Differential Beamforming on Graphs

Gongping Huang, *Student Member, IEEE*, Jacob Benesty, Israel Cohen, *Fellow, IEEE*, and Jingdong Chen, *Senior Member, IEEE*

***Abstract*—In this article, we study differential beamforming from a graph perspective. The microphone array used for differential beamforming is viewed as a graph, where its sensors correspond to the nodes, the number of microphones corresponds to the order of the graph, and linear spatial difference equations among microphones are related to graph edges. Specifically, for the first-order differential beamforming with an array of M microphones, each pair of adjacent microphones are directly connected, resulting in $M-1$ spatial difference equations. On a graph, each of these equations corresponds to a 2-clique. For the second-order differential beamforming, each three adjacent microphones are directly connected, resulting in $M-2$ second-order spatial difference equations, and each of these equations corresponds to a 3-clique. In an analogous manner, the differential microphone array for any order of differential beamforming can be viewed as a graph. From this perspective, we then derive a class of differential beamformers, including the maximum white noise gain beamformer, the maximum directivity factor one, and optimal compromising beamformers. Simulations are presented to demonstrate the performance of the derived differential beamformers.**

***Index Terms*—Microphone arrays, differential beamforming, graphs, Laplacian matrix, incidence matrix, and adjacency matrix.**

I. Introduction

A GRAPH is a set of nodes connected via edges, whose structure is meaningful and different types of graphs can model different geometric structures of signals [1]–[4]. Since much signal/data can be organized around a graph, the development of graph signal processing has emerged and is gaining increasing attention [5]–[8]. Basically, graph signal processing aims at developing tools for processing data in the graph domain, which has demonstrated promises for many applications [9]–[11] as it is useful and efficient for handling data by taking into consideration both the signal and the graph structure [12], [13]. In this work, we attempt to address the problem of microphone array beamforming from the perspective of graph signal processing.

A microphone array is basically an array of sensors of the same type organized in a specified geometry, which is used to sample the sound field with spatial diversity [14]–[17]. The core of a microphone array is beamforming, whose objective is to recover the speech signal of interest from observations, which consist of not only the signal of interest but also additive noise, echo, interference, and reverberation in acoustic environments [18]–[23]. Numerous microphone array beamforming algorithms have been developed in the literature, including the delay-and-sum, filter-and-sum, adaptive [24], [25], and differential beamforming [26]–[30].

Among those, differential beamforming is generally used with closely spaced microphones so that the beamformer output measures the spatial derivatives of the acoustic pressure field [31], [32]. It offers several advantages in comparison with additive beamforming, including but not limited to: 1) it has the potential to achieve high directional gains and frequency-invariant beampatterns, and is therefore attractive for high fidelity sound acquisition [18], [26]; and 2) the associated array is small in size and can be easily integrated into a wide range of devices such as smart speakers, smart phones, smart home systems, and in-car communication systems, to name but a few.

Traditionally, a differential microphone array (DMA) is designed in a multistage manner [33], [34]. If it measures the first-order spatial derivative, then the resulting beamformer is called a first-order differential beamformer and the corresponding microphone array is called a first-order DMA. Similarly, if the beamformer output measures up to an nth-order spatial derivative of the acoustic pressure field, it is called an nth-order differential beamformer and the corresponding microphone array is called an nth-order DMA [18], [35]. However, this traditional DMA design method lacks flexibility in forming different patterns and the resulting DMAs often suffer from white noise amplification, which is very serious, particularly at low frequencies. An alternative design method is the recently-developed one in the short-time-Fourier-transform (STFT) domain based on null constraints [26], [36], in which the design flexibility and robustness of linear DMAs have been significantly improved. Also, some efforts have been devoted to improving the steering flexibility of DMAs and much progress has been made, such as two-dimensional linear DMAs [30], [37], circular DMAs (CDMAs) [38], [39], and concentric circular DMAs (CCDMAs) [29], [40]. Nevertheless, there are still many unsolved important issues with differential beamforming and further efforts are indispensable.

Manuscript received September 1, 2019; revised January 5, 2020; accepted February 7, 2020. Date of publication February 13, 2020; date of current version March 6, 2020. This work was supported in part by the Israel Science Foundation under Grant 576/16 and in part by the ISF-NSFC joint research program under Grants 2514/17 and 61761146001. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Stefan Bilbao. *(Corresponding author: Gongping Huang.)*

Gongping Huang and Israel Cohen are with the Andrew and Erna Viterby Faculty of Electrical Engineering, Technion – Israel Institute of Technology, Technion City, Haifa 3200003, Israel (e-mail: gongpinghuang@gmail.com; icohen@ee.technion.ac.il).

Jacob Benesty is with the INRS-EMT, University of Quebec, Montreal, QC H5A 1K6, Canada (e-mail: benesty@emt.inrs.ca).

Jingdong Chen is with the Center of Intelligent Acoustics and Immersive Communications, Northwestern Polytechnical University, Xi'an, Shaanxi 710072, China (e-mail: jingdongchen@ieee.org).

Digital Object Identifier 10.1109/TASLP.2020.2973795

From a graph signal processing perspective, any sensor array can be well represented by a graph, where each sensor corresponds to a node and connections among microphones in the beamforming process are the edges. This graph perspective of sensor arrays may provide new insights into beamforming as it does for several other signal processing problems. Therefore, this work attempts to explore the application of graph concepts to microphone array processing and the focus is on the design of differential beamformers. In this context, the connection between a DMA and the corresponding graph is as follows. Each DMA sensor corresponds to a graph node, the number of microphones corresponds to the order of the graph, and linear spatial derivatives deduced among microphones are related to graph edges. For instance, with a DMA of M microphones, every two adjacent microphones can form a first-order differential beamformer, resulting in $M-1$ spatial derivatives of the acoustic pressure field, each of which corresponds to an edge (or a 2-clique). This leads to a path graph. Similarly, every three neighboring microphones can form a second-order differential beamformer, resulting in $M-2$ second-order spatial derivatives, each of which corresponds to a 3-clique on a graph. In an analogous manner, a DMA of any order can be modeled as a graph. Given this graph, we can then obtain the corresponding Laplacian, incidence, and adjacency matrices. Different orders of spatial difference can be derived by multiplying the (modified) incidence matrix with the observed signal vector, based on which differential beamformers can be designed. To demonstrate this design process, we discuss in detail the derivation of the maximum white noise gain (MWNG) beamformer, the maximum directivity factor (MDF) beamformer, and the beamformer that can make compromises between a reasonable level of WNG and high DF. These are widely used differential beamformers.

The remainder of this paper is organized as follows. In Section II, we present the signal model, the problem of linear beamforming, and definitions of a few important metrics for beamforming performance evaluation. Some mathematical background on graphs is then provided in Section III. In Section IV, we address the problem of differential beamforming of different orders on graphs. The evaluation of the developed differential beamformers is presented in Section V and, finally, conclusions are made in Section VI.

II. Signal Model, Linear Beamforming, and Performance Measures

We consider an arbitrary sensor array composed of M omnidirectional microphones, which are distributed on a plane as illustrated in Fig. 1. Assuming that the reference point chosen on this plane coincides with the origin of the two-dimensional Cartesian coordinate system and the azimuthal angles are measured anti-clockwise from the x axis, the coordinates corresponding to the positions of the microphones are given by

$$\mathbf{p}_m = p_m \left[\cos\psi_m \ \sin\psi_m\right]^T, \tag{1}$$

for $m = 1, 2, \ldots, M$, where the superscript ^{T} is the transpose operator, p_m is the distance from the mth microphone to the origin point, and ψ_m is the angular position of the mth array element. As a result, the distance between Microphones i and j is

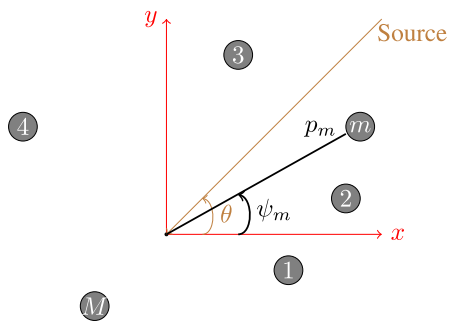


Fig. 1. A two-dimensional array with M omnidirectional microphones.

$$\delta_{ij} = \|\mathbf{p}_i - \mathbf{p}_j\|, \tag{2}$$

for $i, j = 1, 2, \ldots, M$, where $\|\cdot\|$ is the Euclidean norm.

Consider a farfield source signal (plane wave), on the same plane as the array, that propagates in an anechoic acoustic environment at the speed of sound, i.e., $c = 340$ m/s, and impinges on the described array. The direction of the source signal to the array is parameterized by the azimuthal angle θ. The time delay between the mth microphone and the origin is then given by

$$\tau_m(\theta) = \frac{p_m}{c}\cos(\theta - \psi_m), \ m = 1, 2, \ldots, M. \tag{3}$$

With this model, the steering vector of length M is expressed as [14], [15]

$$\mathbf{d}_\theta(f) = \left[e^{\jmath 2\pi f \tau_1(\theta)} \quad e^{\jmath 2\pi f \tau_2(\theta)} \quad \cdots \quad e^{\jmath 2\pi f \tau_M(\theta)}\right]^T, \tag{4}$$

where $f > 0$ is the temporal frequency and $\jmath$ is the imaginary unit.

Let us assume that the desired source signal propagates from the known angle $\theta = \theta_{\rm s}$. Then, the observed signal vector of length M can be expressed in the frequency domain as [17]

$$\begin{aligned}\mathbf{y}(f) &= \left[Y_1(f)\ Y_2(f) \cdots\ Y_M(f)\right]^T \\ &= \mathbf{d}_{\theta_{\rm s}}(f)X(f) + \mathbf{v}(f),\end{aligned} \tag{5}$$

where $Y_m(f)$ is the mth ($m = 1, 2, \ldots, M$) microphone signal of the planar array, $\mathbf{d}_{\theta_{\rm s}}(f)$ is the steering vector at $\theta = \theta_{\rm s}$, $X(f)$ is the (zero-mean) desired signal, and $\mathbf{v}(f)$ is the (zero-mean) additive noise signal vector defined similarly to $\mathbf{y}(f)$. In the rest, in order to simplify the notation, we often drop the dependence on the frequency, f. So, for example, (5) is written as $\mathbf{y} = \mathbf{d}_{\theta_{\rm s}}X + \mathbf{v}$. The covariance matrix of $\mathbf{y}$ is then

$$\mathbf{\Phi}_{\mathbf{y}} = E\left(\mathbf{y}\mathbf{y}^H\right) = \phi_X \mathbf{d}_{\theta_{\rm s}}\mathbf{d}_{\theta_{\rm s}}^H + \mathbf{\Phi}_{\mathbf{v}}, \tag{6}$$

where $E(\cdot)$ denotes mathematical expectation, the superscript ^{H} is the conjugate-transpose operator, $\phi_X = E(|X|^2)$ is the variance of X, and $\mathbf{\Phi}_{\mathbf{v}} = E(\mathbf{v}\mathbf{v}^H)$ is the covariance matrix of $\mathbf{v}$. Assuming that the variance of the noise is the same at all sensors, i.e., $\phi_V = \phi_{V_1} = \phi_{V_2} = \cdots = \phi_{V_M}$, with $\phi_{V_m} = E(|V_m|^2)$, $m = 1, 2, \ldots, M$, we can express (6) as

$$\mathbf{\Phi}_{\mathbf{y}} = \phi_X \mathbf{d}_{\theta_{\rm s}}\mathbf{d}_{\theta_{\rm s}}^H + \phi_V \mathbf{\Gamma}_{\mathbf{v}}, \tag{7}$$

where $\mathbf{\Gamma_v} = \mathbf{\Phi_v}/\phi_V$ is the pseudo-coherence matrix of the noise. In the case of the spherically isotropic (diffuse) noise field, which will always be assumed in this work, (7) becomes

$$\mathbf{\Phi_y} = \phi_X \mathbf{d}_{\theta_s} \mathbf{d}_{\theta_s}^H + \phi_d \mathbf{\Gamma}_d, \tag{8}$$

where ϕ_d is the variance of the diffuse noise and $\mathbf{\Gamma}_d$ is the pseudo-coherence matrix of the diffuse noise, whose (i,j)th $(i,j = 1,2,\ldots,M)$ element is

$$[\mathbf{\Gamma}_d(f)]_{ij} = \text{sinc}\left(\frac{2\pi f \delta_{ij}}{c}\right), \tag{9}$$

with $\text{sinc}(x) = \sin x/x$, and δ_{ij} being defined in (2).

