2016 ICSEE International Conference on the Science of Electrical Engineering

Comparison of binaural multichannel Wiener filters with binaural cue preservation of the interferer

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*—An important objective of binaural speech enhancement algorithms is the preservation of the binaural cues of the sources, in addition to noise reduction. The binaural multichannel Wiener filter (MWF) preserves the binaural cues of the target but distorts the noise binaural cues. To optimally benefit from binaural unmasking and to preserve the spatial impression for the hearing aid user, two extensions of the binaural MWF have therefore been proposed, namely, the MWF with partial noise estimation (MWF-N) and MWF with interference reduction (MWF-IR). In this paper, the binaural cue preservation of these extensions is analyzed theoretically. Although both extensions are aimed at incorporating the binaural cue preservation of the interferer in the binaural MWF cost function, their properties are different. For the MWF-N, while the binaural cues of the target are preserved, there is a tradeoff between the noise reduction and the preservation of the binaural cues of the interferer component. For the MWF-IR, while the binaural cues of the interferer are preserved, those of the target may be slightly distorted. The theoretical results are validated by simulations using binaural hearing aids, demonstrating the capabilities of these beamformers in a reverberant environment.**

I. Introduction

Binaural hearing 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 as a result of binaural unmasking [3], [4]. For directional sources, preservation 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 aimed at preserving the binaural cues have been developed [1], [5]–[16]. In [1], the binaural MWF was presented. It was shown in [1], [8] that the binaural MWF preserves the binaural cues of the target but distorts the binaural cues of the noise, such that both the target and the noise are perceived as arriving from the target direction. Clearly, this is an undesirable phenomenon and in some scenarios (e.g., traffic) can even be dangerous to the hearing aid user.

In this paper, two extensions of the binaural MWF are discussed, which, in addition to minimizing the overall noise output power and speech distortion, are aimed at preserving the binaural cues of the interferer. The first extension, denoted as the MWF-N [8], [9], is aimed to preserve the binaural cues of the noise by mixing the output signals of the binaural MWF with a scaled version of the noisy reference microphone signals. Although in [8], [9], the aim of this beamformer was to preserve the binaural cues of the overall noise (i.e., the interferer plus the background noise), in this study, we focus on the binaural cue preservation of the directional interferer. The second extension of the binaural MWF, denoted as the MWF-IR, is aimed at preserving the binaural cues of the interferer by extending the MWF cost function with a hard constraint that controls the amount of interference reduction [13]–[15]. The binaural cue preservation of these extensions is analyzed theoretically and we compare their properties.

II. Problem Formulation

Consider a binaural hearing aid system consisting of two hearing devices with a total of M microphones and an acoustic scenario comprising one target speech source and one directional interferer in a noisy and reverberant environment. In the frequency domain, the M-dimensional stacked vector of the received microphone signals $\mathbf{y}(\omega)$ can be written as

$$\mathbf{y}(\omega) = \mathbf{x}(\omega) + \mathbf{u}(\omega) + \mathbf{n}(\omega) = \mathbf{x}(\omega) + \mathbf{v}(\omega), \qquad (1)$$

where $\mathbf{x}(\omega)$ is the target component, $\mathbf{u}(\omega)$ the interferer component, and $\mathbf{n}(\omega)$ the background noise component, e.g., diffuse noise. The vector $\mathbf{v}(\omega) = \mathbf{u}(\omega) + \mathbf{n}(\omega)$ is defined as the overall noise component, i.e., the interferer plus background noise component. For brevity, the frequency variable ω is henceforth omitted.

The target and interferer components can be written as $\mathbf{x} = S_{\mathrm{x}}\mathbf{a}$ and $\mathbf{u} = S_{\mathrm{u}}\mathbf{b}$, where S_{x} and S_{u} denote the target and interferer signals and $\mathbf{a}$ and $\mathbf{b}$ denote the acoustic transfer functions (ATFs) relating the microphones and the target and the interferer, respectively.

