Two-Stage Beamforming With Arbitrary Planar Arrays of Differential Microphone Array Units

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

***Abstract*—Differential Microphone Arrays (DMAs) are of great interest in the literature on small-sized microphone arrays, due to their good directivity properties and nearly frequency-invariant spatial responses. Recently developed beamforming techniques combine multiple DMA units to form flexible two-stage spatial filtering systems, where the output of each DMA is fed into a higher-level filter, called virtual filter, for further processing. In this manuscript, we analyze and discuss some properties of a broad class of two-stage beamformers with arbitrary planar geometry. In this context, the DMA units are all assumed to have the same directivity pattern of arbitrary order and can be characterized by a variable number of omnidirectional sensors organized in an arbitrary geometry. For any given choice of the virtual array filter, we introduce a closed-form optimization procedure to design DMA filters that maximize the White Noise Gain (WNG) or the Directivity Factor (DF) of the resulting two-stage beamformer at any frequency. Based on this frequency-dependent design, we propose a frequency-invariant design of the two-stage beamformer and we compare the performance of the two approaches. Finally, we propose two possible computational schemes for the proposed generic two-stage spatial filtering system and discuss their efficiency in performing filtering, steering, and changing beampattern.**

***Index Terms*—Differential Microphone Arrays, DMA, Beamformers.**

I. Introduction

MICROPHONE arrays are widely used to take advantage of the spatial diversity in acoustic environments for advanced audio signal processing and speech enhancement applications [1], [2]. In many of these applications, microphone arrays and sound sources are assumed to be on the same plane. These include hands-free human-machine interfaces, teleconferencing systems, and home and habitat surveillance. All of these systems can benefit from the use of spatial filters (beamformers) to improve their performance. Among the different beamforming techniques, those based on Differential Microphone Arrays (DMAs) are of great interest. DMAs have been extensively studied in the literature for small-sized microphone arrays, due to their good directivity properties, nearly frequency invariant beampatterns, and low computational cost [3]–[10].

In traditional DMA theory, Uniform Linear Arrays (ULAs) are used to construct Nth-order directional responses by combining the derivatives of the acoustic pressure up to order N [1], [6]. Pressure differentials are approximated by differences between signals sensed by microphones. This approximation holds under the assumption that the distance between the sensors is much smaller than the acoustic wavelength of the source signal of interest.

In more recent works, DMAs are implemented in the Short-Time Fourier Transform (STFT) domain [2], [7], [8], [11]–[13], where Nth-order DMAs are designed using different techniques. In addition to traditional ULAs, non-uniform linear arrays [9], circular arrays [8], [11], square arrays [14], and arbitrary planar geometry arrays [12], [13] have also been studied.

Among the STFT domain techniques available in the literature, the one presented in [11] for uniform circular arrays and later extended in [12] to arbitrary planar arrays enables to design spatial filters that approximate beampatterns of traditional DMAs (e.g., cardioid, supercardioid, hypercardioid). The technique, called FIB-LSE [11], is able to steer the beampattern in any direction of the 2D plane without distortion, a feature not available with traditional linear DMAs. The spatial filter taps are derived based on the array geometry, the desired steering angle, and a set of *beam shape* coefficients that fully characterize the target beampattern to be approximated.

Although the current literature on DMAs offers good design flexibility, further beamforming systems with interesting properties can be obtained if we consider arrays of multiple DMAs. The first work in this direction is [15], where multiple first-order steerable DMAs (FOSDMAs) are arranged into an ULA, resulting in a two-stage spatial filtering system. The first stage is a *local* filtering performed by the individual DMA units (which are assumed to steer identical beams in the same direction); the second stage is a Delay-And-Sum (DAS) beamformer combining the outputs of the FOSDMA units, called *virtual* array filtering. The combination of local and virtual array filters yield the derivation of a *global* spatial filter whose coefficients can be applied directly to all the microphone signals to obtain an equivalent single-stage beamformer. The resulting global beamformer can efficiently morph between a DAS-like beamformer and a Super-Directive (SD)-like beamformer by simply varying a scalar parameter that determines the shape of the first-order beams. This approach was later extended to systems combining several FOSDMA units in which both the virtual array and the DMA units are characterized by arbitrary geometries [16]. Other works discuss two-stage spatial filtering techniques on arrays of *identical* (in terms of geometry and spatial filter) DMA units and use a Kronecker Product (KP) description to derive the two-stage beamformer [17], [18]. In particular, the authors in [17] show that for identical DMA units, the global array beampattern is given by the product between the DMA unit

beampattern and the virtual array beampattern. In [18], the authors address KP beamforming using a two-stage approach on ULAs and reframe DMA theory using finite difference operators. More recently in [14] the authors proposed a 3-dimensional two-stage beamforming technique for square DMAs.

State-of-the-art two-stage beamformers using multiple DMAs employ either arbitrary configurations of first-order DMA units with arbitrary geometry or arbitrary configurations of Nth-order DMA units with the same geometry. In this paper, we discuss the properties of a more general class of two-stage beamformers characterized by arbitrary planar configurations of multiple Nth-order DMA units, each one with the same directional response, and where DMA units can have different geometries and different numbers of omnidirectional sensors. Therefore, any two-stage beamformer based on regular distributions of microphones can be derived as a special case of this general approach. We also show that, even for this class of two-stage beamformers, the beampattern of the global array equals the product between the DMA unit beampattern and the virtual array beampattern. Also, the resulting two-stage beamformer is able to perform continuous beam steering in any direction of the plane.

Other contributions of this work relate to design and implementation aspects. As for the spatial filter design, without loss of generality, we consider the DMA units as FIB-LSE beamformers [12], while the virtual array filter is assumed to be either a DAS or SD beamformer. We introduce a closed-form optimization procedure to derive the beam shape coefficients of the local filters that maximize the White Noise Gain (WNG) or the Directivity Factor (DF) of the resulting two-stage beamformer at each frequency and for a given choice of virtual array filter. Based on this frequency-dependent design, we propose a frequency-invariant design of the two-stage beamformer and we compare the performance of the two approaches. Finally, we discuss two possible implementations of the generic two-stage spatial filtering system.

The manuscript is structured as follows. In Section II, the signal model is defined and the metrics used to evaluate the performance of the proposed system are discussed. In Section III the beamformer design for both the DMA units and the virtual array is discussed, together with a brief overview of traditional beamformers (i.e., DAS and SD). Section IV shows that the global array beampattern is the product between the local and virtual array beampatterns. In Section V, we analyse several two-stage differential beamformer designs and compare them to traditional single-stage DAS and SD beamformers using the same set of microphones. We then discuss two implementations of the two-stage beamformer in Section VI. Finally, in Section VII we conclude the manuscript and propose some possible future works.

II. Signal model and metrics

Let us consider K DMA units deployed on the horizontal plane. Each DMA unit can have a different number of sensors M_k, with index $k = 1, \dots, K$. Both the geometry of the DMA units and their distribution on the plane are arbitrary. Fig. 1 shows an illustrative configuration of the system, where the green dots identify the reference points of the DMA units.

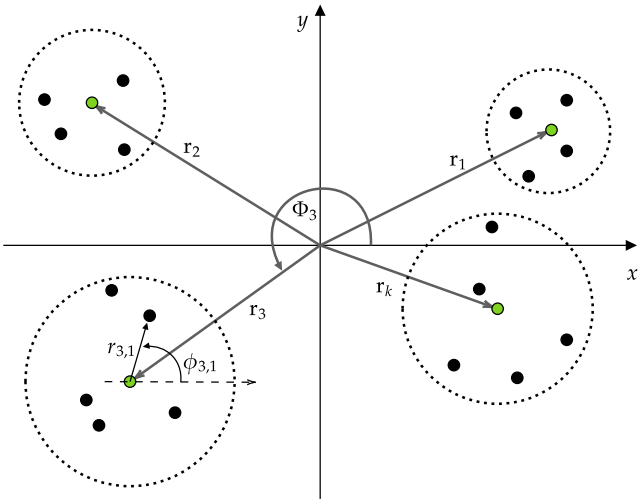

Fig. 1. Example of a planar array of DMA units. The green dots represent the reference points of the DMA units.

In what follows, we refer to the set of reference points of DMA units as *virtual array* to distinguish it from the array of all microphones, which we refer to as *global array*.