Conventional beamforming is performed by applying a complex-valued linear filter, $\mathbf{h}$ of length M, to the observed signal vector, i.e.,

$$Z = \mathbf{h}^H \mathbf{y} = X_{fd} + V_{rn}, \tag{10}$$

where Z is the beamformer output signal (i.e., the estimate of X), $X_{fd} = X\mathbf{h}^H \mathbf{d}_{\theta_s}$ is the filtered desired signal, and $V_{rn} = \mathbf{h}^H \mathbf{v}$ is the residual noise. Since the two terms on the right-hand side of (10) are incoherent, the variance of Z is the sum of two variances:

$$\phi_Z = \mathbf{h}^H \mathbf{\Phi_y} \mathbf{h} = \phi_{X_{fd}} + \phi_{V_{rn}}, \tag{11}$$

where $\phi_{X_{fd}} = \phi_X |\mathbf{h}^H \mathbf{d}_{\theta_s}|^2$ and $\phi_{V_{rn}} = \mathbf{h}^H \mathbf{\Phi_v} \mathbf{h}$. In the context of fixed beamforming, the distortionless constraint is almost always desired, i.e.,

$$\mathbf{h}^H \mathbf{d}_{\theta_s} = 1, \tag{12}$$

meaning that any signal arriving along $\mathbf{d}_{\theta_s}$ will pass through the beamformer undistorted.

Next, we give several important performance measures, which are always used to evaluate all kinds of fixed and differential beamformers.

The beampattern describes the sensitivity of the beamformer to a plane wave (source signal) impinging on the array from the direction θ. Mathematically, it is defined as

$$\mathcal{B}_\theta(\mathbf{h}) = \mathbf{d}_\theta^H \mathbf{h}. \tag{13}$$

From (7), we see that the input signal-to-noise ratio (SNR) is

$$\text{iSNR} = \frac{\phi_X}{\phi_V}. \tag{14}$$

According to (11), the output SNR can be defined as

$$\text{oSNR}(\mathbf{h}) = \frac{\phi_{X_{fd}}}{\phi_{V_{rn}}} = \frac{\phi_X |\mathbf{h}^H \mathbf{d}_{\theta_s}|^2}{\mathbf{h}^H \mathbf{\Phi_v} \mathbf{h}} = \frac{\phi_X}{\phi_V} \times \frac{|\mathbf{h}^H \mathbf{d}_{\theta_s}|^2}{\mathbf{h}^H \mathbf{\Gamma_v} \mathbf{h}}. \tag{15}$$

From the previous definitions of the SNRs, we deduce the array gain:

$$\mathcal{G}(\mathbf{h}) = \frac{\text{oSNR}(\mathbf{h})}{\text{iSNR}} = \frac{|\mathbf{h}^H \mathbf{d}_{\theta_s}|^2}{\mathbf{h}^H \mathbf{\Gamma_v} \mathbf{h}}. \tag{16}$$

A common way to evaluate the sensitivity of the array to some of its imperfections, such as sensor noise, with a specific beamformer, is via the so-called white noise gain (WNG), which is defined by taking $\mathbf{\Gamma_v} = \mathbf{I}_M$ in (16), where $\mathbf{I}_M$ is the $M \times M$ identity matrix, i.e.,

$$\mathcal{W}(\mathbf{h}) = \frac{|\mathbf{h}^H \mathbf{d}_{\theta_s}|^2}{\mathbf{h}^H \mathbf{h}}. \tag{17}$$

The maximization of $\mathcal{W}(\mathbf{h})$ gives the conventional delay-and-sum (DS) beamformer [17]:

$$\mathbf{h}_{DS} = \frac{\mathbf{d}_{\theta_s}}{M}, \tag{18}$$

with $\mathcal{W}(\mathbf{h}_{DS}) = M$. While the DS beamformer maximizes the WNG, it never amplifies the diffuse noise. However, $\mathbf{h}_{DS}$ is not very directional and its beampattern may be very much frequency dependent.

Another important measure, which quantifies the ability of the beamformer to suppress spatial noise from directions other than the look direction is the directivity factor (DF), which is obtained by replacing $\mathbf{\Gamma_v}$ with $\mathbf{\Gamma}_d$ in (16), i.e.,

$$\mathcal{D}(\mathbf{h}) = \frac{|\mathbf{h}^H \mathbf{d}_{\theta_s}|^2}{\mathbf{h}^H \mathbf{\Gamma}_d \mathbf{h}}. \tag{19}$$

The maximization of $\mathcal{D}(\mathbf{h})$ leads to the well-known maximum DF (MDF) beamformer [17], [41]:

$$\mathbf{h}_{MDF} = \frac{\mathbf{\Gamma}_d^{-1} \mathbf{d}_{\theta_s}}{\mathbf{d}_{\theta_s}^H \mathbf{\Gamma}_d^{-1} \mathbf{d}_{\theta_s}}, \tag{20}$$

with $\mathcal{D}(\mathbf{h}_{MDF}) = \mathbf{d}_{\theta_s}^H \mathbf{\Gamma}_d^{-1} \mathbf{d}_{\theta_s}$. While the MDF beamformer maximizes the DF, it may amplify the white noise, especially at low frequencies, i.e., $\mathcal{W}(\mathbf{h}_{MDF}) \leq 1$.

III. Mathematical Background on Graphs

A graph is a mathematical structure composed of nodes (or vertices) and edges. An edge connects a pair of vertices; they are said to be adjacent in this case. A graph can be directed or undirected. The edges of directed graphs (or digraphs) have orientation but those of undirected graphs have no orientation. An undirected graph is connected if it is all in one piece. In the so-called simple graphs, no self-loops are allowed and there is at most one edge from one node to another [1]. Multigraphs have at least two parallel edges between some pair of nodes but have no self-loops [3]. In this paper, we are only interested in connected, undirected simple graphs.

More formally, an undirected simple graph is a pair $\mathsf{G} = (\mathsf{V}, \mathsf{E})$, where $\mathsf{V} = \{\mathsf{v}_1, \mathsf{v}_2, \ldots, \mathsf{v}_M\}$ is a set of M vertices and E is a set of two-elements subsets of V, i.e., $\{\mathsf{v}_i, \mathsf{v}_j\}$ with $\mathsf{v}_i, \mathsf{v}_j \in \mathsf{V}$, $\mathsf{v}_i \neq \mathsf{v}_j$, called edges, and for every $\{\mathsf{v}_i, \mathsf{v}_j\}$, there is at most one edge between v_i and v_j. When v_i is adjacent to v_j, we use the notation $\mathsf{v}_i \sim \mathsf{v}_j$. Again, in an undirected graph, $\mathsf{v}_i \sim \mathsf{v}_j$ and $\mathsf{v}_j \sim \mathsf{v}_i$ are the same. For a set of Q edges, we may write $\mathsf{E} = \{\mathsf{e}_1, \mathsf{e}_2, \ldots, \mathsf{e}_Q\}$. In connected, undirected simple graphs, we always have

$$M - 1 \leq Q \leq \frac{M(M-1)}{2}. \tag{21}$$

Sometimes, it may be more convenient to use the notation $\mathsf{V}(\mathsf{G})$ and $\mathsf{E}(\mathsf{G})$ to emphasize that these two sets are associated with the graph G.

The degree of a vertex v_m of G, denoted $\deg(\mathsf{v}_m)$, is the number of edges incident to v_m. It can be shown that, for any

graph, we have

$$\sum_{m=1}^{M} \deg(\mathsf{v}_m) = 2|\mathsf{E}|, \tag{22}$$

where $|\mathsf{E}| = Q$ is the number of edges, i.e., the size, of G while $|\mathsf{V}| = M$ is the order of the graph. Combining all degrees together, we can form the $M \times M$ diagonal matrix:

$$\mathbf{D} = \text{diag}\left[\deg(\mathsf{v}_1), \deg(\mathsf{v}_2), \ldots, \deg(\mathsf{v}_M)\right], \tag{23}$$

which is called the degree matrix.

The adjacency matrix, $\mathbf{A}$, of G is the $M \times M$ symmetric matrix whose elements are given by

$$A_{ij} = \begin{cases} 1 & \text{if } \mathsf{v}_i \sim \mathsf{v}_j \\ 0 & \text{otherwise.} \end{cases} \tag{24}$$

Mathematically, the graph G can be represented by the matrix $\mathbf{A}$. Then, it is easy to verify that

$$\deg(\mathsf{v}_m) = \sum_{j=1}^{M} A_{mj}. \tag{25}$$

A matrix of great importance in graph analysis is the Laplacian matrix [2] obtained from the two previous matrices:

$$\mathbf{L} = \mathbf{D} - \mathbf{A}. \tag{26}$$

It is clear that $\mathbf{L}$ is symmetric. Let $\mathbf{u}$ be a complex-valued vector of length M, it can be verified that

$$\mathbf{u}^H \mathbf{L} \mathbf{u} = \frac{1}{2} \sum_{i,j,\mathsf{v}_i \sim \mathsf{v}_j} |U_i - U_j|^2, \tag{27}$$

where $U_m,\ m = 1, 2, \ldots, M$ are the components of $\mathbf{u}$. As a consequence, $\mathbf{L}$ is positive semidefinite and all its eigenvalues are real valued and nonnegative. Furthermore, the smallest eigenvalue of $\mathbf{L}$ is equal to 0, and $\mathbf{1}/\sqrt{M}$ is its corresponding eigenvector, where $\mathbf{1}$ is a vector of length M containing only one's.