978-1-5090-2152-9/16/$31.00

Assuming statistical independence between all components in (1), the spatial correlation matrix of the microphone signals $\mathbf{R}_y$ can be written as

$$\mathbf{R}_y = \mathcal{E}\left\{\mathbf{y}\mathbf{y}^H\right\} = \mathbf{R}_x + \mathbf{R}_u + \mathbf{R}_n = \mathbf{R}_x + \mathbf{R}_v, \tag{2}$$

where $\mathbf{R}_x = \mathcal{E}\left\{\mathbf{x}\mathbf{x}^H\right\}$, $\mathbf{R}_u = \mathcal{E}\left\{\mathbf{u}\mathbf{u}^H\right\}$, and $\mathbf{R}_n = \mathcal{E}\left\{\mathbf{n}\mathbf{n}^H\right\}$ denote the target, interferer, and background noise correlation matrices, respectively, and $\mathcal{E}\{\cdot\}$ is the expectation operator. The target and interferer correlation matrices are rank-1 matrices, i.e.,

$$\mathbf{R}_x = P_s\mathbf{a}\mathbf{a}^H, \mathbf{R}_u = P_u\mathbf{b}\mathbf{b}^H, \tag{3}$$

where $P_s = \mathcal{E}\left\{|S_\mathrm{x}|^2\right\}$ and $P_u = \mathcal{E}\left\{|S_\mathrm{u}|^2\right\}$ denote the power spectral density (PSD) of the target and the interferer, respectively.

The reference microphone signals at the left and right hearing devices, selected, e.g., as the microphones closest to the ears, are given by $y_L = \mathbf{e}_L^H\mathbf{y}$ and $y_R = \mathbf{e}_R^H\mathbf{y}$, respectively, where $\mathbf{e}_L$ and $\mathbf{e}_R$ are M-dimensional indicator input vectors with one element equal to 1 and all other elements equal to zero. From (1), the reference microphone signals can be written as

$$y_L = S_\mathrm{x}a_L + S_\mathrm{u}b_L + n_L,\; y_R = S_\mathrm{x}a_R + S_\mathrm{u}b_R + n_R. \tag{4}$$

The output signals at the left and the right hearing devices are given by $z_L = \mathbf{w}_L^H\mathbf{y}$ and $z_R = \mathbf{w}_R^H\mathbf{y}$, respectively, where $\mathbf{w}_L$ and $\mathbf{w}_R$ denote M-dimensional complex-valued weight filter vectors.

The RTF of the target component and the interferer component in the reference microphone of the left and the right hearing aid is defined as the ratio of the respective ATFs, i.e.,

$$\mathrm{RTF}_x^{\mathrm{in}} = \frac{a_L}{a_R}, \qquad \mathrm{RTF}_u^{\mathrm{in}} = \frac{b_L}{b_R}, \tag{5}$$

and the output RTF is defined as the ratio of the filtered target component and the filtered interferer component in the left and the right hearing aid, respectively, i.e.,

$$\mathrm{RTF}_x^{\mathrm{out}} = \frac{\mathbf{w}_L^H\mathbf{a}}{\mathbf{w}_R^H\mathbf{a}}, \qquad \mathrm{RTF}_u^{\mathrm{out}} = \frac{\mathbf{w}_L^H\mathbf{b}}{\mathbf{w}_R^H\mathbf{b}}. \tag{6}$$

The binaural ILD and ITD cues can be calculated from the RTF as

$$\mathrm{ILD} = 10\log_{10}|\mathrm{RTF}|^2, \quad \mathrm{ITD} = \frac{\angle\mathrm{RTF}}{\omega}, \tag{7}$$

with $\angle$ denoting the phase.

The binaural input signal-to-interference-and-noise ratio (SINR) is defined as the ratio of the average power of the target and overall noise in the reference microphones, i.e.,

$$\mathrm{SINR}^{\mathrm{in}} = \frac{\mathbf{e}_L^H\mathbf{R}_x\mathbf{e}_L + \mathbf{e}_R^H\mathbf{R}_x\mathbf{e}_R}{\mathbf{e}_L^H\mathbf{R}_v\mathbf{e}_L + \mathbf{e}_R^H\mathbf{R}_v\mathbf{e}_R}. \tag{8}$$

The binaural output SINR is defined as the ratio of the average power of the target and overall noise in the left and the right hearing aid, i.e.,

$$\mathrm{SINR}^{\mathrm{out}} = \frac{\mathbf{w}_L^H\mathbf{R}_x\mathbf{w}_L + \mathbf{w}_R^H\mathbf{R}_x\mathbf{w}_R}{\mathbf{w}_L^H\mathbf{R}_v\mathbf{w}_L + \mathbf{w}_R^H\mathbf{R}_v\mathbf{w}_R}. \tag{9}$$

III. Binaural Noise Reduction Algorithms

In Section III-A, we briefly review the binaural MWF. Then, in Section III-B and in Section III-C the MWF-N and MWF-IR are described.