A. Local DMA Unit Signal Model

The propagation vector of a plane wave emitted by a far-field source, lying on the same plane of the DMA units, from an angle θ (measured counterclockwise from the x-axis), propagating in an anechoic acoustic environment at the speed of sound $c = 340$ m/s, and incident to the kth DMA unit, is [12], [13]

$$\mathbf{d}_k(\omega, \theta) = [e^{\jmath \overline{\omega}_{k,1} \cos(\theta - \phi_{k,1})}, \dots, e^{\jmath \overline{\omega}_{k,M_k} \cos(\theta - \phi_{k,M_k})}]^T, \quad \text{with} \quad k = 1, \dots, K, \tag{1}$$

where the superscript ^{T} denotes transposition, $\jmath$ is the imaginary unit, $\overline{\omega}_{k,m} = \omega r_{k,m}/c$ with $\omega = 2\pi f$ the angular frequency, $f > 0$ the temporal frequency and $r_{k,m}$, $\phi_{k,m}$, with $m = 1, \dots, M_k$ are the distance and the angular position of the mth microphone in the kth DMA unit w.r.t. its reference point, respectively. The signals acquired by the sensors of the kth DMA unit are modeled by the vector [12], [13]

$$\mathbf{y}_k(\omega) = \mathbf{d}_k(\omega, \theta) X_k(\omega) + \mathbf{v}_k(\omega), \tag{2}$$

where $X_k(\omega)$ is the source signal at the reference point of the kth DMA unit and $\mathbf{v}_k(\omega)$ models the additive noise. To perform spatial filtering, the elements in $\mathbf{y}_k(\omega)$ are multiplied by complex weights $W^*_{k,m}(\omega)$. The weighted sensor outputs are then summed to form the output $Z_k(\omega)$ of the kth DMA unit. By collecting the filter taps in the vector $\mathbf{w}_k(\omega) = [W_{k,1}(\omega), \dots, W_{k,M_k}(\omega)]^T$, we can write the filter output as

$$Z_k(\omega) = \mathbf{w}_k^H(\omega) \mathbf{y}_k(\omega), \tag{3}$$

where $(\cdot)^H$ is the conjugate-transpose operator.

B. Global Array Signal Model

In this work, we consider the farfield assumption to be valid also for the global array. In this case, we can define the propagation vector of the virtual array as

$$\bar{\mathbf{d}}(\omega,\theta) = [e^{j\tilde{\omega}_1 \cos(\theta-\Phi_1)}, \ldots, e^{j\tilde{\omega}_K \cos(\theta-\Phi_K)}]^T, \quad (4)$$

where $\tilde{\omega}_k = \omega \mathrm{r}_k/c$ and r_k, Φ_k are the distance and the angular position of the kth DMA unit reference point, respectively. It follows that the signal model of the global array is expressed as a function of (1) and (4) as

$$\underline{\mathbf{y}}(\omega) = \mathbf{d}(\omega,\theta)X(\omega) + \underline{\mathbf{v}}(\omega), \quad (5)$$

where

$$\begin{aligned}\underline{\mathbf{y}}(\omega) &= [\mathbf{y}_1^T(\omega), \ldots, \mathbf{y}_K^T(\omega)]^T,\\ \underline{\mathbf{v}}(\omega) &= [\mathbf{v}_1^T(\omega), \ldots, \mathbf{v}_K^T(\omega)]^T,\\ \mathbf{d}(\omega,\theta) &= \mathrm{diag}\left(\mathbf{d}_1(\omega,\theta), \ldots, \mathbf{d}_K(\omega,\theta)\right)\bar{\mathbf{d}}(\omega,\theta),\end{aligned}$$

with $\mathrm{diag}(\cdot)$ the operator that forms a rectangular block-diagonal matrix of size $\left(\sum_{k=1}^{K} M_k\right) \times K$ by stacking one per column the column vectors listed in the argument and setting to zero the other entries, and $X(\omega)$ the source signal at the center of the coordinate axes. The output signal of the global array is thus expressed as

$$\overline{Z}(\omega) = \mathbf{g}^H(\omega)\underline{\mathbf{y}}(\omega). \quad (6)$$

where $\mathbf{g}(\omega)$ is called the global array filter and can be conveniently decomposed as follows [16]

$$\mathbf{g}(\omega) = \mathrm{diag}(\mathbf{w}_1(\omega), \ldots, \mathbf{w}_K(\omega))\mathbf{h}(\omega), \quad (7)$$

where $\mathbf{h}(\omega) = [H_1(\omega), \ldots, H_K(\omega)]^T$ is referred to as virtual array filter. A look at eq. (7) reveals that the combined effect of the filter of each DMA unit $\mathbf{w}_k(\omega)$ with the virtual array filter $\mathbf{h}(\omega)$, when used in a two-stage spatial filtering system, is equivalent to the effect of the global filter $\mathbf{g}(\omega)$ acting on all microphone signals in a single-stage beamformer configuration. It is also worth noting that the global array filter expression in eq. (7) is more general than the corresponding Kronecker product-based expression used in [17], since it accounts for DMA units with different geometries and possibly different numbers of sensors. In fact, eq. (7) includes the expression for the global filter in [17] as the special case of a two-stage beamformer employing identical DMA units.

C. Metrics

We now briefly revise important metrics that are commonly used in the literature on microphone arrays [7], and will be used in this manuscript to characterize the spatial response of a beamformer. Two metrics quantify the Signal-to-Noise-Ratio (SNR) gain (i.e., the ratio between the output SNR and the input SNR) achieved by the beamformer. The first, called White Noise Gain (WNG) [7], refers to spatially white noise, and it is often used as a robustness measure against sensor gain or phase errors, and sensor misplacement. The WNG is defined as

$$\mathrm{WNG}[\mathbf{g}(\omega)] = \frac{|\mathbf{g}^H(\omega)\mathbf{d}(\omega,\theta_s)|^2}{\mathbf{g}^H(\omega)\mathbf{g}(\omega)}, \quad (8)$$

where θ_s is the steering angle. A second metric, called Directivity Factor (DF) [7] measures the robustness of the array against diffuse noise (e.g., in a reverberant room) with covariance matrix $\mathbf{\Gamma}_{\mathrm{dn}}(\omega)$,

$$\mathrm{DF}[\mathbf{g}(\omega)] = \frac{|\mathbf{g}^H(\omega)\mathbf{d}(\omega,\theta_s)|^2}{\mathbf{g}^H(\omega)\mathbf{\Gamma}_{\mathrm{dn}}(\omega)\mathbf{g}(\omega)}, \quad [\mathbf{\Gamma}_{\mathrm{dn}}(\omega)]_{i,j} = \frac{\sin[\omega\rho_{ij}/c]}{\omega\rho_{ij}/c} \quad (9)$$

where ρ_{ij} is the Euclidean distance between the ith and the jth microphones. A further metric is the beampattern, i.e., the spatial response of the beamformer, expressed as a function of the direction of arrival θ as [7]

$$\mathcal{B}[\mathbf{g}(\omega),\theta] = \mathbf{g}^H(\omega)\mathbf{d}(\omega,\theta). \quad (10)$$

Finally, the Front-to-Back Ratio (FBR) measures the ability of a spatial filter to attenuate signals coming from the rear of the array with respect to the frontal direction. The "front" and the "back" are defined according to a desired steering angle θ_s. Mathematically, the FBR is expressed as [7]

$$\mathrm{FBR}[\mathbf{g}(\omega)] = \frac{\int_{\theta_s-\pi/2}^{\theta_s+\pi/2} |\mathcal{B}[\mathbf{g}(\omega),\theta]|^2 d\theta}{\int_{\theta_s+\pi/2}^{\theta_s+3\pi/2} |\mathcal{B}[\mathbf{g}(\omega),\theta]|^2 d\theta}. \quad (11)$$

III. Design of the spatial filter

This section first offers a brief overview on the design of traditional spatial filters, namely the DAS and the SD beamformers. We then address the design of the global array filter $\mathbf{g}(\omega)$ in (7) by discussing the design of each DMA unit filter $\mathbf{w}_k(\omega)$ and the virtual array filter $\mathbf{h}(\omega)$ separately. Note that the design of any spatial filter also depends on the desired steering angle θ_s, but we omit this dependence for ease of reading. We assume that the DMA units are implemented using FIB-LSE filters [11], [12], while the virtual array filter is either a DAS or a SD beamformer. Finally, we propose an optimization problem that, set the virtual array filter, provides the beam shape coefficients of the local filters that maximize the WNG or the DF of the resulting two-stage beamformer at each frequency of interest.

