ABSTRACT
This paper presents a region-of-interest (ROI) oriented constant-beamwidth (CB) beamforming approach with rectangular arrays. We decompose a global rectangular beamformer into a Kronecker-product (KP) of two linear sub-beamformers: a uniform constant-beamwidth beamformer along the $y$-axis and a nonuniform robust superdirective (SD) beamformer along the $x$-axis whose topology is optimized for a broadband array directivity criterion. The KP decomposition allows a flexible tradeoff control between the CB threshold frequency and the array directivity by tuning the number of microphones along each axis. The topology optimization considers a continuous ROI and therefore does not require an accurate desired source direction in space. The proposed approach is compared to uniform rectangular arrays with uniform differential sub-beamformers along the $x$-axis. We show that the proposed method is advantageous regarding white noise gain and directivity factor measures, especially when the desired source direction differs from the estimated source direction.

Index Terms— Microphone arrays, region-of-interest beamforming, constant-beamwidth beamformer, rectangular arrays.

1. INTRODUCTION
Beamforming has been a subject of numerous research works in the past few decades, taking advantage of spatial information acquired by a sensor array to estimate signals of interest from noisy observations [1, 2]. Apart from the spatial filter coefficients, that is, the “beamformer”, the array geometry has been shown to play a significant role in the beamforming performance. For example, uniform linear arrays (ULAs) may exhibit either high array directivity or robustness to white noise but typically not both simultaneously. They are also susceptible to the desired source direction of arrival (DOA). Uniform rectangular arrays (URAs), on the other hand, exhibit reduced susceptibility to the desired signal’s DOA and may be designed in a flexible manner by invoking the KP beamforming framework, which potentially enables optimization concerning several design criteria at once [3, 8]. Uniform circular arrays (UCAs) and uniform concentric circular arrays (UCCAs) are known to be invariant to the DOA, with the latter geometry even shown to attain the constant-beamwidth (CB) property [9, 11]. Despite this, none of these geometries can easily be integrated into the KP beamforming framework.

On top of their design flexibility, URAs have been shown valuable for DOA estimation methods [12, 13]. In the context of differential beamforming, when the interelement spacing (along both axes) is small [14, 15]. These approaches improve the array robustness to spatially white noise or allow high directivity beamforming even when the desired signal significantly deviates from the endfire direction. Nevertheless, these two attributes are not attained simultaneously, and neither of these approaches exhibits the CB property.

Using URAs for differential and CB beamforming has recently been proposed to control the white noise gain (WNG), the directivity factor (DF), and the threshold frequency from which the CB property holds [8]. This was achieved by designing a linear CB sub-beamformer [16–18] along the $y$-axis and a linear (uniform) differential sub-beamformer [19–21] along the $x$-axis; then, the URA was obtained by applying a KP between the two linear beamformers. While this approach is indeed flexible, it assumes the desired signal impinges on the array from the endfire direction, resulting from using a uniform differential sub-beamformer along the $x$-axis.

This paper presents a region-of-interest (ROI) oriented CB beamforming approach with rectangular arrays (RAs). We decompose a global rectangular beamformer into a KP of two linear sub-beamformers: a uniform constant-beamwidth beamformer along the $y$-axis and a nonuniform robust superdirective (SD) beamformer along the $x$-axis whose topology is optimized with respect to a broadband array directivity criterion, with the latter inspired by the approach proposed in [23]. The KP decomposition allows a flexible tradeoff control between the CB threshold frequency and the array directivity by tuning the number of microphones along each axis. In contrast, the topology optimization considers a continuous ROI and does not require an accurate desired source direction in space. The proposed approach is compared to URAs with uniform differential sub-beamformers along the $x$-axis. We demonstrate that the proposed approach is preferable regarding WNG and DF measures, notably when the desired signal deviates from the endfire direction or in high frequencies.

The rest of the paper is organized as follows. In Section 2 we present the signal model. In Section 3 we briefly review the KP beamforming. Section 4 presents the proposed nonuniform rectangular beamformer’s array topology optimization and derivation. Finally, Section 5 demonstrates the advantage of the proposed approach compared to URAs.