A. Binaural multi-channel Wiener filter (MWF)

The well-known binaural MWF produces a minimum mean square error (MSE) estimate of the target component at both reference microphones [1]. The MSE cost functions for the filter $\mathbf{w}_L$, estimating the target component x_L at the left hearing aid, and for the filter $\mathbf{w}_R$, estimating the target component x_R at the right hearing aid, are given by

$$\begin{aligned} J_{\mathrm{MWF}}(\mathbf{w}_L) &= \mathcal{E}\{\|[x_L - \mathbf{w}_L^H\mathbf{x}]\|^2 + \mu\|\mathbf{w}_L^H\mathbf{v}\|^2\}, \\ J_{\mathrm{MWF}}(\mathbf{w}_R) &= \mathcal{E}\{\|[x_R - \mathbf{w}_R^H\mathbf{x}]\|^2 + \mu\|\mathbf{w}_R^H\mathbf{v}\|^2\}, \end{aligned} \tag{10}$$

where μ provides a weighting tradeoff between noise reduction and speech distortion. The filter vectors minimizing (10) are equal to [1]

$$\mathbf{w}_{L,\mathrm{MWF}} = P_s a_L^*\tilde{\mathbf{R}}_y^{-1}\mathbf{a}, \quad \mathbf{w}_{R,\mathrm{MWF}} = P_s a_R^*\tilde{\mathbf{R}}_y^{-1}\mathbf{a}, \tag{11}$$

with $\tilde{\mathbf{R}}_y = \mathbf{R}_x + \mu\mathbf{R}_v$ defined as a speech-distortion-weighted correlation matrix. Applying the Woodbury identity to $\tilde{\mathbf{R}}_y^{-1}$ and using (11), the filter vectors can be written as

$$\mathbf{w}_{L,\mathrm{MWF}} = \frac{\rho}{\mu+\rho}\frac{\mathbf{R}_v^{-1}\mathbf{a}}{\sigma_a}a_L^*,\; \mathbf{w}_{R,\mathrm{MWF}} = \frac{\rho}{\mu+\rho}\frac{\mathbf{R}_v^{-1}\mathbf{a}}{\sigma_a}a_R^*, \tag{12}$$

with $\sigma_a = \mathbf{a}^H\mathbf{R}_v^{-1}\mathbf{a}$ and

$$\rho = P_s\mathbf{a}^H\mathbf{R}_v^{-1}\mathbf{a} = P_s\sigma_a. \tag{13}$$

Equation (12) implies that the filter vectors of the binaural MWF $\mathbf{w}_{L,\mathrm{MWF}}$ and $\mathbf{w}_{R,\mathrm{MWF}}$ are parallel and the relation $\mathbf{w}_{L,\mathrm{MWF}} = (\mathrm{RTF}_x^{\mathrm{in}})^*\mathbf{w}_{R,\mathrm{MWF}}$ holds. Hence, the RTF of the target at the output of the binaural MWF is equal to the input RTF, i.e.,

$$\mathrm{RTF}_x^{\mathrm{out}} = \frac{\mathbf{w}_L^H\mathbf{a}}{\mathbf{w}_L^H\mathbf{a}} = \frac{a_L}{a_R} = \mathrm{RTF}_x^{\mathrm{in}}. \tag{14}$$

However, this also implies that *all* sound sources are perceived as arriving from the target direction, which is obviously not a desired phenomenon. Substituting (12) in (9), the output SINR of the binaural MWF is equal to [1], [8]

$$\mathrm{SINR}^{\mathrm{out}} = \rho. \tag{15}$$

B. Binaural MWF with partial noise estimation (MWF-N)

An extension of the binaural MWF that, in addition to preserving the binaural cues of the target component, is also aimed to partially preserve the binaural cues of the noise component has been proposed in [9], denoted as MWF-N. The objective of the MWF-N is to produce a minimum MSE estimate of the sum of the target component and a scaled version of the overall noise component at the reference microphones, i.e.,

