IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 27, 2019, accepted July 29, 2019, date of publication July 31, 2019, date of current version August 29, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2932315

Beamforming Designs Robust to Propagation Model Estimation Errors for Binaural Hearing Aids

HALA AS'AD[1], MARTIN BOUCHARD[1], AND HOMAYOUN KAMKAR-PARSI[2]
[1]School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, ON K1N 6N5, Canada
[2]WS Audiology, 91058 Erlangen, Germany

Corresponding author: Hala As'ad (hasad056@uottawa.ca)

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant.

ABSTRACT This work introduces new binaural beamforming algorithms for hearing aids, with a good robustness to errors in the estimated target source propagation model. Two different binaural beamforming designs are proposed. One design is based on an adaptive null positioning scheme, and the second is based on a combination of the adaptive null positioning scheme and wider beampatterns. Simulations are performed using signals and propagation models obtained from multichannel binaural hearing aids recordings, including some in a mildly reverberant environment. Evaluations are done in terms of noise reduction and target distortion. Evaluation results illustrate the robustness of the two proposed designs to errors between the true and estimated directions of arrival for the target source, and to mismatch between the anechoic propagation models used for the beamformers designs and the reverberant propagation models used to generate the signals at the sensors or beamformer inputs. Both designs surpass the performance of standard binaural Minimum Variance Distortionless Response (MVDR) and binaural Generalized-Side Lobe Canceler (GSC) beamformers.

INDEX TERMS Hearing aids, noise reduction, robust binaural beamformer, source propagation model error, direction of arrival (DOA) estimation error, head related transfer function (HRTF) mismatch, target distortion.

I. INTRODUCTION

Hearing aids are a well-known solution for several pathologies in the auditory system such as sensorineural hearing loss. Hearing aids are good candidates to help their users to perceive audible, intelligible and nearly natural sounds. However, they still produce limited performance under some complex acoustic scenarios [1], [2]. Beamforming algorithms for hearing aids have been extensively researched in the literature [3], [4]. Many of these beamforming algorithms rely on assumptions such as fairly stationary sound sources and the availability of a sophisticated target voice activity detector (VAD) [5], [6]. However, these assumptions may not be valid in some practical situations with multiple speakers and time-varying activity patterns. The presence of background noise and/or reverberation can make the situation even more challenging. When the assumptions are not met, this causes deterioration in the performance of hearing aids, with a negative impact on the resulting processed acoustic signals and their perception by the user.

In binaural hearing aids, i.e., with bidirectional wireless links for signal transmission between the two ears, the binaural beamformer performance can be significantly affected by a mismatch or an error between the target source propagation model assumed for beamformer design and the actual physical source propagation model. This includes errors in the estimated target direction of arrival (DOA) used in the beamformer algorithms, i.e., target DOA mismatch. This kind of mismatch is generated from imperfect DOA estimation schemes, from small head movements of the hearing aid user, and from multipath propagation. Acoustic beamforming methods which attempt to address this problem have been introduced in the literature, although not specifically for binaural hearing aids. These approaches can be divided into two

The associate editor coordinating the review of this article and approving it for publication was Guolong Cui.

categories; the first category is robust Generalized Sidelobe Canceler (GSC) algorithms, where the GSC beamformer has three components: a Fixed Beamformer (FB), a Blocking Matrix (BM), and an Adaptive Noise Canceler unit (ANC). The second category is robust beamformers that do not use the GSC architecture, i.e., without a BM and an ANC unit in the design.

Some of the robust GSC approaches have been designed by using an adaptive BM that uses coefficient-constrained adaptive filters (CCAFs) and a multiple-input canceler (MC) that uses norm-constrained adaptive filters (NCAFs) [7], [8]. These approaches are introduced to minimize the target leakage at the output of the BM and to prevent the undesired target cancelation in case of target leakage from the BM. However, these designs are based on a detection of high-SNR periods, which requires a target VAD. A target VAD can be very difficult to design in complex environments with multiple dynamic sources, reverberation, and high levels of noise. Moreover, these approaches are more suitable for suppressing directional interferers, and their performance can significantly degrade under isotropic ambient noise. Another robust GSC beamforming approach with enhanced suppression abilities for both isotropic ambient noise and directional interferers has been proposed based on a sound-source presence probability in the BM [9]. The sound-source presence probability measures the probability of the target activity at each frequency bin in the output of the FB. However, this design also relies on a target VAD. Other approaches are based on the independent component analysis (ICA) method [10]. However, it is known that ICA methods cannot cope well with environments having dynamic moving sources and variable activity patterns, as well as overdetermined problems (more sources than sensors). In some robust GSC designs such as the design in [11], the ANC is replaced by a crosstalk-resistant adaptive noise canceler. The crosstalk-resistant adaptive noise canceler is based on two adaptation step sizes, which depend on the instantaneous SNR. However, like the target VAD problem, such an instantaneous SNR estimation can also be difficult to achieve in low SNR multi-talker environments.

Beamforming designs without the BM and ANC units and robust to propagation model estimation errors have also been introduced in the literature. A steerable binaural MVDR beamformer with a target DOA estimator was proposed in [12] in order to improve the robustness of the binaural MVDR. A discriminative Support Vector Machine (SVM) was used as a binaural classifier to estimate the target DOA. The classifier is trained under acoustic scenarios of target speakers from different directions and diffuse noise only (without directional interferers). However, this will lead to degradation in the DOA estimator/ classifier performance in the presence of directional interferers.

Another attempt to enhance the robustness of the MVDR beamformer to propagation model estimation errors was introduced by applying a White Noise Gain (WNG) constraint in the MVDR design. The MVDR beamformer is sensitive to noise uncorrelated between different sensors, leading to noise amplification or WNG. By designing a WNG-constrained MVDR beamformer, reduction in the uncorrelated noise amplification is achieved, and at the same time the designed beamformer has shown a better robustness under mismatch conditions [13]. However, it is difficult to choose the value of the tunable parameters that control the WNG amplification in order to achieve an acceptable level of robustness to mismatch, as the level of propagation model mismatch does not have a direct relation with the uncorrelated noise. Several other approaches were introduced to improve the robustness of the MVDR beamformer such as the work in [14], [15]. Alternative approaches based on the constrained least-squares beamformer [16], [17] and the eigenvector constrained minimum power beamformer [18], [19] were also investigated and developed to have robustness to the mismatch conditions. Another recently proposed algorithm is based on sparsely selecting the position of the sensors for each frequency bin [20]. However, this approach is not suitable for hearing aids with small size and limited number of microphones.

Overall, without using sophisticated components difficult to achieve in real-life challenging environments (such as target VAD, noise-only correlation matrix estimation, instantaneous SNR estimation, speech probability presence estimation, etc.), the methods proposed so far in the literature have not shown a capability to produce a good performance robust to target DOA mismatch for the conditions under which binaural hearing aids need to operate: small number of microphones available, wide variety of possible acoustic scenarios, and for each of those scenarios a good reduction is required both for competing talkers (directional interferers) and for diffuse-like background noise. Therefore, for binaural hearing aids there is still a need to develop a practical beamforming algorithm robust to target DOA mismatch and adaptive to the acoustic environment, which does not require information or estimates often not available in difficult acoustic scenarios. This is the main contribution of this work.

In this work, two different binaural beamformers with improved robustness to target DOA mismatch are developed. A minimal target attenuation is achieved in a zone around the estimated/assumed target DOA, and a significant attenuation of interferers and noise is achieved outside the target zone, for different target directions and for different frequencies. An estimate of the target DOA is assumed to be known in this paper, with ± 10 degrees accuracy. Several approaches have been proposed in the literature in order to estimate the DOA of directional sources such as the work in [21]–[27]. However, target DOA estimation is outside the scope of this paper, and this paper does not use any target DOA estimation scheme. The proposed beamformers are also shown to be robust to another type of error in the source propagation model, which is the error or mismatch between the anechoic propagation models used for the beamformers designs and the reverberant propagation models used to generate the signals at the sensors or beamformer inputs.

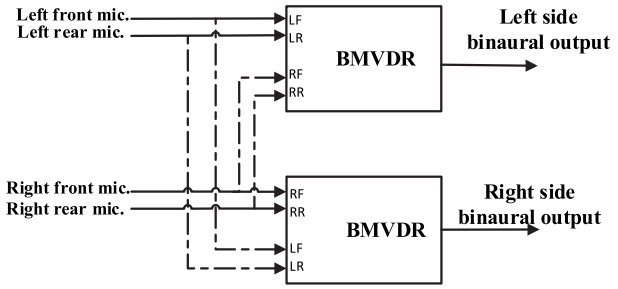

FIGURE 1. 2+2 microphone configuration. Dashed lines represent the wireless links.

This paper is organized as follows. Section II provides the system notations that will be used throughout this paper. Section III provides the mathematical derivation of a beamformer to be used as a baseline for comparison with the developed robust beamforming designs, and introduces the two developed beamformers robust to target DOA mismatch. Section IV provides detailed experimental evaluations of the proposed beamformers. Finally, Section V provides a conclusion.

II. SYSTEM NOTATIONS

Binaural hearing aids with two microphones at each ear, i.e., four microphones in total, are considered as shown in Fig. 1. Binaural wireless links between the hearing aid at each ear allow bidirectional transmission of the microphone signals between the two sides. Ideal links are considered in this work (no jitter, delay, packet loss, etc.). The input noisy microphone signals in the time-frequency (T-F) domain can be written as:

$$y_m(f,t) = x_{\text{in},m}(f,t) + v_{\text{in},m}(f,t) \tag{1}$$

where, m is the microphone index such that the input noisy signal at the Front Left (FL) microphone is $y_1(f,t)$, the input noisy signal at the Rear Left (RL) is $y_2(f,t)$, the input noisy signal at the Front Right (FR) microphone is $y_3(f,t)$, and the input noisy signal at the Rear Right (RR) microphone is $y_4(f,t)$. At the m^{th} microphone, $x_{\text{in},m}(f,t)$ and $v_{\text{in},m}(f,t)$ are respectively the input target speaker component and the sum of all directional interfering speakers combined with the diffuse-like background noise component. f is the frequency index and t is the time index, i.e., frame index or sub-band downsampled rate time index. Equation (1) can be re-written as:

$$\mathbf{y}(f,t) = \mathbf{x}(f,t) + \mathbf{v}(f,t) \tag{2}$$

where, $\mathbf{y}(f,t) = [y_1(f,t), y_2(f,t), y_3(f,t), y_4(f,t)]^T$, $\mathbf{x}(f,t) = [x_{\text{in},1}(f,t), x_{\text{in},2}(f,t), x_{\text{in},3}(f,t), x_{\text{in},4}(f,t)]^T$, and $\mathbf{v}(f,t) = [v_{\text{in},1}(f,t), v_{\text{in},2}(f,t), v_{\text{in},3}(f,t), v_{\text{in},4}(f,t)]^T$.

