Differential Kronecker Product Beamforming

Israel Cohen, Jacob Benesty, and Jingdong Chen

Abstract—Differential beamformers have attracted much interest over the past few decades. In this paper, we introduce differential Kronecker product beamformers that exploit the structure of the steering vector to perform beamforming differently from the well-known and studied conventional approach. We consider a class of microphone arrays that enable to decompose the steering vector as a Kronecker product of two steering vectors of smaller virtual arrays. In the proposed approach, instead of directly designing the differential beamformer, we break it down following the decomposition of the steering vector, and show how to derive differential beamformers using the Kronecker product formulation. As demonstrated, the Kronecker product decomposition facilitates further flexibility in the design of differential beamformers and in the trade-off control between the directivity factor and the white noise gain.

Index Terms—Differential beamforming, Kronecker product, microphone array, superdirective beamforming, beampattern, white noise gain, directivity factor, front-to-back ratio.

I. INTRODUCTION

Kronecker product decompositions have been recently investigated in the context of acoustic imaging [1], [2], linear system identification [3], [4], and beamforming [5], [6]. Ribeiro and Nascimento [1] considered the problem of acoustic imaging using planar microphone arrays. They introduced the Kronecker array transform, which enables the acceleration of acoustic imaging techniques, essentially by reorganizing the matrix-vector multiplication structure for a special case of separable arrays. They showed that for planar arrays with a rectangular grid of microphones, the Kronecker array transform enables to reconstruct acoustic images more accurately and with higher resolutions than standard methods. This transform, which was originally designed for acoustic camera applications, was later generalized for compressive beamforming by Nascimento et al. [2].

Recently, Paleologu et al. [4] addressed the problem of linear system identification by using approximate Kronecker product (KP) decompositions of the impulse response. They proposed to estimate optimal KP decompositions of the impulse response based on an iterative Wiener filter, which compared to the conventional Wiener filter, is computationally more efficient and yields better performances using less amount of data to estimate the statistics. They showed that finding an optimal KP decomposition of the impulse response is generally useful for real-world acoustic echo cancellation applications.

Differential beamformers [7]–[9] are a particular important class of fixed beamformers, which yield the highest gains in diffuse noise. Differential microphone arrays are generally small in size, and their beampatterns are almost frequency invariant, which is critical when we deal with broadband signals such as speech. Consequently, optimal designs of differential beamformers have attracted much interest over the past few decades.

In this paper, we introduce differential KP beamformers that take advantage of the steering vector structure to perform beamforming differently from the well-known and studied conventional approach. The conventional way of doing beamforming is by applying in the frequency domain a complex-valued linear filter, \( h_2 \), of length \( M \) to the observation signal vector, where \( M \) is the number of microphones in the array. With the conventional beamforming approach, \( M \) coefficients of \( h_2 \) need to be estimated for each frequency bin. We consider a particular class of microphone arrays that enable to decompose the steering vector as a Kronecker product of two steering vectors of smaller virtual arrays [1], [5], [10]. That is, arrays of \( M = M_1 M_2 \) microphones that can be configured from replications of one virtual array (containing \( M_1 \) microphones) to the microphone positions of the other virtual array (containing \( M_2 \) microphones). In this case, the steering vector (of length \( M \)) can be decomposed as

\[
d = d_1 \otimes d_2,
\]

where \( \otimes \) is the Kronecker product, \( d_1 \) is the steering vector (of length \( M_1 \)) corresponding to a virtual array of \( M_1 \) microphones, and \( d_2 \) is the steering vector (of length \( M_2 \)) corresponding to another virtual array of \( M_2 \) microphones. Note that (1) is satisfied whenever the physical array can be obtained from replications of one virtual array to the microphone positions of the other virtual array (see e.g., Fig. 1).

In the proposed approach, instead of directly designing the filter \( h_2 \) of length \( M \), we break it down following the decomposition of the global steering vector as

\[
h_2 = h_1 \otimes h_2,
\]

where \( h_1 \) and \( h_2 \) are two complex-valued linear filters of lengths \( M_1 \) and \( M_2 \), respectively. With this method, we need to estimate \( M_1 + M_2 \) coefficients [\( M_1 \) for \( h_1 \) and \( M_2 \) for \( h_2 \)] instead of \( M = M_1 M_2 \) for the conventional technique. This generally implies smaller matrices to invert (increasing robustness) [8], [11].

There are many ways to optimize the coefficients of \( h_1 \) and \( h_2 \) depending on what we want and the application at hand. We introduce differential KP beamformers, and show...
how to derive such beamformers, as well as new approaches through the designs of $b_1$ and $b_2$. We show that using the KP decomposition, the global beampattern can be expressed as the product of the beampatterns of the two virtual arrays. Also, the white noise gain of the global array can be expressed as the product of the white noise gains of the two virtual arrays, while the directivity factor and the front-to-back ratio of the global array cannot be factorized. We demonstrate designs of very flexible differential KP beamformers by exploiting these interesting properties.

The rest of the paper is organized as follows. In Section II we introduce the KP decomposition of a physical array into two virtual arrays and present the signal model. In Section III we formulate the problem of KP beamforming. In Section IV we define some important performance measures by using KP filters. Then, in Section V we show how to derive differential beamformers using the KP formulation.