2. SIGNAL MODEL
Consider a farfield signal of interest propagating from an azimuth angle $\phi$ and an elevation angle $\theta$ in an anechoic acoustic environment at the speed of sound, i.e., $c = 340$ m/s. The plane wave impinges on a two-dimensional (2-D) microphone array located on the $x$-$y$ plane, which contains $M_x$ and $M_y$ omnidirectional microphones along the $x$-axis and $y$-axis, respectively. The 2-D array is uniform concerning the $y$-axis, with an interelement spacing $\delta_y$, and nonuniform concerning the $x$-axis. Defining the microphone located at $(0, 0)$ as the origin of the Cartesian coordinate, the array
steering vector of length $M_M M_F$ is expressed by [2]:

$$d_{\theta,\phi}(\omega) = \begin{bmatrix} B_{\theta,\phi,1}(\omega) a_{\theta,\phi}(\omega) \\ \vdots \\ B_{\theta,\phi,M_M}(\omega) a_{\theta,\phi}(\omega) \end{bmatrix}^T = b_{\theta,\phi}(\omega) \odot a_{\theta,\phi}(\omega),$$

where

$$a_{\theta,\phi}(\omega) = \begin{bmatrix} A_{\theta,\phi,1}(\omega) \\ \vdots \\ A_{\theta,\phi,M_M}(\omega) \end{bmatrix}^T = \begin{bmatrix} e^{j\theta_{\theta,x}(\omega)} \\ \vdots \\ e^{j\theta_{\theta,\phi,M_M}(\omega)} \end{bmatrix}^T$$

is the steering vector associated with the $x$-axis with $\{x_m\}_{m=1}^{M_M}$ denoting the microphone locations, and

$$b_{\theta,\phi}(\omega) = \begin{bmatrix} B_{\theta,\phi,1}(\omega) & B_{\theta,\phi,2}(\omega) & \cdots & B_{\theta,\phi,M_M}(\omega) \end{bmatrix}^T = \begin{bmatrix} 1 & e^{j\theta_{\phi,y}(\omega)} & \cdots & e^{j(M_M-1)\theta_{\phi,y}(\omega)} \end{bmatrix}^T$$

is the steering vector associated with the $y$ axis, where

$$\omega_{\theta,\phi,x}(\omega) = \omega \sin \theta \cos \phi,$$

$$\omega_{\theta,\phi,y}(\omega) = \omega \sin \theta \sin \phi.$$

The superscript $^T$ denotes the transpose operator, $\odot$ is the KP operator, $j = \sqrt{-1}$ is the imaginary unit, $\omega = 2\pi f$ is the angular frequency and $f > 0$ is the temporal frequency.

The observed signal vector of length $M_M M_F$ can be expressed in the frequency domain as [22]:

$$y(\omega) = \begin{bmatrix} y_1^T(\omega) \\ \vdots \\ y_{M_M}^T(\omega) \end{bmatrix}^T = x(\omega) + v(\omega) = d_{\theta,\phi}(\omega) X(\omega) + v(\omega),$$

where $X(\omega)$ is the zero-mean desired source signal, $v(\omega)$ is the zero-mean additive noise signal vector, and

$$y_{m_{y}}(\omega) = \begin{bmatrix} Y_{m_{y},1}(\omega) \\ \vdots \\ Y_{m_{y},M_M}(\omega) \end{bmatrix}^T = \begin{bmatrix} x_{m_{y}}(\omega) + v_{m_{y}}(\omega) \end{bmatrix}^T = B_{\theta,\phi,M_M}(\omega) a_{\theta,\phi}(\omega) X(\omega) + v_{m_{y}}(\omega),$$

for $m_{y} = 1, 2, \ldots, M_F$, is the observed signal vector of length $M_M$ of the $m_{y}$-th nonuniform linear array parallel to the $x$ axis. Denoting the desired signal incident angle by $(\theta_0, \phi_0)$ and dropping the dependence on $\omega$, [4] becomes:

$$y = (b_{\theta_0,\phi_0} \odot a_{\theta_0,\phi_0}) X + v,$$

where $b_{\theta_0,\phi_0} \odot a_{\theta_0,\phi_0} = d_{\theta_0,\phi_0}$ is the steering matrix at $(\theta_0, \phi_0)$, and the covariance matrix of $y$ is

$$\Phi_y = E(yyyy^H) = px d_{\theta_0,\phi_0} d_{\theta_0,\phi_0}^H + \Phi_v,$$

where $E(\cdot)$ denotes mathematical expectation, the superscript $^H$ is the conjugate-transpose operator, $px = E(|X|^2)$ is the variance of $X$, and $\Phi_v = E(\cdot\cdot\cdot v v^H)$ is the covariance matrix of $v$. Assuming the noise variance is approximately the same at all sensors, we can express [7] as

$$\Phi_y = px d_{\theta_0,\phi_0} d_{\theta_0,\phi_0}^H + p v \Gamma_v,$$

where $p v$ is the noise variance at the reference microphone (i.e., the microphone at the origin of the Cartesian coordinate system) and $\Gamma_v = \Phi_v / p v$ is the pseudo-coherence matrix of the noise. From [8], we deduce that the input signal-to-noise ratio (SNR) is

$$iSNR = \frac{px d_{\theta_0,\phi_0} d_{\theta_0,\phi_0}^H}{px p v} = \frac{px}{pf \Gamma_v f},$$

where $tr(\cdot)$ denotes the trace of a square matrix.

### 3. Kronecker-Product Beamforming

In this section, we present a global rectangular beamformer $f$ of length $M_M M_F$ as a KP of two linear sub-beamformers designed with respect to each axis of the RA. Hence, $f$ is of the form:

$$f = w \odot h,$$

where $h$ is a linear sub-beamformer of length $M_F$ and $w$ is a linear sub-beamformer of length $M_M$. Then, the beamformer output signal is

$$Z = f^H y = X_{fd} + V_{in},$$

where $Z$ is an estimate of $X$,

$$X_{fd} = \left( w^H b_{\theta_0,\phi_0} \right) h^H a_{\theta_0,\phi_0} X,$$

is the filtered desired signal, and

$$V_{in} = (w \odot h)^H v,$$

is the residual noise. In addition, it is clear that a distortionless constraint is satisfied by

$$h^H a_{\theta_0,\phi_0} = 1, \ w^H b_{\theta_0,\phi_0} = 1.$$

The output SNR and the gain in SNR are, respectively,

$$oSNR(f) = \frac{px}{pv} \times \frac{|f^H d_{\theta_0,\phi_0}|^2}{f^H \Gamma_v f},$$

and

$$G(f) = \frac{oSNR(f)}{iSNR} = \frac{|f^H d_{\theta_0,\phi_0}|^2}{f^H \Gamma_v f},$$

from which we deduce the WNG:

$$\mathcal{W}(f) = \frac{|f^H d_{\theta_0,\phi_0}|^2}{f^H f} = \frac{|w^H b_{\theta_0,\phi_0}|^2}{w^H w} \times \frac{|h^H a_{\theta_0,\phi_0}|^2}{h^H h},$$

and the DF:

$$D(f) = \frac{|f^H d_{\theta_0,\phi_0}|^2}{f^H \Gamma_{fd} f},$$

where $\Gamma_{fd}$ is the pseudo-coherence matrix of the spherically isotropic (diffuse) noise field $\mathcal{W}(\cdot)$.

The beampattern is given by

$$B_{\theta,\phi}(f) = f^H d_{\theta,\phi} = \left( w^H b_{\theta,\phi} \right) h^H a_{\theta,\phi} = B_{\theta,\phi}(w) B_{\theta,\phi}(h),$$

where $B_{\theta,\phi}(w) = w^H b_{\theta,\phi}$ is the beampattern of $w$ and $B_{\theta,\phi}(h) = h^H a_{\theta,\phi}$ is the beampattern of $h$. 
4. OPTIMAL ROI-ORIENTED CONSTANT-BEAMWIDTH BEAMFORMING