$$J_{\text{MWF-N}}(\mathbf{w}_L) = \mathcal{E}\{\|[x_L - \mathbf{w}_L^H \mathbf{x}]\|^2 + \mu\|\eta_{\text{N}} v_L - \mathbf{w}_L^H \mathbf{v}\|^2\},$$
$$J_{\text{MWF-N}}(\mathbf{w}_R) = \mathcal{E}\{\|[x_R - \mathbf{w}_R^H \mathbf{x}]\|^2 + \mu\|\eta_{\text{N}} v_R - \mathbf{w}_R^H \mathbf{v}\|^2\}, \quad (16)$$

where $0 \leq \eta_{\text{N}} \leq 1$ denotes the overall noise parameter, which provides a tradeoff between noise reduction and the preservation of the binaural cues of the noise component. The filter vectors minimizing (16) are equal to [8], [9]

$$\begin{aligned} \mathbf{w}_{L,\text{MWF-N}} &= (1-\eta_{\text{N}}) \frac{\rho}{\mu+\rho} \frac{\mathbf{R}_v^{-1}\mathbf{a}}{\sigma_a} a_L^* + \eta_{\text{N}} \mathbf{e}_L \\ &= (1-\eta_{\text{N}}) \, \mathbf{w}_{\text{MWF},L} + \eta_{\text{N}} \mathbf{e}_L, \\ \mathbf{w}_{R,\text{MWF-N}} &= (1-\eta_{\text{N}}) \frac{\rho}{\mu+\rho} \frac{\mathbf{R}_v^{-1}\mathbf{a}}{\sigma_a} a_R^* + \eta_{\text{N}} \mathbf{e}_R \\ &= (1-\eta_{\text{N}}) \, \mathbf{w}_{\text{MWF},R} + \eta_{\text{N}} \mathbf{e}_R. \end{aligned} \quad (17)$$

Hence, the output signals of the MWF-N are equal to the sum of the output signals of the binaural MWF (weighted with $(1-\eta_{\text{N}})$) and the noisy reference microphone signals (weighted with η_{N}). As shown in [8], for the MWF-N the output RTF of the target component is equal to the input RTF for all tradeoff parameters η_{N}, i.e.,

$$\text{RTF}_x^{\text{out}} = \frac{(1-\eta_{\text{N}}) \frac{\rho}{(\mu+\rho)} a_L + \eta_{\text{N}} a_L}{(1-\eta_{\text{N}}) \frac{\rho}{(\mu+\rho)} a_R + \eta_{\text{N}} a_R} = \frac{a_L}{a_R} = \text{RTF}_x^{\text{in}}. \quad (18)$$

Substituting (17) in (6), it can be shown that the output RTF of the interferer is equal to

$$\begin{aligned} \text{RTF}_u^{\text{out}} = & \frac{(1-\eta_{\text{N}}) \frac{\rho_{ab}}{\mu+\rho}}{(1-\eta_{\text{N}}) \frac{\rho_{ab}}{\mu+\rho} + \frac{b_R}{a_R}\eta_{\text{N}}} \text{RTF}_x^{\text{in}} \\ & + \frac{\frac{b_R}{a_R}\eta_{\text{N}}}{(1-\eta_{\text{N}}) \frac{\rho_{ab}}{\mu+\rho} + \frac{b_R}{a_R}\eta_{\text{N}}} \text{RTF}_u^{\text{in}}, \end{aligned} \quad (19)$$

where $\rho_{ab} = P_s \sigma_{ab}$ and $\sigma_{ab} = \mathbf{a}^H \mathbf{R}_v^{-1} \mathbf{b}$.

Equation (19) shows that output RTF of the interferer is a weighted sum of the input RTF of the target and the input RTF of the interferer. If $\eta_{\text{N}} = 0$, the output RTF of the interferer is equal to the input RTF of the target while on the other hand, if $\eta_{\text{N}} = 1$, the output RTF of the interferer is equal to the input RTF of the interferer.

Substituting (17) in (9), the output SINR for the MWF-N is equal to[1]

$$\text{SINR}_{\text{MWF-N}}^{\text{out}} = \frac{(\mu\eta_{\text{N}} + \rho)^2 \, \text{SINR}^{\text{in}}}{(1-\eta_{\text{N}})(2\mu\eta_{\text{N}} + \rho + \rho\eta_{\text{N}}) \, \text{SINR}^{\text{in}} + (\mu+\rho)^2 \eta_{\text{N}}^2}, \quad (20)$$

where

$$\text{SINR}^{\text{ in}} = \frac{P_s \left(|a_L|^2 + |a_R|^2\right)}{\mathbf{e}_L^H \mathbf{R}_v \mathbf{e}_L + \mathbf{e}_R^H \mathbf{R}_v \mathbf{e}_R}. \quad (21)$$

[1] Similar derivation for the output SINRs for the left and the right hearing aids are obtained in [8].