Another very useful matrix in graphs is the incidence matrix [4], $\mathbf{B}$, which is of size $M \times Q$, where its qth column corresponds to the edge e_q and has exactly two nonzero values: $+1$ in the row corresponding to one vertex and -1 in the row corresponding to the other vertex. The signs here are completely arbitrary but it is important that each column has one positive value and one negative one. The incidence matrix quantifies the variation of a signal on a graph. Then, one can check that

$$\mathbf{L} = \mathbf{B}\mathbf{B}^T. \tag{28}$$

It is seen that all these matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{D}$, and $\mathbf{L}$ are somehow related.

To end this section, let us give two more definitions. We say that H is a subgraph of G if $\mathsf{V}(\mathsf{H}) \subseteq \mathsf{V}(\mathsf{G})$ and $\mathsf{E}(\mathsf{H}) \subseteq \mathsf{E}(\mathsf{G})$. A clique is a complete[1] subgraph of a graph and we define an i-clique as a clique with i vertices.

[1]In a complete graph, each pair of distinct vertices are adjacent.

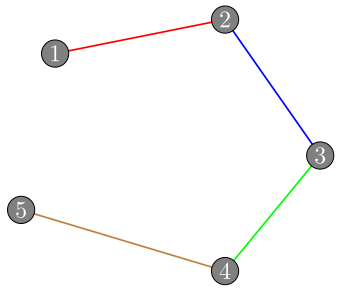


Fig. 2. First-order differential beamforming on a graph with $M = 5$.

IV. Differential Beamforming on Graphs

In our context, we may consider the microphone array as a graph, whose set of nodes corresponds to the sensors composing the array with the order of the graph being the number of microphones, i.e., M. The set of edges, as it will be explained below, corresponds to the number of linear spatial difference equations among microphones and their orders. In the rest of this section, in order to be as clear as possible, we often illustrate the concepts with the example of $M = 5$. Generalization to any value of M is straightforward. Next, we discuss differential beamforming of different orders separately.

A. First Order

In first-order differential beamforming, Microphones m and $m+1$ are directly connected for $m = 1, 2, \ldots, M-1$, resulting in $M-1$ spatial difference equations. On a graph, each one of these equations corresponds to an edge (or to a 2-clique). Fig. 2 shows an example with $M = 5$, where each color corresponds to a different first-order difference equation connecting two consecutive microphones. As a matter of fact, this simple graph is the well-known path graph, P_M. We recall that a path graph is a connected graph, with $M-1$ edges, where 2 vertices are of degree 1 and the other $M-2$ vertices are of degree 2.

The adjacency matrix (for $M = 5$) is

$$\mathbf{A}_{(1)} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \tag{29}$$

where the subscript $_{(1)}$ denotes first order. We deduce that the Laplacian matrix is

$$\mathbf{L}_{(1)} = \mathbf{D}_{(1)} - \mathbf{A}_{(1)}, \tag{30}$$

where

$$\mathbf{D}_{(1)} = \text{diag}(1, 2, 2, 2, 1). \tag{31}$$

Since we have four edges, the incidence matrix (of size 5×4) is

$$\mathbf{B}_{(1)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix} \tag{32}$$

and, of course, $\mathbf{L}_{(1)} = \mathbf{B}_{(1)}\mathbf{B}_{(1)}^T$. Now, if we multiply $\mathbf{B}_{(1)}^T$ by the observed signal, $\mathbf{y}$, we find that

$$\begin{aligned}\mathbf{B}_{(1)}^T\mathbf{y} &= \left[\, Y_1 - Y_2 \; Y_2 - Y_3 \; Y_3 - Y_4 \; Y_4 - Y_5 \,\right]^T \\ &= \left[\, Y_{(1),1} \; Y_{(1),2} \; Y_{(1),3} \; Y_{(1),4} \,\right]^T \\ &= \mathbf{y}_{(1)}. \end{aligned} \tag{33}$$

Therefore, for any M, our new observed signal (of length $M-1$) is

$$\begin{aligned}\mathbf{y}_{(1)} &= \mathbf{B}_{(1)}^T\mathbf{y} \\ &= \mathbf{x}_{(1)} + \mathbf{v}_{(1)} \\ &= \mathbf{B}_{(1)}^T\mathbf{d}_{\theta_\mathrm{s}}X + \mathbf{B}_{(1)}^T\mathbf{v}, \end{aligned} \tag{34}$$

where the mth ($m = 1, 2, \ldots, M-1$) component of $\mathbf{y}_{(1)}$ is $Y_{(1),m} = Y_m - Y_{m+1}$, and the first-order differential beamformer output signal is

$$Z_{(1)} = \mathbf{h}_{(1)}^H\mathbf{y}_{(1)}, \tag{35}$$

where $\mathbf{h}_{(1)}$ is the first-order differential filter of length $M-1$. As a result, the beampattern, WNG, and DF are, respectively,

$$\mathcal{B}_\theta\left(\mathbf{h}_{(1)}\right) = \mathbf{d}_\theta^H\mathbf{B}_{(1)}\mathbf{h}_{(1)}, \tag{36}$$

$$\mathcal{W}\left(\mathbf{h}_{(1)}\right) = \frac{\left|\mathbf{h}_{(1)}^H\mathbf{B}_{(1)}^T\mathbf{d}_{\theta_\mathrm{s}}\right|^2}{\mathbf{h}_{(1)}^H\mathbf{B}_{(1)}^T\mathbf{B}_{(1)}\mathbf{h}_{(1)}}, \tag{37}$$

$$\mathcal{D}\left(\mathbf{h}_{(1)}\right) = \frac{\left|\mathbf{h}_{(1)}^H\mathbf{B}_{(1)}^T\mathbf{d}_{\theta_\mathrm{s}}\right|^2}{\mathbf{h}_{(1)}^H\mathbf{B}_{(1)}^T\mathbf{\Gamma}_\mathrm{d}\mathbf{B}_{(1)}\mathbf{h}_{(1)}}. \tag{38}$$

From the maximization of the WNG, we get the first-order differential maximum WNG (MWNG) beamformer:

$$\mathbf{h}_{(1),\mathrm{MWNG}} = \frac{\left(\mathbf{B}_{(1)}^T\mathbf{B}_{(1)}\right)^{-1}\mathbf{B}_{(1)}^T\mathbf{d}_{\theta_\mathrm{s}}}{\mathbf{d}_{\theta_\mathrm{s}}^H\mathbf{B}_{(1)}\left(\mathbf{B}_{(1)}^T\mathbf{B}_{(1)}\right)^{-1}\mathbf{B}_{(1)}^T\mathbf{d}_{\theta_\mathrm{s}}}. \tag{39}$$

From the maximization of the DF, we obtain the first-order differential maximum DF (MDF) beamformer:

$$\mathbf{h}_{(1),\mathrm{MDF}} = \frac{\left(\mathbf{B}_{(1)}^T\mathbf{\Gamma}_\mathrm{d}\mathbf{B}_{(1)}\right)^{-1}\mathbf{B}_{(1)}^T\mathbf{d}_{\theta_\mathrm{s}}}{\mathbf{d}_{\theta_\mathrm{s}}^H\mathbf{B}_{(1)}\left(\mathbf{B}_{(1)}^T\mathbf{\Gamma}_\mathrm{d}\mathbf{B}_{(1)}\right)^{-1}\mathbf{B}_{(1)}^T\mathbf{d}_{\theta_\mathrm{s}}}. \tag{40}$$

The MDF beamformer will lead to a high value of the DF but at the expense of some white noise amplification, while the MWNG beamformer will lead to a large value of the WNG but at the expense of some low value of the DF. Therefore, in practice, it is important to be able to compromise between WNG and DF. To do so, we propose to exploit the joint diagonalization technique of two Hermitian matrices.

The two Hermitian matrices $\mathbf{B}_{(1)}^T\mathbf{d}_{\theta_\mathrm{s}}\mathbf{d}_{\theta_\mathrm{s}}^H\mathbf{B}_{(1)}$ and $\mathbf{B}_{(1)}^T\mathbf{\Gamma}_\mathrm{d}\mathbf{B}_{(1)}$, which appear in the definition of the DF, can be jointly diagonalized as follows [42]:

$$\mathbf{T}_{(1)}^H\mathbf{B}_{(1)}^T\mathbf{d}_{\theta_\mathrm{s}}\mathbf{d}_{\theta_\mathrm{s}}^H\mathbf{B}_{(1)}\mathbf{T}_{(1)} = \mathbf{\Lambda}_{(1)}, \tag{41}$$

$$\mathbf{T}_{(1)}^H\mathbf{B}_{(1)}^T\mathbf{\Gamma}_\mathrm{d}\mathbf{B}_{(1)}\mathbf{T}_{(1)} = \mathbf{I}_{M-1}, \tag{42}$$

where

$$\mathbf{T}_{(1)} = \left[\, \mathbf{t}_{(1),1} \; \mathbf{t}_{(1),2} \cdots \; \mathbf{t}_{(1),M-1} \,\right] \tag{43}$$

is a full-rank square matrix of size $(M-1)\times(M-1)$,

$$\begin{aligned}\mathbf{t}_{(1),1} &= \frac{\left(\mathbf{B}_{(1)}^T\mathbf{\Gamma}_\mathrm{d}\mathbf{B}_{(1)}\right)^{-1}\mathbf{B}_{(1)}^T\mathbf{d}_{\theta_\mathrm{s}}}{\sqrt{\mathbf{d}_{\theta_\mathrm{s}}^H\mathbf{B}_{(1)}\left(\mathbf{B}_{(1)}^T\mathbf{\Gamma}_\mathrm{d}\mathbf{B}_{(1)}\right)^{-1}\mathbf{B}_{(1)}^T\mathbf{d}_{\theta_\mathrm{s}}}} \\ &= \sqrt{\lambda_{(1),1}}\mathbf{h}_{(1),\mathrm{MDF}} \end{aligned} \tag{44}$$

is the first eigenvector of the matrix $(\mathbf{B}_{(1)}^T\mathbf{\Gamma}_\mathrm{d}\mathbf{B}_{(1)})^{-1}\mathbf{B}_{(1)}^T\mathbf{d}_{\theta_\mathrm{s}}\mathbf{d}_{\theta_\mathrm{s}}^H\mathbf{B}_{(1)}$,

$$\mathbf{\Lambda}_{(1)} = \mathrm{diag}\left(\lambda_{(1),1}, 0, \ldots, 0\right) \tag{45}$$

is a diagonal matrix of size $(M-1)\times(M-1)$,

$$\lambda_{(1),1} = \mathbf{d}_{\theta_\mathrm{s}}^H\mathbf{B}_{(1)}\left(\mathbf{B}_{(1)}^T\mathbf{\Gamma}_\mathrm{d}\mathbf{B}_{(1)}\right)^{-1}\mathbf{B}_{(1)}^T\mathbf{d}_{\theta_\mathrm{s}} \tag{46}$$

is the only nonnull eigenvalue of $(\mathbf{B}_{(1)}^T\mathbf{\Gamma}_\mathrm{d}\mathbf{B}_{(1)})^{-1}\mathbf{B}_{(1)}^T\mathbf{d}_{\theta_\mathrm{s}}\mathbf{d}_{\theta_\mathrm{s}}^H\mathbf{B}_{(1)}$, whose corresponding eigenvector is $\mathbf{t}_{(1),1}$, and $\mathbf{I}_{M-1}$ is the $(M-1)\times(M-1)$ identity matrix. The vectors $\mathbf{t}_{(1),i}$ ($i \neq 1$) are eigenvectors corresponding to the null eigenvalue. It can be checked from (41) that

$$\mathbf{t}_{(1),i}^H\mathbf{B}_{(1)}^T\mathbf{d}_{\theta_\mathrm{s}} = 0, \; i = 2, 3, \ldots, M-1. \tag{47}$$

