EXTENSIONS OF THE BINAURAL MWF WITH INTERFERENCE REDUCTION PRESERVING THE BINAURAL CUES OF THE INTERFERING SOURCE

Elior Hadad[1], Daniel Marquardt[2], Simon Doclo[2], and Sharon Gannot[1]

[1] Faculty of Engineering, Bar-Ilan University, Ramat-Gan, Israel
[2] University of Oldenburg, Department of Medical Physics and Acoustics and Cluster of Excellence Hearing4All, Oldenburg, Germany

{elior.hadad,sharon.gannot}@biu.ac.il {daniel.marquardt,simon.doclo}@uni-oldenburg.de

ABSTRACT

Recently, an extension of the binaural multichannel Wiener filter (BMWF), referred to as BMWF-IR$_0$, was presented in which an interference rejection constraint was added to the BMWF cost function. Although the BMWF-IR$_0$ aims to entirely suppress the interfering source, residual interfering sources (as well as unconstrained noise sources) are undesirably perceived as impinging the array from the desired source direction. In this paper, we propose two extensions of the BMWF-IR$_0$ that address this issue by preserving the spatial impression of the interfering source. In the first extension, the binaural cues of the interfering source are preserved, while those of the desired source may be slightly distorted. In the second extension, the binaural cues of both the desired and interfering sources are preserved. Simulation results show that the noise reduction performance of both proposed extensions is comparable to the BMWF-IR$_0$.

Index Terms— Hearing aids, Binaural cues, LCMV Beamforming, MWF, Noise and interference reduction.

1. INTRODUCTION

Binaural hearing aid devices consisting of a hearing aid mounted on each ear of a hearing-impaired person, are known to outperform their monaural counterparts in terms of noise reduction performance and their capability to preserve the binaural cues and hence the spatial impression of the acoustical scene [1, 2]. By preserving the binaural cues, in addition to improving sound localization, a better speech intelligibility in noisy environments can be achieved due to binaural unmasking [3, 4]. For directional sources, preserving of the interaural level difference (ILD) and the interaural time difference (ITD) cues can be achieved by preserving the so-called relative transfer function (RTF), which is defined as the ratio of the acoustical transfer functions relating the source and the two ears.

In the last decade, several binaural speech enhancement algorithms aiming to preserve the binaural cues have been developed [1, 5–16]. In [1], the binaural multichannel Wiener filter (BMWF) was presented. It was shown in [1, 11] that the BMWF preserves the binaural cues of the desired source but distorts the binaural cues of the noise, such that both the desired source and the noise are perceived as arriving from the desired source direction. To optimally exploit the benefits of binaural unmasking and optimize the spatial awareness of the hearing aid user, several extensions of the BMWF have been proposed, which aim to also preserve the binaural cues of the residual noise by including cue preservation terms in the binaural cost function [11, 12, 14, 17]. If the desired source must be processed without distortion, the binaural minimum variance distortionless response (BMVDR) beamformer can be applied [15]. However, similarly to the BMWF, a drawback of the BMVDR is the fact that the binaural cues of the noise are not preserved.

Many acoustic scenarios consist of a desired source corrupted by one or more directional interfering sources (e.g. competing speakers) and additive noise (e.g. diffuse background noise). In order to control both the suppression and the binaural cue preservation of the directional interfering sources, the binaural linearly constrained minimum variance (BLCMV) beamformer was proposed in [13, 16]. In the BLCMV criterion, a hard constraint controlling the amount of interference reduction was added to the BMVDR cost function, with a parameter η denoting the *cue gain factor*. When η is larger than 0, it was shown that the BLCMV beamformer is able to preserve the binaural cues of both the desired and interfering sources.

Recently, it has been proposed to add a similar interference rejection constraint (with $\eta = 0$) to the BMWF cost function, referred to as the binaural multichannel Wiener filter with interference rejection (BMWF-IR$_0$) [18]. It was shown that the BMWF-IR$_0$ can be decomposed into the BLCMV beamformer (with $\eta = 0$) and a single-channel Wiener postfilter. Since in addition to spatial information, the BMWF-IR$_0$ also exploits the spectral characteristics of the sources, it is able to provide additional noise reduction sacrificing speech distortion. However, similarly to the BMWF and the BMVDR beamformer, in the BMWF-IR$_0$ *all* sources are perceived as coming from the desired source direction. This implies that in practice, due to unavoidable estimation errors, the residual interfering source will also be perceived as arriving from the desired source direction. Clearly, this is an undesirable phenomenon and in some scenarios (e.g. traffic) even dangerous to the hearing aid user.