C. Binaural MWF with interference reduction (MWF-IR)

The second extension of the binaural MWF, denoted as the MWF-IR, is aimed to better control the suppression and binaural cue preservation of the interferer by extending the MWF cost function in (10) with a hard interference reduction (IR) constraint that controls the amount of interference reduction [15], i.e.,

$$\begin{aligned} \min_{\mathbf{w}_L} J_{\text{MWF}}(\mathbf{w}_L) \quad & \text{subject to} \quad \mathbf{w}_L^H \mathbf{b} = \eta_{\text{IR}} b_L, \\ \min_{\mathbf{w}_R} J_{\text{MWF}}(\mathbf{w}_R) \quad & \text{subject to} \quad \mathbf{w}_R^H \mathbf{b} = \eta_{\text{IR}} b_R, \end{aligned} \quad (22)$$

where $0 \leq \eta_{\text{IR}} \ll 1$ denotes the interference parameter, which provides a tradeoff between interference reduction and binaural cue preservation. The filter vectors minimizing (22) are equal to

$$\begin{aligned} \mathbf{w}_{L,\text{MWF-IR}} = \quad & P_s a_L^* \left(\tilde{\mathbf{R}}_y^{-1} \mathbf{a} - \frac{\lambda_{ab}^* - \frac{\eta_{\text{IR}} b_L^*}{P_s a_L^*}}{\lambda_b} \tilde{\mathbf{R}}_y^{-1} \mathbf{b} \right), \\ \mathbf{w}_{R,\text{MWF-IR}} = \quad & P_s a_R^* \left(\tilde{\mathbf{R}}_y^{-1} \mathbf{a} - \frac{\lambda_{ab}^* - \frac{\eta_{\text{IR}} b_R^*}{P_s a_R^*}}{\lambda_b} \tilde{\mathbf{R}}_y^{-1} \mathbf{b} \right). \end{aligned} \quad (23)$$

with $\lambda_{ab} = \mathbf{a}^H \tilde{\mathbf{R}}_y^{-1} \mathbf{b}$ and $\lambda_b = \mathbf{b}^H \tilde{\mathbf{R}}_y^{-1} \mathbf{b}$. For the special case of $\eta_{\text{IR}} = 0$ (i.e., denoted as MWF-IR$_0$), the filter vectors are equal to [13], [14]

$$\begin{aligned} \mathbf{w}_{L,\text{MWF-IR}_0} &= P_s a_L^* \left[\tilde{\mathbf{R}}_y^{-1} \mathbf{a} - \frac{\lambda_{ab}^*}{\lambda_b} \tilde{\mathbf{R}}_y^{-1} \mathbf{b} \right], \\ \mathbf{w}_{R,\text{MWF-IR}_0} &= P_s a_R^* \left[\tilde{\mathbf{R}}_y^{-1} \mathbf{a} - \frac{\lambda_{ab}^*}{\lambda_b} \tilde{\mathbf{R}}_y^{-1} \mathbf{b} \right], \end{aligned} \quad (24)$$

such that the filter vectors $\mathbf{w}_{L,\text{MWF-IR}_0}$ and $\mathbf{w}_{R,\text{MWF-IR}_0}$ are parallel, namely, $\mathbf{w}_{L,\text{MWF-IR}_0} = (\text{RTF}_x^{\text{in}})^* \mathbf{w}_{R,\text{MWF-IR}_0}$, and the RTF of the target at the output of the MWF-IR$_0$ is equal to the input RTF.