II. SIGNAL MODEL

We consider two different virtual uniform linear arrays (ULAs) on the same line denoted 1 and 2. ULA 1 has $M_1$ microphones and an interelement spacing equal to $M_2\delta$, where $\delta$ is some unit spacing, and ULA 2 has $M_2$ microphones and an interelement spacing equal to $\delta$ (see Fig. 1). Assume that a desired source signal (plane wave), in the farfield, propagates from the azimuth angle, $\theta$, in an anechoic acoustic environment at the speed of sound, i.e., $c = 340 \text{ m/s}$, and impinges on the above described virtual arrays. In this scenario, the corresponding steering vectors (of lengths $M_1$ and $M_2$, respectively) are

$$d_{1,\theta}(\omega) = \left[ 1 \quad e^{-jM_2\varpi_1} \cos \theta \quad \ldots \quad e^{-j(M_1-1)M_2\varpi_1} \cos \theta \right]^T,$$

$$d_{2,\theta}(\omega) = \left[ 1 \quad e^{-j\varpi_2 \cos \theta} \quad \ldots \quad e^{-j(M_2-1)\varpi_2 \cos \theta} \right]^T,$$

where the superscript $^T$ is the transpose operator, $j$ is the imaginary unit,

$$\varpi = \frac{\omega \delta}{c},$$

$\omega = 2\pi f$ is the angular frequency, and $f > 0$ is the temporal frequency. In this study, we are interested in the physical linear array whose associated steering vector is of the form:

$$d_0(\omega) = d_{1,\theta}(\omega) \otimes d_{2,\theta}(\omega).$$

The corresponding physical ULA consists of $M = M_1M_2$ microphones with an interelement spacing equal to $\delta$. We will refer to this array as the global or whole ULA. We consider doing beamforming with small values of $\delta$ [7], [8], [13], where the main lobe is at the angle $\theta = 0$ (endfire direction) and the desired signal propagates from the same angle. Beamforming under these conditions has the potential to lead to large array gains. Our objective is to study differential beamforming with the proposed ULAs.

Since the source propagates from the angle $\theta = 0$, the observation signal vector of length $M_1M_2$ can be expressed in the frequency domain as

$$y(\omega) = \left[ Y_1(\omega) \quad Y_2(\omega) \quad \ldots \quad Y_M(\omega) \right]^T = x(\omega) + v(\omega) = d_0(\omega)X(\omega) + v(\omega),$$

where $Y_m(\omega)$ is the $m$th microphone signal, $x(\omega) = d_0(\omega)X(\omega)$, $X(\omega)$ is the zero-mean desired source signal, $d_0(\omega) = d_{1,0}(\omega) \otimes d_{2,0}(\omega)$ is the steering vector at $\theta = 0$ (direction of the desired source), $v(\omega)$ is the zero-mean additive noise signal vector defined similarly to $y(\omega)$, and $X(\omega)$ and $v(\omega)$ are uncorrelated. In the rest, in order to simplify the notation, we drop the dependence on the angular frequency, $\omega$. So, for example, (7) is written as $y = d_0X + v$. We deduce that the covariance matrix of $y$ is

$$\Phi_y = E(yy^H) = \Phi_x + \Phi_v,$$

where $E(\cdot)$ denotes mathematical expectation, the superscript $^H$ is the conjugate-transpose operator,

$$\Phi_x = \phi_x d_0 d_0^H$$

$$= \phi_x (d_{1,0} \otimes d_{2,0}) (d_{1,0} \otimes d_{2,0})^H$$

$$= \phi_x (d_{1,0} d_{1,0}^H) \otimes (d_{2,0} d_{2,0}^H)$$

is the covariance matrix of $x$, with $\phi_x = E\left(|X|^2\right)$ being the variance of $X$, and $\Phi_v = E(vv^H)$ is the covariance matrix of $v$. Assuming that the first microphone is the reference, we can express (5) as

$$\Phi_y = \phi_x d_0 d_0^H + \phi_v \Gamma_v,$$

where $\phi_v = E\left(|V_1|^2\right)$ is the variance of the noise at the reference microphone and $\Gamma_v = \Phi_v/\phi_v$ is the pseudo-coherence matrix of the noise. In the case of the spherically isotropic (diffuse) noise field, which will often be assumed here, (10) becomes

$$\Phi_y = \phi_x d_0 d_0^H + \phi \Gamma,$$

where $\phi$ is the variance of the diffuse noise and

$$\Gamma = \frac{1}{2} \int_0^\pi d_0 d_0^H \sin \theta d\theta.$$
It can be verified that the elements of the $M \times M$ matrix $\mathbf{\Gamma} (\omega)$ are
\[
[\mathbf{\Gamma} (\omega)]_{mn} = \frac{\sin ((m-n)\omega)}{(m-n)\omega} = \text{sinc}[((m-n)\omega),
\]
with $[\mathbf{\Gamma} (\omega)]_{mm} = 1, \ m = 1, 2, \ldots, M$.

One of our main objectives in this study is to take advantage of the global ULA steering vector structure in the particular case of $M = M_1 M_2$ to perform beamforming differently from the well-known and studied conventional approach. It will be demonstrated that the new technique is very flexible.

### III. KRONCKER PRODUCT BEAMFORMING

The conventional way of doing beamforming is by applying a complex-valued linear filter, $\mathbf{h}$, of length $M$ to the observation signal vector, $\mathbf{y}$. This processing is \[11\]
\[Z = \mathbf{h}^H \mathbf{y}, \]
where $Z$ is the estimate of the desired signal, $X$. In order to fully exploit the structure of the global steering vector, let us consider the complex-valued filters having the form:
\[\mathbf{h} = \mathbf{h}_1 \otimes \mathbf{h}_2, \]
where $\mathbf{h}_1$ and $\mathbf{h}_2$ are two complex-valued linear filters of lengths $M_1$ and $M_2$, respectively. In other words, the global beamformer, $\mathbf{h}$, follows the decomposition of the global steering vector, $\mathbf{d}_\theta$. In the proposed approach, beamforming is performed by applying $\mathbf{h}$ [as defined in \[15\]] to $\mathbf{y}$ [from \[7\]]. We get
\[Z = \mathbf{h}^H \mathbf{y}, \]
\[= \mathbf{h}^H \mathbf{d}_0 \mathbf{X} + \mathbf{h}^H \mathbf{v} = X_{6l} + V_{rn}, \]
where $Z$ is the estimate of the desired signal, $\mathbf{X}$,
\[X_{6l} = (\mathbf{h}_1 \otimes \mathbf{h}_2)^H (\mathbf{d}_{1,0} \otimes \mathbf{d}_{2,0}) \mathbf{X} = (\mathbf{h}_1^H \mathbf{d}_{1,0}) (\mathbf{h}_2^H \mathbf{d}_{2,0}) \mathbf{X} \]
\[= \text{is the filtered desired signal, and} \]
\[V_{rn} = (\mathbf{h}_1 \otimes \mathbf{h}_2)^H \mathbf{v} \]
\[= \text{is the residual noise. We deduce that the variance of Z is} \]
\[\phi_Z = \phi_{\mathbf{X}} \left[ \mathbf{h}_1^H \mathbf{d}_{1,0} \right] \left[ \mathbf{h}_2^H \mathbf{d}_{2,0} \right] + \phi_{\mathbf{v}_1} (\mathbf{h}_1 \otimes \mathbf{h}_2)^H \mathbf{\Gamma}_v (\mathbf{h}_1 \otimes \mathbf{h}_2). \]