Let us define the matrix of size $(M-1)\times P$:

$$\mathbf{T}_{(1),1:P} = \left[\, \mathbf{t}_{(1),1} \; \mathbf{t}_{(1),2} \cdots \; \mathbf{t}_{(1),P} \,\right], \tag{48}$$

with $1 \leq P \leq M-1$. The beamformer that we consider has the form:

$$\mathbf{h}_{(1),1:P} = \mathbf{T}_{(1),1:P}\boldsymbol{\alpha}, \tag{49}$$

where

$$\boldsymbol{\alpha} = \left[\, \alpha_1 \; \alpha_2 \cdots \; \alpha_P \,\right]^T \neq \mathbf{0} \tag{50}$$

is a vector of length P. Substituting (49) into (35), we find that

$$\begin{aligned}Z_{(1)} &= \boldsymbol{\alpha}^H\mathbf{T}_{(1),1:P}^H\mathbf{B}_{(1)}^T\mathbf{d}_{\theta_\mathrm{s}}X + \boldsymbol{\alpha}^H\mathbf{T}_{(1),1:P}^H\mathbf{B}_{(1)}^T\mathbf{v} \\ &= \alpha_1^*\sqrt{\lambda_{(1),1}}X + \boldsymbol{\alpha}^H\mathbf{T}_{(1),1:P}^H\mathbf{B}_{(1)}^T\mathbf{v}, \end{aligned} \tag{51}$$

where the superscript * is the complex-conjugate operator. Since the distortionless constraint is desired, it is clear from the previous expression that we always choose

$$\alpha_1 = \frac{1}{\sqrt{\lambda_{(1),1}}}. \tag{52}$$

Now, we need to determine the other elements of $\boldsymbol{\alpha}$. We can express the WNG as

$$\begin{aligned}\mathcal{W}\left(\mathbf{h}_{(1),1:P}\right) &= \frac{\left|\mathbf{h}_{(1),1:P}^H\mathbf{B}_{(1)}^T\mathbf{d}_{\theta_\mathrm{s}}\right|^2}{\mathbf{h}_{(1),1:P}^H\mathbf{B}_{(1)}^T\mathbf{B}_{(1)}\mathbf{h}_{(1),1:P}} \\ &= \frac{\left|\boldsymbol{\alpha}^H\mathbf{T}_{(1),1:P}^H\mathbf{B}_{(1)}^T\mathbf{d}_{\theta_\mathrm{s}}\right|^2}{\boldsymbol{\alpha}^H\mathbf{T}_{(1),1:P}^H\mathbf{B}_{(1)}^T\mathbf{B}_{(1)}\mathbf{T}_{(1),1:P}\boldsymbol{\alpha}}. \end{aligned} \tag{53}$$

Now, if we want to compromise between WNG and DF, we need to maximize the WNG given above, from which we easily

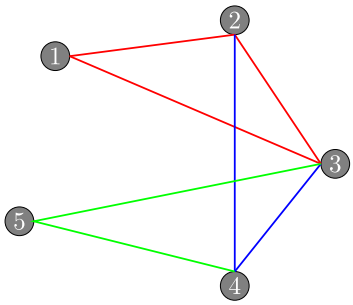


Fig. 3. Second-order differential beamforming on a graph with $M = 5$.

deduce that the optimal beamformer is

$$\mathbf{h}_{(1),1:P} = \frac{\left(\mathbf{B}_{(1)}^T\mathbf{B}_{(1)}\right)^{-1}\mathbf{B}_{(1)}^T\mathbf{P}_{\mathbf{T}_{(1),1:P}}\mathbf{d}_{\theta_\mathrm{s}}}{\mathbf{d}_{\theta_\mathrm{s}}^H\mathbf{P}_{\mathbf{T}_{(1),1:P}}\mathbf{d}_{\theta_\mathrm{s}}}, \tag{54}$$

where

$$\mathbf{P}_{\mathbf{T}_{(1),1:P}} = \mathbf{B}_{(1)}\mathbf{T}_{(1),1:P} \times \left(\mathbf{T}_{(1),1:P}^H\mathbf{B}_{(1)}^T\mathbf{B}_{(1)}\mathbf{T}_{(1),1:P}\right)^{-1}\mathbf{T}_{(1),1:P}^H\mathbf{B}_{(1)}^T. \tag{55}$$

For $P = 1$, we get

$$\mathbf{h}_{(1),1:1} = \frac{\mathbf{t}_{(1),1}}{\mathbf{d}_{\theta_\mathrm{s}}^H\mathbf{B}_{(1)}\mathbf{t}_{(1),1}} = \mathbf{h}_{(1),\mathrm{MDF}}, \tag{56}$$

which is the first-order differential MDF beamformer, and for $P = M - 1$, we obtain

$$\mathbf{h}_{(1),1:M-1} = \frac{\left(\mathbf{B}_{(1)}^T\mathbf{B}_{(1)}\right)^{-1}\mathbf{B}_{(1)}^T\mathbf{d}_{\theta_\mathrm{s}}}{\mathbf{d}_{\theta_\mathrm{s}}^H\mathbf{B}_{(1)}\left(\mathbf{B}_{(1)}^T\mathbf{B}_{(1)}\right)^{-1}\mathbf{B}_{(1)}^T\mathbf{d}_{\theta_\mathrm{s}}} = \mathbf{h}_{(1),\mathrm{MWNG}}, \tag{57}$$

which is the first-order differential MWNG beamformer. Therefore, by adjusting the positive integer P, we can obtain different beamformers whose performances are in between the performances of $\mathbf{h}_{(1),\mathrm{MDF}}$ and $\mathbf{h}_{(1),\mathrm{MWNG}}$.

B. Second Order

In second-order differential beamforming, microphones m, $m+1$, and $m+2$ are directly connected for $m = 1, 2, \ldots, M-2$, resulting in $M-2$ second-order spatial difference equations. On a graph, each one of these equations corresponds to a 3-clique (whose number is equal to $M-2$). Fig. 3 shows an example with $M = 5$, where the number of 3-cliques is equal to 3.

Since we have 5 microphones and 7 edges, where each edge connects a different pair of microphones, the incidence matrix (of size 5×7) is

$$\mathbf{B}_{(2)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & -1 \end{bmatrix}. \tag{58}$$

In order to get the second-order difference equations from each one of the 3-cliques, we need to define the matrix of size 7×3:

$$\bar{\mathbf{I}}_{3\mathrm{C}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{59}$$

Then, the matrix of interest[2] is

$$\overline{\mathbf{B}}_{(2)} = \mathbf{B}_{(2)}\bar{\mathbf{I}}_{3\mathrm{C}} = \begin{bmatrix} 2 & 0 & 0 \\ -1 & 2 & 0 \\ -1 & -1 & 2 \\ 0 & -1 & -1 \\ 0 & 0 & -1 \end{bmatrix}, \tag{60}$$

where the main diagonal contains the degrees (equal to 2) of the vertices of the 3-cliques, each columns represents a 3-clique of the graph, and the sum of the elements of each column is equal to 0. In the general case, this matrix is of size $M \times (M-2)$, where $M-2$ corresponds to the number of the 3-cliques. Therefore, $\overline{\mathbf{B}}_{(2)}$ contains all the important information of the 3-cliques composing the graph. Now, if we multiply $\overline{\mathbf{B}}_{(2)}^T$ by the observed signal, $\mathbf{y}$, we find that

$$\begin{aligned}\overline{\mathbf{B}}_{(2)}^T\mathbf{y} &= \begin{bmatrix} 2Y_1 - Y_2 - Y_3 & 2Y_2 - Y_3 - Y_4 \\ & \cdots \quad 2Y_3 - Y_4 - Y_5 \end{bmatrix}^T \\ &= \begin{bmatrix} Y_{(2),1} & Y_{(2),2} & Y_{(2),3} \end{bmatrix}^T \\ &= \mathbf{y}_{(2)}.\end{aligned} \tag{61}$$

Therefore, for any M, our new observed signal (of length $M-2$) is

$$\begin{aligned}\mathbf{y}_{(2)} &= \overline{\mathbf{B}}_{(2)}^T\mathbf{y} \\ &= \mathbf{x}_{(2)} + \mathbf{v}_{(2)} \\ &= \overline{\mathbf{B}}_{(2)}^T\mathbf{d}_{\theta_\mathrm{s}}X + \overline{\mathbf{B}}_{(2)}^T\mathbf{v},\end{aligned} \tag{62}$$

where the mth ($m = 1, 2, \ldots, M-2$) component of $\mathbf{y}_{(2)}$ is $Y_{(2),m} = 2Y_m - Y_{m+1} - Y_{m+2}$, and the second-order differential beamformer output signal is

$$Z_{(2)} = \mathbf{h}_{(2)}^H\mathbf{y}_{(2)}, \tag{63}$$

where $\mathbf{h}_{(2)}$ is the second-order differential filter of length $M-2$. The beampattern, WNG, and DF are defined similarly to the first-order case.

[2]From a rigorous mathematical perspective, the columns of $\overline{\mathbf{B}}_{(2)}$ do not fully correspond to second-order difference equations but because of their resemblance to the true ones, we make an abuse of language and call them second-order differentiators. The same is true for higher orders.

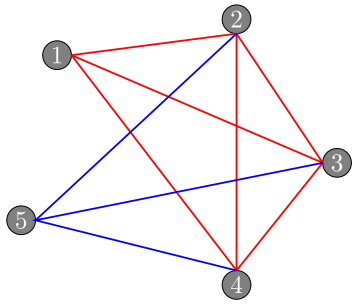


Fig. 4. Third-order differential beamforming on a graph with $M = 5$.

Then, it's not hard to find that the second-order differential MWNG and MDF beamformers are, respectively,

$$\mathbf{h}_{(2),\mathrm{MWNG}} = \frac{\left(\overline{\mathbf{B}}_{(2)}^T \overline{\mathbf{B}}_{(2)}\right)^{-1} \overline{\mathbf{B}}_{(2)}^T \mathbf{d}_{\theta_\mathrm{s}}}{\mathbf{d}_{\theta_\mathrm{s}}^H \overline{\mathbf{B}}_{(2)} \left(\overline{\mathbf{B}}_{(2)}^T \overline{\mathbf{B}}_{(2)}\right)^{-1} \overline{\mathbf{B}}_{(2)}^T \mathbf{d}_{\theta_\mathrm{s}}} \tag{64}$$

and

$$\mathbf{h}_{(2),\mathrm{MDF}} = \frac{\left(\overline{\mathbf{B}}_{(2)}^T \mathbf{\Gamma}_\mathrm{d} \overline{\mathbf{B}}_{(2)}\right)^{-1} \overline{\mathbf{B}}_{(2)}^T \mathbf{d}_{\theta_\mathrm{s}}}{\mathbf{d}_{\theta_\mathrm{s}}^H \overline{\mathbf{B}}_{(2)} \left(\overline{\mathbf{B}}_{(2)}^T \mathbf{\Gamma}_\mathrm{d} \overline{\mathbf{B}}_{(2)}\right)^{-1} \overline{\mathbf{B}}_{(2)}^T \mathbf{d}_{\theta_\mathrm{s}}}. \tag{65}$$

Also, using the same steps as in the previous subsection, we can find the beamformer $\mathbf{h}_{(2),1:P}$ (where $1 \leq P \leq M-2$) that can compromise between WNG and DF.