For η_{IR} greater than zero, the filter vectors of the MWF-IR in (23) can be decomposed as a combination of two beamformers, i.e.,

$$\mathbf{w}_{L,\text{MWF-IR}} = \mathbf{w}_{L,\text{MWF-IR}_0} + \eta_{\text{IR}} \mathbf{w}_{L,\text{MVDR-U}}, \quad (25)$$
$$\mathbf{w}_{R,\text{MWF-IR}} = \mathbf{w}_{R,\text{MWF-IR}_0} + \eta_{\text{IR}} \mathbf{w}_{R,\text{MVDR-U}}, \quad (26)$$

where on the one hand, $\mathbf{w}_{L,\text{MWF-IR}_0}$ and $\mathbf{w}_{R,\text{MWF-IR}_0}$ are the filter vectors of the MWF-IR$_0$, and on the other hand, $\mathbf{w}_{L,\text{MVDR-U}}$ and $\mathbf{w}_{R,\text{MVDR-U}}$ are the filter vectors of a minimum variance distortionless response (MVDR) steered towards the interferer, which is denoted as the MVDR-U beamformer, i.e.,

$$\mathbf{w}_{L,\text{MVDR-U}} = \frac{\tilde{\mathbf{R}}_y^{-1} \mathbf{b}}{\lambda_b} b_L^*, \quad \mathbf{w}_{R,\text{MVDR-U}} = \frac{\tilde{\mathbf{R}}_y^{-1} \mathbf{b}}{\lambda_b} b_R^*, \quad (27)$$

such that the filter vectors $\mathbf{w}_{L,\text{MVDR-U}}$ and $\mathbf{w}_{R,\text{MVDR-U}}$ are parallel, namely, $\mathbf{w}_{L,\text{MVDR-U}} = (\text{RTF}_u^{\text{in}})^* \mathbf{w}_{R,\text{MVDR-U}}$, and the RTF of the interferer at the output of the MVDR-U is equal to the input RTF. The MWF-IR$_0$ is steering a null toward the interferer. However, when RTF estimation errors occur, the interferer is not entirely suppressed and the residual interference will also be perceived as arriving from the target direction. To resolve this issue, it is required to set the scaling parameter η_{IR} to a value greater than zero.

In general, the RTF of the target at the output of the MWF-IR is equal to [15]

$$\text{RTF}_x^{\text{out}} = \frac{\mathbf{w}_L^H \mathbf{a}}{\mathbf{w}_R^H \mathbf{a}} = \frac{P_s \lambda_a \left(a_L (1 - \Lambda) + \eta_{\text{IR}} b_L \frac{\Lambda}{P_s \lambda_{ab}} \right)}{P_s \lambda_a \left(a_R (1 - \Lambda) + \eta_{\text{IR}} b_R \frac{\Lambda}{P_s \lambda_{ab}} \right)}, \quad (28)$$

with $\lambda_a = \mathbf{a}^H \tilde{\mathbf{R}}_y^{-1} \mathbf{a}$ and $0 \leq \Lambda = \frac{|\lambda_{ab}|^2}{\lambda_a \lambda_b} \leq 1$. For η_{IR} greater than zero, since the MWF-IR satisfies the constraints in (22) for the interferer, the output RTF of the interferer is equal to the input RTF, i.e.,

$$\text{RTF}_u^{\text{out}} = \frac{\mathbf{w}_L^H \mathbf{b}}{\mathbf{w}_R^H \mathbf{b}} = \frac{\eta_{\text{IR}} b_L}{\eta_{\text{IR}} b_R} = \text{RTF}_u^{\text{in}}. \quad (29)$$

Hence, the MWF-IR preserves the binaural cues of the interferer. For $\eta_{\text{IR}} = 0$, the output RTF of the interferer is undefined.

It was shown in [13] that the output SINR for the MWF-IR$_0$ is equal to

$$\text{SINR}_{\text{MWF-IR}}^{\text{out}} = \varrho_{\text{IR}} \text{SINR}_{\text{MWF}}^{\text{out}}, \quad (30)$$

with

$$\varrho_{\text{IR}} = \frac{1 + \Lambda^2 - 2\Lambda}{1 + \nu \Lambda^2 - 2\Lambda}, \quad \nu = \frac{(\mu + \rho)^2}{\mu^2 \Sigma} - \frac{\rho^2 + 2\mu\rho}{\mu^2}, \quad (31)$$