We see that with this method, we only need to estimate $M_1 + M_2$ coefficients ($M_1$ for $\mathbf{h}_1$ and $M_2$ for $\mathbf{h}_2$) instead of $M_1 M_2$ for the conventional technique. This generally implies smaller matrices to invert (increasing robustness) \[8\], \[11\]. Notice that this way of doing beamforming may not be completely new. A similar approach was proposed in \[8\], \[11\], but in the context of multiple-input and multiple-output (MIMO) radar applications.

In our context, the distortionless constraint in the direction of the desired source, i.e., $\theta = 0$, is often required, i.e.,
\[\mathbf{h}_1^H \mathbf{d}_{1,0} = (\mathbf{h}_1^H \mathbf{d}_{1,0}) (\mathbf{h}_2^H \mathbf{d}_{2,0}) = 1. \]

Therefore, we may choose $\mathbf{h}_1^H \mathbf{d}_{1,0} = \mathbf{h}_2^H \mathbf{d}_{2,0} = 1$, so that \[19\] is satisfied. In some applications, it may be useful to design $\mathbf{h}_1$ and $\mathbf{h}_2$ such that $\mathbf{h}_1^H \mathbf{d}_{1,0}$ varies as a function of frequency, while \[20\] is still satisfied, but we do not consider this here.

### IV. PERFORMANCE MEASURES

In this section, we define some important performance measures by using KP filters.

The first useful measure discussed in this section is the beampattern or directivity pattern, which describes the sensitivity of the beamformer to a plane wave (source signal) impinging on the global ULA from the direction $\theta$. Mathematically, it is defined as
\[B_\theta (\mathbf{h}) = \mathbf{d}^H \mathbf{h} \]
\[= (\mathbf{d}_{1,0}^H \mathbf{h}_1) (\mathbf{d}_{2,0}^H \mathbf{h}_2) = B_{1,\theta} (\mathbf{h}_1) \times B_{2,\theta} (\mathbf{h}_2), \]
where
\[B_{i,\theta} (\mathbf{h}_i) = \mathbf{d}_{i,0}^H \mathbf{h}_i \]
\[= \sum_{m=1}^{M_i} H_{i,m} e^{j(m-1)\omega_i \cos \theta} \]
for $i = 1, 2$ are the beampatterns of the virtual ULAs, with $\omega_1 = M_2 \omega$, $\omega_2 = \omega$, and $H_{i,m}, \ m = 1, 2, \ldots, M_i$, being the coefficients of $\mathbf{h}_i$. Let $Z_i = e^{j\omega_i \cos \theta}$ for $i = 1, 2$, we can express the global beampattern as a polynomial in two variables, which is the product of two polynomials (of degrees $M_1 - 1$ and $M_2 - 1$) in one variable each, i.e.,
\[B (Z_1, Z_2) = B_1 (Z_1) \times B_2 (Z_2) \]
\[= \left( \sum_{m_1=1}^{M_1} H_{1,m_1} Z_{1,m_1-1} \right) \left( \sum_{m_2=1}^{M_2} H_{2,m_2} Z_{2,m_2-1} \right). \]

From this perspective, we can see that this beampattern has at most $M_1 + M_2 - 2$ distinct nulls (between $0$ and $\pi$), while the beampattern with the conventional approach has at most $M_1 M_2 - 1$ distinct nulls (between $0$ and $\pi$). The fact that the global beampattern, $B_\theta (\mathbf{h})$, can be expressed as the product of two beamformer beampatterns is an interesting property that can be exploited in the design of very flexible global beamformers.

Given that the first microphone is the reference, we can define the input signal-to-noise ratio (SNR) with respect to this reference as
\[\text{iSNR} = \frac{\phi_{\mathbf{X}}}{\phi_{\mathbf{v}_1}} \]
The output SNR is defined [from the variance of $Z$, see \[19\]] as
\[\text{oSNR} (\mathbf{h}) = \phi_{\mathbf{X}} \left[ \mathbf{h}_1^H \mathbf{d}_{1,0} \right] \left[ \mathbf{h}_2^H \mathbf{d}_{2,0} \right] \]
\[= \phi_{\mathbf{X}} \times \frac{\mathbf{h}_1^H \mathbf{d}_{1,0}}{\mathbf{h}_1^H \mathbf{\Gamma}_v \mathbf{h}}. \]
The definition of the gain in SNR is obtained from the previous definitions, i.e.,

$$G(h) = \frac{oSNR(h)}{iSNR} = \frac{\left| h^H d_0 \right|^2}{h^H \Gamma h}. \tag{26}$$

