processing; space-time audio processing; geometrical acoustics; music information extraction and music modeling. He served in the IEEE Technical Committee on Audio and Acoustics Signal Processing for two terms. He served as Associate Editor of IEEE/ACM Tr. On Audio Speech and Language Processing, and as Senior Area Editor of IEEE Signal Processing Letters, and in 2017 he received the "Outstanding Editorial Board Member Award" by the IEEE Signal Processing Society. He co-chaired the IEEE Intl. Conf. on Advanced Video and Signal based Surveillance (AVSS-05); he chairman the Digital Audio Effects conference (DAFx-09); and he co-chaired the IEEE Intl. Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA-19). He was in the organizing committees of numerous International Conferences, including IEEE ICASSP-14; the ACM/IEEE Intl. Conf. on Distributed Smart Cameras (ICDSC-09); the IEEE Intl. Workshop on Haptic Audio-Visual Environment and Game (HAVE-09); and the European Signal Processing Conference (EUSIPCO-2018). He is currently serving in the EURASIP board of directors.