The input target signal $\mathbf{x}(f,t)$ can be written in terms of a source signal $s_x(f,t)$ coming from an angle θ_s. The directivity vector $\mathbf{d}(f,\theta_s)$, which is the frequency response between the target source and each microphone, can represent the propagation model and be written as in (3):

$$\mathbf{x}(f,t) = \mathbf{d}(f,\theta_s)s_x(f,t) \tag{3}$$

where, the target directivity vector $\mathbf{d}(f,\theta_s)$ is $[d_1(f,\theta_s), ..., d_4(f,\theta_s)]$. In hearing aids, Head-Related Transfer Functions (HRTFs) measured in an anechoic ("dry") environment are often used as the directivity vectors, since they include the effect of the head, ears and torso in the beamformer design. The FL and FR microphones are normally used as the reference microphone on their respective side for the beamformer design. Therefore, we will refer to the target directivity vector at the reference microphone as $d_{ref,l}(f,\theta_S) = d_1(f,\theta_S)$ if the FL microphone is used as a reference, and as $d_{ref,r}(f,\theta_S) = d_3(f,\theta_S)$ if the FR microphone is used as a reference.

The correlation matrix of the input noisy signals can be written as in (4) and (5) in terms of the individual input components (assuming uncorrelated sources):

$$\mathbf{R}_y(f,t) = E\{\mathbf{y}(f,t)\mathbf{y}^H(f,t)\} \tag{4}$$

$$\begin{aligned}\mathbf{R}_y(f,t) &= \mathbf{R}_x(f,t)+\mathbf{R}_v(f,t)\\ &= E\{\mathbf{x}(f,t)\mathbf{x}^H(f,t)\} + E\{\mathbf{v}(f,t)\mathbf{v}^H(f,t)\}\\ &= \mathbf{d}(f,\theta_s)\mathbf{d}^H(f,\theta_s)E\{|s_x(f,t)|^2\}+E\{\mathbf{v}(f,t)\mathbf{v}^H(f,t)\}\end{aligned} \tag{5}$$

The superscript H refers to "Hermitian", which is the complex conjugate transpose.

III. BINAURAL BEAMFORMER ALGORITHMS

A. BASELINE DESIGN

In order to design a binaural beamformer which is robust to errors in the estimated target DOA, it is required to have a good baseline beamformer that performs well under ideal conditions, i.e., for cases without errors in the target DOA estimation or in the source propagation model (HRTF). For example, errors in the source propagation model (HRTF mismatch) come from the mismatch between anechoic HRTFs used in designing the binaural beamformer and the reverberant HRTFs which generate the beamformer input signals. It should be noted that what we call reverberant HRTFs in this paper is the frequency response of what is often referred to as Binaural Room Impulse Responses (BRIRs) in the literature. For this baseline design, the Binaural Minimum Variance Distortionless response (BMVDR) with the microphone configuration in Fig. 1 is expected to lead to a good noise reduction. The BMVDR is a practical state-of-the-art algorithm [28, first alternative]. We call the BMVDR beamformer structure in Fig.1 a "2+2" structure, as it uses two microphones from each ear. The BMVDR design is based on a constrained minimization of the beamformer output power. In the BMVDR, the response of the beamformer in the direction of the target signal is constrained to be equal to the target signal at the reference microphone at each ear, leading to (6) and (7). In the rest of this subsection and in the following subsection, time and frequency indices have been discarded

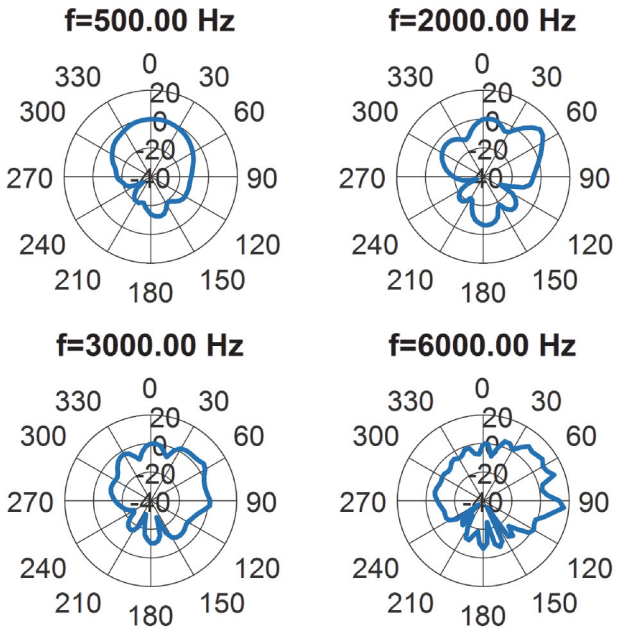

FIGURE 2. Beampatterns for the 2+2 BMVDR with a frontal target at 0 degree, left side, under 2-D diffuse noise conditions.

for simplicity.

$$\min_{\mathbf{W}} \mathbf{w}_l^H \mathbf{R}_y \mathbf{w}_l \quad \text{s.t. } \mathbf{w}_l^H \mathbf{d}(\theta_s) = d_{\text{ref},l}(\theta_s) \tag{6}$$

$$\min_{\mathbf{W}} \mathbf{w}_r^H \mathbf{R}_y \mathbf{w}_r \quad \text{s.t. } \mathbf{w}_r^H \mathbf{d}(\theta_s) = d_{\text{ref},r}(\theta_s) \tag{7}$$

To solve the constrained minimization problems in (6) and (7), the derivation in [28] is used to get the left and right beamformer coefficients as in (8) and (9), respectively.

$$\mathbf{w}_l = \frac{\mathbf{R}_y^{-1}\mathbf{d}(\theta_S)}{\mathbf{d}^H(\theta_S)\mathbf{R}_y^{-1}\mathbf{d}(\theta_S)} d_{\text{ref},l}^H(\theta_S) \tag{8}$$

$$\mathbf{w}_r = \frac{\mathbf{R}_y^{-1}\mathbf{d}(\theta_S)}{\mathbf{d}^H(\theta_S)\mathbf{R}_y^{-1}\mathbf{d}(\theta_S)} d_{\text{ref},r}^H(\theta_S) \tag{9}$$

Unlike the BVMDR in [28], the noisy signal correlation matrix $\mathbf{R}_y$ is used in (8) and (9) instead of the correlation matrix for the noise-only components $\mathbf{R}_v$. This is to avoid the need of a sophisticated online target VAD system to estimate the second order statistics of the noise-only components, which is not available in practice for complicated acoustic environments, as previously discussed. However, the disadvantage of using the noisy correlation matrix $\mathbf{R}_y$ is the increased sensitivity to target DOA mismatch and HRTF mismatch [29].

The BMVDR uses only one constraint in the estimated target direction, leading to a limited control of the beampattern, which affects the beamformer robustness to target DOA mismatch. The weak robustness to target DOA mismatch for the binaural beamformer is generated from the larger microphone spacing (distance between ears), which leads to narrower beams especially for high frequency components. Fig. 2 illustrates the narrow beams (and therefore potential for target DOA mismatch sensitivity issue) in the beampattern

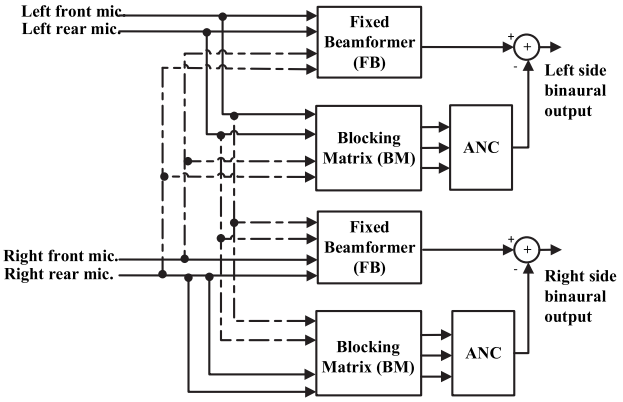

FIGURE 3. Generalized-side lobe canceler (GSC). Dashed lines represent the wireless links.

response of a BMVDR beamformer with a target at 0 degree and 2-D diffuse noise conditions, using HRTFs measured from behind-the-ear (BTE) hearing aid units on a mannequin in an anechoic environment. The beampatterns $BP(\theta)$ are computed as in (10)- (12):

$$\mathbf{R}_n = \frac{1}{N}\sum_{n=1}^{N} \mathbf{d}(\theta_n)\mathbf{d}^H(\theta_n) \tag{10}$$

where,$\mathbf{R}_n$ is the diffuse noise correlation matrix computed using 72 dry anechoic HRTFs with direction θ_n varying from 0 degree to 355 degrees (steps of 5 degrees) [30].$\mathbf{w}_i$ in (11) is a vector with the beamforming coefficients,θ_S is at 0 degree, and i is an index representing either the left or right side (l, r):

$$\mathbf{w}_i = \frac{\mathbf{R}_n^{-1}\mathbf{d}(\theta_S)}{\mathbf{d}^H(\theta_S)\mathbf{R}_n^{-1}\mathbf{d}(\theta_S)} d_{\text{ref},i}^H(\theta_S) \tag{11}$$

$$BP_i(\theta) = |\mathbf{w}_i^H \mathbf{d}(\theta)|^2. \tag{12}$$

It is noticeable that the beampatterns start to be narrower in the direction of the target (0 degrees here) as the frequency increases. A narrow beam response means sensitivity to target DOA mismatch, leading to target attenuation in the binaural output signals and a reduction of the array gain or signal to noise ratio improvement in the case of DOA mismatch.