A. Traditional Single-Stage Filtering Approaches

In traditional microphone array processing, DAS and SD beamformers are certainly among the most commonly used. Considering all the sensors in the system of Fig. 1, and the corresponding propagation vector $\mathbf{d}(\omega,\theta)$, a spatial filter $\mathbf{f}(\omega)$ can be derived by solving an optimization problem whose solution is a general form that includes both the DAS and SD beamformers as special cases. The optimization problem aims to reduce unwanted noise and interference while keeping the signal from a desired direction of arrival θ_s unmodified, and it can be written as [19]

$$\begin{aligned}\underset{\mathbf{f}(\omega)}{\arg\min} \quad & \mathbf{f}^H(\omega)\mathbf{R}(\omega)\mathbf{f}(\omega)\\ \text{subject to} \quad & \mathbf{f}^H(\omega)\mathbf{d}(\omega,\theta_s) = 1\end{aligned} \quad (12)$$

where $\mathbf{R}(\omega) = \mathbb{E}[\mathbf{v}(\omega)\mathbf{v}^H(\omega)]$ is the noise covariance matrix and $\mathbb{E}[\cdot]$ the expectation operator. The solution to (12) is given by

$$\mathbf{f}(\omega) = \frac{\mathbf{R}^{-1}(\omega)\mathbf{d}(\omega,\theta_s)}{\mathbf{d}^H(\omega,\theta_s)\mathbf{R}^{-1}(\omega)\mathbf{d}(\omega,\theta_s)}. \quad (13)$$

Setting $\mathbf{R}$ either equal to the identity matrix $\mathbf{I}$ or $\mathbf{\Gamma}_{\text{dn}}(\omega)$, obtained as in eq. (9), gives the DAS or the SD beamformers, respectively.

B. Two-Stage Filtering Approach

1) DMA Unit Filter: The local filters of the DMA units are designed as FIB-LSE beamformers, which allow us to design continuously steerable beampatterns that approximate a target beampattern (e.g., cardioid, supercardioid, hypercardioid) of any order N. The filter of the kth DMA unit is derived as [12]

$$\mathbf{w}_k(\omega) = \mathbf{\Psi}_k^{\dagger}(\omega)\mathbf{\Upsilon}^*(\theta_s)\mathbf{b}_{2N}, \tag{14}$$

where $\mathbf{b}_{2N}$ is a column vector collecting the $2N+1$ beam shape coefficients characterizing the target beampattern we wish to approximate, $\mathbf{\Psi}_k^{\dagger}(\omega) = \mathbf{\Psi}_k^H(\omega)\left[\mathbf{\Psi}_k(\omega)\mathbf{\Psi}_k^H(\omega)\right]^{-1}$ is the right Moore-Penrose inverse of $\mathbf{\Psi}_k(\omega)$, which is defined as

$$\mathbf{\Psi}_k(\omega) = \begin{bmatrix} (-j)^{-N}\boldsymbol{\psi}_{k,-N}^H(\omega) \\ \vdots \\ \boldsymbol{\psi}_{k,0}^H(\omega) \\ \vdots \\ (-j)^{N}\boldsymbol{\psi}_{k,N}^H(\omega) \end{bmatrix}, \tag{15}$$

with

$$\boldsymbol{\psi}_{k,n}(\omega) = \begin{bmatrix} J_n\left(\frac{\omega r_{k,1}}{c}\right)e^{-jn\phi_{k,1}} \\ \vdots \\ J_n\left(\frac{\omega r_{k,M_k}}{c}\right)e^{-jn\phi_{k,M_k}} \end{bmatrix}, \tag{16}$$

where J_n is the nth-order Bessel function of the first kind, and finally

$$\mathbf{\Upsilon}(\theta_s) = \text{diag}\left(e^{jN\theta_s}, \ldots, 1, \ldots, e^{-jN\theta_s}\right). \tag{17}$$

This filter design approach can be thought of as a modal matching problem based on circular harmonics, where the goal is to derive the spatial filter taps $\mathbf{w}_k(\omega)$ that yield a least-square optimal approximation of a target beampattern, given the array geometry, the steering angle θ_s, and the beam shape coefficients in vector $\mathbf{b}_{2N}$.

In the case of a symmetric target beampattern only a reduced set of $N+1$ beam shape coefficients $\mathbf{a}_N = [a_{0,N}, \ldots, a_{N,N}]^T$ is sufficient to fully characterize the target spatial response [11], [12]. The relationship between the two sets of beam shape coefficients is given by

$$\mathbf{b}_{2N} = \mathbf{L}\mathbf{a}_N \tag{18}$$

where L is a $(2N+1)\times(N+1)$ matrix of the type

$$\mathbf{L} = \begin{bmatrix} 0 & 0 & \ldots & \frac{1}{2} \\ \vdots & \vdots & \iddots & \vdots \\ 0 & \frac{1}{2} & \ldots & 0 \\ 1 & 0 & \ldots & 0 \\ 0 & \frac{1}{2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \frac{1}{2} \end{bmatrix}. \tag{19}$$

From now on, with the term beam shape coefficients we will refer to the reduced set of coefficients $\mathbf{a}_N$. The beam shape coefficients for generating the most common low-order symmetric beampatterns are well-known in the literature on DMAs and are summarized in [20].

2) Virtual Array Filter: In this work, we assume the virtual array filter to be either a DAS or a SD beamformer, although other spatial filters may also be employed. It is known that the DAS beamformer achieves the highest possible WNG, while the SD beamformer maximizes the DF.

3) Global Array Filter: Once the design of the local and virtual array filters is chosen, the resulting global array filter can generally be computed using (7). Since all the DMA units are assumed to be characterized by the same beampattern, after substituting the DMA filter equation (14) into (7), the global filter equation can be rewritten as

$$\mathbf{g}(\omega) = \mathbf{\Lambda}(\omega)\mathbf{\Upsilon}^*(\theta_s)\mathbf{b}_{2N} = \mathbf{\Lambda}(\omega)\mathbf{\Upsilon}^*(\theta_s)\mathbf{L}\mathbf{a}_N, \tag{20}$$

where

$$\mathbf{\Lambda}(\omega) = \begin{bmatrix} H_1(\omega)\mathbf{\Psi}_1^{\dagger}(\omega) \\ H_2(\omega)\mathbf{\Psi}_2^{\dagger}(\omega) \\ \vdots \\ H_K(\omega)\mathbf{\Psi}_K^{\dagger}(\omega) \end{bmatrix}.$$

C. Design of DMA Unit Beam Shape Coefficients

In this subsection we discuss a possible strategy to derive the DMA unit beam shape coefficients $\mathbf{a}_N$ which, leaving fixed the virtual array filter, maximize the WNG or the DF of the resulting global filter. Since the optimization is performed for any frequency ω, we express the vector of beam shape coefficients as a function of frequency, i.e., $\mathbf{a}_N(\omega)$. Such an optimization can be recast to a constrained minimization problem in the form

$$\begin{aligned} &\arg\min_{\mathbf{a}_N(\omega)} \mathbf{g}(\omega)^H\mathbf{R}(\omega)\mathbf{g}(\omega) \\ &\text{subject to } \mathbf{g}^H(\omega)\mathbf{d}(\omega,\theta_s) = 1 \end{aligned} \tag{21}$$

where $\mathbf{R}$ is set equal to either $\mathbf{I}$ or $\mathbf{\Gamma}_{\text{dn}}(\omega)$ depending on whether we maximize the WNG or the DF of the two-stage beamformer, and the constraint imposes a unitary gain in the steering direction. We can rewrite the optimization problem in (21) by making explicit the dependence of the global filter on the beam shape parameters $\mathbf{a}_N(\omega)$ of each DMA unit. Using (20), the minimization problem becomes

$$\begin{aligned} &\arg\min_{\mathbf{a}_N(\omega)} \mathbf{a}_N(\omega)^T\mathbf{A}(\omega,\theta_s)\mathbf{a}_N(\omega) \\ &\text{subject to } \mathbf{a}_N(\omega)^T\mathbf{B}(\omega,\theta_s) = 1 \end{aligned} \tag{22}$$

where

$$\begin{aligned} \mathbf{A}(\omega,\theta_s) &= \mathbf{L}^T\mathbf{\Upsilon}^T(\theta_s)\mathbf{\Lambda}^H(\omega)\mathbf{R}(\omega)\mathbf{\Lambda}(\omega)\mathbf{\Upsilon}^*(\theta_s)\mathbf{L} \\ \mathbf{B}(\omega,\theta_s) &= \mathbf{L}^T\mathbf{\Upsilon}^T(\theta_s)\mathbf{\Lambda}^H(\omega)\mathbf{d}(\omega,\theta_s). \end{aligned} \tag{23}$$

Since the problem in (22) is a quadratic minimization problem with linear constraints, it has a closed-form solution in the form [19]

$$\mathbf{a}_N(\omega) = \frac{\mathbf{A}(\omega,\theta_s)^{-1}\mathbf{B}(\omega,\theta_s)}{\mathbf{B}(\omega,\theta_s)^H\mathbf{A}(\omega,\theta_s)^{-1}\mathbf{B}(\omega,\theta_s)}. \tag{24}$$