In this paper, we introduce two extensions of the BMWF-IR$_0$ that address this issue and optimally exploit the benefits of binaural unmasking by preserving the binaural cues of the interfering source. In the first extension, a hard constraint is added to the BMWF cost function, which controls the amount of interference reduction (similarly to the BLCMV). It is proven that the binaural cues of the interfering source are preserved while the binaural cues of the desired source may be slightly distorted. The second extension comprises a combination of two binaural beamformers, i.e. the BMWF-IR$_0$ steering a null towards the interfering source, and a binaural beamformer, controlling the amount of interference reduction while steer-

This work was partially supported by a joint Lower Saxony-Israeli Project, funded by the State of Lower Saxony in Germany, a Grant from the German-Israeli Foundation for Scientific Research and Development, and the Cluster of Excellence 1077 Hearing4All, funded by the German Research Foundation (DFG).

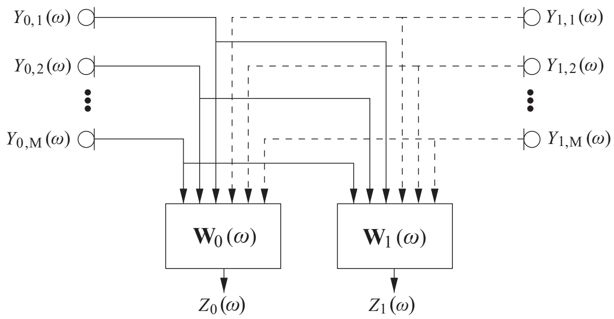

Fig. 1: General binaural processing scheme.

ing a null towards the desired source. It is proven that for the second extension, the binaural cues of both the desired and interfering sources are preserved. Experimental validations in an office scenario demonstrate the performance of the proposed beamformers.

2. PROBLEM FORMULATION

We consider an acoustic scenario comprising one desired source and one directional interfering source in a noisy and reverberant environment. The binaural hearing device, consisting of two hearing aids each equipped with M microphones, is depicted in Fig. 1. All microphone signals can be stacked in the $2M$-dimensional vector $\mathbf{Y}(\omega)$ in the frequency-domain as

$$\mathbf{Y}(\omega) = \mathbf{X}(\omega) + \mathbf{U}(\omega) + \mathbf{N}(\omega) = \mathbf{X}(\omega) + \mathbf{V}(\omega), \quad (1)$$

with $\mathbf{Y}(\omega) = [Y_{0,1}(\omega) \ldots Y_{0,M}(\omega)\ Y_{1,1}(\omega) \ldots Y_{1,M}(\omega)]^T$, $\mathbf{X}(\omega)$ the desired source component, $\mathbf{U}(\omega)$ the interfering source component, and $\mathbf{N}(\omega)$ the additional background noise. The vector $\mathbf{V}(\omega) = \mathbf{U}(\omega) + \mathbf{N}(\omega)$ is defined as the total undesired component as received by the microphones, i.e. the interfering source component plus background noise. $\mathbf{X}(\omega)$, $\mathbf{U}(\omega)$, $\mathbf{V}(\omega)$, and $\mathbf{N}(\omega)$ are defined similarly to $\mathbf{Y}(\omega)$. The variable ω will henceforth be omitted for brevity. We can further write $\mathbf{X} = S_{\mathrm{d}}\mathbf{A}$ and $\mathbf{U} = S_{\mathrm{u}}\mathbf{B}$, where S_{d} and S_{u} are the desired and interfering source signals and $\mathbf{A}$ and $\mathbf{B}$ are the acoustic transfer functions (ATFs) relating the microphones and the desired and interfering source, respectively. Assuming the directional sources and the noise are uncorrelated, the spatial correlation matrix of the noisy microphone signals can be written as

$$\mathbf{R}_y = \mathbf{R}_x + \mathbf{R}_u + \mathbf{R}_n = \mathbf{R}_x + \mathbf{R}_v, \quad (2)$$

where $\mathbf{R}_x = \mathcal{E}\{\mathbf{X}\mathbf{X}^H\} = P_s\mathbf{A}\mathbf{A}^H$, $\mathbf{R}_u = \mathcal{E}\{\mathbf{U}\mathbf{U}^H\} = P_u\mathbf{B}\mathbf{B}^H$, and $\mathbf{R}_n = \mathcal{E}\{\mathbf{N}\mathbf{N}^H\}$ are the desired source, interfering source, and noise correlation matrices, respectively. $\mathcal{E}\{\cdot\}$ denotes the expectation operator and $P_s = \mathcal{E}\{|S_{\mathrm{d}}|^2\}$ and $P_u = \mathcal{E}\{|S_{\mathrm{u}}|^2\}$ denote the power spectral densities (PSDs) of the desired source and the interfering source, respectively. Without loss of generality, the first microphone on the left and the right hearing aid are selected as the reference microphones. For conciseness, the reference microphone signals $Y_{0,1}$ and $Y_{1,1}$ on the left and the right hearing aid are denoted as Y_0 and Y_1, and are equal to

$$Y_0 = S_{\mathrm{d}}A_0 + S_{\mathrm{u}}B_0 + N_0, Y_1 = S_{\mathrm{d}}A_1 + S_{\mathrm{u}}B_1 + N_1. \quad (3)$$

The input RTFs of the desired and the interfering source between the reference microphones on the left and the right hearing aid are defined as the ratio of the ATFs, i.e.

$$\mathrm{RTF}_{\mathrm{x}}^{\mathrm{in}} = \frac{A_0}{A_1}, \qquad \mathrm{RTF}_{\mathrm{u}}^{\mathrm{in}} = \frac{B_0}{B_1}. \quad (4)$$

Note that the RTF is a complex-valued frequency-dependent scalar from which the binaural ILD and ITD cues [11] can be computed as

$$\mathrm{ILD} = 20\log_{10}(|\mathrm{RTF}|), \qquad \mathrm{ITD} = \frac{\angle(\mathrm{RTF})}{\omega}, \quad (5)$$

with $\angle$ denoting the phase.