The Generalized-Side Lobe Canceler (GSC) proposed in [31] is also used as a baseline design for comparison. The GSC beamformer structure consists of three parts: a Fixed Beamformer (FB), a Blocking Matrix (BM), and an Adaptive Noise Canceler unit (ANC), as Fig. 3 shows. In [31], a traditional delay-and-sum beamformer was used for the FB. In this work, we follow a similar approach, but since the signal propagation is not in free field we first "equalize" the signals from the different input microphones such that the target component in all the signals has the same magnitude and phase at each frequency. This is achieved by multiplying the different input microphone signals with the inverse of the acoustic transfer function ratios for the different microphones (HRTF ratios, relative transfer functions), with a ratio relative to a reference microphone and in the target direction. Adding the different equalized signals then

becomes a coherent sum for the target component (a.k.a. matched filter beamforming), and subtracting two equalized signals (from two microphones) produces a resulting signal with the target component removed (blocking matrix output). Consequently, for the traditional GSC in this paper, the FB is simply the average of the equalized microphone signals, and for the BM the equalized signals from the $M-1$ non-reference microphones are subtracted from the reference microphone signal to produce $M-1$ BM output signals.

Finally, the ANC unit uses the fixed beamformer output as a "desired" signal to be linearly predicted and the blocking matrix outputs as the input or reference signals for the ANC unit. The residual signal from the linear prediction becomes the GSC beamformer output, as shown in Fig. 3. More details on the ANC unit will be provided later. The binaural GSC structure has an equivalence with the BMVDR using a single constraint, or more generally the binaural GSC has an equivalence with the Binaural Linearly Constraint Minimum Variance (BLCMV), which uses multiple constraints [32].

B. PROPOSED ROBUST DESIGN BASED ON ADAPTIVE BLOCKING MATRIX

In this work, an approach to achieve a design that is robust to errors in the estimated target DOA is proposed. The structure of the Generalized Sidelobe Canceler (GSC) beamformer [31] in Fig. 3 is considered, but with modifications on the FB and BM blocks to make them robust to target DOA mismatch. As it was mentioned previously, the traditional GSC structure can be seen as an adaptive implementation equivalent to the BMVDR or BLCMV [32]. However, the GSC algorithm proposed in this paper for increased robustness to target DOA mismatch does not have a direct BMVDR or BLCMV equivalent, and the FB and BM output signals are no longer orthogonal in the proposed algorithm. For this reason, in this work we will refer to the proposed GSC beamformer as a modified GSC beamformer. The use of a binaural beamformer in the FB block in Fig. 3 is not robust to target DOA mismatch, since at higher frequencies it has a narrow beam in the assumed target DOA direction in the beampattern response (as in Fig. 2), because of the large microphone distance. A simple solution to avoid this is to use a frontal microphone noisy signal instead of the FB output, as shown in Fig. 4. A more sophisticated robust approach for the FB is presented in the next sub-section.

For the BM block, an Adaptive BM (ABM) is designed with a narrow "notch" in the target direction in the beampattern, using an adaptive null positioning algorithm. In the ABM, a 2+2 Binaural Linearly Constrained Minimum Variance (BLCMV) in the target cancelling mode is used. The BLCMV in the target canceling mode aims to cancel the target while preserving the noise components, i.e., directional interferers and background diffuse noise components, in order to provide a good estimate of these noise components. Assuming that the target could come from any direction within $\pm$ 10 degrees of the estimated/assumed target direction, the BLCMV uses three target level preserving

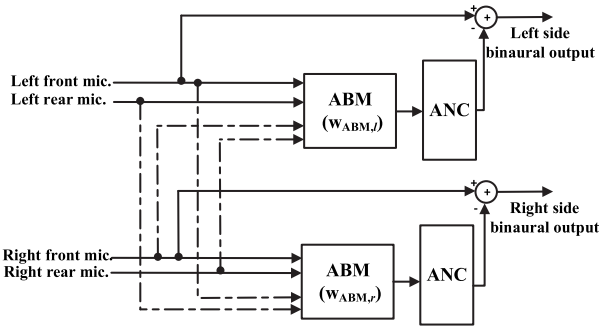

FIGURE 4. Robust beamforming design based on adaptive null modified GSC (ANGSC). Dashed lines represent the wireless links.

("unit gain") constraints in the half hemisphere which is opposite to the estimated target direction. Using three unit gain constraints in the opposite half-hemisphere helps to ensure that if the beamformer creates an adaptive null or notch with the remaining one degree of freedom, it will be positioned in the same half-hemisphere as the target. If no null or notch is detected within $\pm$10 degrees from the estimated target direction, a fallback design will be used, as discussed in more detail later.

As previously mentioned, the BLCMV [33] is a general case of the BMVDR described earlier, with multiple constraints instead of one constraint. Therefore, the BLCMV is based on a constrained minimization of the beamformer output power as in (13) and (14):

$$\min_{\mathbf{w}_{\mathrm{BM},l}} \mathbf{w}_{\mathrm{BM},l}^H \mathbf{R}_y \mathbf{w}_{\mathrm{BM},l} \quad \text{subject to } C^H \mathbf{w}_{\mathrm{BM},l} = \mathbf{g}_l^H \tag{13}$$

$$\min_{\mathbf{w}_{\mathrm{BM},r}} \mathbf{w}_{\mathrm{BM},R}^H \mathbf{R}_y \mathbf{w}_{\mathrm{BM},r} \quad \text{subject to } C^H \mathbf{w}_{\mathrm{BM},r} = \mathbf{g}_r^H. \tag{14}$$

The constraint matrix $\mathbf{C}$ includes the directivity vectors (or HRTFs) of each constraint direction, such that $\mathbf{C} = [\mathbf{d}(\theta_{v_1}), \mathbf{d}(\theta_{v_2}), \mathbf{d}(\theta_{v_3})]$, where θ_{v_1}, θ_{v_2} and θ_{v_3} are located in the half-hemisphere opposite to the estimated target direction. Specific values of θ_{v_1}, θ_{v_2} and θ_{v_3} are presented in Section IV. The left and right gain vectors $\mathbf{g}_l$ and $\mathbf{g}_r$, respectively, are as in (15) and (16):

$$\mathbf{g}_l = [d_{\mathrm{ref},l}(\theta_{v_1}), d_{\mathrm{ref},l}(\theta_{v_2}), d_{\mathrm{ref},l}(\theta_{v_3})] \tag{15}$$

$$\mathbf{g}_r = [d_{\mathrm{ref},r}(\theta_{v_1}), d_{\mathrm{ref},r}(\theta_{v_2}), d_{\mathrm{ref},r}(\theta_{v_3})]. \tag{16}$$

Using the complex Lagrangian multiplier to solve the constrained optimization problems in (13) and (14), the resulting left and right binaural ABM beamformer coefficients $\mathbf{w}_{\mathrm{BM},l}$ and $\mathbf{w}_{\mathrm{BM},r}$ are as in (17) and (18), respectively:

$$\mathbf{w}_{\mathrm{BM},l} = \mathbf{R}_y^{-1}\mathbf{C}(\mathbf{C}^H\mathbf{R}_y^{-1}\mathbf{C})^{-1}\mathbf{g}_l^H \tag{17}$$

$$\mathbf{w}_{\mathrm{BM},r} = \mathbf{R}_y^{-1}\mathbf{C}(\mathbf{C}^H\mathbf{R}_y^{-1}\mathbf{C})^{-1}\mathbf{g}_r^H. \tag{18}$$

$\mathbf{R}_y$ in (4) to (18) is adaptively estimated using a moving average lowpass first order recessive filter across consecutive frames with a forgetting factor of 0.985, in each frequency bin. The adaptive estimation of $\mathbf{R}_y$ in (17) and (18) leads to adaptive target cancellation and adaptive beampatterns, e.g., for a single directional source in the half-hemisphere of the

estimated target direction an adaptive null would be produced in the beampattern.

The aim of the BM is to produce signals to serve as good estimates of the noise components, to be used as references (inputs) for the ANC unit. The proposed adaptive null positioning process needs to ensure that a null is positioned by the BM beamformer and that it is located in the direction of the target. This null positioning process relies on dynamic evaluations of the beampattern at different frequencies and time frames, to detect the presence or absence of a null at the expected/estimated target direction. In case of absence of a null near the expected/estimated target direction, the method reverts to a fallback solution, for example a design with a fixed but wide notch in the estimated target direction, or a previous version (set of coefficients) of the adaptive beamformer solution. In this section, the adaptive null positioning algorithm is presented in general; however, this algorithm should be repeated for the left and the right sides in a binaural hearing aid, to compute the left and right beamforming coefficients. The adaptive null positioning algorithm is achieved as follows:

1. Compute the beampattern at each T-F bin as in (19):

$$BP(f,t,\theta) = |\mathbf{w}_{\mathrm{BM}}^{H}(f,t)\mathbf{d}(f,\theta)|^{2}. \tag{19}$$

   The beamforming coefficients $\mathbf{w}_{\mathrm{BM}}(f,t)$ of the BLCMV in target cancelling mode are computed as in (17) and (18).
2. Measure the depth of the null in the expected/estimated target zone $BP_{\text{target-zone}}$(e.g., $\theta_{\text{target-zone}}$ between -10 degrees to 10 degrees in case of an estimated frontal target around 0 degree).

$$BP_{\text{target-zone}} = \min_{\theta}(BP(f,t,\theta_{\text{target-zone}})) \tag{20}$$

$$\theta_{\text{null}} = \arg\min_{\theta}(BP(f,t,\theta_{\text{target-zone}})) \tag{21}$$

3. Compute the beamforming coefficients $\mathbf{w}_{\mathrm{ABM}}(f,t)$ of the ABM using the following method: If the detected null θ_{null} in the expected target zone is deep enough (20 dB lower than the max. gain in the half-hemisphere of the estimated target direction), the beamformer coefficients $\mathbf{w}_{\mathrm{BM}}(f,t)$ are copied to $\mathbf{w}_{\mathrm{ABM}}(f,t)$ and saved to be used as a fallback scenario in future time frames at same frequency component. If the detected null in the expected target zone is not deep enough, then the method reverts to a fallback solution $\mathbf{w}_{\mathrm{rev}}(f,t)$:

$$\mathbf{w}_{\mathrm{ABM}}(f,t) = \begin{Bmatrix} \mathbf{w}_{\mathrm{BM}}(f,t) & \text{deep null in} \theta_{\text{target-zone}} \\ \mathbf{w}_{\mathrm{rev}}(f,t) & \text{otherwise} \end{Bmatrix}. \tag{22}$$