Unfortunately, this minimization problem may lead to DMA unit beampatterns having amplified sidelobes (i.e., sidelobes with gain greater than one) in directions other than the steering direction θ_s. In this regard, ill-designs of beampatterns are known since early work on higher-order DMA [21]. In our two-stage beamforming context, such an undesired result is due to the zeros in the beampattern of the fixed virtual array filter, which cause the DMA unit sidelobes in the same directions as the zeros of the virtual array to grow in an unbounded way. This fact will become clearer in the light of the considerations drawn in the next section, where it is shown that the global beampattern is the product between the local DMA unit beampattern and the virtual array beampattern. In scenarios in which the outputs of DMA units are not directly used, local DMA beampatterns with amplified sidelobes are still acceptable as they lead to a global beamformer with a single maximum in the steering direction θ_s. However, in case the outputs of the DMA units are of some use, it is better to avoid amplified sidelobes in directions other than the steering direction. For the latter case, we propose a modified version of (22) that addresses the amplified sidelobes problem by bounding the squared norm of the coefficient vector $\mathbf{a}_N(\omega)$, through the addition of a term in the cost function. A scalar coefficient ϵ within the range $[0; 1]$ is used to weight the contribution of the additional term with respect to the original cost function. The solution to the modified minimization problem can be computed using (24) where

$$\mathbf{A}(\omega,\theta_s) = (1-\epsilon)\mathbf{L}^T\boldsymbol{\Upsilon}^T(\theta_s)\boldsymbol{\Lambda}^H(\omega)\mathbf{R}(\omega)\boldsymbol{\Lambda}(\omega)\boldsymbol{\Upsilon}^*(\theta_s)\mathbf{L}+ \epsilon\mathbf{I}_{N+1} \tag{25}$$

and $\mathbf{B}(\omega,\theta_s)$ remains unchanged.

IV. Product of Beampatterns

In this section we show that, under the assumption of DMA units with identical beampatterns, the global beampattern turns out to be the product of the beampattern of the DMA units and the beampattern of the virtual array. This property has already been proved using the KP beamformer formalism in [17]. We extend this result to our more general framework, where the DMA units can have different geometries and different numbers of sensors. We can safely assume that the DMA units have the same beampattern, since FIB-LSE filters have high accuracy in approximating arbitrary target beampatterns, unless inappropriate array geometries are used [12].

We can express the beampattern of the global filter by exploiting the decomposition into local and virtual array filters from eq. (7), as

$$\begin{aligned}\mathcal{B}[\mathbf{g}(\omega),\theta] &= \mathbf{g}^H(\omega)\mathbf{d}(\omega,\theta)\\ &= \mathbf{h}^H(\omega)[\text{diag}(\mathbf{w}_1(\omega),\ldots,\mathbf{w}_K(\omega))]^H\mathbf{d}(\omega,\theta).\end{aligned} \tag{26}$$

If we expand accordingly the steering vector of the global array, we obtain

$$\begin{aligned}\mathcal{B}[\mathbf{g}(\omega),\theta] = \mathbf{h}^H(\omega)[\text{diag}(\mathbf{w}_1(\omega),\ldots,\mathbf{w}_K(\omega))]^H\times\\ \text{diag}(\mathbf{d}_1(\omega,\theta),\ldots,\mathbf{d}_K(\omega,\theta))\bar{\mathbf{d}}(\omega,\theta).\end{aligned} \tag{27}$$

As a result, the global array beampattern can be expressed as

$$\begin{aligned}&\mathcal{B}[\mathbf{g}(\omega),\theta] =\\ &\quad\mathbf{h}^H(\omega)\text{diag}(\mathcal{B}[\mathbf{w}_1(\omega),\theta],\ldots,\mathcal{B}[\mathbf{w}_K(\omega),\theta])\bar{\mathbf{d}}(\omega,\theta).\end{aligned} \tag{28}$$

Since all the beampatterns of the DMA units are assumed to be equal, we have that $\mathcal{B}[\mathbf{w}_1(\omega),\theta] = \cdots = \mathcal{B}[\mathbf{w}_K(\omega),\theta] = \mathcal{B}_{\text{DMA}}(\omega,\theta)$, and eq. (28) simplifies to

$$\mathcal{B}[\mathbf{g}(\omega),\theta] = \mathbf{h}^H(\omega)\mathcal{B}_{\text{DMA}}(\omega,\theta)\bar{\mathbf{d}}(\omega,\theta). \tag{29}$$

Finally, since $\mathcal{B}_{\text{DMA}}(\omega,\theta)$ is a scalar quantity, we obtain

$$\begin{aligned}\mathcal{B}[\mathbf{g}(\omega),\theta] &= \mathcal{B}_{\text{DMA}}(\omega,\theta)\mathbf{h}^H(\omega)\bar{\mathbf{d}}(\omega,\theta)\\ &= \mathcal{B}_{\text{DMA}}(\omega,\theta)\mathcal{B}[\mathbf{h}(\omega),\theta],\end{aligned} \tag{30}$$

which shows that the global array beampattern is the product of the DMA unit beampattern and the virtual array beampattern.

Note, however, that DMA units with different geometries are characterized by different approximations to the same target beampattern, thus $\mathcal{B}[\mathbf{w}_1(\omega),\theta] \approx \cdots \approx \mathcal{B}[\mathbf{w}_K(\omega),\theta]$. Therefore, the zeros in each DMA directional response are shifted around their ideal positions. We will now show that even in this more general case, the beampattern approximation errors introduced by the DMA units are negligible when propagated to the global beamformer. As an example, Fig. 2 illustrates the beampattern at 1500Hz of a two-stage beamformer consisting of an arbitrary configuration of 10 DMA units with different geometries and different numbers of microphones, where the microphones are randomly placed as described in Section V-A. At the top of Fig. 2, we show the virtual array filter, designed as a DAS, and an *ideal* target beampattern to be approximated by the DMA units through eq. (14), i.e., a first-order hypercardioid $\mathcal{B}_{\text{ideal}}(\theta) = \sum_{n=0}^{N} a_{n,N}\cos(n\theta)$ with $\mathbf{a}_N = [0.3339, 0.6661]^T$. Below in Fig. 2 we show a comparison between the global array beampattern $\mathcal{B}[\mathbf{g}(\omega),\theta]$ and its *ideal* counterpart, which is given by the product $\mathcal{B}_{\text{ideal}}\mathcal{B}[\mathbf{h}(\omega),\theta]$. The result shows that the difference between the ideal beampattern product and the global array beampattern is negligible.

To further evaluate the accuracy with which the proposed two-stage beamformer matches the ideal beampattern product, we define the Mean Square Error (MSE) as

$$\begin{aligned}&\text{MSE}(\omega,\theta) =\\ &\qquad\frac{1}{N_r}\sum_{r=1}^{N_r}(|\mathcal{B}[\mathbf{g}^{(r)}(\omega),\theta]| - |\mathcal{B}_{\text{ideal}}(\theta)\mathcal{B}[\mathbf{h}^{(r)}(\omega),\theta]|)^2,\end{aligned}$$

where N_r is the number of realizations of arbitrary geometries, and the superscript $^{(r)}$ is used to denote the rth realization. Fig. 3 shows the MSE as a function of both frequency ω and angle θ. From the figure, it can be seen that the MSE is larger at the angles corresponding to the nulls of the target DMA unit beampattern. This effect is due to the error introduced by the DMA units in approximating the target beampattern at higher frequencies, where the nulls deviate slightly from their ideal positions. Considering that the beampattern values are dimensionless scalars in the range [0;1], the MSE lies in the same interval and we can see from the figure that the MSE

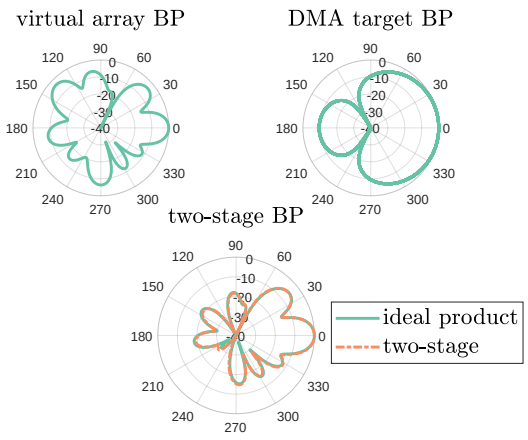

Fig. 2. Virtual array beampattern (top left), target DMA units beampattern (top-right), ideal beampattern product and resulting global array beampattern (bottom), at 1500 Hz. The virtual array filter is a DAS, while the target beampattern is a first-order hypercardioid.