C. Third Order

Similarly to what we have explained in the two previous subsections, in third-order differential beamforming, Microphones m, $m+1$, $m+2$, and $m+3$ are directly connected for $m = 1, 2, \ldots, M-3$, resulting in $M-3$ third-order spatial difference equations. On a graph, each one of these equations corresponds to a 4-clique (whose number is equal to $M-3$). Continuing with our example of $M = 5$, third-order differential beamforming leads to two equations whose on a graph correspond to two 4-cliques as shown in Fig. 4.

We deduce that the matrix of interest (for $M = 5$) is

$$\begin{aligned} \overline{\mathbf{B}}_{(3)} &= \mathbf{B}_{(3)} \bar{\mathbf{I}}_{4\mathrm{C}} \\ &= \begin{bmatrix} 3 & 0 \\ -1 & 3 \\ -1 & -1 \\ -1 & -1 \\ 0 & -1 \end{bmatrix}, \end{aligned} \tag{66}$$

where the main diagonal contains the degrees (equal to 3) of the vertices of the 4-cliques, each columns represents a 4-clique of the graph, and the sum of the elements of each column is equal to 0. The matrices $\mathbf{B}_{(3)}$ (of size 5×9) and $\bar{\mathbf{I}}_{4\mathrm{C}}$ (of size 9×2) are easily defined as

$$\mathbf{B}_{(3)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & -1 \end{bmatrix} \tag{67}$$

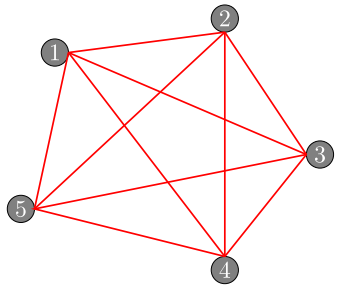


Fig. 5. Fourth-order differential beamforming on a graph with $M = 5$.

and

$$\bar{\mathbf{I}}_{4\mathrm{C}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{68}$$

As a consequence,

$$\begin{aligned} \overline{\mathbf{B}}_{(3)}^T \mathbf{y} &= \begin{bmatrix} 3Y_1 - Y_2 - Y_3 - Y_4 & 3Y_2 - Y_3 - Y_4 - Y_5 \end{bmatrix}^T \\ &= \begin{bmatrix} Y_{(3),1} & Y_{(3),2} \end{bmatrix}^T \\ &= \mathbf{y}_{(3)}. \end{aligned} \tag{69}$$

Therefore, for any M, our new observed signal (of length $M-3$) is

$$\begin{aligned} \mathbf{y}_{(3)} &= \overline{\mathbf{B}}_{(3)}^T \mathbf{y} \\ &= \mathbf{x}_{(3)} + \mathbf{v}_{(3)} \\ &= \overline{\mathbf{B}}_{(3)}^T \mathbf{d}_{\theta_\mathrm{s}} X + \overline{\mathbf{B}}_{(3)}^T \mathbf{v}, \end{aligned} \tag{70}$$

where the mth ($m = 1, 2, \ldots, M-3$) component of $\mathbf{y}_{(3)}$ is $Y_{(3),m} = 3Y_m - Y_{m+1} - Y_{m+2} - Y_{m+3}$, and the third-order differential beamformer output signal is

$$Z_{(3)} = \mathbf{h}_{(3)}^H \mathbf{y}_{(3)}, \tag{71}$$

where $\mathbf{h}_{(3)}$ is the third-order differential filter of length $M-3$. We observe that as the order increases, the less flexibility we have in designing this filter. The beampattern, WNG, and DF are defined similarly to the first-order case. Obviously, the third-order differential MWNG and MDF beamformers resemble the second-order differential MWNG and MDF beamformers given in the previous subsection.

D. Highest Order

The highest order corresponds to $M-1$. The $(M-1)$th-order differential beamforming leads to one spatial difference equation connecting all microphones. Obviously, we have a complete graph, K_M, as shown in Fig. 5 with $M = 5$ (fourth-order differential beamforming).

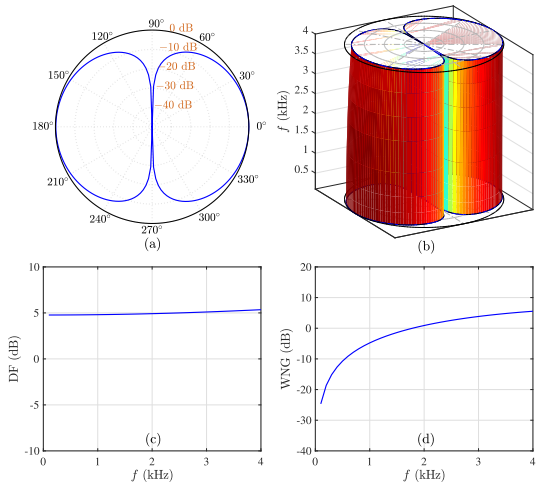


Fig. 6. Performance of the proposed first-order differential MWNG beamformer with a ULA: (a) beampattern at $f = 1.0$ kHz, (b) beampattern versus frequency, (c) DF, and (d) WNG. Conditions: $M = 5$, $\delta = 1.0$ cm, and $\theta_{\rm s} = 0^\circ$.

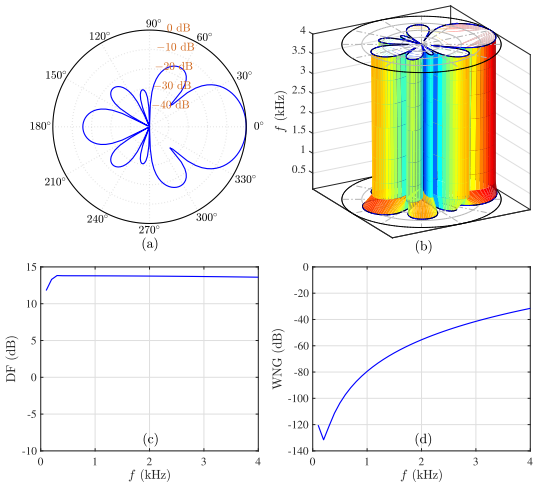


Fig. 7. Performance of the proposed first-order differential MDF beamformer with a ULA: (a) beampattern at $f = 1.0$ kHz, (b) beampattern versus frequency, (c) DF, and (d) WNG. Conditions: $M = 5$, $\delta = 1.0$ cm, and $\theta_{\rm s} = 0^\circ$.

In this context and for any M, the matrix of interest, which is actually a vector of length M, is

$$\overline{\mathbf{B}}_{(M-1)} = \left[M-1 \; -1 \; -1 \cdots -1 \right]^T, \tag{72}$$

where its first component is the degree (equal to $M-1$) of the vertices of K_M. Therefore, our new observed scalar signal is

$$\begin{aligned} Y_{(M-1)} &= \overline{\mathbf{B}}_{(M-1)}^T \mathbf{y} \\ &= X_{(M-1)} + V_{(M-1)} \\ &= \overline{\mathbf{B}}_{(M-1)}^T \mathbf{d}_{\theta_{\rm s}} X + \overline{\mathbf{B}}_{(M-1)}^T \mathbf{v}, \end{aligned} \tag{73}$$

where $Y_{(M-1)} = (M-1)Y_1 - Y_2 - \cdots - Y_M$, and the $(M-1)$th-order differential beamformer output signal is

$$Z_{(M-1)} = H_{(M-1)} Y_{(M-1)}, \tag{74}$$

where $H_{(M-1)}$ is the $(M-1)$th-order differential gain. We see that with the highest order, we do not have any flexibility except for the distortionless constraint. In this case, we take

$$H_{(M-1)} = \frac{1}{\overline{\mathbf{B}}_{(M-1)}^T \mathbf{d}_{\theta_{\rm s}}}. \tag{75}$$

V. Simulations

Having explained the fundamentals of differential beamforming on graphs and derived a family of differential beamformers, we study in this section their performance in terms of beampattern, WNG, and DF. Also, we compare them to the recently developed null-constrained differential beamformers [26], [36].

A. Performance of Differential Beamformers on Graphs

We first study the performance of the developed differential beamformers. We consider a uniform linear array (ULA) consisting of five microphones and with an interelement spacing of 1.0 cm. The desired source signal propagates from the endfire direction, i.e., $\theta_{\rm s} = 0^\circ$. Fig. 6 plots the performance of the first-order differential MWNG beamformer, i.e., beampattern (at $f = 1.0$ kHz), beampattern versus frequency, and DF and WNG as a function of the frequency. As seen, the designed beampattern has a null at 90°, and a one (maximum) at 0° and 180°, which corresponds to the first-order dipole [18]. It is clearly seen that the beamformer has a frequency-invariant beampattern and achieves an almost frequency-invariant DF of approximately 5 dB.

Fig. 7 plots the performance of the first-order differential MDF beamformer. It is also clearly seen that this beamformer has frequency-invariant beampattern and DF. Compared to the MWNG, the MDF beamformer has a higher value of DF but a lower value of WNG, which indicates more serious white noise amplification. While the MWNG and MDF beamformers are derived from the same path graph, which corresponds to the family of first-order DMAs, their performance is different since we use different criteria.

While ULAs are widely used in practice, other geometries are also important. So, we also show an example of a differential beamformer on graphs with a microphone array of arbitrary geometry, which also consists of five omnidirectional microphones. The coordinates of the microphones are random numbers generated with the uniform distribution, with the constraints of $1 \leq r_m \leq 2$ cm and $-\pi < \psi_m \leq \pi$. The resulting coordinates of the five microphones are, respectively, (0.5, 0) cm, (0.92, 0.77) cm, $(-1.0, 0.0)$ cm, $(-0.84, 1.0)$ cm, and $(0.31, -0.95)$ cm. To show the steering ability, we assume that the desired look direction is 30°, i.e., $\theta_{\rm s} = 30^\circ$. The performance of the MWNG beamformer is shown in Fig. 8. We observe that the beampattern is (almost) frequency invariant, the main beam points in the direction of 30°, and the DF does not vary much with frequency. Comparing Figs. 6 and 8, one can see that the DFs for both the ULA and the array with randomly generated microphone positions are almost the same; but the WNG for the array with random geometry is smaller than that of the ULA. The underlying reason for this is that the array with random geometry has a smaller interelement spacing. Generally, with the same number of sensors and same DF, the smaller are the

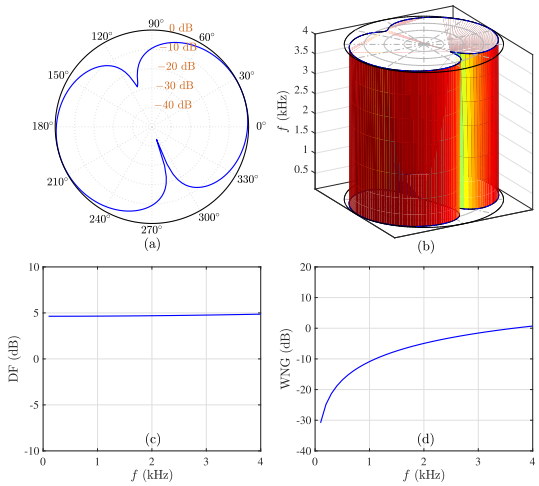


Fig. 8. Performance of the proposed first-order differential MWNG beamformer with an array of arbitrary geometry: (a) beampattern at $f = 1.0$ kHz, (b) beampattern versus frequency, (c) DF, and (d) WNG. Conditions: $M = 5$, $\delta = 1.0$ cm, and $\theta_{\rm s} = 30^\circ$.