3. BINAURAL NOISE REDUCTION ALGORITHMS

In Section 3.1, we first briefly review the BMWF and BMWF-IR$_0$. In Section 3.2 and Section 3.3, two extensions of the BMWF-IR$_0$ are introduced with a term related to interfering source binaural cues.

3.1. BMWF/ BMWF with interference rejection (BMWF-IR$_0$)

The well-known BMWF produces a minimum mean square error (MSE) estimate of the desired source component at both reference microphones [1]. The MSE cost functions for the filter $\mathbf{W}_0$, estimating the desired source component X_0 at the left hearing aid, and for the filter $\mathbf{W}_1$, estimating the desired source component X_1 at the right hearing aid, are given by

$$\begin{aligned} J_{\mathrm{BMWF}}(\mathbf{W}_0) &= \mathcal{E}\{\|[X_0 - \mathbf{W}_0^H\mathbf{X}]\|^2 + \mu\|\mathbf{W}_0^H\mathbf{V}\|^2\}, \\ J_{\mathrm{BMWF}}(\mathbf{W}_1) &= \mathcal{E}\{\|[X_1 - \mathbf{W}_1^H\mathbf{X}]\|^2 + \mu\|\mathbf{W}_1^H\mathbf{V}\|^2\}, \end{aligned} \quad (6)$$

where μ provides a trade-off between noise reduction and speech distortion. The filters minimizing $J_{\mathrm{BMWF}}(\mathbf{W}_0)$ and $J_{\mathrm{BMWF}}(\mathbf{W}_1)$ are given by [1]:

$$\begin{aligned} \mathbf{W}_{0,\mathrm{BMWF}} &= P_s A_0^* \tilde{\mathbf{R}}_y^{-1}\mathbf{A}, \\ \mathbf{W}_{1,\mathrm{BMWF}} &= P_s A_1^* \tilde{\mathbf{R}}_y^{-1}\mathbf{A}, \end{aligned} \quad (7)$$

with $\tilde{\mathbf{R}}_y = \mathbf{R}_x + \mu\mathbf{R}_v$.

In order to reject a directional interfering source, an extension of the BMWF cost function with a constraint related to the interfering source component was proposed, resulting in the BMWF-IR$_0$ criteria [17, 18], i.e.

$$\begin{aligned} &\min_{\mathbf{W}_0} J_{\mathrm{BMWF}}(\mathbf{W}_0) \quad \text{subject to} \quad \mathbf{W}_0^H\mathbf{B} = 0, \\ &\min_{\mathbf{W}_1} J_{\mathrm{BMWF}}(\mathbf{W}_1) \quad \text{subject to} \quad \mathbf{W}_1^H\mathbf{B} = 0. \end{aligned} \quad (8)$$

Solving (8), the left and the right filters of the BMWF-IR$_0$ are given by [18]:

$$\begin{aligned} \mathbf{W}_{0,\mathrm{BMWF\text{-}IR}_0} &= P_s A_0^* \left[\tilde{\mathbf{R}}_y^{-1}\mathbf{A} - \frac{\lambda_{ab}^*}{\lambda_b}\tilde{\mathbf{R}}_y^{-1}\mathbf{B}\right], \\ \mathbf{W}_{1,\mathrm{BMWF\text{-}IR}_0} &= P_s A_1^* \left[\tilde{\mathbf{R}}_y^{-1}\mathbf{A} - \frac{\lambda_{ab}^*}{\lambda_b}\tilde{\mathbf{R}}_y^{-1}\mathbf{B}\right], \end{aligned} \quad (9)$$

with $\lambda_a = \mathbf{A}^H\tilde{\mathbf{R}}_y^{-1}\mathbf{A}$, $\lambda_{ab} = \mathbf{A}^H\tilde{\mathbf{R}}_y^{-1}\mathbf{B}$ and $\lambda_b = \mathbf{B}^H\tilde{\mathbf{R}}_y^{-1}\mathbf{B}$. For both the filters of the BMWF in (7) and the BMWF-IR$_0$ in (9), $\mathbf{W}_0 = (\mathrm{RTF}_{\mathrm{x}}^{\mathrm{in}})^*\mathbf{W}_1$. This implies that $\mathbf{W}_0$ and $\mathbf{W}_1$ are parallel, such that the RTF of the desired source at the output of both the BMWF and BMWF-IR$_0$ is equal to the input RTF, i.e.

$$\mathrm{RTF}_{\mathrm{x}}^{\mathrm{out}} = \frac{A_0}{A_1} = \mathrm{RTF}_{\mathrm{x}}^{\mathrm{in}}. \quad (10)$$

However, this also implies that *all* sound sources are perceived as arriving from the desired source direction. The BMWF-IR$_0$ is steering a null towards the interfering source. However, when RTF estimation errors occur, the interfering source is not entirely suppressed and the residual interference will also be perceived as arriving from the desired source direction. To resolve this issue, in the following subsections we introduce two extensions that simultaneously aim to reduce the interfering source and preserve the binaural cues of the interfering source.