Consider deriving a rectangular version of the SD beamformer [22], which is not necessarily a KP beamformer. That is, we are interested in solving

\[
\min_f \mathbf{f}^H \Gamma_d \mathbf{f} \quad \text{s. t.} \quad \mathbf{f}^H \mathbf{d}_{\theta_0, \phi_0} = 1, \tag{20}
\]

whose solution is obtained by

\[
\mathbf{f}_{SD} = \frac{\Gamma_{d,1}^{-1} \mathbf{d}_{\theta_0, \phi_0}}{\mathbf{d}_{\theta_0, \phi_0}^H \Gamma_{d,1} \mathbf{d}_{\theta_0, \phi_0}} = \left[ \sum_{\mu=1}^{M} \mathbf{J}_{\mu,p} \otimes \Gamma_{d,p} \right]^{-1} \mathbf{d}_{\theta_0, \phi_0}, \tag{21}
\]

where \( \Gamma_{d,p} \) is the \( p \)-th \( M_x \times M_x \) block in the top block row of \( \Gamma_d \) and

\[
(\mathbf{J}_{\mu,p})_{ij} = \begin{cases} 1, & |i-j| = p \\ 0, & |i-j| \neq p \end{cases}, \tag{22}
\]

is a binary matrix of size \( M_y \times M_y \) with ones on the \( \pm p \)-th diagonals and zeros elsewhere. In particular, \( \mathbf{J}_{\mu,0} = \mathbf{I}_{M_y} \), which is the identity matrix of size \( M_y \times M_y \). Assuming

\[
|x_{m_i} - x_{m_j}| \leq \delta_y, \quad \forall m_i, m_j \in [1, M_x], \tag{23}
\]

Figure 1: Optimal array topologies for two distinct array ranges. (a) \( M_x = 4, M_y = 11 \) and (b) \( M_x = 6, M_y = 7 \).

\[
\theta_0 = \pi/2, \text{ where } \phi_H \text{ is taken as half of the mainlobe beamwidth of } \mathbf{w}_{\text{rect}}. \text{ This is performed in the same manner described in [23] and yields } \text{the set of optimal microphone locations } \{x_{m_i}, y_{m_j}\} \text{ with the first microphone always set to } x_1 = 0. \text{ Then, as a second step, we optimize a (normalized) robust SD beamformer } \mathbf{h}_{SD, opt, \epsilon} \text{ whose WNG in (17) is designed according to the desirable WNG of the rectangular beamformer, considering the optimal microphone locations along the x-axis, and the approach proposed in [24]. Note that the latter was used in the coefficient post-processing performed in [23]. However, in our work, this derivation is only performed once concerning the endfire direction. Therefore, the exact desired signal location in space is not required to be known in advance nor estimated as long as it is within the ROI: the linear array topology is optimized considering the entire ROI and the SD sub-beamformer is optimized concerning the endfire direction, assuming the desired signal is more likely to be found there. Finally, the rectangular ROI-oriented CB beamformer is given by:

\[
\mathbf{f}_{ROI/CB} = \mathbf{w}_{\text{rect}} \otimes \mathbf{h}_{SD, opt, \epsilon}. \tag{27}
\]

It is worthwhile comparing the proposed approach to the approach suggested in [8]. While both approaches guarantee the CB property to hold, the latter provides much flexibility in setting either the WNG or DF and controlling their inherent tradeoff, yet it suffers from several drawbacks. Notably, the desired-source location is assumed to be known and specific, that is, on the x-y plane in the endfire direction (\( \theta_0 = \pi/2, \phi_0 = 0 \)). Any deviation from this direction may significantly deteriorate the array gains. In addition, the rectangular array topology is uniform concerning both axes and, therefore, not optimized with respect, for example, to the array directivity.