   As previously mentioned, the beamforming coefficients $\mathbf{w}_{\mathrm{ABM}}(f,t)$ are computed for the left and the right sides. The 20 dB threshold for the depth of the null in the estimated target zone was experimentally adjusted based on the resulting noise reduction (SNR gain) and target distortion (SDmag) for different acoustic scenarios.
4. If the detected null in the expected target zone is not deep enough, the method finds whether the null has been detected or not in a previous time frame for this specific frequency. If the null in the target zone has been detected previously, then the previously saved beamformer coefficients $\mathbf{w}_{\mathrm{BM}}(f,t-i), i>0$ are used. If the null in the target zone has not been detected in one of the previous time frames at this specific frequency components, then some "initial condition" coefficients are used, i.e., a fallback design $\mathbf{w}_{\mathrm{IC}}(f,t)$ with a fixed 2+2 constraint-based design (explained below) and a wide notch in the estimated target direction.

$$\mathbf{w}_{\mathrm{rev}}(f,t) = \begin{Bmatrix} \mathbf{w}_{\mathrm{BM}}(f,t-i), i>0 \\ \quad \text{last previous coefficients} \\ \quad \text{computed with null detected} \\ \quad \text{in target zone, if available} \\ \mathbf{w}_{\mathrm{IC}}(f) \\ \quad \text{if no previous } \mathbf{w}_{\mathrm{BM}}(f,t-i), i>0 \text{ available} \end{Bmatrix} \tag{23}$$

The wide notch beamforming coefficients for the initial condition design $\mathbf{w}_{\mathrm{IC}}(f,t)$ are computed as in (24):

$$\mathbf{w}_{\mathrm{IC}}(f) = \mathrm{C}_{\mathrm{IC}}^{H}(f)^{-1}\mathbf{g}^{H}(f) \tag{24}$$

where, $\mathbf{C}_{\mathrm{IC}}$ is a (square) constraint matrix such that $\mathbf{C}_{\mathrm{IC}}(f) = [\mathbf{d}(f,\theta_s+\Delta), \mathbf{d}(f,\theta_s-\Delta), \mathbf{d}(f,\theta_{v_1}), \mathbf{d}(f,\theta_{v_2})]$, and $\mathbf{g}$ is a gain vector such that θ_{v_1}, θ_{v_2} are located in the half-hemisphere opposite to the estimated target direction θ_s. Specific values of θ_{v_1}, θ_{v_2} and Δ are presented in Section IV. For this fixed design, the number of constraints is equal to the number of microphones. To compute the left beamforming initial condition coefficients $\mathbf{w}_{\mathrm{IC},l}(f,t)$ and the right beamforming initial condition coefficients $\mathbf{w}_{\mathrm{IC},r}(f,t)$, the left and right gains vectors are $\mathbf{g}_l(f) = [0, 0, d_{\mathrm{ref},l}(f,\theta_{v_1}), d_{\mathrm{ref},l}(f,\theta_{v_2})]$ and $\mathbf{g}_r = [0, 0, d_{\mathrm{ref},r}(f,\theta_{v_1}), d_{\mathrm{ref},r}(f,\theta_{v_2})]$, respectively.

The initial solution is only used if no null is currently detected in the estimated target zone and no beamformer in target canceling mode has previously been found (since the initialization and start of the algorithm) with a null in the estimated target zone at that frequency. Therefore, with the potential exception of frequency bins for which a null is never detected in the estimated target zone, the use of the initial solution only occurs at the early stages after the initialization and start of the algorithm.

The initial condition design is a fixed solution that depends only on the target DOA, as eq. 24 shows. Under a stationary DOA condition (as in this paper), the initial condition coefficients are computed only once, but for non-stationary DOA conditions the initial condition solution could be updated if there are frequency bins where a null is never detected in the estimated target zone.

A good noise estimation ability is expected to be achieved by the BM with an adaptive notch, as it aims to have a narrow notch in the direction of the target, i.e., the interferers and

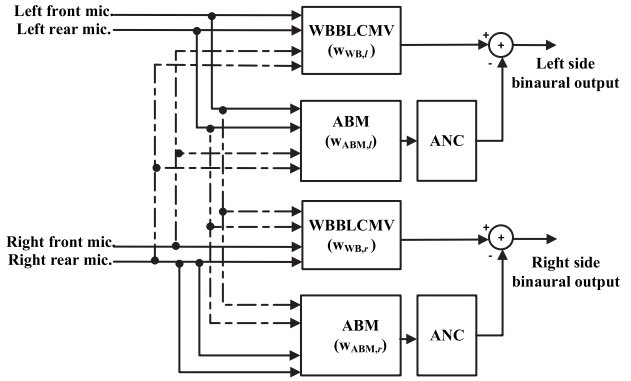

FIGURE 5. Robust beamforming design based on Adaptive Null modified GSC with Wide Beam BLCMV (ANGSC-WBBLCMV). Dashed lines represent the wireless links.

noise components should still be in the BM output. In the GSC structure of Fig. 4, the ANC uses the front microphone noisy signal as a "desired" signal and the BM output as a "reference" input signal, in a classic linear prediction setup. If a FB is used, the output of the FB is used as the "desired" signal, as we will see later. The ANC is therefore an adaptive filtering system to minimize the noise and interferer components in the desired signal, using the noise and interference components in the BM output which are correlated to the desired signal. The ANC coefficients can be adaptively updated using a Least Mean Square (LMS) algorithm, or using other adaptive filtering algorithms. In this paper, we consider a BM with only one output. This can theoretically limit the noise reduction achievable by the ANC unit, however if the acoustic sources are sparse in the T-F domain, a dominant interferer or noise component will normally be found at the output of the adaptive BM for each T-F bin, and good noise reduction can still be achieved overall, as will be shown in the simulation results. In this work, the LMS algorithm is used to adaptively update the ANC coefficients w_{ANC} as in (25) and (26):

$$z(f,t) = z_{\mathrm{FB}}(f,t) - w^*_{\mathrm{ANC}}(f,t) z_{\mathrm{BM}}(f,t) \quad (25)$$

$$w_{\mathrm{ANC}}(f,t+1) = w_{\mathrm{ANC}}(f,t) + \mu z_{\mathrm{BM}}(f,t) z^*(f,t) \quad (26)$$

where, for the left or the right side, z is the beamformer output, z_{FB} is the FB output, z_{BM} is the BM output (ANC input), and μ is a step size tuned to $0.5/\sigma^2_{z_{BM}}$, where $\sigma^2_{z_{BM}}$ is the estimated power of z_{BM}.

C. PROPOSED ROBUST DESIGNS BASED ON WIDE BEAM AND ADAPTIVE BM

As an attempt to enhance the performance of the robust design introduced in the previous sections, a binaural beamformer with wide beampatterns is used as a FB combined with the ABM of the previous sub-section, as shown in Fig. 5.

A binaural beamformer with wider beams is used to have a FB beamformer more robust to target DOA mismatch. This can be achieved by using more than one constraint near the estimated target direction in the design. Consequently, a 2+2 Binaural Linearly Constrained Minimum Variance

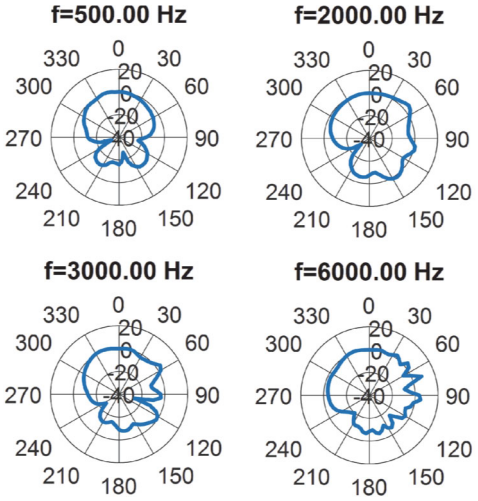

FIGURE 6. Beampatterns for the 2+2 BLCMV with two constraints around a frontal target at 0 degree, left side, under 2-D diffuse noise conditions.

(2+2 BLCMV) beamformer is used. Assuming that the target could come from any direction within ± 10 degrees of the estimated/assumed target direction, two non-zero target level preserving constraints in the middle of this target zone can be used as in [34], leading to a wider beam width in the target zone. Fig. 6 shows the beampatterns for some selected frequencies of the fixed 2+2 BLCMB with two constraints at +5 and -5 degrees under 2-D diffuse noise conditions. The 2+2 BLCMV has a wider beam around the 0 degree target direction, unlike the 2+2 BMVDR beampatterns in Fig. 2. At the same time, Fig. 6 shows that the 2+2 BLCMV attenuates some of the noise components, unlike using the raw noisy signals from the front microphone. The left and right binaural wide beam beamformer coefficients $\mathbf{w}_{\mathrm{WB},l}$ and $\mathbf{w}_{\mathrm{WB},r}$ are computed as in (17) and (18), respectively. In this case, the constraint matrix $\mathbf{C}$ includes the directivity vectors (or HRTFs) of each constraint direction around the estimated target direction, such that $\mathbf{C} = [\mathbf{d}(\theta_s + \Delta), \mathbf{d}(\theta_s - \Delta)]$, where $\theta_s + \Delta$ and $\theta_s - \Delta$ are in the middle of the target zone. The left and right gain vectors $\mathbf{g}_l$ and $\mathbf{g}_r$, respectively, are $\mathbf{g}_l = [d_{\mathrm{ref},l}(\theta_s + \Delta), d_{\mathrm{ref},l}(\theta_s - \Delta)]$ and

$\mathbf{g}_r = [d_{\mathrm{ref},r}(\theta_s + \Delta), d_{\mathrm{ref},r}(\theta_s - \Delta)]$.

Although three constraints could be used in the assumed target zone, using two constraints in the 2+2 BLCMV provides more degree of freedom for noise reduction (two degrees of freedom), i.e., potential for two adaptive nulls to provide better noise reduction.

IV. SYSTEM EVALUATION

A. EXPERIMENTAL SETUP

Two binaural BTE hearing aids worn by a KEMAR$^{\mathrm{TM}}$ mannequin were used to measure anechoic HRTFs, used as directivity vectors for the beamformer designs. A different set of HRTFs (or BRIRs) were also measured in a mildly reverberant environment (T60 $\approx$130ms), and these HRTFs were used to generate the different signals from directional sources in our simulations. Speech signals were

used for the directional sources, i.e., target and interferers. The distance between a loudspeaker source and the center of the head, which is used for the reverberant and the anechoic HRTFs measurements, was 1 m. Multichannel diffuse-like background noise was generated using the same BTE units and KEMAR mannequin, by playing babble noise recordings at eight loudspeakers on a circle with a radius of 1 m around the KEMAR mannequin. A filter bank with 48 one-sided uniform bands and a down-sampling factor of 24 was used for T-F analysis and synthesis. The recordings were sampled at 24 kHz. The two sets of HRTFs, the diffuse noise recordings, and the filter bank were provided by a hearing aid manufacturer.