One convenient way to evaluate the sensitivity of the global ULA to some of its imperfections (e.g., sensor noise, small deviations of the microphone characteristics and positions) is via the so-called white noise gain (WNG), which is defined by taking $\Gamma = I_M$ in (26), where $I_M$ is the $M \times M$ identity matrix, i.e.,

$$W(h) = \frac{\left| h^H d_0 \right|^2}{h^H h} = \frac{\left| h^H d_{1,0} \right|^2}{h^H h_1} \times \frac{\left| h^H d_{2,0} \right|^2}{h^H h_2} = W_1(h_1) \times W_2(h_2), \tag{27}$$

where

$$W_i(h_i) = \frac{\left| h^H d_{i,0} \right|^2}{h^H h_i} \tag{28}$$

for $i = 1, 2$. Obviously, the WNG of the global ULA is simply the product of the WNGs of the first and second ULAs described in Section II. It is easy to check that

$$W(h) \leq M_1 M_2, \forall h. \tag{29}$$

Another important measure, which quantifies how the microphone array performs in the presence of reverberation (which can be assumed as diffuse noise) is the directivity factor (DF). Considering the spherically isotropic (diffuse) noise field, the DF is defined by taking $\Gamma = \Gamma$ in (26):

$$D(h) = \frac{\left| h^H d_0 \right|^2}{h^H \Gamma h}, \tag{30}$$

where $\Gamma$ is defined in (12). It is clear that

$$D(h) \leq d_0^H \Gamma^{-1} d_0, \forall h. \tag{31}$$

We observe that contrary to the beampattern and the WNG, the DF of the global ULA cannot be factorized as the product of the DFs of the two virtual ULAs, i.e.,

$$D(h) \neq D_1(h_1) \times D_2(h_2), \tag{32}$$

where

$$D_i(h_i) = \frac{\left| h^H d_{i,0} \right|^2}{h^H \Gamma_i h_i}, \tag{33}$$

with

$$\Gamma_i = \frac{1}{2} \int_0^\pi d_{i,\theta} d_{i,\theta}^* \sin \theta d\theta, \tag{34}$$

for $i = 1, 2$. The elements of the $M_i \times M_i$ matrix $\Gamma_i(\omega)$ are given by

$$[\Gamma_i(\omega)]_{mn} = \text{sinc} \left( (m-n)\frac{\pi}{2} \right) \tag{35}$$

with $[\Gamma_i(\omega)]_{mm} = 1, m = 1, 2, \ldots, M_i$. In fact, one can verify that $\Gamma \neq \Gamma_1 \otimes \Gamma_2$, that is why (32) is true in general. One can check that

$$h_1 \otimes h_2 = (h_1 \otimes I_{M_2}) h_2 \tag{36} = (I_{M_1} \otimes h_2) h_1, \tag{37}$$

where $I_{M_i}$ is the identity matrix of size $M_i \times M_i$ ($i = 1, 2$). Basically, the previous expressions, which separate $h_2$ and $h_1$ into matrix-vector products, are very helpful and convenient to use in the derivations of hypercardioid and supercardioid beamformers (see Sections V-C and V-D). When $h_2$ is fixed, given, and satisfies the distortionless constraint, i.e., $h_2^H d_{2,0} = 1$; then, thanks to (37), we can write the DF as

$$D(h_1|h_2) = \frac{\left| B_{1,0}(h_1) \right|^2}{h_1^H \Gamma_1 h_1} \times \frac{\left| B_{2,0}(h_2) \right|^2}{h_2^H \Gamma_2 h_2} \tag{38}$$

where

$$\Gamma_{h_2} = \frac{1}{2} \int_0^\pi d_{1,\theta} d_{1,\theta}^* |B_{2,0}(h_2)|^2 \sin \theta d\theta \tag{39} = (I_{M_1} \otimes h_2)^H \Gamma (I_{M_1} \otimes h_2).$$

In the same way, when $h_1$ is fixed, given, and satisfies the distortionless constraint, i.e., $h_1^H d_{1,0} = 1$; then, thanks to (36), we can express the DF as

$$D(h_2|h_1) = \frac{\left| B_{2,0}(h_2) \right|^2}{h_2^H \Gamma_2 h_2} \times \frac{\left| B_{1,0}(h_1) \right|^2}{h_1^H \Gamma_1 h_1} \tag{40}$$

where

$$\Gamma_{h_1} = \frac{1}{2} \int_0^\pi d_{2,\theta} d_{2,\theta}^* |B_{1,0}(h_1)|^2 \sin \theta d\theta \tag{41} = (h_1 \otimes I_{M_2})^H \Gamma (h_1 \otimes I_{M_2}).$$

Another measure of interest in this study is the front-to-back ratio (FBR), which is defined as the ratio of the power of the output of the array to signals propagating from the front-half plane to the output power for signals arriving from the rear-half plane [15]. This ratio, for the spherically isotropic (diffuse) noise field, is mathematically defined as [15]

$$F(h) = \frac{\int_0^{\pi/2} |B_{1,0}(h)|^2 \sin \theta d\theta}{\int_0^{\pi/2} |B_{1,0}(h)|^2 \sin \theta d\theta} \tag{42}$$

$$= \frac{\int_0^{\pi/2} |B_{1,0}(h_1)|^2 |B_{2,0}(h_2)|^2 \sin \theta d\theta}{\int_0^{\pi/2} |B_{1,0}(h_1)|^2 |B_{2,0}(h_2)|^2 \sin \theta d\theta}$$

$$= \frac{h_1^H \Gamma_1 h_1}{h_2^H \Gamma_2 h_2}, \tag{43}$$
where
\[ \Gamma_f = \int_0^{\pi/2} d\theta d\phi^H \sin \theta \theta, \quad (43) \]
\[ \Gamma_b = \int_{\pi/2}^{\pi} d\theta d\phi^H \sin \theta \theta. \quad (44) \]