3.2. BMWF-IR with controllable interference reduction (BMWF-IR$_\eta$)

In the first extension of the BMWF-IR$_0$, we propose to preserve the binaural cues of the interfering source by extending the BMWF cost function in (6) with a hard constraint that controls the amount of interference reduction (similarly to the BLCMV [13, 16]), i.e.

$$\begin{array}{lll} \min_{\mathbf{W}_0} J_{\text{BMWF}}(\mathbf{W}_0) & \text{subject to} & \mathbf{W}_0^H\mathbf{B} = \eta B_0, \\ \min_{\mathbf{W}_1} J_{\text{BMWF}}(\mathbf{W}_1) & \text{subject to} & \mathbf{W}_1^H\mathbf{B} = \eta B_1, \end{array} \tag{11}$$

where $0 \leq \eta \leq 1$ is a real-valued scalar, defined as the *cue gain factor* [16] that provides a trade-off between interference reduction and binaural cue preservation. As can be observed, the BMWF-IR$_0$ criterion in (8) is a special case of (11) with $\eta = 0$.

The Lagrangian for the left filter cost function is equal to

$$\begin{aligned} \mathcal{L}(\mathbf{W}_0) = \mathbf{W}_0^H\tilde{\mathbf{R}}_y\mathbf{W}_0 - \mathbf{W}_0^H\mathbf{r}_{x,0} - \mathbf{r}_{x,0}^H\mathbf{W}_0 + P_s|A_0|^2 + \\ \lambda(\mathbf{W}_0^H\mathbf{B} - \eta B_0) - \lambda^*(\mathbf{B}^H\mathbf{W}_0 - (\eta B_0)^*), \end{aligned} \tag{12}$$

where λ is a Lagrange multiplier and $\mathbf{r}_{x,0} = P_s\mathbf{A}A_0^*$. Setting the derivative with respect to $\mathbf{W}_0^H$ to 0 yields

$$\mathbf{W}_0 = \tilde{\mathbf{R}}_y^{-1}(\mathbf{r}_{x,0} - \lambda\mathbf{B}). \tag{13}$$

By applying the constraint in (11), the Lagrange multiplier λ can be computed as $\mathbf{B}^H\mathbf{W}_0 = \eta B_0^* = \mathbf{B}^H\tilde{\mathbf{R}}_y^{-1}(\mathbf{r}_{x,0} - \lambda\mathbf{B})$, yielding

$$\lambda = \frac{\mathbf{B}^H\tilde{\mathbf{R}}_y^{-1}\mathbf{r}_{x,0} - \eta B_0^*}{\mathbf{B}^H\tilde{\mathbf{R}}_y^{-1}\mathbf{B}} = \frac{P_s\lambda_{ab}^*A_0^* - \eta B_0^*}{\lambda_b}. \tag{14}$$

The left filter is derived by substituting (14) into (13). The right filter is similarly derived. Hence, the left and the right filters, which we will refer to as BMWF-IR$_\eta$, are given by

$$\begin{aligned} \mathbf{W}_{0,\text{BMWF-IR}_\eta} &= P_sA_0^*\left(\tilde{\mathbf{R}}_y^{-1}\mathbf{A} - \frac{\lambda_{ab}^* - \frac{\eta B_0^*}{P_sA_0^*}}{\lambda_b}\tilde{\mathbf{R}}_y^{-1}\mathbf{B}\right), \\ \mathbf{W}_{1,\text{BMWF-IR}_\eta} &= P_sA_1^*\left(\tilde{\mathbf{R}}_y^{-1}\mathbf{A} - \frac{\lambda_{ab}^* - \frac{\eta B_1^*}{P_sA_1^*}}{\lambda_b}\tilde{\mathbf{R}}_y^{-1}\mathbf{B}\right). \end{aligned} \tag{15}$$

By comparing (9) and (15), it can be observed that the filters of the BMWF-IR$_0$ and BMWF-IR$_\eta$ are quite similar (and equal for $\eta = 0$). Note however that the BMWF-IR$_\eta$ filters in (15) are in general not parallel[1].