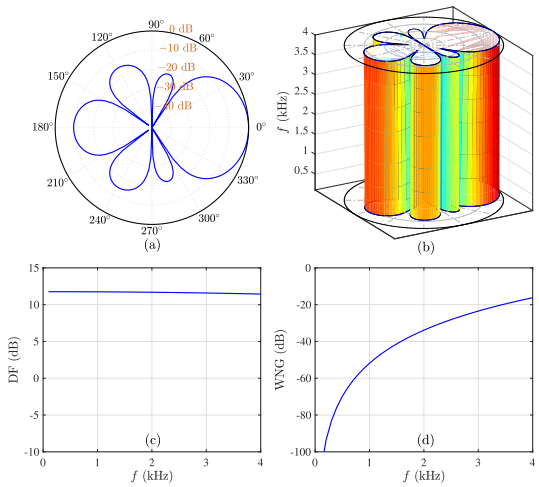


Fig. 9. Performance of the proposed second-order differential MDF beamformer with a ULA: (a) beampattern at $f = 1.0$ kHz, (b) beampattern versus frequency, (c) DF, and (d) WNG. Conditions: $M = 5$, $\delta = 1.0$ cm, and $\theta_{\rm s} = 0^\circ$.

interelement spacing and aperture of the array, the smaller is the WNG.

We then study the performance of the proposed higher-order differential beamformers with the same conditions as in the first simulation. The results for the second-order differential MDF (beampattern at $f = 1.0$ kHz, beampattern versus frequency, and DF and WNG as a function of frequency) are plotted in Fig. 9. It is clearly seen that the beampatterns and DFs are almost frequency invariant. One can see that the second-order differential MDF beamformer's beampattern has three nulls; but the first-order differential MDF beamformer's beampattern has four nulls. Also, the first-order differential MDF beamformer has higher DF values and lower WNG values than the second-order differential MDF beamformer. This is a major difference with the traditional DMA. In the new approach, the differential beamforming follows the structure of a clique, which does not necessarily correspond to the conventional differentiation.

B. Performance Comparison With the Conventional and Robust Null-Constrained Differential Beamformers

It is worth comparing the new differential beamformers on graphs to the recently developed conventional and robust null-constrained differential beamformers [26], [36].

The null-constrained differential beamformer is derived from the principle that the designed beampattern is as close as possible to the desired frequency-independent directivity pattern [26]. Generally, the frequency-independent target directivity pattern can be written as a linear combination of a series of cosine functions of different orders. Since the target directivity pattern is determined by its null information, the design of a differential beamformer is equivalent to finding an optimal filter, such that the resulting beampattern has the same nulls as the desired frequency-independent directivity pattern. Briefly, if we assume that the Nth-order directivity pattern has N nulls $\theta_{N,1}, \theta_{N,2}, \ldots, \theta_{N,N}$ in the range of $(0, 180^\circ]$ [18], the linear differential beamforming filter is computed by solving the following linear system of equations [26]:

$$\mathbf{D}\mathbf{h} = \mathbf{i}_1, \tag{76}$$

where

$$\mathbf{D} = \begin{bmatrix} \mathbf{d}_{\theta_{\rm s}}^H \\ \mathbf{d}_{\theta_{N,1}}^H \\ \vdots \\ \mathbf{d}_{\theta_{N,N}}^H \end{bmatrix} \tag{77}$$

is a constraint matrix of size $(N+1) \times M$ and $\mathbf{i}_1$ is a vector of length $N+1$, whose first element is 1 and all its other components are 0. It is assumed that there are N distinct nulls by default. For the design of a beampattern with multiple nulls in the same direction, we can add multiple constraints into the partial derivative of the beampattern [36]. The conventional differential beamformer is designed with $M = N+1$, so the solution to (76) is [26]

$$\mathbf{h}_{\rm DMA} = \mathbf{D}^{-1}\mathbf{i}_1. \tag{78}$$

To improve the robustness, more microphones are employed, i.e., $M > N+1$, which leads to the robust differential beamformer [26]:

$$\mathbf{h}_{\rm RDMA} = \mathbf{D}^H \left(\mathbf{D}\mathbf{D}^H\right)^{-1} \mathbf{i}_1. \tag{79}$$

As seen from the previous study, the proposed MWNG beamformer's beampattern corresponds to the dipole, so the target beampattern of the conventional differential beamformers in this simulation is also chosen as the dipole for fair comparison. Simulation conditions are the same as in the first simulation, i.e., with a ULA, $M = 5$, $\delta = 1$ cm, and $\theta_{\rm s} = 0^\circ$. For brevity, we refer to the conventional null-constrained differential beamformer, robust null-constrained differential beamformer, and the proposed differential beamformer as conventional DMA, robust DMA, and graphs DMA, respectively. Fig. 10 plots the WNG and DF of the aforementioned beamformers of the first order as a function of frequency. It should be noted that in Fig. 10(a), three lines (conventional DMA, robust DMA, and graphs DMA with $P = 4$) overlap, and in Fig. 10(b), two lines (robust DMA

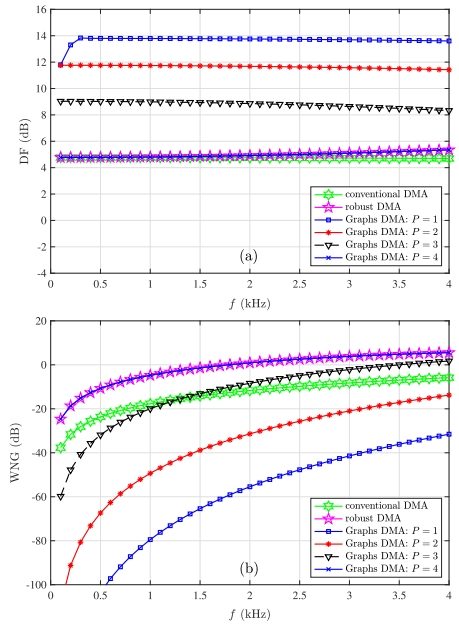


Fig. 10. DF and WNG of the proposed differential beamformers on graphs and the null-constrained differential beamformers of first order as a function of the frequency, f, with a ULA: (a) DF and (b) WNG. Conditions: $M = 5$, $\delta = 1.0$ cm, and $\theta_s = 0^\circ$.

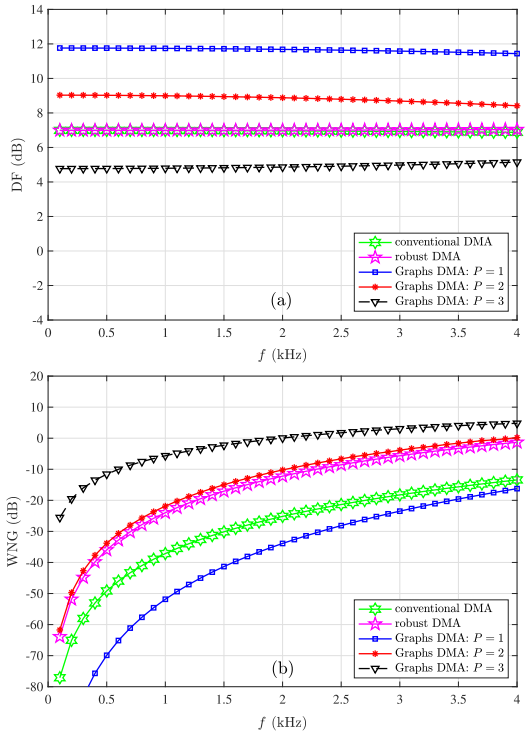


Fig. 11. DF and WNG of the proposed differential beamformers on graphs and the null-constrained differential beamformers of second order as a function of the frequency, f, with a ULA: (a) DF and (b) WNG. Conditions: $M = 5$, $\delta = 1.0$ cm, and $\theta_s = 0^\circ$.

and graphs DMA with $P = 4$) overlap. From Fig. 10, we make the following observations. 1) The conventional and robust null-constrained differential beamformers have the same DF, but the robust differential beamformer has much higher value of WNG than that of the conventional differential beamformer. 2) The robust null-constrained differential beamformer has the same performance (in terms of DF and WNG) as the optimal compromising beamformer with $P = 4$, which corresponds to the MWNG beamformer. 3) The DF of the optimal compromising beamformer decreases while the WNG increases with the value of P. Consequently, one can achieve a good compromise between DF and WNG through choosing a proper value of P.

Fig. 11 plots the performance of the second-order differential beamformer under the same conditions. Note that in Fig. 11(a), two lines (conventional DMA and robust DMA) overlap. Similarly, with different values of P, the proposed differential beamformer can nicely compromise between DF and WNG. This shows that the developed differential beamformer provides more flexibility to control the compromise between two performances. We also observe that the optimal compromising differential beamformer on graphs with $P = 2$ not only has a higher DF, but also a higher WNG than the robust null-constrained differential beamformer.