It can be verified that the elements of the \( M \times M \) matrices \( \Gamma_f(\omega) \) and \( \Gamma_b(\omega) \) are given, respectively, by
\[ [\Gamma_f(\omega)]_{mn} = \frac{e^{j(n-m)\pi}}{j(n-m)\pi} \quad (45) \]
and
\[ [\Gamma_b(\omega)]_{mn} = \frac{1 - e^{-j(n-m)\pi}}{j(n-m)\pi}, \quad (46) \]
with \([\Gamma_f(\omega)]_{mm} = [\Gamma_b(\omega)]_{mm} = 1, m = 1, 2, \ldots, M\). Same as the DF, the FBR of the global ULA is not equal to the product of the FBRs of the two virtual ULAs, i.e.,
\[ F(h) \neq F_1(h_1) \times F_2(h_2), \quad (47) \]
where
\[ F_i(h_i) = \int_0^{\pi/2} |B_{i,\theta}(h_i)|^2 \sin \theta \theta, \quad (48) \]
\[ \Gamma_{f,i} = \int_0^{\pi/2} d_1 d_\theta d_{\phi}^H \sin \theta \theta, \quad \Gamma_{b,i} = \int_{\pi/2}^{\pi} d_1 d_\theta d_{\phi}^H \sin \theta \theta, \quad (49, 50) \]
for \( i = 1, 2 \). The elements of the \( M_i \times M_i \) matrices \( \Gamma_{f,i}(\omega) \) and \( \Gamma_{b,i}(\omega) \) are given by
\[ [\Gamma_{f,i}(\omega)]_{mn} = \frac{e^{j(n-m)\pi}}{j(n-m)\pi}, \quad (51) \]
\[ [\Gamma_{b,i}(\omega)]_{mn} = \frac{1 - e^{-j(n-m)\pi}}{j(n-m)\pi}, \quad (52) \]
with \([\Gamma_{f,i}(\omega)]_{mm} = [\Gamma_{b,i}(\omega)]_{mm} = 1, m = 1, 2, \ldots, M_i\).

When \( h_2 \) is fixed and given, and thanks to (35), we can write the FBR as
\[ F(h_2|h_1) = \frac{h_2^H \Gamma_{f,h_1} h_2}{h_1^H \Gamma_{b,h_1} h_2}, \quad (56) \]
where
\[ \Gamma_{f,h_1} = \int_0^{\pi/2} d_1 d_{\theta} d_{\phi}^H |B_{1,\theta}(h_1)|^2 \sin \theta \theta, \quad (57) \]
\[ \Gamma_{b,h_1} = \int_{\pi/2}^{\pi} d_1 d_{\theta} d_{\phi}^H |B_{1,\theta}(h_1)|^2 \sin \theta \theta, \quad (58) \]

V. Differential Beamformers

In this section, we derive some useful examples of differential KP beamformers. Of course, many more can be deduced depending on the applications at hand.

A. Cardioid

The \((M_2-1)\)-th order cardioid has a unique null of multiplicity \( M_2-1 \) in the direction \( \pi \). Therefore, the \( i \)th derivative, with \( i = 0, 1, \ldots, M_2-2 \), of the beampattern of \( h_2 \) with respect to \( \cos \theta \) is equal to 0 at \( \cos \pi = -1 \), i.e.,
\[ \left. \frac{d^i B_{2,\theta}(h_2)}{d \cos \theta} \right|_{\cos \theta = -1} = B_{2,\pi}^{(i)}(h_2) = 0, \quad (59) \]
with
\[ B_{2,\pi}^{(0)}(h_2) = B_{2,\pi}(h_2). \]

We easily find that
\[ B_{2,\pi}^{(i)}(h_2) = (j\omega \delta/c)^i (\Sigma^H d_{2,\pi})^H h_2, \quad (60) \]
where
\[ \Sigma = \text{diag}(0, 1, \ldots, M_2 - 1) \]
is a diagonal matrix of size \( M_2 \times M_2 \). Combining the distortionless constraint, i.e.,
\[ B_{2,0}(h_2) = d_{2,0}^H h_2 = 1, \quad (62) \]
with the \( M_2-1 \) equations from (59), we obtain a linear system of \( M_2 \) equations with \( M_2 \) unknowns:
\[ D_{2,\pi}^H h_2 = i, \quad (63) \]
where
\[ D_{2,\pi}^H = \begin{bmatrix} \Sigma \Sigma^H d_{2,\pi}^H \\ \Sigma \Sigma^H d_{2,\pi}^H \\ \vdots \\ \Sigma \Sigma^H d_{2,\pi}^H \end{bmatrix} \quad (64) \]
and \( i \) is the first column of \( I_{M_2} \). Therefore, the cardioid of order \( M_2-1 \) at the second ULA is
\[ h_{2,C} = D_{2,\pi}^{-H} i. \quad (65) \]
values of $M$ as a function of frequency for $h_{2,C}$, (solid line with circles) and the third-order KP cardioid, $h_C$, as a function of frequency for $\delta = 1 \text{ cm}, M_2 = 4$, and several values of $M_1$: $M_1 = 2$ (dashed line with asterisks), $M_1 = 3$ (dotted line with squares), and $M_1 = 4$ (dash-dot line with triangles). (a) DF and (b) WNG.

```
DF (dB)
-50
-40
-30
-20
-10
0
10
20
30
40
50
60
70
80
90
100
110
f (kHz)
0
1
2
3
4
5
6
7
8
```

(b) MN cardioid with $M_1 = M_2 = 4$, and (d) KP cardioid with $M_1 = 2$ and $M_2 = 8$. Conditions: $\delta = 1 \text{ cm}$ and $f = 3 \text{ kHz}$.

```
DF (dB)
-50
-40
-30
-20
-10
0
10
20
30
40
50
60
70
80
90
100
110
f (kHz)
0
1
2
3
4
5
6
7
8
```

Furthermore, the DFs and WNGs of the KP cardioids increase as we increase the numbers of microphones in ULA 1.