B. PERFORMANCE MEASUREMENT

The evaluation of the proposed beamforming process is based on two criteria: noise reduction and target distortion. The Signal to Noise Ratio gain (SNR gain) is used to evaluate the performance of the beamforming process in terms of noise reduction. For each frequency, to compute the SNR gain, we first find the frequency dependent SNR of the input signal as:

$$SNR_{\text{input}} = 10\log\left(\frac{\Gamma_{x_{\text{ref}},x_{\text{ref}}}}{\Gamma_{v_{\text{ref}},v_{\text{ref}}}}\right) \tag{27}$$

where $\Gamma_{x_{\text{ref}},x_{\text{ref}}}$ is the auto power spectrum density (auto-PSD) of the target components at the reference microphone signal, and $\Gamma_{v_{\text{ref}},v_{\text{ref}}}$ is the auto-PSD of the noise components, i.e., sum of the directional interferers and diffuse-like background noise components at the reference microphone signal. The frequency dependent SNR of the output signal is computed as in (28):

$$\text{SNR}_{\text{output}} = 10\log\left(\frac{\Gamma_{x_{\text{out}},x_{\text{out}}}}{\Gamma_{v_{\text{out}},v_{\text{out}}}}\right) \tag{28}$$

where, $\Gamma_{x_{\text{out}},x_{\text{out}}}$ is the auto-PSD of the target component at the output of a beamformer and $\Gamma_{v_{\text{out}},v_{\text{out}}}$ is the auto-PSD of all the noise/interferers components at the output of the beamformer. Finally, the frequency dependent SNR gain is computed as in (29):

$$\text{SNRgain(dB)} = \text{SNR}_{\text{output}}\text{(dB)} - \text{SNR}_{\text{input}}\text{(dB)}. \tag{29}$$

To measure the target distortion generated from the beamforming process, for each frequency the Speech Distortion Magnitude (SDmag) is used. This measurement is computed by finding the absolute difference (in dB) between the auto-PSD of the target component at the reference microphone $\Gamma_{x_{\text{ref}},x_{\text{ref}}}$ and the auto-PSD of the target component at the beamformer output $\Gamma_{x_{\text{out}},x_{\text{out}}}$ as in (30):

$$\text{SDmag (dB)} = |10\log(\Gamma_{x_{\text{ref}},x_{\text{ref}}}) - 10\log(\Gamma_{x_{\text{out}},x_{\text{out}}})|. \tag{30}$$

A "shadow filtering" method was used in the simulations by processing all the signal components individually with the same time-variant filter coefficients, in order to obtain the target components and the noise components in the beamformer outputs. The SNR gain and SDmag metrics can be computed

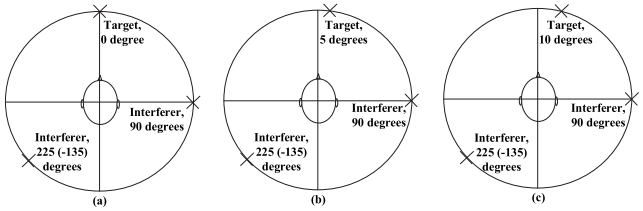

FIGURE 7. Near-frontal target with true DOA at 0, 5 or 10 degrees, interferers at 90 and 225 degrees (same level as target), and diffuse noise (level 14 dB lower than the target and interferer levels).

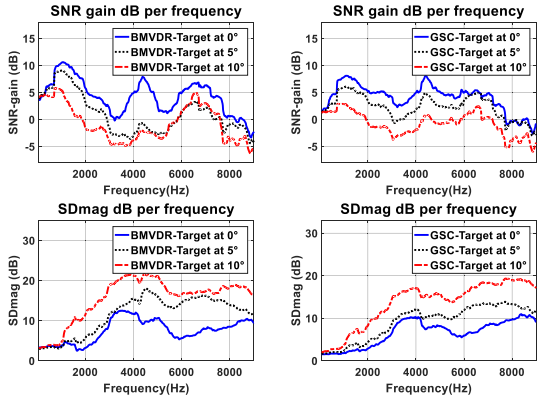

FIGURE 8. SNR gain and SDmag of baseline BMVDR (left) and GSC (right) with increasing target DOA mismatch, under near-frontal target acoustic scenarios.

for the left and the right binaural beamformers, but in this paper we present the values for the "better ear" side, i.e., the most important side with the better broadband (time domain) input SNR.

C. SIMULATION RESULTS

In this work, we refer to the robust design based on the ABM in Fig. 4 as the Adaptive Null modified GSC (ANGSC), and we refer to the robust design based on the combination of ABM and the wide beampattern in Fig. 5 as the Adaptive Null modified GSC with Wide Beam BLCMV (ANGSC-WBBLCMV).

First, the performances of the BMVDR and GSC benchmarks are characterized using three near-frontal target acoustic scenarios, with and without target DOA mismatch, with a target estimate DOA at 0 degree. Fig. 7 shows the three acoustic scenarios with the true target DOA at 0, 5, or 10 degrees, interferers at 90 and 225 degrees (each with the same level as the target source), and diffuse noise (level 14 dB lower than the target level).

As the target DOA mismatch increases, a significant deterioration in the performance of the benchmarks can be noticed in Fig. 8. The SNR gain decreases as the target DOA mismatch increases from 0 degree to 10 degrees. The reduction in the SNR gain reaches more than 5 dB for some frequency components. The SDmag also significantly increases with the increase of target DOA mismatch. This increase in the SDmag is an indication of a significant target attenuation introduced from the beamforming processing.

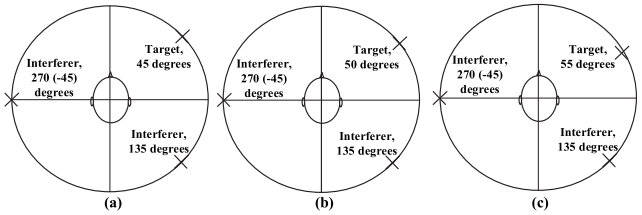

FIGURE 9. Non-frontal target with true DOA at 45, 50 or 55 degrees, interferers at 270 (-90) and 135 degrees (same level as target), and diffuse noise (level 14 dB lower than the target level).

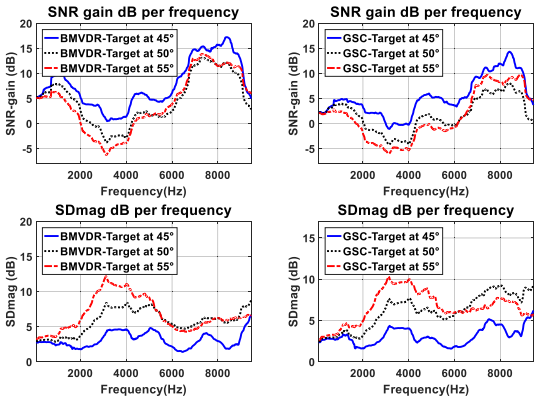

FIGURE 10. SNR gain and SDmag of baseline BMVDR (left) and GSC (right) with increasing target DOA mismatch, for non-frontal target acoustic scenarios.

It is noticeable that even without target DOA mismatch, i.e., the acoustic scenario in Fig. 7(a), the SDmag is not zero. This result is justified by the HRTF mismatch mentioned in Section III, i.e., mismatch between the anechoic HRTFs used in the beamformer designs and the reverberant HRTFs used in generating the target and interferer signals in the simulated acoustic scenarios. It is also noticeable that the overall performance and sensitivity to DOA mismatch of the benchmark GSC beamformer and the benchmark BMVDR beamformer in terms of SNR gain and SDmag are overall similar, with differences for some frequency components.

Further characterizations of the BMVDR and GSC benchmarks are done using three non-frontal target acoustic scenarios as in Fig. 9. Those three acoustic scenarios have a target signal with true DOA at 45, 50, or 55 degrees, while the BMVDR and GSC benchmarks assume that the target direction is at 45 degrees. The resulting performance metrics in Fig. 10 show a reduction in the SNR gain as the target DOA mismatch increases for both the BMVDR and the GSC. This reduction reaches 4-5 dB for some frequencies. The target distortion also increases with DOA mismatch, as also shown in Fig. 10 in terms of SDmag.

For acoustic scenarios with a near fontal target, where it is assumed that the target true DOA is located within $\pm$ 10 degrees of the estimated/assumed target direction, the benchmark BMVDR uses a single unity gain (reference microphone target level preserving) constraint at 0 degree as shown in Fig. 11. The benchmark GSC also assumes the target to be at 0 degree. Instead, the wide beam beamforming

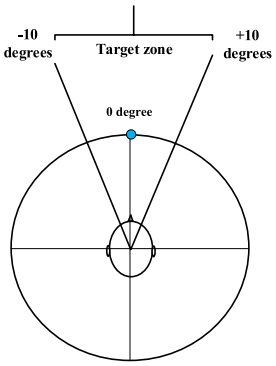

FIGURE 11. Design criteria for BMVDR, near frontal target acoustic scenarios.

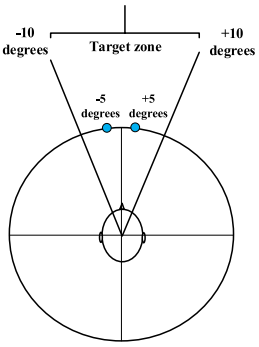

FIGURE 12. Design criteria for wide beam BLCMV in ANGSC-WBBLCMV, near frontal target acoustic scenarios.

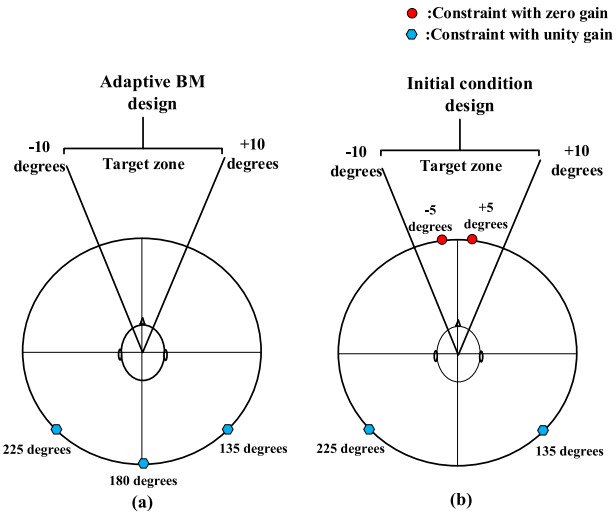

FIGURE 13. Design criteria of a) ABM design and b) initial condition fallback design in ANGSC and ANGSC-WBBLCMV, near frontal target acoustic scenarios.