C. Performance in Reverberant Acoustic Environments

Finally, we evaluate the performance of the graph DMAs in a reverberant room of size $6 \times 5 \times 3$ m simulated using the image model [43]. A diagram of the floor layout of the setup is shown in Fig. 12. All the setup parameters are described in Fig. 12. For simplicity, we assume that the reflection coefficients of all the six walls are identical. We study eight reverberation conditions by setting room reflection coefficients to 0, 0.5, 0.6, 0.7, 0.8, 0.85, 0.9, and 0.95, respectively. The corresponding reverberation

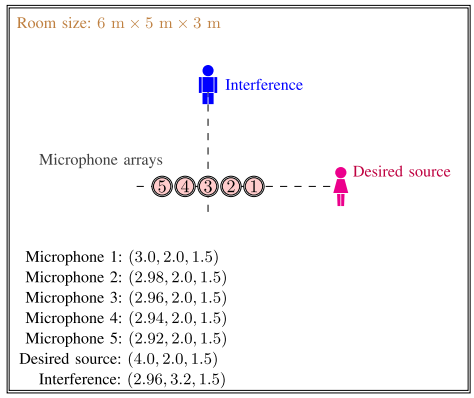


Fig. 12. Floor layout of the simulation setup in a reverberant room.

times, T_{60}, are approximately 0 ms, 100 ms, 130 ms, 170 ms, 370 ms, 480 ms, 700 ms, and 1000 ms. We first generate the acoustic channel impulse responses from the desired source and interference positions to the microphones. Then, the microphone signals are generated by convolving the source signal (clean speech prerecorded from a female speaker) and interference signal (clean speech prerecorded from a male speaker) with the obtained impulse responses, respectively, and adding the results together. In our study, we decompose the desired signal (and impulse response) into the direct path, early reverberation (the first 40 ms), and late reverberation for better evaluation [45]. Finally, we add some diffuse noise (with an input SNR of 10 dB) and white Gaussian noise (with an input SNR of 20 dB). The signal length is 30 seconds and the sampling frequency is

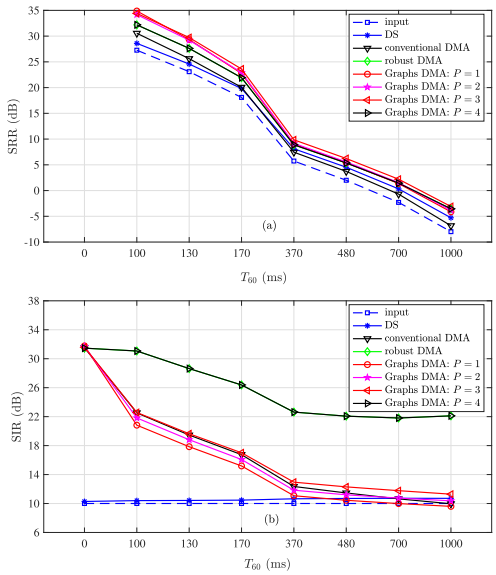


Fig. 13. SRR and SIR of the DS, conventional DMA, robust DMA, and graphs DMA for different reverberation conditions: (a) SRR and (b) SIR.

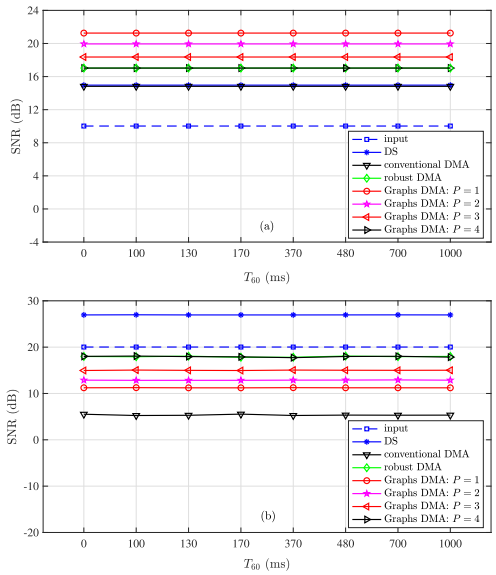


Fig. 14. SNR of the DS, conventional DMA, robust DMA, and graphs DMA for different reverberation conditions: (a) in diffuse noise and (b) in white noise.

16 kHz. The beamforming process is implemented in the STFT domain (the first microphone is chosen as the reference) and the processed signals are transformed back to the time domain using the inverse STFT. For all beamformers, the matrix $\mathbf{\Gamma}_{\mathrm{d}}$ is regularized with a small number of 0.001 to avoid numerical instabilities.

We study the graphs DMA beamformers with $P = 1, 2, 3, 4$ of the first order, and compare them with the delay-and-sum (DS), conventional DMA, and robust DMA beamformers (also first order). We are primarily concerned with the robustness, i.e., WNG, and suppressing (late) reverberation, interference, and diffuse noise. For that, we define the signal-to-reverberation ratio (SRR), signal-to-interference ratio (SIR), and signal-to-noise ratio (SNR) for the output signal, respectively, as

$$\mathrm{SRR} = \frac{E\left[x_{\mathrm{fd}}^2(k)\right]}{E\left[x_{\mathrm{fr}}^2(k)\right]}, \tag{80}$$

$$\mathrm{SIR} = \frac{E\left[x_{\mathrm{fd}}^2(k)\right]}{E\left[x_{\mathrm{fi}}^2(k)\right]}, \tag{81}$$

$$\mathrm{SNR} = \frac{E\left[x_{\mathrm{fd}}^2(k)\right]}{E\left[v_{\mathrm{fn}}^2(k)\right]}, \tag{82}$$

where $x_{\mathrm{fd}}(k)$ is the filtered desired signal, $x_{\mathrm{fr}}(k)$ is the filtered late reverberation, $x_{\mathrm{fi}}(k)$ is the filtered interference, and $v_{\mathrm{fn}}(k)$ is the filtered noise. In a similar way, one can define the input SRR, SIR, and SNR before the beamforming process.

Fig. 13(a) plots the results of the SRR. It is seen that the SRR increases with the reverberation time and all beamformers can achieve a higher SRR than the input signal. The graphs DMAs with $P = 1, 2, 3$ have the best performance in dealing with reverberation. The graphs DMA with $P = 4$ has the same performance as the robust DMA. The DS beamformer has the worst performance in low reverberation, but the conventional DMA yields worse performance than DS in strong reverberation. Fig. 13(b) plots the results of the SIR. It is seen that the graphs DMA with $P = 4$ and the robust DMA have the best performance in suppressing the interference. The DS beamformer can only slightly reduce the interference, and the graphs DMAs with $P = 1, 2, 3$ achieve a compromised performance.

The results of the SNR in diffuse noise are plotted in Fig. 14(a). In comparison, the graphs DMA with $P = 1$ yields the highest SNR in diffuse noise, $P = 2$ yields the second-best performance, the conventional DMA and DS beamformers have the lowest SNR in diffuse noise. However, it should be noted that the conventional DMA is designed with two microphones while the DS beamformer is designed with five microphones. The results of the SNR in white noise are shown in Fig. 14(b). As seen, the SNR of the filtered signal is lower than the input signal except for the DS beamformer. This indicates that the white noise is amplified. In practice, we need to control the WNG to a reasonable range, which depends on the sensors' self noise level as well as imperfections of the designed microphone array. It is seen that the SNR does not vary with T_{60}. This is due to the fact that the diffuse noise and white noise are independent with reverberation. We are generally more concerned with suppressing reverberation, interferences, and diffuse noise, so the proposed methods are advantageous in practice.

VI. Conclusion

In this paper, the concept of graphs has been explored for the design of differential beamformers. In this context, the microphone array is considered as a graph, where the nodes correspond to the sensors, the order of the graph corresponds to the number of microphones, and the set of edges corresponds to some linear spatial difference equations among microphones with proper orders. Based on this, we discussed the fundamentals of differential beamforming from a graph perspective and developed a family of new differential beamformers. The performance of the developed beamformers was demonstrated with a ULA and an array with arbitrary geometry. The graph approach developed in this paper does not only provide new perspectives

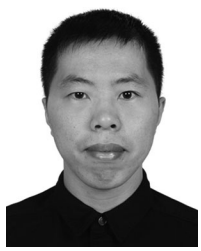
on differential beamforming, but also enables to develop new differential beamformers with any geometry of arrays.

ACKNOWLEDGMENT

The authors thank the associate editor and the anonymous reviewers for their constructive comments and useful suggestions.

REFERENCES

[1] D. B. West *Introduction to Graph Theory*, 2nd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 2001.
[2] C. Godsil and G. Royle, *Algebraic Graph Theory*. New York, NY, USA: Springer-Verlag, 2001.
[3] R. Diestel, *Graph Theory*. New York, NY, USA: Springer-Verlag, 2000.
[4] A. Agaskar and Y. Lu, "A spectral graph uncertainty principle," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4338–4356, Jul. 2013.
[5] A. Ortega, P. Frossard, J. Kovačević, J. Moura, and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications," *Proc. IEEE*, vol. 106, no. 5, pp. 808–828, May 2018.
[6] A. Sakiyama, Y. Tanaka, T. Tanaka, and A. Ortega, "Efficient sensor position selection using graph signal sampling theory," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2016, pp. 6225–6229.
[7] A. Gavili and X. Zhang, "On the shift operator, graph frequency, and optimal filtering in graph signal processing," *IEEE Trans. Signal Process.*, vol. 65, no. 23, pp. 6303–6318, Dec. 2017.
[8] J. Mei and J. Moura, "Signal processing on graphs: Causal modeling of unstructured data," *IEEE Trans. Signal Process.*, vol. 65, no. 8, pp. 2077–2092, Apr. 2016.
[9] A. Sandryhaila and J. Moura, "Discrete signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1644–1656, Apr. 2013.
[10] H. Loeliger, "An introduction to factor graphs," *IEEE Signal Process. Mag.*, vol. 21, no. 1, pp. 28–41, Jan. 2004.
[11] E. Pavez and A. Ortega, "Generalized laplacian precision matrix estimation for graph signal processing," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2016, pp. 6350–6354.
[12] A. Dimakis, S. Kar, J. Moura, M. Rabbat, and A. Scaglione, "Gossip algorithms for distributed signal processing," *Proc. IEEE*, vol. 98, no. 11, pp. 1847–1864, Nov. 2010.
[13] N. Perraudin and P. Vandergheynst, "Stationary signal processing on graphs," *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3462–3477, Jul. 2017.
[14] R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays*. Raleigh, NC, USA: SciTech, 1980.
[15] H. V. Trees, *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. New York, NY, USA: Wiley, 2002.
[16] M. Brandstein and D. Ward, *Microphone Arrays: Signal Processing Techniques and Applications*. Berlin, Germany: Springer, 2001.
[17] J. Benesty, I. Cohen, and J. Chen, *Fundamentals of Signal Enhancement and Array Signal Processing*. Singapore: Wiley, 2018.
[18] G. Elko and J. Meyer, "Microphone arrays," in *Springer Handbook of Speech Processing*, J. Benesty, M. M. Sondhi, and Y. Huang, Eds., Berlin, Germany: Springer-Verlag, 2008, ch. 48, pp. 1021–1041.
[19] K. Andreas, R. Hendriks, H. Richard, and J. Jesper, "Relaxed binaural LCMV beamforming," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 25, no. 1, pp. 137–152, Jan. 2017.
[20] T. Higuchi, N. Ito, S. Araki, T. Yoshioka, M. Delcroix, and T. Nakatani, "Online MVDR beamformer based on complex Gaussian mixture model with spatial prior for noise robust ASR," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 25, no. 4, pp. 780–793, Apr. 2017.
[21] W. Herbordt, H. Buchner, S. Nakamura, and W. Kellermann, "Multichannel bin-wise robust frequency-domain adaptive filtering and its application to adaptive beamforming," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 15, no. 4, pp. 1340–1351, May 2007.
[22] O. L. Frost III, "An algorithm for linearly constrained adaptive array processing," *Proc. IEEE*, vol. 60, pp. 926–935, Aug. 1972.
[23] B. D. V. Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Mag.*, vol. 5, no. 2, pp. 4–24, Apr. 1988.
[24] M. Crocco and A. Trucco, "Design of robust superdirective arrays with a tunable tradeoff between directivity and frequency-invariance," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2169–2181, May 2011.
[25] S. Doclo and M. Moonen, "Superdirective beamforming robust against microphone mismatch," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 15, no. 2, pp. 617–631, Feb. 2007.
[26] J. Benesty and J. Chen, *Study and Design of Differential Microphone Arrays*. Berlin, Germany: Springer-Verlag, 2012.
[27] M. Kolundzija, C. Faller, and M. Vetterli, "Spatiotemporal gradient analysis of differential microphone arrays," *J. Audio Eng. Soc.*, vol. 59, no. 1/2, pp. 20–28, 2011.
[28] G. Huang, J. Chen, and J. Benesty, "Design of planar differential microphone arrays with fractional orders," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 28, no. 1, 2020.
[29] G. Huang, J. Chen, and J. Benesty, "Insights into frequency-invariant beamforming with concentric circular microphone arrays," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 26, no. 12, pp. 2305–2318, Dec. 2018.
[30] A. Bernardini, F. Antonacci, and A. Sarti, "Wave digital implementation of robust first-order differential microphone arrays," *IEEE Signal Process. Lett.*, vol. 25, no. 2, pp. 253–257, 2018.
[31] H. F. Olson, "Gradient microphones," *J. Acoust. Soc. Amer.*, vol. 17, no. 3, pp. 192–198, 1946.
[32] G. W. Elko, "Superdirectional microphone arrays," in *Acoustic Signal Processing for Telecommunication*. Berlin, Germany: Springer, 2000, pp. 181–237.
[33] G. W. Elko, "Microphone array systems for hands-free telecommunication," *Speech Commun.*, vol. 20, no. 3, pp. 229–240, 1996.
[34] G. W. Elko, "Differential microphone arrays," in *Audio Signal Processing for Next-Generation Multimedia Communication Systems*. Berlin, Germany: Springer, 2004, pp. 11–65.
[35] E. D. Sena, H. Hacihabiboglu, and Z. Cvetkovic, "On the design and implementation of higher-order differential microphones," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 20, no. 1, pp. 162–174, Jan. 2012.
[36] J. Chen, J. Benesty, and C. Pan, "On the design and implementation of linear differential microphone arrays," *J. Acoust. Soc. Amer.*, vol. 136, pp. 3097–3113, Dec. 2014.
[37] R. M. Derkx and K. Janse, "Theoretical analysis of a first-order azimuth-steerable superdirective microphone array," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 17, no. 1, pp. 150–162, Jan. 2009.
[38] G. Huang, J. Benesty, and J. Chen, "On the design of frequency-invariant beampatterns with uniform circular microphone arrays," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 25, no. 5, pp. 1140–1153, May 2017.
[39] J. Benesty, J. Chen, and I. Cohen, *Design of Circular Differential Microphone Arrays*. Berlin, Germany: Springer-Verlag, 2015.
[40] G. Huang, J. Benesty, and J. Chen, "Design of robust concentric circular differential microphone arrays," *J. Acoust. Soc. Amer.*, vol. 141, no. 5, pp. 3236–3249, 2017.
[41] H. Cox, R. M. Zeskind, and T. Kooij, "Practical supergain," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-34, pp. 393–398, 1986.
[42] J. N. Franklin, *Matrix Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1968.
[43] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating smallroom acoustics," *J. Acoust. Soc. Amer.*, vol. 65, pp. 943–950, Apr. 1979.
[44] F. Jacobsen, *The Diffuse Sound Field: Statistical Considerations Concerning the Reverberant Field in the Steady State*. Acoustics Laboratory, Technical University of Denmark, 1979.
[45] D. A. Berkley and J. B. Allen, "Normal listening in typical rooms—The physical and psychophysical correlates of reverberation," in *Acoustical Factors Affecting Hearing Aid Performance*, G. A. Studebaker and I. Hochberg, Eds., Allyn & Bacon, Needham Height, 1993.