Clearly, the increase in both the DF and WNG is achieved at the expense of increasing $M$, the total number of microphones in the array. Therefore, we also compare the KP cardioids to the minimum-norm (MN) cardioid [8], [11], which employs $M$ microphones ($M > M_2$) to design an $\{M_2 - 1\}$th-order cardioid. Figure 3 shows plots of the DFs and WNGs of the third-order MN cardioid with $M_2 = 16$, third-order KP cardioid with $M_1 = M_2 = 4$, and third-order KP cardioid with $M_1 = 2$ and $M_2 = 8$. Figure 4 displays the patterns of the corresponding third-order cardioids for $f = 3 \text{ kHz}$. We observe that the MN cardioid achieves the maximum WNG, but for low frequencies its DF drops even below the traditional cardioid (which employs only $M = 4$ microphones). The KP cardioids, on the other hand, provide a more moderate increase in the WNG compared with the traditional cardioid, while also increasing the DF. Compared with the MN cardioid, the KP cardioids enable more flexibility in compromising the trade-off between WNG and DF.

To get a more robust beamformer, with higher WNG, we may optimize

$$
\min_{\mathbf{h}_2} \frac{1}{\mathbf{h}_2} \left(\mathbf{D}_{2,\pi}' \cdot \mathbf{D}_{2,\pi}^{H} + \mathbf{c}_2 \mathbf{I}_{M_2} \right) \mathbf{h}_2 \quad \text{subject to} \quad \mathbf{C}_{2,\pi}^{H} \mathbf{h}_2 = \mathbf{i}_c,
$$

(68)

where

$$
\mathbf{D}_{2,\pi}' = \left[ \begin{array}{cccc} \mathbf{S}_{1,\pi} & \mathbf{S}_{2,\pi} & \cdots & \mathbf{S}_{M_2-2,\pi} \end{array} \right]
$$

(69)

is a matrix of size $M_2 \times (M_2 - 2)$, $\epsilon_2 \geq 0$ is the regularization parameter,

$$
\mathbf{C}_{2,\pi}^{H} = \left[ \begin{array}{c} \mathbf{d}_{2,\pi}^{H} \mathbf{d}_{2,\pi} \end{array} \right]
$$

(70)

For the first filter, we may take the DS beamformer, i.e.,

$$
\mathbf{h}_{1,DS} = \frac{\mathbf{d}_{1,0}}{M_1},
$$

(66)

which maximizes the WNG. As a result, the $(M_2 - 1)$th-order KP cardioid is

$$
\mathbf{h}_C = \mathbf{h}_{1,DS} \otimes \mathbf{h}_{2,C}.
$$

(67)

Figure 2 shows plots of the DFs and WNGs of the traditional third-order cardioid, $h_{2,C}$, and the third-order KP cardioid, $h_C$, as a function of frequency for $\delta = 1 \text{ cm}, M_2 = 4$, and several values of $M_1$. We observe that the DFs and WNGs of the KP cardioids are higher than those of the traditional cardioid.
is the constraint matrix of size $2 \times M_2$, and
\[
i_c = \begin{bmatrix} 1 & 0 \end{bmatrix}^T \tag{71}
\]
is a vector of length 2. We see that with $C_{2,\pi}$, the two main constraints are fulfilled, i.e., the distortionless one and a null in the direction $\pi$. We find that the optimal filter is
\[
h_{2,C,\varepsilon_2} = \left( D_{2,\pi} H H + \varepsilon_2 I M_2 \right)^{-1} C_{2,\pi} \times \left[ C_{2,\pi}^{-1} D_{2,\pi} H H + \varepsilon_2 I M_2 \right]^{-1} i_c. \tag{72}
\]
Therefore, the robust KP cardioid is
\[
h_{C,\varepsilon_2} = h_{1,DS} \otimes h_{2,C,\varepsilon_2}. \tag{73}
\]

Figure 5 shows plots of the DFs and WNGs of the third-order robust KP cardioid, $h_{C,\varepsilon_2}$, as a function of frequency for $\delta = 1$ cm, $M_1 = 4$, $M_2 = 4$, and several values of $\varepsilon_2$: $\varepsilon_2 = 0.001$ (solid line with circles), $\varepsilon_2 = 0.01$ (dashed line with asterisks), $\varepsilon_2 = 0.1$ (dotted line with squares), and $\varepsilon_2 = 1$ (dash-dot line with triangles). (a) DF and (b) WNG.