BLCMV design used in the ANGSC-WBBLCMV has symmetric unity constraints at 5 and -5 degrees, in the middle of the allowed frontal target zone, as shown in Fig. 12.

In the ANGSC and ANGSC-WBBLCMV beamforming designs, the ABM uses the BLCMV in target cancelling mode, with fallback designs. Fig. 13 shows the locations of the resulting constraints for the ABM and the fallback initial condition design when the assumed target direction is

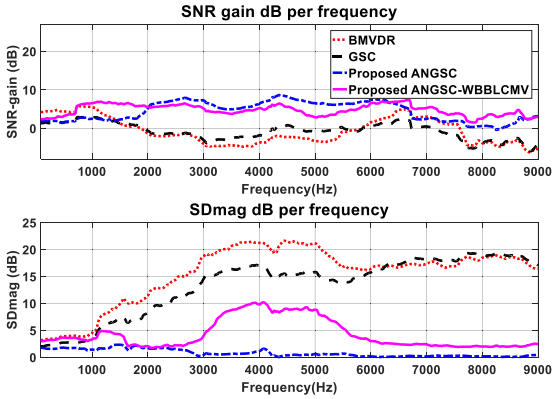

FIGURE 14. Performance of benchmark BMVDR, benchmark GSC, proposed ANGSC design, and proposed ANGSC-WBBLCMV design, with target at 10 degrees (10 degrees DOA mismatch).

between $\pm$10 degrees. To help the BLCMV in target cancelling mode to position an adaptive null in the same half-hemisphere as the estimated target direction, three constraints of unity gain at 135, 180, and 225 degrees are used for the ABM in the half-hemisphere opposite to the estimated target direction (Fig. 13 (a)). For the fallback initial conditions design (Fig. 13 (b)), to have a wide notch in the estimated target DOA direction (0 degree here) two constraints of zero gains at $\pm$ 5 degrees (in the middle of the target zone) and two constraints of unity gain around 180 degrees (opposite direction) are used.

Fig. 14 compares the performances of the two proposed beamformers robust to target DOA mismatch (ANGSC and ANGSC-WBBLCMV) with the benchmarks BMVDR and GSC. The acoustic scenario with a near frontal target at a true DOA of 10 degrees in Fig. 7 (c) is used for the evaluation, with all the beamformers assuming that the target is at DOA 0 degree. The proposed robust designs show a better overall noise reduction except for low frequencies (e.g. below 700 Hz for the ANGSC-WBBLCMV), where target DOA mismatch is not a concern (no narrow beam, e.g. Fig. 2 at 500 Hz) and the BMVDR has more adaptive nulls available for noise reduction (positioning of more nulls). The ANGSC beamforming design provided a higher SNR gain than the AGCS-WBBLCMV for frequency components between 2.2 kHz to 6.2 kHz. The ANGSC and AGCS-WBBLCMV both introduce lower target distortion and attenuation, i.e., lower SDmag, compared to the BMVDR and GSC benchmarks, as Fig. 14 shows. Compared to the ANGSC design, the AGCS-WBBLCMV provided a higher noise reduction for some frequency components, but at the cost of some increase in the target distortion over all frequency components.

To further illustrate the performance of the ABM used in the proposed ANGSC design and the proposed ANGSC-WBBLCMV design, a few beampatterns are shown in Fig. 15 under the acoustic scenario in Fig. 7 (c) with 10 degrees DOA mismatch. The first beampattern in Fig. 15 is an example where the null is deep and inside the target zone, for a subband with a center frequency of 1750 Hz.

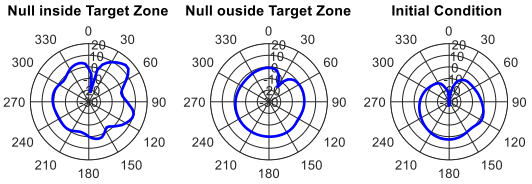

FIGURE 15. Beampatterns from the ABM under the acoustic scenario with target at 10 degrees.

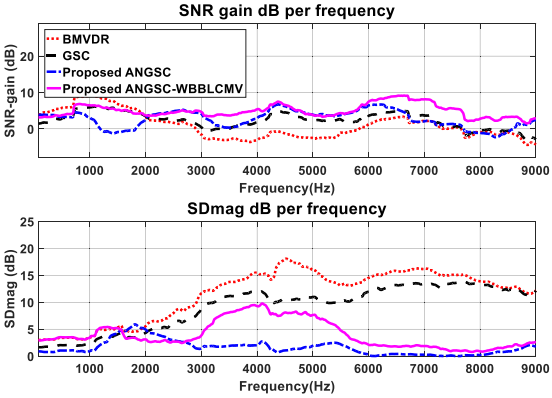

FIGURE 16. Performance of benchmark BMVDR, benchmark GSC proposed ANGSC design, and proposed ANGSC-WBBLCMV design, with target at 5 degrees (5 degrees DOA mismatch).

The second beampattern in Fig. 15 shows an example when the null is outside the target zone, for a subband with center frequency 1000 Hz. The third beampattern in Fig. 15 is for the fixed initial condition design with a wide notch region, for a subband with a center frequency 1000 Hz. Under this acoustic scenario, the nulls were detected inside the estimated target zone for 65.4 % of T-F bins and outside the estimated target zone for 34.4% of the T-F bins.

The proposed robust designs have been also evaluated under the acoustic scenario in Fig. 7 (b) with a lower DOA mismatch of 5 degrees. Similar to the case of 10 degrees DOA mismatch, the proposed robust designs in Fig. 16 show a better overall noise reduction and lower target distortion in comparison with the benchmark beamformers. The AGCS-WBBLCMV beamforming design provides a higher SNR gain than the AGCS for all frequency components for this scenario, while the ANGSC provides overall the lowest target distortion (SDmag) over frequencies.

Fig. 17 compares the performances of the two proposed beamforming designs and the two benchmarks designs, for a case with no DOA mismatch. The acoustic scenario with a frontal target at 0 degree in Fig. 7 (a) is used, with all the beamformers assuming that the target is at DOA 0 degree. As expected, for this case with no DOA mismatch the benchmark algorithms performed well, although the proposed beamforming designs still delivered an overall performance similar to the benchmarks in terms of SNR gain, and a better overall performance in terms of SDmag. Some average values will be presented later in this section to highlight this.

The proposed robust designs have also been tested under a non-frontal acoustic scenario as in Fig. 9 (c), with a true target

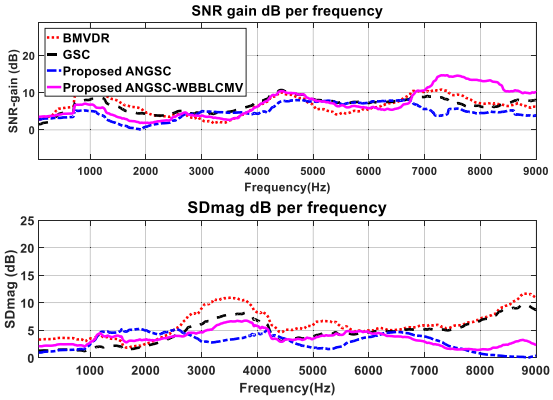

FIGURE 17. Performance of benchmark BMVDR, benchmark GSC proposed ANGSC design, and proposed ANGSC-WBBLCMV design, with target at 0 degree (without DOA mismatch).

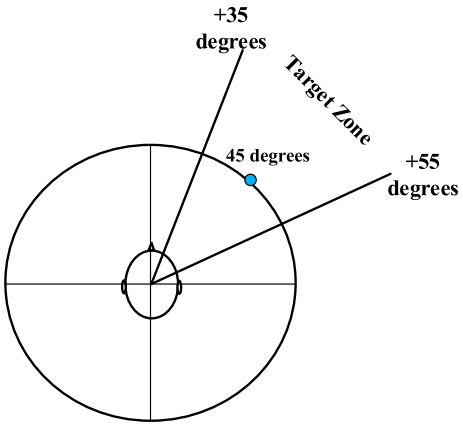

FIGURE 18. Design criteria for BMVDR, for near 45 degrees target acoustic scenarios.

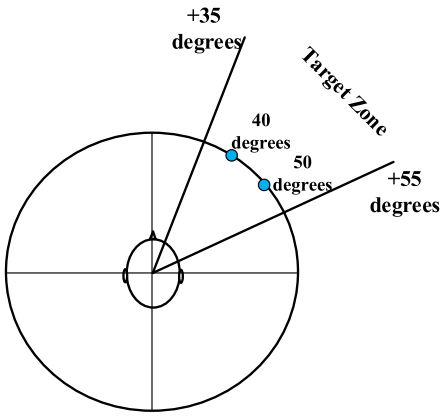

FIGURE 19. Design criteria for wide beam BLCMV in ANGSC-WBBLCMV, for near 45 degrees target acoustic scenarios.

DOA at 55 degrees and beamformers assuming that the target is at DOA 45 degrees. For this acoustic scenario with a target speaker near 45 degrees, the locations of the constraints of the BMVDR and the BLCMV in the ANGSC-WBBLCMV are as shown in Fig. 18 and Fig. 19, respectively. Following the same logic in designing the ABM beamformer as for the previous near frontal acoustic scenarios, the design criteria

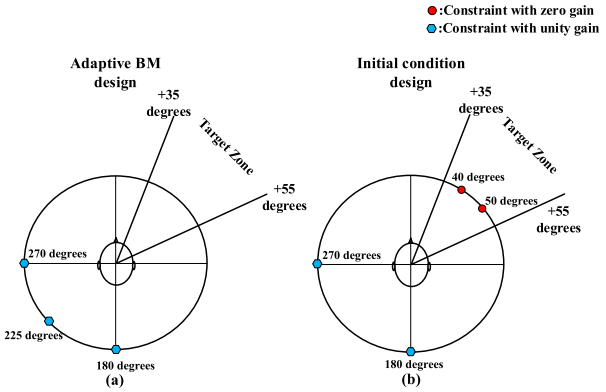

FIGURE 20. Design criteria for a) ABM design and b) initial condition design in ANGSC and ANGSC-WBBLCMV, for near 45 degrees target acoustic scenarios.