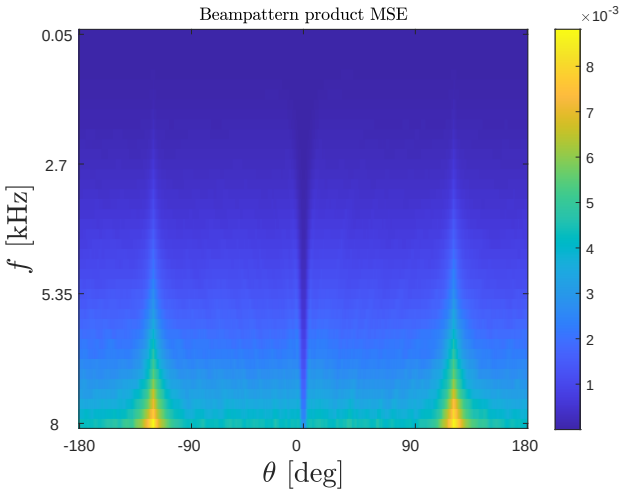

Fig. 3. MSE of the beampattern as a function of frequency $f = \omega/(2\pi)$ and angle θ, obtained through $N_r = 1000$ realizations of arbitrary geometry. The virtual array filter is a DAS, while the DMA units approximate a first-order hypercardioid through (14).

is in the order of 10^{-3} at all frequencies except for localized peaks of 10^{-2} in proximity of the nulls at high frequencies.

In accordance to the shown experiment, when considering two-stage beamformers with different DMA units approaching the same target beampattern, the difference between the global array beampattern and the ideal beampattern product is generally small. As a result, if we have both the DMA unit filter and the virtual array filter, we can estimate the global array beampattern with good accuracy by performing the product between the target DMA beampattern and the virtual array beampattern. Moreover, the fact that a null in the local array beampattern causes a null to appear in the global array beampattern simplifies the attenuation of unwanted interferences placed at known angular positions.

V. Performance Analysis

In this section, we analyze through simulations different design choices for the proposed two-stage beamforming system and evaluate them using the metrics discussed in Section II-C. In all the following analyses, we set the desired steering angle to $\theta_s = 0°$. We derive the beam shape coefficients of the DMA units according to the optimization strategy discussed in Subsection III-C, in which the matrix $\mathbf{A}(\omega, \theta_s)$ is computed according to eq. (25) with $\epsilon = 0.05$. The virtual array filter is either the DAS or the SD beamformer. We evaluate the WNG, the DF, and the FBR of the designed two-stage beamformers and we compare them with the same metrics of traditional single-stage DAS and SD beamformers.

A. Two-stage Beamformers with First-Order DMA Units

As a first subclass of two-stage beamformers, let us consider those with DMA units characterized by first-order beampatterns. Depending on the filter chosen for the virtual array (DAS or SD) and the optimization criterion used to design the beam shape coefficients of the DMA units (WNG maximization or DF maximization), we use the following naming convention to identify the two-stage beamformer designs: DAS-maxWNG, DAS-maxDF, SD-maxWNG, SD-maxDF.

We consider a system consisting of $K = 10$ randomly distributed DMA units, each characterized by an arbitrary geometry. The virtual array geometry is generated by distributing the reference points of DMA units according to the two uniform distributions $\mathrm{r}_k \sim \mathcal{U}(5\mathrm{cm}, 25\mathrm{cm})$ and $\Phi_k \sim \mathcal{U}(0, 2\pi)$, where symbol $\mathcal{U}(a, b)$ is used here to denote a continuous uniform distribution with boundaries a and b. Also, the DMA reference points are placed at least 10 cm apart. The microphones of the DMA units are randomly placed according to $r_{k,m} \sim \mathcal{U}(0.3\mathrm{cm}, 1.5\mathrm{cm})$ and $\phi_{k,m} \sim \mathcal{U}(0, 2\pi)$, with the sensors at least 3 mm apart. In addition, each DMA unit contains an arbitrary number of sensors drawn from the discrete uniform distribution $M_k \sim \mathrm{U}(4, 6)$. All the following analyses consider 1000 realizations of arbitrary geometry.

Fig. 4 shows the average values and the standard deviations of the beampatterns of the considered two-stage beamformers for all the four designs at some frequencies of interest (i.e., 500 Hz, 1 kHz, 4 kHz). The average beampatterns are represented with solid lines, while the standard deviations around the mean are shown by dashed lines. Fig. 5 shows WNG, DF, and FBR of the same beamformers as a function of frequency. The solid and dashed lines in the figure represent the averaged values, while the shaded areas represent the standard deviations of the results around the mean values. The results in Fig. 5 show that the considered two-stage beamformer is generally a good compromise between a DAS beamformer achieving the maximum WNG and the SD beamformer achieving the maximum DF at almost any frequency. As far as the FBR is concerned, on the other hand, the proposed beamforming method outperforms the traditional designs, especially at higher frequencies and for designs in which the WNG is maximized. Observing Fig. 4 and Fig. 5, we see that the distribution of the beampattern and the quality metrics around the mean value is narrow. This suggests that the placement of the microphones on the plane has a small impact on performance. We note that this is also true for the virtual array filter, whose contribution to the metrics is very limited compared to the choice of DMA unit filters, especially

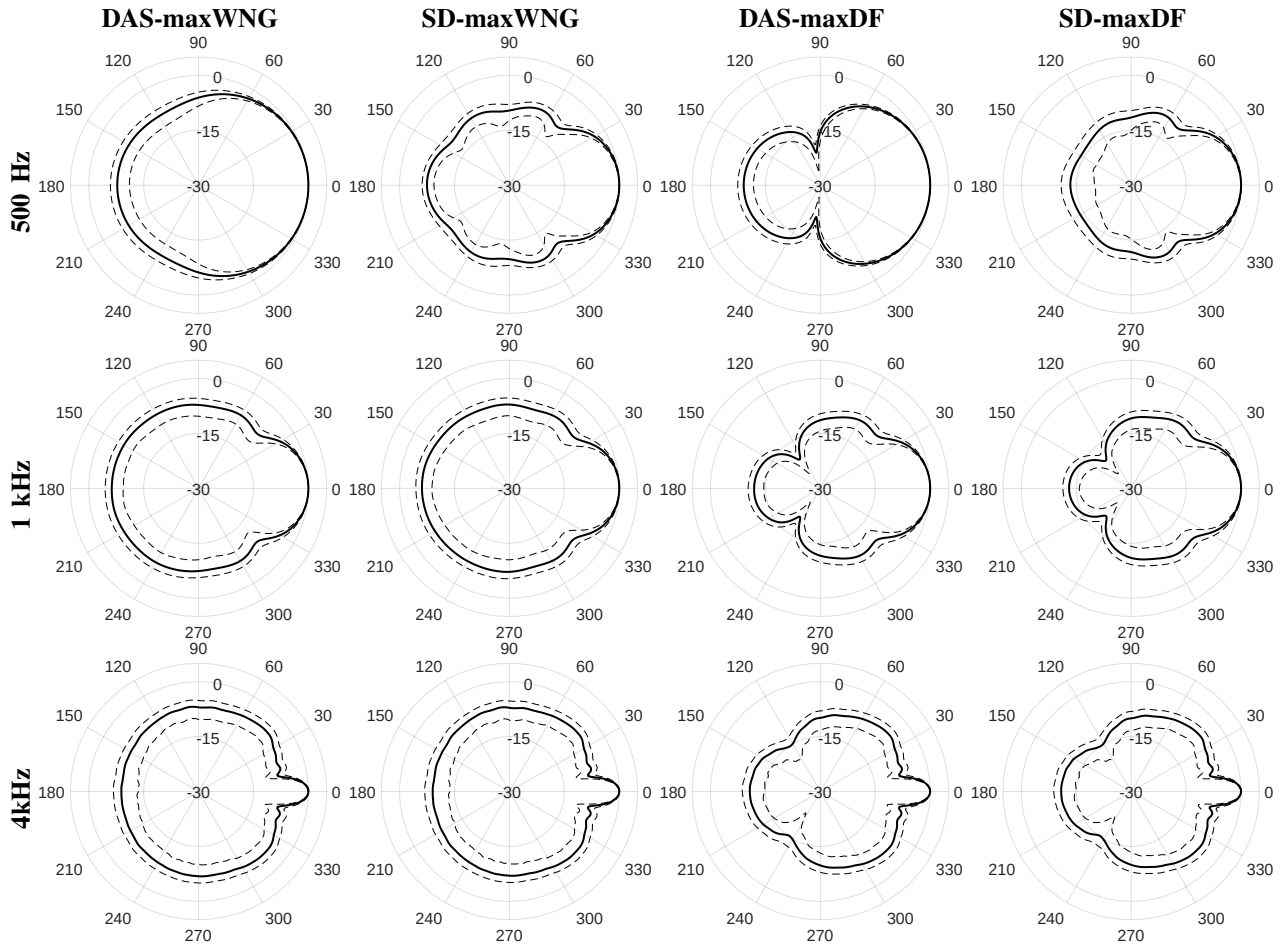

Fig. 4. Average beampattern of the first-order of two different designs of the two-stage beamformer with arbitrary geometry at three different frequency, i.e., 500 Hz, 1 kHz, 4 kHz.