Gongping Huang (Student Member, IEEE) received the bachelor's degree in electronics and information engineering and the Ph.D. degree in information and communication engineering from the Northwestern Polytechnical University, Xian, China, in 2012 and 2019, respectively.

From 2015 to 2017, he was a Visiting Researcher with the University of Quebec, INRS-EMT, Montreal, QC, Canada. He is a Postdoctoral Research Fellow of electrical engineering with the Technion—Israel Institute of Technology, Haifa, Israel. His research interests include microphone array signal processing, speech enhancement, audio and speech processing, and graph signal processing. He received the Andrew and Erna Finci Viterbi PostDoctoral Fellowship's Award in 2019. He was a Reviewer for ten scientific journals include the IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, IEEE SIGNAL PROCESSING LETTER, JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA, IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, SPEECH COMMUNICATION, etc.

Jacob Benesty received the master's degree in microwaves from Pierre & Marie Curie University, Paris, France, in 1987, and a Ph.D. degree in control and signal processing from Orsay University, France, in April 1991. During his Ph.D. (from November 1989 to April 1991), he worked on adaptive filters and fast algorithms with the Centre National d'Etudes des Telecomunications (CNET), Paris, France.

From January 1994 to July 1995, he was with Telecom Paris University on multichannel adaptive filters and acoustic echo cancellation. From October 1995 to May 2003, he was first a Consultant and then a member of the Technical Staff at Bell Laboratories, Murray Hill, NJ, USA. In May 2003, he joined the University of Quebec, INRS-EMT, in Montreal, QC, Canada, as a Professor. He is also a Visiting Professor with the Technion, Haifa, Israel, an Adjunct Professor with Aalborg University, Denmark, and a Guest Professor with Northwestern Polytechnical University, Xi'an, China. His research interests include signal processing, acoustic signal processing, and multimedia communications. He is the Inventor of many important technologies. In particular, he was the Lead Researcher with Bell Labs who conceived and designed the world-first real-time hands-free full-duplex stereophonic teleconferencing system. Also, he conceived and designed the world-first PC-based multi-party hands-free full-duplex stereo conferencing system over IP networks. He is the editor of the book series *Springer Topics in Signal Processing*. He was the General Chair and the Technical Chair of many international conferences and a member of several IEEE technical committees. Four of his journal papers were awarded by the IEEE Signal processing Society and in 2010 he received the Gheorghe Cartianu Award from the Romanian Academy. He has coauthored and coedited/coauthored numerous books in the area of acoustic signal processing.

Israel Cohen (Fellow, IEEE) received the B.Sc. (*summa cum laude*), M.Sc., and Ph.D. degrees in electrical engineering from the Technion—Israel Institute of Technology, Israel, in 1990, 1993, and 1998, respectively.

From 1990 to 1998, he was a Research Scientist with RAFAEL Research Laboratories, Haifa, Israel Ministry of Defense. From 1998 to 2001, he was a Postdoctoral Research Associate with the Computer Science Department, Yale University, New Haven, CT, USA. In 2001, he joined the Electrical Engineering Department, Technion. He is currently a Professor of electrical engineering with the Technion – Israel Institute of Technology, Haifa, Israel. He is also a Visiting Professor with Northwestern Polytechnical University, Xian, China. He is a coeditor of the Multichannel Speech Processing Section of the *Springer Handbook of Speech Processing* (Springer, 2008), and a coauthor of *Fundamentals of Signal Enhancement and Array Signal Processing* (Wiley-IEEE Press, 2018). His research interests include array processing, statistical signal processing, analysis and modeling of acoustic signals, speech enhancement, noise estimation, microphone arrays, source localization, blind source separation, system identification and adaptive filtering.

Dr. Cohen received the Norman Seiden Prize for Academic Excellence in 2017, the SPS Signal Processing Letters Best Paper Award in 2014, the Alexander Goldberg Prize for Excellence in Research in 2010, and the Muriel and David Jacknow Award for Excellence in Teaching in 2009. He was an Associate Member of the IEEE Audio and Acoustic Signal Processing Technical Committee, and a Distinguished Lecturer of the IEEE Signal Processing Society. He was an Associate Editor of the IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING and IEEE SIGNAL PROCESSING LETTERS, and as a member of the IEEE Audio and Acoustic Signal Processing Technical Committee and the IEEE Speech and Language Processing Technical Committee.

Jingdong Chen (Senior Member, IEEE) received the Ph.D. degree in pattern recognition and intelligence control from the Chinese Academy of Sciences, Beijing, China, in 1998.

From 1998 to 1999, he was with ATR Interpreting Telecommunications Research Laboratories, Kyoto, Japan, where he conducted research on speech synthesis, speech analysis, as well as objective measurements for evaluating speech synthesis. He then joined the Griffith University, Brisbane, Australia, where he was engaged in research on robust speech recognition and signal processing. From 2000 to 2001, he was with ATR Spoken Language Translation Research Laboratories on robust speech recognition and speech enhancement. From 2001 to 2009, he was a member of Technical Staff at Bell Laboratories, Murray Hill, NJ, USA, working on acoustic signal processing for telecommunications. He subsequently joined WeVoice, Inc., in New Jersey, serving as the Chief Scientist. He is currently a Professor with the Northwestern Polytechnical University, Xi'an, China. His research interests include array signal processing, adaptive signal processing, speech enhancement, adaptive noise/echo control, signal separation, speech communication, and artificial intelligence.

Dr. Chen was an Associate Editor of the IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING from 2008 to 2014 and as a technical committee (TC) member of the IEEE Signal Processing Society (SPS) TC on Audio and Electroacoustics from 2007 to 2009. He is currently a member of the IEEE SPS TC on Audio and Acoustic Signal Processing, and a member of the editorial advisory board of the *Open Signal Processing Journal*. He was the General Co-Chair of ACM WUWNET 2018 and IWAENC 2016, the Technical Program Chair of IEEE TENCON 2013, a Technical Program Co-Chair of IEEE WASPAA 2009, IEEE ChinaSIP 2014, IEEE ICSPCC 2014, and IEEE ICSPCC 2015, and helped organize many other conferences. He coauthored 12 monograph books including *Array Processing–Kronecker Product Beamforming*, (Springer, 2019), *Fundamentals of Signal Enhancement and Array Signal Processing*, (Wiley, 2018), *Fundamentals of Differential Beamforming*, (Springer, 2016), *Design of Circular Differential Microphone Arrays* (Springer, 2015), *Noise Reduction in Speech Processing* (Springer, 2009), *Microphone Array Signal Processing* (Springer, 2008), and *Acoustic MIMO Signal Processing* (Springer, 2006), etc. He received the 2008 Best Paper Award from the IEEE Signal Processing Society (with Benesty, Huang, and Doclo), the Best Paper Award from the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics in 2011 (with Benesty), the Bell Labs Role Model Teamwork Award twice, respectively, in 2009 and 2007, the NASA Tech Brief Award twice, respectively, in 2010 and 2009, and the Young Author Best Paper Award from the 5th National Conference on Man-Machine Speech Communications in 1998. He is a coauthor of a paper for which C. Pan received the IEEE R10 (Asia-Pacific Region) Distinguished Student Paper Award (First Prize) in 2016. He was also a recipient of the Japan Trust International Research Grant from the Japan Key Technology Center in 1998 and the "Distinguished Young Scientists Fund" from the National Natural Science Foundation of China in 2014.