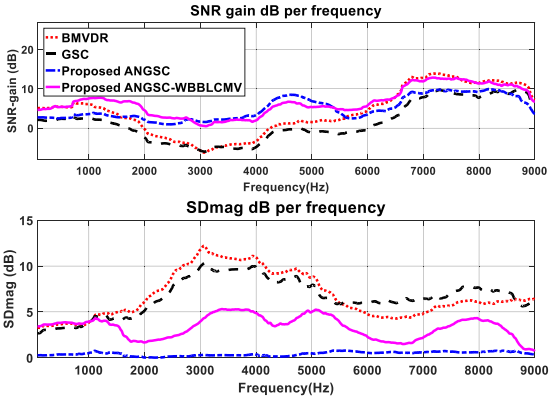

FIGURE 21. Performance of BMVDR and GSC benchmarks, proposed ANGSC design, and proposed ANGSC-WBBLCMV design, with target at 55 degrees (10 degrees DOA mismatch).

for the ABM and the fallback initial condition design used in the ANGSC and ANGSC-WBBLCMV are shown in Fig. 20.

The resulting performance metrics in Fig. 21 illustrate the better performance of the proposed ANGCS in terms of SDmag, providing the lowest target distortion with 10 degrees target DOA mismatch. The ANGSC-WBBLCMV design also provides significant target distortion improvements over both the BMVDR and GSC benchmarks. The robust designs, i.e., ANGSC and ANGSC-WBBLCMV, both outperform the BMVDR and GSC benchmarks in terms of noise reduction, and the ANGSC-WBBLCMV provides improvements over a slightly wider frequency range.

The performance of the proposed designs as well as the benchmark designs were then evaluated for the non-frontal near 45 degrees target acoustic scenarios with a lower DOA mismatch of 5 degrees (Fig. 9(b)), and without DOA mismatch (Fig. 9 (a)). For a DOA mismatch of 5 degrees (Fig. 22), the proposed robust designs show a better overall noise reduction and a lower overall target distortion in comparison with the benchmark beamformers, with the AGCS-WBBLCMV providing a better SNR gain over a wider frequency range, and the AGCS providing the best SDmag performance. For the case with no DOA mismatch

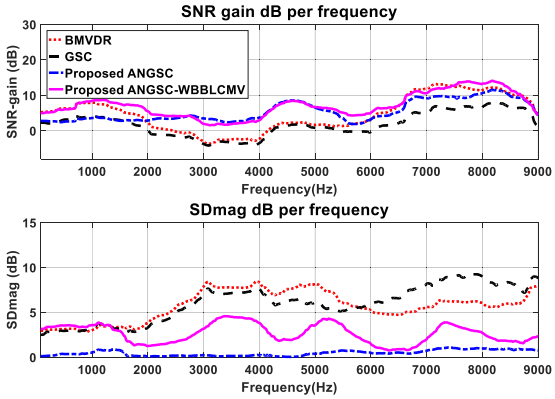

FIGURE 22. Performance of BMVDR and GSC benchmarks, proposed ANGSC design, and proposed ANGSC-WBBLCMV design, with target at 50 degrees (5 degrees DOA mismatch).

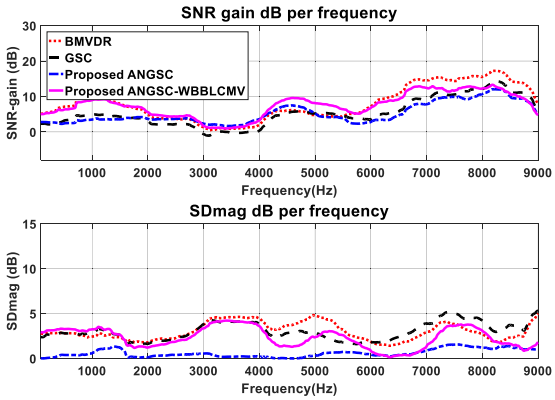

FIGURE 23. Performance of BMVDR and GSC benchmarks, proposed ANGSC design, and proposed ANGSC-WBBLCMV design, with target at 45 degrees (without DOA mismatch).

(Fig. 23), as expected for this case the benchmark algorithms performed well, especially the BMVDR, but the proposed AGCS-WBBLCMV still delivered an overall performance similar to the BMVDR benchmark in terms of SNR gain, and the proposed ANGSC still provided a better overall performance in terms of SDmag. Some average values are presented later to highlight this.

To analyze the single effect of HRTF mismatch in the performance of the proposed robust designs as well as in the BMVDR and GSC benchmarks, acoustic scenarios with HRTF mismatch and no target DOA mismatch are considered. These acoustic scenarios use either signals generated using anechoic HRTFs (unlike all previous cases) or signals generated using reverberant HRTFs (like all previous cases). Simulations are performed under a frontal target acoustic scenario with a target at 0 degree and interferers at 225 and 90 degrees as in Fig. 7 (a), a non-frontal target acoustic scenario with a target at 45 degrees and interferers at 270 and 135 degrees as in Fig. 9 (a), and a new lateral target acoustic scenario with a target at 90 degrees and interferers at 315 and 225 degrees. For all these acoustic scenarios, no target DOA mismatch is considered. All the interferers are at the same level as the target signal, and diffuse noise is added with a level 14 dB lower than the target and interferers levels.

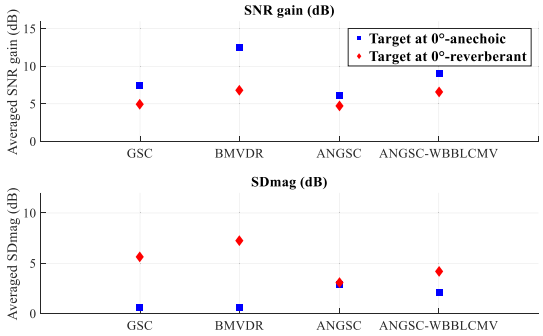

FIGURE 24. Performance of BMVDR and GSC benchmarks, proposed ANGSC design, and proposed ANGSC-WBBLCMV design, with target at 0 degree (no DOA mismatch) in anechoic and reverberant environments.

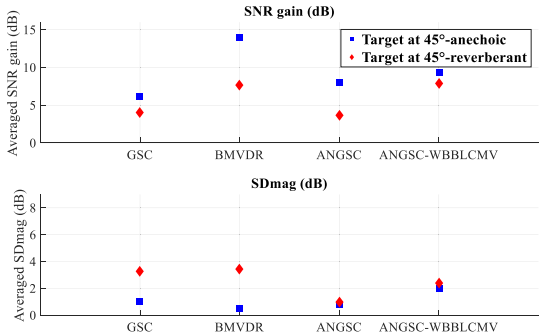

FIGURE 25. Performance of BMVDR and GSC benchmarks, proposed ANGSC design, and proposed ANGSC-WBBLCMV design, with target at 45 degrees (no DOA mismatch) in anechoic and reverberant environments.

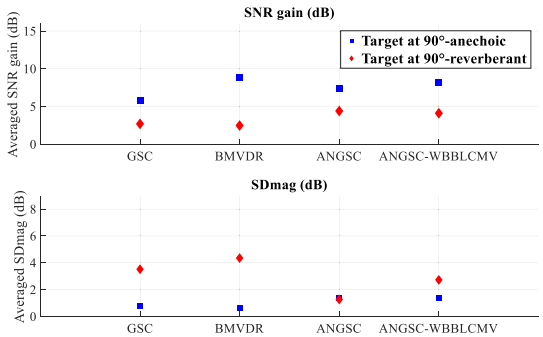

FIGURE 26. Performance of BMVDR and GSC benchmarks, proposed ANGSC design, and proposed ANGSC-WBBLCMV design, with target at 90 degrees (no DOA mismatch) in anechoic and reverberant environments.

For the lateral target at 90 degrees, design criteria for the benchmark BMVDR, BLCMV in ANGSC-WBBLCMV, and ABM in ANGSC or ANGSC-WBBLCMV follow the same logic as for the previous near frontal acoustic scenario (Fig. 11 - Fig. 13) and non-frontal target acoustic scenario (Fig. 18 - Fig. 20).

For the SNR gain, an average or broadband SNR gain value can be computed by replacing the auto-PSDs in (27), (28) with time domain power measurements. An average SDmag value can also be defined as the average over frequencies of the frequency dependent SDmag in (30). The average SNR gain and SDmag performance metrics are shown in

Fig. 24 to Fig. 26. Under ideal conditions without either target DOA mismatch or HRTF mismatch, the benchmark BMVDR and GSC outperform the proposed robust designs in terms of SDmag, as expected since this is an ideal scenario for the benchmark designs. The benchmark BMVDR also outperforms the proposed designs in terms of SNR under all tested ideal acoustic scenarios. However, under the more realistic conditions of HRTF mismatch, both proposed designs outperform the benchmark BMVDR and GSC in terms of target distortion (even though there is no DOA mismatch here). And for noise reduction under these conditions of HRTF mismatch and no DOA mismatch, the proposed ANGSC- WBBLCMV performs either similarly or better than the benchmark algorithm with the best performance (BMVDR).

V. CONCLUSION

In this work, two binaural beamformer designs robust to target DOA mismatch are introduced for binaural hearing aids. These designs are the ANGSC design, which is an adaptive blocking matrix based design, and the ANGSC- WBBLCMV designs, which is a combination of an adaptive blocking matrix based design and a wide beam-based design. The proposed robust designs have been evaluated and compared with BMVDR and GSC benchmarks. The results show that the proposed robust designs overall outperform the BMVDR and GSC benchmarks in terms of noise reduction and target distortion in the presence of target DOA mismatch of up to 10 degrees, which is a realistic figure for many real-life scenarios. The proposed beamformers also show robustness to HRTF mismatch, i.e., mismatch between the anechoic HRTFs used for the beamformer designs and the reverberant HRTFs which generate the signals at the sensors or beamformer inputs. For further development and validation of the proposed algorithms in the future, testing under environments with higher levels of reverberation as well as cases of time-varying DOA scenarios and time-varying source activity patterns will need to be considered.

REFERENCES

[1] W. M. Whitmer, K. F. Wright-Whyte, J. A. Holman, and M. A. Akeroyd, "Hearing aid validation," in *Hearing Aids*, vol. 56, 1st ed. Cham, Switzerland: Springer, 2016, pp. 291–321.

[2] G. R. Popelka and B. C. J. Moore, "Future directions for hearing aid development," in *Hearing Aids*, vol. 56, 1st ed. Cham, Switzerland: Springer, 2016, pp. 331–341.