5. SIMULATIONS

Let us demonstrate the performance of the proposed approach compared to the uniform-differential CB beamformer [8,25]:

\[
\mathbf{f}_{diff/CB} = \mathbf{w}_{\text{rect}} \otimes \mathbf{h}_{SD, \epsilon}. \tag{28}
\]
where the CB sub-beamformer $w_{\text{rect}}$ is identical to $f_{\text{ROI/CB}}$ and $h_{\text{ROI/CB}}$, a robust SD beamformer \cite{24} optimized with respect to a uniform differential array with an interelement spacing of 5 mm. In addition, we set the ROI’s upper edge to $\phi_H = 20^\circ$, the frequency range to $\{\omega_L, \omega_H\} = [1, 6]$ kHz, the mainlobe beamwidth to $40^\circ$, $\delta_\phi = 4$ cm, the (maximal) aperture of the x-axis to 8 cm with 17 equally-spaced grid points, and the minimal value of the WNG to $-10$ dB. Note that here, we limit the scope of our work to signals of interest located on the x-y plane that is, $\theta_0 = \pi/2$, but this can be easily generalized to any $\theta_0$. We simulate two distinct array settings with each of the two referred beamformers: $M_x = 4, M_y = 11$ and $M_x = 6, M_y = 7$.

Figure 2 depicts the two optimal array topologies for the simulated settings. We observe that in both cases, the optimal topology is not uniform but symmetric around the center of the actual aperture, which may be interpreted as an attempt to maximize the array directivity (narrower-spaced microphones) and the white noise amplification (wider-spaced microphones). In addition, the maximal directivity (narrower-spaced microphones) and the white noise amplitude, which may be interpreted as an attempt to maximize the array directivity, which is well known to be maximized with closely-spaced microphones, particularly concerning the endfire direction \cite{22}. Note that the condition in (23) does not strictly hold for all the microphones. This merely implies that the proposed approach differs from the rectangular SD beamformer.

Figure 2 shows the WNG and DF for each of the two RAs with $M_x = 4, M_y = 11$. Finally, we note that both beamformers exhibit a beampattern’s discontinuity which originates like the modified rectangular beamformer $w_{\text{rect}}$ \cite{10}

Figure 3 shows the beampatterns of the two RAs for the two addressed array settings. It is clear that $M_y$ alone determines the CB threshold frequency, above which the mainlobe beamwidth remains constant. This implies that considering the CB property, the optimal and the uniform differential beamformers behave similarly. Finally, comparing the sidelobe levels of each beamformer concerning itself but with the complementary settings, we infer that the higher $M_y$ is, the higher the array directivity. This implies that when the total number of microphones in the RA $M_x M_y$ is roughly constant, the relative values of $M_x$ and $M_y$ control the tradeoff between the array directivity and the CB threshold frequency.

6. CONCLUSIONS

We have introduced an ROI-oriented CB beamforming approach for RAs that maximizes broadband array directivity. We designed a CB ULA along the y-axis and a nonuniform linear array along the x-axis whose topology was optimized for the broadband array directivity of the RA. Then, the RA was obtained as a KP of the two linear arrays. Since the proposed approach considers a continuous ROI rather than a single direction in space, acquiring the desired source’s precise direction is unnecessary. Our method was demonstrated to allow a flexible tuning between the CB threshold frequency by increasing $M_y$ and improving the array directivity by increasing $M_x$. Finally, we have compared the proposed URA approach with uniform differential sub-beamformers along the x-axis. The advantages of the proposed approach were demonstrated in terms of WNG and DF measures, either when the DOA deviates from the endfire direction or in high frequencies.

Figure 2: WNG and DF in the ROI for the optimal and uniform-differential array topologies. (a) WNG with $f_{\text{ROI/CB}}$, (b) WNG with $f_{\text{diff/CB}}$, (c) DF with $f_{\text{ROI/CB}}$, and (d) DF with $f_{\text{diff/CB}}$. Array settings: $M_x = 6, M_y = 7$.

Figure 3: Beampatterns as a function of the frequency and the azimuth angle $\phi$ for different values of $M_x$ and $M_y$. Red dashed lines indicate the desirable mainlobe beamwidth. (a) $f_{\text{ROI/CB}}$: $M_x = 4, M_y = 11$, (b) $f_{\text{diff/CB}}$: $M_x = 4, M_y = 11$, (c) $f_{\text{ROI/CB}}$: $M_x = 6, M_y = 7$, and (d) $f_{\text{diff/CB}}$: $M_x = 6, M_y = 7$. 
7. REFERENCES