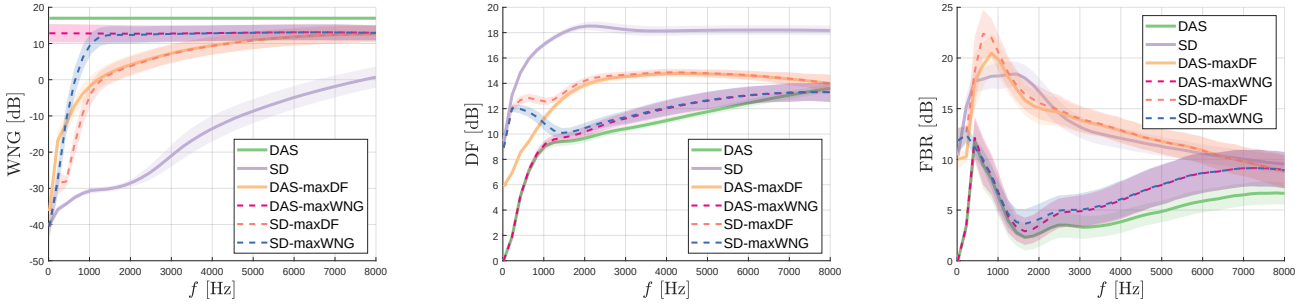

Fig. 5. Averaged global beamformer metrics (i.e., WNG, DF and FBR) of two-stage spatial filters using randomly distributed first-order DMA units with arbitrary geometry and based on four different designs.

at higher frequencies. For this reason, in the following we will only consider designs where the virtual array filter is a DAS beamformer because its implementation is generally less computationally expensive than the one of a SD beamformer.

B. Two-Stage Beamformers with Higher-Order DMA Units

In this subsection, we examine how the performance of the global beamformer benefits from the use of higher order DMA units. For comparison, we also consider zero-order DMA units computed as FIB-LSE beamformers with $N = 0$, which provide an approximation to an omnidirectional microphone regardless of local geometry. We consider the same system configuration with $K = 10$ DMAs as in the previous subsection, except that we increase the number of microphones in each DMA unit to allow for higher order DMA filters. Therefore, the number of sensors in each DMA unit is taken from the distribution $M_k \sim \mathrm{U}(8, 10)$. In this analysis, the virtual array filter is set as DAS beamformer. The spatial filter of the individual DMA units is optimized by maximizing the WNG or the DF, as in Sec. III-C.

Fig. 6 and Fig. 7 show the results in terms of WNG, DF, and FBR as a function of frequency. All the metrics are averaged

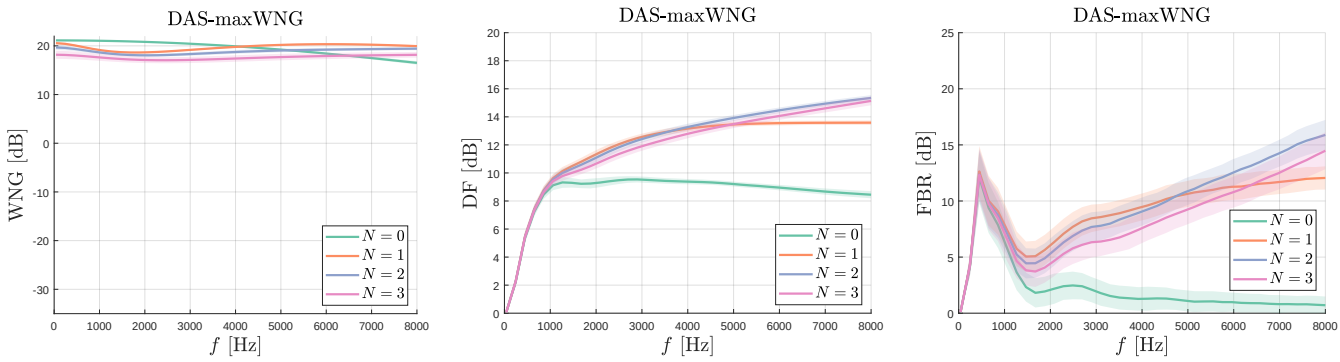

Fig. 6. Averaged global beamformer metrics (i.e., WNG, DF and FBR) for the maximum WNG two-stage beamformer design with a DAS as the virtual array, and N-order DMA units.

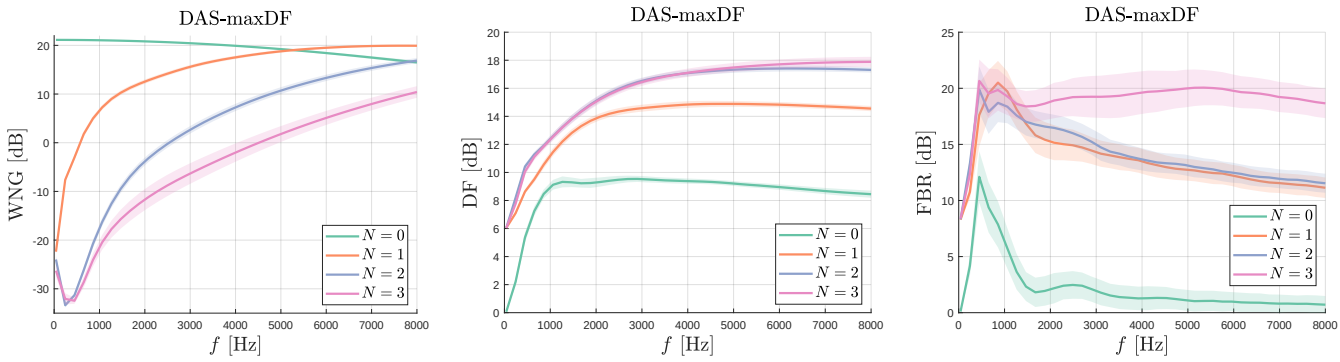

Fig. 7. Averaged global beamformer metrics (i.e., WNG, DF and FBR) for the maximum DF two-stage beamformer design with a DAS as the virtual array, and Nth-order DMA units.

quantities obtained over 1000 realisations of arbitrary geometries. Solid lines in the figures represent the averaged values, while the shaded areas represent the standard deviations of the results around the mean. It can be noticed that increasing the order generally lowers the WNG but increases the DF, and this effect is more pronounced for two-stage beamformers designed to maximize the DF. Conversely, the designs that maximize the WNG exhibit a performance difference that is less pronounced when the order of the DMA units is increased. Note also that zero-order designs do not change using different optimization criteria, as expected.

C. Frequency-Independent approach

The solution $\mathbf{a}_N(\omega)$ to the optimization problem in (21) depends on the angular frequency ω. Here, we propose a method to obtain frequency-independent DMA unit beam shape coefficients that yields results similar to a frequency-dependent design with a small performance reduction. Such a frequency-independent approach may be useful in contexts where the proposed two-stage beamformer is implemented in resource-constrained systems (e.g., embedded systems), since it requires fewer operations and a smaller number of coefficient updates.

There are several ways to obtain a vector $\bar{\mathbf{a}}_N$ of frequency-independent beam shape coefficients, e.g., by solving a modified optimization problem that considers all frequencies of interest at once, or by averaging the results of the frequency-dependent design over all frequencies. In this work, we opt for the latter option, since it leads to a two-stage beamformer that well approximates the performance of the frequency-dependent case. A further justification to this choice is illustrated in Fig. 8, where we show how the beam shape coefficients resulting from the optimization problem in (22) vary with frequency. In particular, Fig. 8 shows the coefficients in $\mathbf{a}_N(\omega)$ for the cases in which DMA units are first- and second-order beamformers maximizing either the WNG or the DF of the global two-stage spatial filter, and the virtual array filter is a DAS beamformer. Fig. 8 shows average results over 1000 realizations, considering the same simulation parameters of the previous subsection.

Interestingly when the DF is maximized, it can be seen from Fig. 8 that the coefficients approach a constant value above 1.5 kHz. Thus, we may investigate the extent to which a frequency-independent approach differs from a frequency-dependent approach in terms of array metrics.