[3] S. Doclo, S. Gannot, M. Moonen, and A. Spriet, "Acoustic beamforming for hearing aid applications," in *Handbook on Array Processing and Sensor Networks*. Hoboken, NJ, USA: Wiley, 2008, pp. 269–302.

[4] I. A. McCowan, "Robust speech recognition using microphone arrays," Ph.D. dissertation, Queensland Univ. Technol., Brisbane, QLD, Australia, 2001. [Online]. Available: http://www.aplu.ch/home/download/microphone_array.pdf

[5] B. Cornelis, M. Moonen, and J. Wouters, "Comparison of frequency domain noise reduction strategies based on multichannel Wiener filtering and spatial prediction," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2009, pp. 129–132.

[6] B. Cornelis, M. Moonen, and J. Wouters, "Performance analysis of multichannel Wiener filter-based noise reduction in hearing aids under second order statistics estimation errors," *IEEE Trans. Audio, Speech, Language Process.*, vol. 19, no. 5, pp. 1368–1381, Jul. 2011.

[7] O. Hoshuyama, A. Sugiyama, and A. Hirano, "A robust adaptive beamformer for microphone arrays with a blocking matrix using constrained adaptive filters," *IEEE Trans. Signal Process.*, vol. 47, no. 10, pp. 2677–2684, Oct. 1999.

[8] W. Herbordt and W. Kellermann, "Computationally efficient frequency-domain robust generalized sidelobe canceller," in *Proc. Int. Workshop Acoustic Echo Noise Control*, 2001, pp. 51–55.

[9] B.-J. Yoon, I. Tashev, and A. Acero, "Robust adaptive beamforming algorithm using instantaneous direction of arrival with enhanced noise suppression capability," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, vol. 1, Apr. 2007, pp. I-133–I-136.

[10] C. Choi, D. Kong, J. Kim, and S. Bang, "Speech enhancement and recognition using circular microphone array for service robots," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, vol. 4, Oct. 2003, pp. 3516–3521.

[11] L. Lepauloux, P. Scalart, and C. Marro, "Computationally efficient and robust frequency-domain GSC," in *Proc. 12th IEEE Int. Workshop Acoustic Echo Noise Control*, Tel-Aviv, Israel, Aug./Sep. 2010, pp. 1–4.

[12] D. Marquardt and S. Doclo, "Performance comparison of bilateral and binaural MVDR-based noise reduction algorithms in the presence of DOA estimation errors," in *Proc. ITG Symp. Speech Commun.*, Oct. 2016, pp. 1–5.

[13] S. Doclo and M. Moonen, "Superdirective beamforming robust against microphone mismatch," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 15, no. 2, pp. 617–631, Feb. 2007.

[14] M. R. P. Thomas, J. Ahrens, and I. J. Tashev, "Beamformer design using measured microphone directivity patterns: Robustness to modelling error," in *Proc. Asia–Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*, Dec. 2012, pp. 1–4.

[15] X. Wang, M. Amin, and X. Wang, "Robust sparse array design for adaptive beamforming against DOA mismatch," *Signal Process.*, vol. 146, pp. 41–49, May 2018.

[16] E. Mabande, A. Schad, and W. Kellermann, "Design of robust superdirective beamformers as a convex optimization problem," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICAASP)*, Apr. 2009, pp. 77–80.

[17] H. Barfuss, C. Huemmer, G. Lamani, A. Schwarz, and W. Kellermann, "HRTF-based robust least-squares frequency-invariant beamforming," in *Proc. IEEE Workshop Appl. Signal Process. Audio Acoust. (WASPAA)*, Oct. 2015, pp. 1–5.

[18] J. Yang, Y. Zhang, and Z. Chen, "Robust adaptive wideband beamforming with combined frequency response invariance and eigenvector constraints," in *Proc. Progr. Electromagn. Res. Symp. (PIERS)*, Aug. 2016, pp. 1374–1378.

[19] P. Chen, Y. Yang, Y. Wang, Y. Ma, and L. Yang, "Robust covariance matrix reconstruction algorithm for time-domain wideband adaptive beamforming," *IEEE Trans. Veh. Technol.*, vol. 68, no. 2, pp. 1405–1416, Feb. 2019.

[20] Y. Buchris, A. Amar, J. Benesty, and I. Cohen, "Incoherent synthesis of sparse arrays for frequency-invariant beamforming," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 27, no. 3, pp. 482–495, Mar. 2019.

[21] C. Blandin, E. Vincent, and A. Ozerov, "Multi-source TDOA estimation using SNR-based angular spectra," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2011, pp. 2616–2619.

[22] M. Zohourian, G. Enzner, and R. Martin, "On the use of beamforming approaches for binaural speaker localization," in *Proc. ITG Symp. Speech Commun.*, Oct. 2016, pp. 1–5.

[23] D. Marquardt and S. Doclo, "Noise power spectral density estimation for binaural noise reduction exploiting direction of arrival estimates," in *Proc. IEEE Workshop Appl. Signal Process. Audio Acoust. (WASPAA)*, Oct. 2017, pp. 234–238.

[24] S. Chakrabarty and E. A. P. Habets, "Broadband DOA estimation using convolutional neural networks trained with noise signals," in *Proc. IEEE Workshop Appl. Signal Process. Audio Acoust. (WASPAA)*, Oct. 2017, pp. 136–140.

[25] R. Venkatesan and A. B. Ganesh, "Deep recurrent neural networks based binaural speech segregation for the selection of closest target of interest," *Multimed Tools Appl.*, vol. 77, no. 15, pp. 20129–20156, Aug. 2018.

[26] M. Yiwere and E. J. Rhee, "Distance estimation and localization of sound sources in reverberant conditions using deep neural networks," *Int. J. Appl. Eng. Res.*, vol. 12, no. 22, pp. 12384–12389, 2017.

[27] X. Xiao, S. Zhao, X. Zhong, D. L. Jones, E. S. Chng, and H. Li, "A learning-based approach to direction of arrival estimation in noisy and reverberant environments," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2015, pp. 2814–2818.

[28] E. Hadad, D. Marquardt, S. Doclo, and S. Gannot, "Theoretical analysis of binaural transfer function MVDR beamformers with interference cue preservation constraints," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 23, no. 12, pp. 2449–2464, Dec. 2015.
[29] H. Cox, "Resolving power and sensitivity to mismatch of optimum array processors," *J. Acoust. Soc. Amer.*, vol. 54, no. 3, pp. 771–785, Sep. 1973.
[30] T. Lotter and P. Vary, "Dual-channel speech enhancement by superdirective beamforming," *EURASIP J. Adv. Signal Process.*, vol. 2006, Dec. 2006, Art. no. 63297.
[31] L. J. Griffiths and C. W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Antennas Propag.*, vol. 30, no. 1, pp. 27–34, Jan. 1982.
[32] B. R. Breed and J. Strauss, "A short proof of the equivalence of LCMV and GSC beamforming," *IEEE Signal Process. Lett.*, vol. 9, no. 6, pp. 168–169, Jun. 2002.
[33] E. Hadad, S. Doclo, and S. Gannot, "The binaural LCMV beamformer and its performance analysis," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 24, no. 3, pp. 543–558, Mar. 2016.
[34] H. As'ad, M. Bouchard, and H. Kamkar-Parsi, "A robust target linearly constrained minimum variance beamformer with spatial cues preservation for binaural hearing aids," *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 27, no. 10, pp. 1549–1563, Oct. 2019.

HALA AS'AD received the B.A.Sc. degree in biomedical engineering from the Jordan University of Science and Technology, Jordan, in 2010, and the M.A.Sc. degree in electrical engineering with a specialization in audio and speech processing from the University of Ottawa, Ottawa, ON, Canada, in 2015, where she is currently pursuing the Ph.D. degree in electrical engineering. Her Ph.D. research focuses on robust binaural beamforming, binaural cues preservation, and the source direction of arrival detection in hearing aids. Her research interests include applied signal processing and machine learning with emphases on audio and speech processing, array signal processing, beamforming, speech enhancement, acoustic source localization, and hearing aids. Her scholarships and awards include Natural Sciences and Engineering Research Council (NSERC) Scholarship, Excellence Scholarship, and Ontario Graduate Scholarship.

MARTIN BOUCHARD received the B.Ing., M.Sc.A., and Ph.D. degrees in electrical engineering from the Université de Sherbrooke, Sherbrooke, QC, Canada, in 1993, 1995, and 1997, respectively. In January 1998, he joined the School of Electrical Engineering and Computer Science, Faculty of Engineering, University of Ottawa, Ottawa, ON, Canada, where he is currently a Professor. In 1996, he co-founded SoftdB Inc., Quebec, QC, Canada, which is still active today. Over the years, he has conducted research activities and consulting activities with more than 20 private sector and governmental partners, supervised more than 50 graduate students and postdoctoral fellows, and authored or coauthored more than 40 journal papers and 85 conference papers. His current research interests include signal processing methods in general and machine learning, with an emphasis on speech, audio, acoustics, hearing aids, and biomedical engineering applications. He served as a member for the Speech and Language Technical Committee of the IEEE Signal Processing Society from 2009 to 2011. He is a member of the Ordre des Ingénieurs du Québec. He served as an Associate Editor for the *EURASIP Journal on Audio, Speech and Music Processing* from 2006 to 2011 and for the IEEE TRANSACTIONS ON NEURAL NETWORKS from 2008 to 2009.

HOMAYOUN KAMKAR-PARSI received the B.A.Sc., M.A.Sc., and Ph.D. degrees in electrical engineering from the School of Information Technology and Engineering, University of Ottawa, Ottawa, ON, Canada, in 2001, 2004, and 2009, respectively. During his bachelor's studies, he has obtained the highest standing in his graduating class in electrical engineering and the Silver medal for the second highest standing in the entire Faculty of Engineering. His graduate scholarships included the Natural Sciences and Engineering Research Council (NSERC) scholarship and the Ontario Graduate Scholarship. Since 2009, he has been with Siemens Audiologische Technik GmbH (renamed as Sivantos GmbH in 2015 and as WS Audiology in 2019), Erlangen, Germany, where his main work and research include speech/audio signal processing with applications in speech enhancement, advanced multi-microphone beamforming for binaural hearing aids including remote external microphones (e.g., from smartphone), source localization and tracking, advanced scene analysis, and machine learning (neural networks). In 2018, he was selected as one of the top inventors at Sivantos.