Since the BMWF-IR$_\eta$ satisfies the constraints in (11) for the interfering source, the RTF of the interfering source at the output of the BMWF-IR$_\eta$ is equal to the input RTF, i.e.

$$\text{RTF}_\text{u}^\text{out} = \frac{\mathbf{W}_0^H\mathbf{B}}{\mathbf{W}_1^H\mathbf{B}} = \frac{\eta B_0}{\eta B_1} = \text{RTF}_\text{u}^\text{in}. \tag{16}$$

[1]For $\eta \neq 0$, the BMWF-IR$_\eta$ filters are only parallel if $A_0/A_1 = B_0/B_1$, i.e. $\text{RTF}_\text{x}^\text{in} = \text{RTF}_\text{u}^\text{in}$.

Hence, the BMWF-IR$_\eta$ preserves the binaural cues of the interfering source. On the other hand, the RTF of the desired source at the output of the BMVDR-IR$_\eta$ is equal to

$$\text{RTF}_\text{x}^\text{out} = \frac{\mathbf{W}_0^H\mathbf{A}}{\mathbf{W}_1^H\mathbf{A}} = \frac{P_s\lambda_a\left(A_0(1-\Lambda) + \eta B_0\frac{\Lambda}{P_s\lambda_{ab}}\right)}{P_s\lambda_a\left(A_1(1-\Lambda) + \eta B_1\frac{\Lambda}{P_s\lambda_{ab}}\right)}, \tag{17}$$

with $\Lambda = \frac{|\lambda_{ab}|^2}{\lambda_a\lambda_b} \leq 1$. Hence, for the desired source the output RTF is not equal to the input RTF, such that the binaural cues are distorted. Note however that for (very) small values of η, the output RTF of the desired source is (very) close to the input RTF.

3.3. BMWF-IR with additional filter for the interfering source (BMWF-IR$_\eta$-B)

The second proposed extension of the BMWF-IR$_0$ is related to the BLCMV beamformer [13, 16]. The BLCMV beamformer is designed to reproduce the desired source component of both reference microphone signals without distortion, while minimizing the overall noise power and reducing the directional interfering source by the same amount in both hearing aids, i.e.

$$\begin{array}{lll} \min_{\mathbf{W}_0} \mathbf{W}_0^H\mathbf{R}_v\mathbf{W}_0 & \text{subject to} & \mathbf{C}^H\mathbf{W}_0 = \mathbf{b}_0, \\ \min_{\mathbf{W}_1} \mathbf{W}_1^H\mathbf{R}_v\mathbf{W}_1 & \text{subject to} & \mathbf{C}^H\mathbf{W}_1 = \mathbf{b}_1, \end{array} \tag{18}$$

with the constraint set equal to

$$\mathbf{C} = \begin{bmatrix}\mathbf{A} & \mathbf{B}\end{bmatrix},\ \mathbf{b}_0 = \begin{bmatrix}A_0^* \\ \eta B_0^*\end{bmatrix},\ \mathbf{b}_1 = \begin{bmatrix}A_1^* \\ \eta B_1^*\end{bmatrix}, \tag{19}$$

where η is the *cue gain factor* as defined above. The filters solving (18) are given by:

$$\mathbf{W}_0 = \mathbf{R}_v^{-1}\mathbf{C}\left[\mathbf{C}^H\mathbf{R}_v^{-1}\mathbf{C}\right]^{-1}\mathbf{b}_0, \mathbf{W}_1 = \mathbf{R}_v^{-1}\mathbf{C}\left[\mathbf{C}^H\mathbf{R}_v^{-1}\mathbf{C}\right]^{-1}\mathbf{b}_1. \tag{20}$$

The left and the right filters of the BLCMV beamformer can be decomposed as a combination of two beamformers, i.e.

$$\mathbf{W}_0 = \mathbf{W}_{X,0} + \eta\mathbf{W}_{U,0}, \quad \mathbf{W}_1 = \mathbf{W}_{X,1} + \eta\mathbf{W}_{U,1}. \tag{21}$$

On the one hand, $\mathbf{W}_{X,0}$ and $\mathbf{W}_{X,1}$ denote the left and the right filters of the so-called desired source BLCMV (D-BLCMV) beamformer, which reproduce the desired source while entirely canceling the interfering source, i.e. solving (18) with

$$\mathbf{b}_0 = \begin{bmatrix}A_0^* \\ 0\end{bmatrix},\ \mathbf{b}_1 = \begin{bmatrix}A_1^* \\ 0\end{bmatrix}. \tag{22}$$

On the other hand, $\mathbf{W}_{U,0}$ and $\mathbf{W}_{U,1}$ denote the left and the right filters of the so-called undesired source BLCMV (U-BLCMV) beamformer, which reproduce the interfering source while entirely canceling the desired source, i.e. solving (18) with

$$\mathbf{b}_0 = \begin{bmatrix}0 \\ B_0^*\end{bmatrix},\ \mathbf{b}_1 = \begin{bmatrix}0 \\ B_1^*\end{bmatrix}. \tag{23}$$

In [18], it was proven that the BMWF-IR$_0$ can be decomposed into the D-BLCMV beamformer (i.e. the BLCMV beamformer with $\eta = 0$) followed by a single-channel Wiener postfilter, i.e.