The results in Fig. 9 and Fig. 10 show that the performance difference between the frequency-dependent design and frequency-independent design in terms of WNG and DF is small. Both configurations with first- and second-order DMA units are considered. The subscript "fi" is introduced to indicate the frequency-independent design. Results show that the frequency-independent approximation is close to the frequency-dependent design, especially when the optimization objective is to maximize the DF. Therefore, the frequency-independent approach is desirable whenever we want to sacrifice little performance in favor of efficient implementations

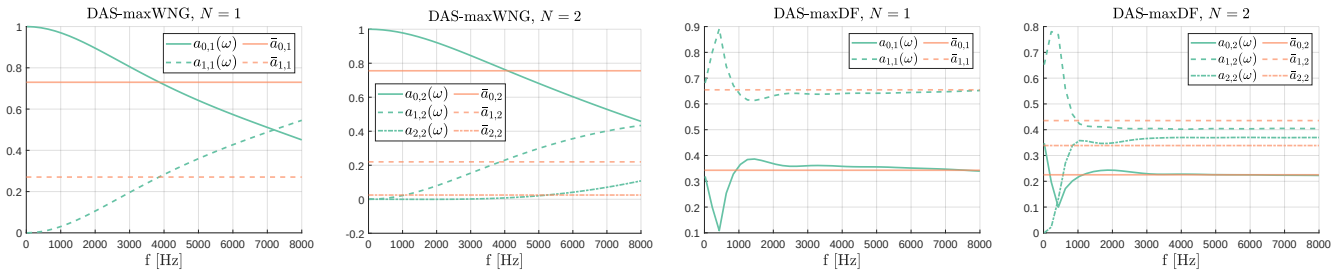

Fig. 8. Comparison between frequency-dependent beam shape coefficients resulting from the optimization problem in (21) and their average values over all frequencies.

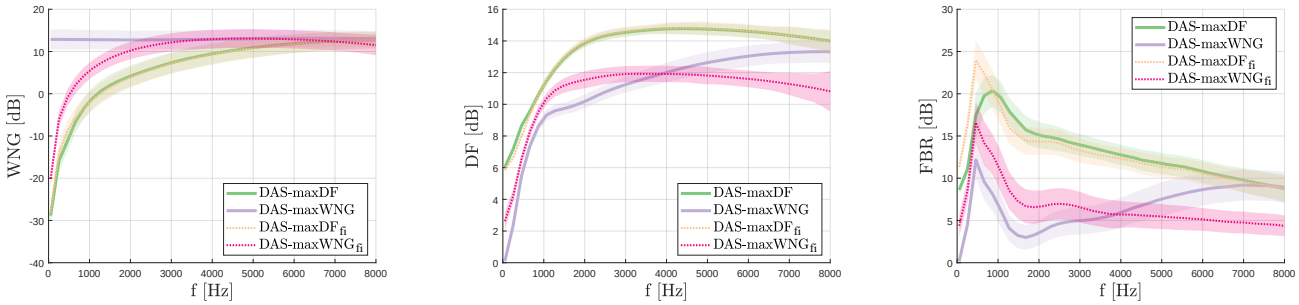

Fig. 9. Comparison between frequency-dependent and frequency-independent first-order two-stage beamformer metrics. Pedix f_i is used to denote frequency-independent designs obtained according to subsection V-C.

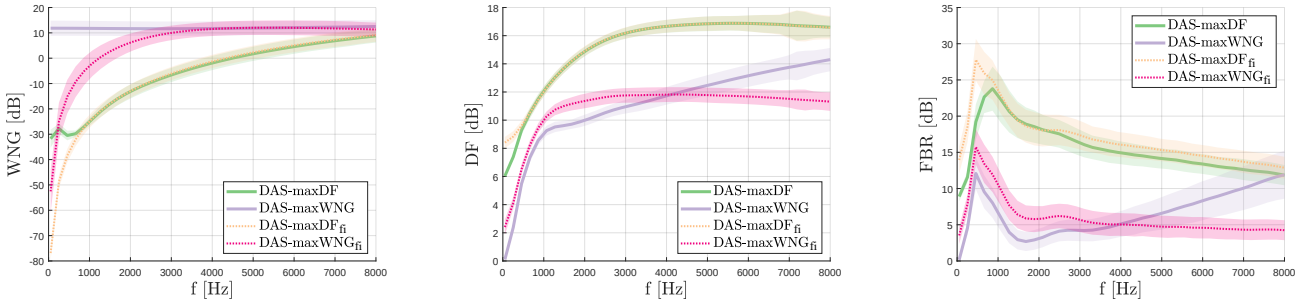

Fig. 10. Comparison between frequency-dependent and frequency-independent second-order two-stage beamformer metrics. Subscript "fi" is used to denote frequency-independent designs obtained according to subsection V-C

of the two-stage beamformer. In the next section we discuss some possible implementations of the sort.

VI. Implementations

In this section, we propose two possible implementations of the two-stage beamformer and compare them in terms of their computational complexity. In evaluating the computational cost, we consider the tasks of spatial filtering, steering, and beampattern switching. The latter consists of changing the directivity of the beamformer by updating the beam shape coefficients (e.g., using the optimal design strategies discussed in Sec. III-C). In this analysis, we estimate the computational cost as the number of multiplications required for each task. We also make the following assumptions:

- spatial filtering is performed in the frequency domain;
- only the frequency-independent design is considered;
- optimal beam shape coefficients $\mathbf{a}_N$ are computed in advance and stored in memory;
- DMA units use the same number of microphones, i.e., $M_1 = \cdots = M_k$;
- the virtual array filter is a DAS beamformer.

Two possible implementations of the two-stage beamformer are shown in Fig. 11, Fig. 12 and are referred to as *implementation A* and *implementation B*, respectively. Implementation A is based on the straightforward application of the global filter $\mathbf{g}(\omega)$ to all microphone signals. This implementation allows for efficient spatial filtering, which is trivially performed by computing the dot product $\mathbf{g}^H(\omega)\mathbf{y}(\omega)$ for each frequency bin. However, this comes at the cost of a high number of operations to recalculate the filters in order to change beampattern shape or to perform steering in a different direction. This is because to steer or change the beampattern, the global filter $\mathbf{g}(\omega)$ must be recomputed for each frequency bin, which in turn means that the local filters and the virtual filter must be updated in order to apply eq. (7).

Instead, implementation B is based on the global filter formulation in eq. (20), which collects the vector of beam

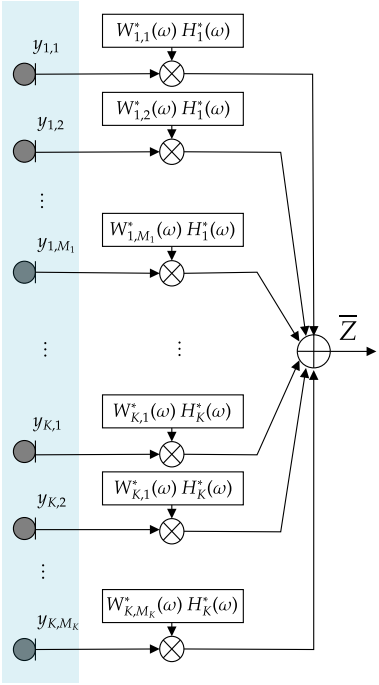

Fig. 11. Implementation A of the proposed two-stage beamformer.

shape coefficients and the matrix $\boldsymbol{\Upsilon}$, since they are the same for all DMA units. At the cost of higher computational cost for the filtering task, implementation B is more efficient at steering and changing beampattern, and it allows us to switch from a maximum WNG design to a maximum DF design by updating just $N+1$ coefficients. As for the steering, we only need to update $2N+1$ coefficients in $\boldsymbol{\Upsilon}(\theta_{\mathbf{s}})$ for steering all DMA units and then recalculate the virtual array filter for each frequency bin. Note that this implementation of the steering task does not require the beam shape coefficients to be updated. However, in case the beam shape coefficients are obtained by the optimization problem in Sec. III-C, they also depend on the steering angle θ_s. Therefore, we would use sub-optimal beam shape coefficients after steering. Through experiments we found that these sub-optimal beam shape coefficients are almost identical to those obtained by solving the optimization problems of Sec. III-C for the new desired steering angle. Thus, we can say that the beam shape coefficients are practically invariant with respect to the steering angle and we can avoid updating them.

We now propose a numerical analysis of the computational complexity of both implementations of the beamformer system by examining the cost of performing the aforementioned tasks. In this analysis, we consider $K = 10$ DMA units consisting of $2N+1$ microphones, with $N = 1, 2, 3$, and an FFT of size 128. The results are shown in Table I, where complexity is given as the estimated number of multiplications required for each task. The results show the trade-off between the cost of filtering and the cost of beampattern steering/switching in the two implementations. Indeed, implementation A is computationally cheaper than implementation B in terms of spatial filtering, but requires more computations for beampattern steering/switching and vice versa. Note also that implementation B has a fixed cost for the steering task, which is also independent of the order, and that multiplications are not required to change the beampattern.

In future application scenarios, we can envision DMA units with their own processors, so that the cost of local filtering is distributed in all DMAs and the execution time required for filtering is reduced. In this scenario, the execution time required by implementation B becomes comparable to that of implementation A, while maintaining the advantage of more efficient steering and beampattern switch. This implementation can be useful in any application where steering is performed frequently, such as source localization and tracking algorithms.

Note that the implementations of Fig. 11 and Fig. 12 are not the only options, and other computational schemes can be derived, since all the operations involved in the evaluation of the global filter output are linear and we can swap the order of such operations without affecting the final result. It is also worth noting that the proposed computational schemes can accommodate both frequency-dependent and frequency-independent beam shape DMA unit parameters.