We get
\[
B_{2,\pi/2}^{[1]}(h_2) = (j \omega \delta/c)^{i} (\Sigma' d_{2,\pi/2})^{H} h_2, \tag{75}
\]
where $\Sigma$ is defined in [61]. Combining the distortionless constraint with the $M_2 - 1$ constraints from [74], we have
\[
D_{2,\pi/2}^{H} h_2 = i, \tag{76}
\]
where
\[
D_{2,\pi/2}^{H} = \begin{bmatrix} d_{2,0}^{H} \\ (\Sigma' d_{2,\pi/2})^{H} \\ (\Sigma' d_{2,\pi/2})^{H} \\ \vdots \\ (\Sigma^{M_2-2} d_{2,\pi/2})^{H} \end{bmatrix}. \tag{77}
\]
As a result, the dipole of order $M_2 - 1$ at the second ULA is
\[
h_{2,D} = D_{2,\pi/2}^{H} i. \tag{78}
\]
Another feature of the dipole is that it has a 1 in the direction $\pi$. To ensure that the global beampattern has also a 1 at $\pi$, we must add this constraint in the design of the first filter. We deduce that the constraint equation for $h_1$ is
\[
C_{1,\pi}^{H} h_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \tag{79}
\]
where
\[
C_{1,\pi}^{H} = \begin{bmatrix} d_{1,0}^{H} \\ d_{1,\pi}^{H} \end{bmatrix} \tag{80}
\]
is the constraint matrix of size $2 \times M_1$. Since we want to maximize the WNG of $h_1$ subject to (79), we find the MN beamformer:
\[
h_{1,MN} = C_{1,\pi} (C_{1,\pi}^{H} C_{1,\pi})^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \tag{81}
\]
Now, that the two filters are derived, we deduce that the robust $(M_2 - 1)$th-order KP dipole is
\[
h_{D} = h_{1,MN} \otimes h_{2,D}. \tag{82}
\]
As we did for the cardioid, we can derive a robust KP dipole:
\[
h_{D,\varepsilon_2} = h_{1,MN} \otimes h_{2,D,\varepsilon_2}, \tag{83}
\]
where
\[
h_{2,D,\varepsilon_2} = \left( D_{2,\pi/2}^{H} + \varepsilon_2 I M_2 \right)^{-1} C_{2,\pi/2} \times \left[ C_{2,\pi/2}^{-1} D_{2,\pi/2}^{H} + \varepsilon_2 I M_2 \right]^{-1} i_c, \tag{84}
\]
with
\[
D_{2,\pi/2}^{H} = \begin{bmatrix} \Sigma_{1}^{H} d_{2,\pi/2} \\ \Sigma_{2}^{H} d_{2,\pi/2} \\ \vdots \\ \Sigma^{M_2-2} d_{2,\pi/2} \end{bmatrix} \tag{85}
\]
and
\[
C_{2,\pi/2}^{H} = \begin{bmatrix} d_{2,0}^{H} \\ d_{2,\pi/2}^{H} \end{bmatrix}. \tag{86}
\]
Figure 6 shows plots of the DFs and WNGs of different types of robust second-order dipoles as a function of frequency, whose maximization with respect to $h$ in (41), we get

$$\Gamma_{h_1^{(1)}} = \left( h_1^{(1)} \otimes I_{M_2} \right)^H \Gamma \left( I_{M_1} \otimes h_1^{(1)} \right).$$

(91)

As a result, the DF in (40) is

$$\mathcal{D} \left( h_2^{(1)} | h_1^{(1)} \right) = \frac{\left( h_2^{(1)} \otimes I_{M_2} \right)^H \Gamma \left( I_{M_1} \otimes h_2^{(1)} \right)}{\left( h_2^{(1)} \otimes I_{M_2} \right)^H \Gamma \left( I_{M_1} \otimes h_2^{(1)} \right)}.$$  

(92)

whose maximization with respect to $h_2^{(1)}$ gives

$$h_2^{(1)} = \frac{\Gamma_{h_2^{(1)}}^{-1} d_2.0}{d_2.0 \Gamma_{h_2^{(1)}}^{-1} d_2.0}.$$  

(93)

Continuing the iterations up to the iteration $n$, we easily get for the first filter:

$$h_1^{(n)} = \frac{\Gamma_{h_1^{(n-1)}}^{-1} d_1.0}{d_1.0 \Gamma_{h_1^{(n-1)}}^{-1} d_1.0},$$  

(94)

with

$$\Gamma_{h_1^{(n-1)}} = \left( I_{M_1} \otimes h_2^{(n-1)} \right)^H \Gamma \left( I_{M_1} \otimes h_2^{(n-1)} \right),$$  

(95)

and for the second filter:

$$h_2^{(n)} = \frac{\Gamma_{h_2^{(n)}}^{-1} d_2.0}{d_2.0 \Gamma_{h_2^{(n)}}^{-1} d_2.0},$$  

(96)

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TASLP.2019.2895241, IEEE/ACM Transactions on Audio, Speech, and Language Processing.
Finally, we deduce that the KP hypercardioid is at iteration $n$: \[ h_{1(n)}^t = h_{1(0)}^t \otimes h_{2(n)}^t. \] (98)

Figure 8 shows plots of the DFs and WNGs of the KP hypercardioid, $h_{2(n)}^{(0)}$, as a function of frequency for $M_1 = M_2 = 3$, $\delta = 5$ mm, and several values of $n$: $n = 1$ (solid line with circles), $n = 2$ (dashed line with asterisks), $n = 3$ (dotted line with squares), and $n = 10$ (dash-dot line with triangles). (a) DF and (b) WNG.

with

\[ \Gamma_{h_1^{(n)}} = \left( h_{1(0)}^{(n)} \otimes I_{M_2} \right)^H \Gamma_{h_1^{(0)}} \left( h_{1(0)}^{(n)} \otimes I_{M_2} \right). \] (97)

Finally, we deduce that the KP hypercardioid is at iteration $n$: \[ h_{1(n)}^t = h_{1(0)}^t \otimes h_{2(n)}^t. \] (98)

Figure 8 shows plots of the DFs and WNGs of the KP hypercardioid, $h_{2(n)}^{(0)}$, as a function of frequency for $M_1 = M_2 = 3$, $\delta = 5$ mm, and several values of $n$. We observe that the DF of the KP hypercardioid increases at each iteration, and roughly converges after five iterations, while the WNG decreases at each iteration.

**D. Supercardioid**

The supercardioid is obtained by maximizing the FBR. To fully maximize the FBR in (142), we need to derive an iterative algorithm.