$$\mathbf{W}_{0,\text{BMWF-IR}_0} = \frac{\rho}{\mu+\rho}\mathbf{W}_{X,0},\ \mathbf{W}_{1,\text{BMWF-IR}_0} = \frac{\rho}{\mu+\rho}\mathbf{W}_{X,1}, \tag{24}$$

where $\rho = P_s\gamma_a(1-\Gamma)$ is the output signal-to-noise ratio (SNR) of the D-BLCMV beamformer, with $\gamma_a = \mathbf{A}^H\mathbf{R}_v^{-1}\mathbf{A}$, $\gamma_{ab} = \mathbf{A}^H\mathbf{R}_v^{-1}\mathbf{B}$, $\gamma_b = \mathbf{B}^H\mathbf{R}_v^{-1}\mathbf{B}$, and $\Gamma = \frac{|\gamma_{ab}|^2}{\gamma_a\gamma_b}$. Since adding an identical postfilter to both sides of the D-BLCMV beamformer will not affect the binaural cues, the binaural cues of the output signal of the BMWF-IR$_0$ and the D-BLCMV beamformer are equal.

Contrary to the BLCMV beamformer, which only utilizes spatial information, we now propose to additionally exploit the spectral characteristics of the sources by replacing the D-BLCMV beamformer in (21) with the BMWF-IR$_0$, i.e. the D-BLCMV beamformer followed by a single-channel postfilter, yielding

$$\begin{aligned} \mathbf{W}_{0,\text{BMWF-IR}_\eta\text{-B}} &= \mathbf{W}_{0,\text{BMWF-IR}_0} + \eta\mathbf{W}_{U,0}, \\ \mathbf{W}_{1,\text{BMWF-IR}_\eta\text{-B}} &= \mathbf{W}_{1,\text{BMWF-IR}_0} + \eta\mathbf{W}_{U,1}, \end{aligned} \tag{25}$$

which we will refer to as BMWF-IR$_\eta$-B.

Since the BMWF-IR$_0$ preserves the binaural cues of the desired source and since the U-BLCMV beamformer steers a null towards the desired source, the RTF of the desired source at the output of the BMWF-IR$_\eta$-B is equal to the respective input RTF, i.e.

$$\text{RTF}_\text{x}^\text{out} = \frac{A_0+0}{A_1+0} = \text{RTF}_\text{x}^\text{in}. \tag{26}$$

In addition, since the BMWF-IR$_0$ steers a null towards the interfering source and since the U-BLCMV beamformer satisfies the constraints for the interfering source, the RTF of the interfering source at the output of the BMWF-IR$_\eta$-B is equal to the respective input RTF, i.e.

$$\text{RTF}_\text{u}^\text{out} = \frac{0+\eta B_0}{0+\eta B_1} = \text{RTF}_\text{u}^\text{in}. \tag{27}$$

Hence, the BMWF-IR$_\eta$-B preserves the RTFs of both the desired and interfering sources.

4. EXPERIMENTAL VALIDATION

The performance of the algorithms was evaluated using Behind-The-Ear impulse responses (IRs) from [19] at a sampling frequency of 16 kHz. To verify the theoretical analysis, we used actual IRs and artificial sources, hence circumventing any estimation errors issues[2]. All experiments were carried out using $M = 4$ microphones, i.e. two microphones on each hearing aid. The acoustic scenario comprised one desired source at $\theta_x = -30^o$ and 1m from the listener, one interfering source at $\theta_v = 45^o$ and 1m from the listener[3], and diffuse noise. The reverberation time was approximately 400 ms. Two different stationary signals with speech-shaped PSDs P_s and P_u were selected as the desired and interference input signals. A cylindrically isotropic noise field was simulated by averaging the anechoic IRs from [19]. The noise PSD was modelled as speech-shaped noise calculated by averaging multiple speech PSDs. For all algorithms, the trade-off parameter μ was set to 1. The BMWF-IR$_\eta$ and BMWF-IR$_\eta$-B were evaluated for η equal to 0.1 and 0.2.

We compared the performance of the considered algorithms in terms of the (wide-band) signal-to-interference-and-noise ratio (SINR), which is defined as the ratio of the average PSDs of the desired and the total undesired sources (i.e. interference plus background noise) over all frequency bands. The results for various input SNRs and SIRs are summarized in Table 1. It can be observed that the BMWF yields a large SINR for all scenarios compared with the other considered algorithms. By setting the input SIR to 0 dB, the SINR of the BMWF-IR$_0$ is comparable for both BMWF-IR$_\eta$, and BMWF-IR$_\eta$-B for $\eta = 0.1$. The SINR of both extensions degrades for $\eta = 0.2$ (due to lower interference reduction). By setting the input SIR to 10 dB, the SINR of the BMWF-IR$_0$ degrades compare to both BMWF-IR$_\eta$, and BMWF-IR$_\eta$-B, due to the null constraint directed towards the interfering source.