VII. Conclusions and Future Works

In this article we have described some properties of a broad class of two-stage beamformers characterized by arbitrary planar configurations of multiple DMA units with arbitrary geometry and order, all having the same beampattern. A global spatial filter is created by combining the local spatial filters of the DMA units and a virtual array filter applied to their outputs. Decomposing a global array into local and virtual arrays has several advantages, both in terms of design flexibility and implementation.

Beampattern design flexibility stems from the separate design of the local and the virtual array filters, since the product of the corresponding beampatterns yields the beampattern of the global array, greatly simplifying the design of the latter. Also, fixed the virtual array beamformer as either DAS or SD, we presented a closed-form solution to an optimization problem aimed at deriving the beam shape coefficients of the DMA units that maximize the WNG or the DF of the resulting global beamformer at each frequency. We have shown that the choice of virtual filter type (DAS or SD) has little effect on the resulting global beamformer metrics, as does the actual microphone placement.

As far as the implementation of the presented beamforming systems is concerned, we have proposed an approach in which frequency-invariant beam shape coefficients of the DMA units are computed as averaged values over frequency. Moreover, we have discussed two computational schemes, one more efficient for spatial filtering and the other more efficient for beampattern steering/switching. It is also worth noticing that, as discussed in [16], if we restrict ourselves to first-order DMA units, it is possible to employ implementations of DMA units in both the frequency domain and the discrete-time domain that are computationally lighter than the ones used in this work.

As for future developments, we would like to extend the technique to DMA-based multi-stage beamformers with more

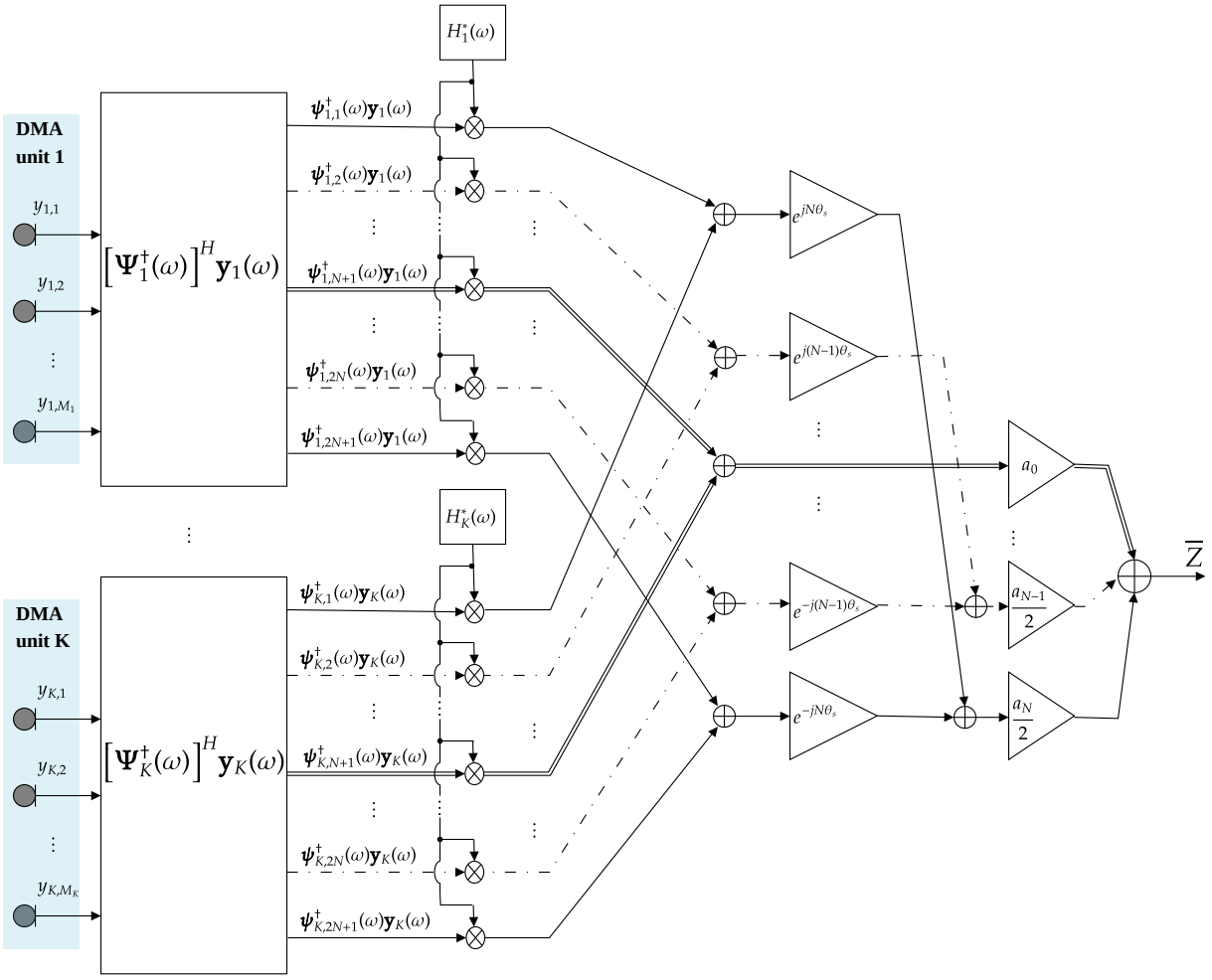

Fig. 12. Implementation B of the proposed two-stage beamformer. Terms $\psi_{k,i}^\dagger(\omega)$ are used to indicate the ith row of matrix $[\mathbf{\Psi}_k^\dagger(\omega)]^H$, with $1 \leq i \leq 2N+1$.

TABLE I
COMPUTATIONAL COSTS FOR THE SPATIAL FILTERING (FILTER), BEAMPATTERN STEERING (BP. STEER), AND BEAMPATTERN SWITCHING (BP. SWITCH) TASKS FOR BOTH IMPLEMENTATIONS A AND B. THE COSTS ARE EXPRESSED AS THE ESTIMATED NUMBER OF MULTIPLICATIONS.

N=1	Filter	BP. Steer	BP. Switch
Imp. A	3840	16644	15362
Imp. B	15872	1536	-

N=2	Filter	BP. Steer	BP. Switch
Imp. A	6400	39688	38404
Imp. B	39296	1536	-

N=3	Filter	BP. Steer	BP. Switch
Imp. A	8960	72972	71686
Imp. B	72960	1536	-

than two stages and explore their potential advantages in spatial audio processing applications.

REFERENCES

[1] G. W. Elko and J. Meyer, "Microphone arrays," in *Springer handbook of speech processing*. Springer, Jan. 2008, pp. 1021–1041.

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

[3] G. W. Elko and Anh-Tho N. P., "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.

[4] H. Teutsch and G. W. Elko, "First- and second- order adaptive differential microphone arrays," in *Proc. IWAENC*, Sep. 2001, pp. 35–38.

[5] M. Buck, "Aspects of first-order differential microphone arrays in the presence of sensor imperfections," *European Transactions on Telecommunications*, vol. 13, no. 1, pp. 115–122, Mar.-Apr. 2002.

[6] G. W. Elko, "Differential microphone arrays," in *Audio Signal Processing for Next-Generation Multimedia Communication Systems, Y. Huang and J. Benesty*, M. K. Norwell, Ed., 2004.

[7] 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.

[8] 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.

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

[10] 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.

[11] 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.

[12] 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.

[13] F. Borra, A. Bernardini, F. Antonacci, and A. Sarti, "Efficient implementations of first-order steerable differential microphone arrays with arbitrary planar geometry," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 1755–1766, 2020.

[14] X. Zhao, G. Huang, J. Benesty, J. Chen, and I. Cohen, "On the design of square differential microphone arrays with a multistage structure," in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 746–750.

[15] 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.

[16] F. Borra, A. Bernardini, I. Bertuletti, F. Antonacci, and A. Sarti, "Arrays of first-order steerable differential microphones," in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 751–755.

[17] G. Huang, I. Cohen, J. Benesty, and J. Chen, "Kronecker product beamforming with multiple differential microphone arrays," in *2020 IEEE 11th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2020, pp. 1–5.

[18] G. Itzhak, J. Benesty, and I. Cohen, "On the design of differential kronecker product beamformers," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 1397–1410, 2021.

[19] P. Stoica and R. Moses, "Spectral analysis of signals," *Prentice Hall*, p. 354, Jan. 2005.

[20] E. De 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, January 2012.

[21] G. W. Elko, "Microphone Array Systems for Hands-Free Telecommunication," *Speech Communication*, vol. 20, no. 3-4, pp. 229–240, Dec. 1996.