At iteration 0, we may take the supercardioid of order $M_2 = 1$ at the second ULA, i.e.,

\[ h_{2}^{(0)} = \frac{t_2}{d_{2,0}^H t_2}. \] (99)

where $t_2$ is the eigenvector corresponding to the maximum eigenvalue of the matrix $\Gamma_{S,2}^{-1} \Gamma_{f,2}$. Substituting $h_{2}^{(0)}$ into (103) and (55), we get

\[ \begin{aligned} \Gamma_{f,h_2^{(0)}} &= \left( I_{M_1} \otimes h_{2}^{(0)} \right)^H \Gamma_{f} \left( I_{M_1} \otimes h_{2}^{(1)} \right), \quad \text{(100)} \\ \Gamma_{b,h_2^{(0)}} &= \left( I_{M_1} \otimes h_{2}^{(0)} \right)^H \Gamma_{b} \left( I_{M_1} \otimes h_{2}^{(0)} \right). \quad \text{(101)} \end{aligned} \]

Now, plugging these expressions into the FBR in (55), we obtain at iteration 1:

\[ \begin{aligned} \mathcal{F} \left( h_1^{(1)} | h_2^{(0)} \right) &= \frac{\Gamma_{f,h_2^{(0)}} h_1^{(1)} \left( h_1^{(1)} \right)^H}{\Gamma_{b,h_2^{(0)}} h_1^{(1)} \left( h_1^{(0)} \right)^H}. \quad \text{(102)} \end{aligned} \]

The maximization of $\mathcal{F} \left( h_1^{(1)} | h_2^{(0)} \right)$ with respect of $h_1^{(1)}$ leads to

\[ h_1^{(1)} = \frac{t_1^{(0)}}{d_{1,0}^H t_1^{(0)}}, \] (103)

where $t_1^{(0)}$ is the eigenvector corresponding to the maximum eigenvalue of the matrix $\Gamma_{b,h_2^{(0)}} \Gamma_{f,h_2^{(0)}}$. Using $h_1^{(1)}$ in (57) and (58), we get

\[ \begin{aligned} \Gamma_{f,h_1^{(1)}} &= \left( h_1^{(1)} \otimes I_{M_2} \right)^H \Gamma_{f} \left( h_1^{(0)} \otimes I_{M_2} \right), \quad \text{(104)} \\ \Gamma_{b,h_1^{(1)}} &= \left( h_1^{(1)} \otimes I_{M_2} \right)^H \Gamma_{b} \left( h_1^{(0)} \otimes I_{M_2} \right). \quad \text{(105)} \end{aligned} \]

As a result, the FBR in (56) is

\[ \mathcal{F} \left( h_2^{(1)} | h_1^{(1)} \right) = \frac{\Gamma_{f,h_2^{(1)}} h_2^{(1)}}{\Gamma_{b,h_1^{(1)}} h_1^{(1)}}, \] (106)

whose maximization with respect to $h_2^{(1)}$ gives

\[ h_2^{(1)} = \frac{t_2^{(1)}}{d_{2,0}^H t_2^{(0)}}, \] (107)

where $t_2^{(1)}$ is the eigenvector corresponding to the maximum eigenvalue of the matrix $\Gamma_{b,h_1^{(1)}} \Gamma_{f,h_1^{(1)}}^{-1}$. Continuing to iterate up to iteration $n$, we easily get for the first filter:

\[ h_1^{(n)} = \frac{t_1^{(n-1)}}{d_{1,0}^H t_1^{(n-1)}}, \] (108)

where $t_1^{(n-1)}$ is the eigenvector corresponding to the maximum eigenvalue of the matrix $\Gamma_{b,h_2^{(n-1)}} \Gamma_{f,h_2^{(n-1)}}$, with

\[ \begin{aligned} \Gamma_{f,h_2^{(n-1)}} &= \left( I_{M_1} \otimes h_{2}^{(n-1)} \right)^H \Gamma_{f} \left( I_{M_1} \otimes h_{2}^{(n-1)} \right), \quad \text{(109)} \\ \Gamma_{b,h_2^{(n-1)}} &= \left( I_{M_1} \otimes h_{2}^{(n-1)} \right)^H \Gamma_{b} \left( I_{M_1} \otimes h_{2}^{(n-1)} \right). \quad \text{(110)} \end{aligned} \]

and for the second filter:

\[ h_2^{(n)} = \frac{t_2^{(n)}}{d_{2,0}^H t_2^{(n)}}, \] (111)

where $t_2^{(n)}$ is the eigenvector corresponding to the maximum eigenvalue of the matrix $\Gamma_{b,h_1^{(n)}} \Gamma_{f,h_1^{(n)}}$, with

\[ \begin{aligned} \Gamma_{f,h_1^{(n)}} &= \left( h_1^{(n)} \otimes I_{M_2} \right)^H \Gamma_{f} \left( h_1^{(n)} \otimes I_{M_2} \right), \quad \text{(112)} \\ \Gamma_{b,h_1^{(n)}} &= \left( h_1^{(n)} \otimes I_{M_2} \right)^H \Gamma_{b} \left( h_1^{(n)} \otimes I_{M_2} \right). \quad \text{(113)} \end{aligned} \]

Finally, we deduce that the KP supercardioid is at iteration $n$:

\[ h_S^{(n)} = h_1^{(n)} \otimes h_2^{(n)} \] (114)

Figure 8 shows plots of the DFs, WNGs, and FBRs of the KP supercardioid, $h_S^{(n)}$, as a function of frequency for $M_1 = M_2 = 2$, $\delta = 5$ mm, and several values of $n$. We observe that the FBR of the KP supercardioid increases at each iteration, and roughly converges after three iterations, while the DF also increases at each iteration and the WNG remains almost the same.
VI. CONCLUSIONS

We have introduced differential Kronecker product beamforming based on Kronecker product decompositions of a physical array into two virtual arrays. We defined performance measures by using Kronecker product filters, and showed how to derive some useful examples of differential beamformers, namely the cardioid, dipole, hypercardioid, and supercardioid, using the Kronecker product formulation. We showed that the beampattern and the WNG of the global array can be expressed as the product of the counterparts of the two virtual arrays; but the DF and FBR cannot be factorized into a similar form and as a result these two measures cannot be maximized directly. Accordingly, we developed iterative algorithms for their maximization. The proposed approach enables much more flexibility in the design of differential beamformers, compared to the well-known and studied conventional approach. Furthermore, the proposed approach can be extended to other array structures and other beamformers, depending on the application at hand.

ACKNOWLEDGEMENTS

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

REFERENCES