		SINR improvement					
SNR$_\text{IN}$	SIR$_\text{IN}$	MWF	IR$_0$	IR$_\eta$ (0.1)	IR$_\eta$ (0.2)	IR$_\eta$-B (0.1)	IR$_\eta$-B (0.2)
0	0	10.65	9.89	9.78	8.87	9.78	8.82
0	10	8.37	6.60	6.87	6.91	6.88	6.90

Table 1: SINR improvement (in dB) relative to the left signal obtained by the BMWF, BMWF-IR$_0$, BMWF-IR$_\eta(\eta)$, and BMWF-IR$_\eta$-B(η).

		ILD$_u$ error		ILD$_x$ error			
SNR$_\text{IN}$	SIR$_\text{IN}$	MWF	IR$_0$	IR$_\eta$ (0.1)	IR$_\eta$ (0.2)	IR$_\eta$-B (0.1)	IR$_\eta$-B (0.2)
0	0	15.9	NaN	1.67	2.7	0	0
0	10	15.9	NaN	1.67	2.7	0	0
		ITD$_u$ errors		ITD$_x$ errors			
SNR$_\text{IN}$	SIR$_\text{IN}$	MWF	IR	IR$_\eta$ (0.1)	IR$_\eta$ (0.2)	IR-B (0.1)	IR-B (0.2)
0	0	$307\mu s$	NaN	$162\mu s$	$157\mu s$	0	0
0	10	$307\mu s$	NaN	$162\mu s$	$157\mu s$	0	0

Table 2: ILD error (in dB) and ITD error (in μs) obtained by the BMWF, BMWF-IR$_0$, BMWF-IR$_\eta(\eta)$, and BMWF-IR$_\eta$-B(η).

The binaural cue preservation performance was analyzed by calculating the ILD and ITD errors, averaged over all frequencies, for the desired and interfering sources. The ILD and ITD cues were calculated according to (5) and are summarized in Table 2. Both the BMWF and BMWF-IR$_0$ preserve the binaural cues of the desired source, as well as the BMWF-IR$_\eta$ and BMWF-IR$_\eta$-B, preserve the binaural cues of the interfering source and therefore omitted from the table due to space constraints. Since the BMWF filters are parallel, hence imposing on all sources the binaural cues of the desired source, the resulting ILD and ITD errors for the interfering source are very high. In contrast, both the BMWF-IR$_\eta$ and BMWF-IR$_\eta$-B preserve the binaural cues of the interfering source. The binaural cues of the desired source are preserved for the BMWF-IR$_\eta$-B, while small ILD and ITD errors are obtained for the desired source when the BMWF-IR$_\eta$ is applied.

5. CONCLUSION

In this paper, we proposed two extensions of the BMWF-IR$_0$ designed to estimate the desired source and control the amount of interference reduction in a noisy and reverberant environment. Both proposed extensions are capable of preserving the binaural cues of the interfering source. While the BMWF-IR$_\eta$-B preserves the binaural cues of the desired source, the BMWF-IR$_\eta$ slightly distorts it. The noise reduction performance of both extensions is comparable to the original BMWF-IR$_0$.

[2]Note that for implementing the algorithms it is sufficient to estimate the relative ATFs rather than the ATFs of the desired and interfering sources. Relative ATF estimation procedures can be found in [20–23].

[3]The proposed beamformers can be generalized to multiple interferers similar to [16].

6. REFERENCES

[1] S. Doclo, S. Gannot, M. Moonen, and A. Spriet, "Acoustic beamforming for hearing aid applications," *Handbook on Array Processing and Sensor Networks*, pp. 269–302, 2008.

[2] J. Wouters, S. Doclo, R. Koning, and T. Francart, "Sound processing for better coding of monaural and binaural cues in auditory prostheses," *Proc. IEEE*, vol. 101, no. 9, pp. 1986–1997, Sep. 2013.

[3] R. Beutelmann and T. Brand, "Prediction of speech intelligibility in spatial noise and reverberation for normal-hearing and hearing-impaired listeners," *J. of the Acoustical Society of America*, vol. 120, no. 1, pp. 331–342, 2006.

[4] M. Lavandier and J. F. Culling, "Prediction of binaural speech intelligibility against noise in rooms," *J. of the Acoustical Society of America*, vol. 127, no. 1, pp. 387–399, 2010.

[5] T. Lotter and P. Vary, "Dual-channel speech enhancement by superdirective beamforming," *EURASIP Journal on Advances in Signal Proc.*, vol. 2006, pp. 175–175, Jan. 2006.

[6] R. Aichner, H. Buchner, M. Zourub, and W. Kellermann, "Multi-channel source separation preserving spatial information," in *Proc. IEEE International Conference on Acoustics speech and Signal Processing (ICASSP)*, Honolulu HI, USA, Apr. 2007, pp. 5–8.

[7] K. Reindl, Y. Zheng, and W. Kellermann, "Speech enhancement for binaural hearing aids based on blind source separation," in *4th IEEE Int. Symp. on Communications, Control and Signal Processing (ISCCSP)*, 2010, pp. 1–6.

[8] A.H. Kamkar-Parsi and M. Bouchard, "Instantaneous binaural target PSD estimation for hearing aid noise reduction in complex acoustic environments," *IEEE Trans. on Instrumentation and Measurement*, vol. 60, no. 4, pp. 1141–1154, 2011.

[9] Y. Suzuki, S. Tsukui, F. Asano, and R. Nishimura, "New design method of a binaural microphone array using multiple constraints," *IEICE Tran. on Fundamentals of Elect., Comm. and Comp. Sci.*, vol. 82, no. 4, pp. 588–596, 1999.

[10] T. Takatani, S. Ukai, T. Nishikawa, H. Saruwatari, and K. Shikano, "Evaluation of SIMO separation methods for blind decomposition of binaural mixed signals," in *Proc. Int. Workshop on Acoustic Signal Enhancement (IWAENC)*, Eindhoven, The Netherlands, Sep. 2005, pp. 233–236.

[11] B. Cornelis, S. Doclo, T. Van dan Bogaert, M. Moonen, and J. Wouters, "Theoretical analysis of binaural multimicrophone noise reduction techniques," *IEEE Trans. Audio, Speech and Language Proc.*, vol. 18, no. 2, pp. 342 –355, Feb. 2010.

[12] T. Klasen, T. Van den Bogaert, M. Moonen, and J. Wouters, "Binaural noise reduction algorithms for hearing aids that preserve interaural time delay cues," *IEEE Trans. Signal Proc.*, vol. 55, no. 4, pp. 1579–1585, Apr. 2007.

[13] E. Hadad, S. Gannot, and S. Doclo, "Binaural linearly constrained minimum variance beamformer for hearing aid applications," in *Proc. Int. Workshop on Acoustic Signal Enhancement (IWAENC)*, Aachen, Germany, Sep. 2012, pp. 117–120.

[14] D. Marquardt, V. Hohmann, and S. Doclo, "Interaural coherence preservation in multi-channel wiener filtering-based noise reduction for binaural hearing aids," *IEEE Trans. Audio, Speech and Language Proc.*, vol. 23, no. 12, pp. 2162–2176, Dec. 2015.

[15] E. Hadad, D. Marquardt, S. Doclo, and S. Gannot, "Theoretical analysis of binaural transfer function MVDR beamformers with interference cue preservation constraints," *IEEE Trans. Audio, Speech and Language Proc.*, vol. 23, no. 12, pp. 2449–2464, Dec. 2015.

[16] E. Hadad, S. Doclo, and S. Gannot, "The binaural LCMV beamformer and its performance analysis," *IEEE Trans. Audio, Speech and Language Proc.*, in press.

[17] D. Marquardt, E. Hadad, S. Gannot, and S. Doclo, "Theoretical analysis of linearly constrained multi-channel Wiener filtering algorithms for combined noise reduction and binaural cue preservation in binaural hearing aids," *IEEE Trans. Audio, Speech and Language Proc.*, vol. 23, no. 12, pp. 2384–2397, Dec. 2015.

[18] E. Hadad, D. Marquardt, S. Doclo, and S. Gannot, "Binaural multichannel Wiener filter with directional interference rejection," in *Proc. IEEE International Conference on Acoustics speech and Signal Processing (ICASSP)*, Brisbane, Australia, Apr. 2015, pp. 644–648.

[19] H. Kayser, S. Ewert, J. Anemüller, T. Rohdenburg, V. Hohmann, and B Kollmeier, "Database of multichannel In-Ear and Behind-The-Ear Head-Related and Binaural Room Impulse Responses," *EURASIP Journal on Advances in Signal Processing*, p. 10, 2009.

[20] S. Gannot, D. Burshtein, and E. Weinstein, "Signal enhancement using beamforming and nonstationarity with applications to speech," *Signal Processing*, vol. 49, no. 8, pp. 1614–1626, Aug. 2001.

[21] S. Markovich, S. Gannot, and I. Cohen, "Multichannel eigenspace beamforming in a reverberant environment with multiple interfering speech signals," *IEEE Trans. Audio, Speech and Language Proc.*, vol. 17, no. 6, pp. 1071–1086, Aug. 2009.

[22] S. Markovich-Golan and S. Gannot, "Performance analysis of the covariance subtraction method for relative transfer function estimation and comparison to the covariance whitening method," in *Proc. IEEE International Conference on Acoustics speech and Signal Processing (ICASSP)*, Brisbane, Australia, Apr. 2015, pp. 544–548.

[23] X. Li, L. Girin, R. Horaud, and S. Gannot, "Estimation of relative transfer function in the presence of stationary noise based on segmental power spectral density matrix subtraction," in *Proc. IEEE International Conference on Acoustics speech and Signal Processing (ICASSP)*, Brisbane, Australia, Apr. 2015, pp. 320–324.