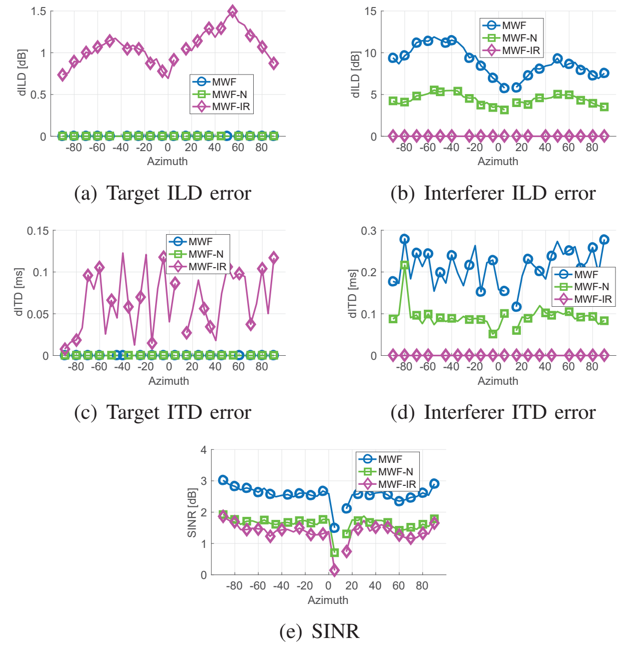
with $0 \leq \Sigma = \frac{|\sigma_{ab}|^2}{\sigma_a \sigma_b} \leq 1$, and $\sigma_b = \mathbf{b}^H \mathbf{R}_v^{-1} \mathbf{b}$. Since ϱ_{IR} is always smaller than or equal to one [13], the output SINR of the MWF-IR$_0$ is always smaller than or equal to the output SINR of the binaural MWF. Hence,

$$\text{SINR}_{\text{MWF}}^{\text{out}} \geq \text{SINR}_{\text{MWF-IR}_0}^{\text{out}}. \quad (32)$$

IV. Experimental Study

In this section, we present simulation results for comparing the performance of the binaural MWF, MWF-N, and the MWF-IR using Behind-The-Ear Impulse Responses (BTE-IR) measured in a reverberant office environment, as described in [17]. To verify the theoretical analysis, we use measured ATFs ($\mathbf{a}$ and $\mathbf{b}$) and artificial signals, hence circumventing estimation error issues. All experiments were conducted using $M = 3$ microphones, i.e., two microphones on the left hearing aid and one microphone on the right hearing aid, at a sampling frequency of 16 kHz. The acoustic scenario comprised one target at $\theta_x = 10°$ and 1 m from the artificial head, one interferer at different angles[2] (also 1 m from the artificial head), and diffuse background noise. The angle $\theta = 0^o$ corresponds to the frontal direction and $\theta = 90^o$ corresponds to the right side of the head. The reverberation time is approximately 300 ms. The PSDs of the target and the interferer P_s and P_u were calculated from two different speech signals (Welch method using an FFT size of 512 and a Hann window). For the background noise a cylindrically isotropic noise field was assumed, where

[2]Note that the interferer at $10°$ was not evaluated.



(a) Target ILD error (b) Interferer ILD error

(c) Target ITD error (d) Interferer ITD error

(e) SINR

Fig. 1. Performance measures for the binaural MWF, MWF-N, and MWF-IR for a target at $10°$ and different angles of the interferer. The global input SNR and SIR are equal to 0 dB.

the spatial coherence matrix was calculated using anechoic ATFs of the same database; the PSD of the background noise was equal to the PSD of speech-shaped noise. The tradeoff parameters η_{N} and η_{IR} were both set to 0.1, and μ was set to 1 for all algorithms.

For the objective validation, we calculated the global performance measures by averaging the logarithmic values of the output SINR in (9) over all frequencies. In order to evaluate the binaural cue preservation performance, we calculated the ILD and ITD error, averaged over all frequencies for the target and the interferer, i.e., [8]

$$\text{dILD} = \frac{1}{K} \sum_{k=1}^{K} \left| \text{ILD}^{\text{out}}(\omega_k) - \text{ILD}^{\text{in}}(\omega_k) \right| \quad (33)$$

$$\text{dITD} = \frac{1}{K} \sum_{k=1}^{K} \left| \text{ITD}^{\text{out}}(\omega_k) - \text{ITD}^{\text{in}}(\omega_k) \right|, \quad (34)$$

with ILD and ITD as defined in (7), ω_k denoting the k-th frequency, and K the total number of frequencies.

The ILD and ITD errors of the target and the interferer components are depicted in Fig. 1(a)–Fig. 1(d). On the one hand, for the target, the MWF-IR introduces a small ILD error (up to 1.5 dB) and a very small ITD error (up to 0.13 ms), depending on the position of the interferer, while the binaural MWF and the MWF-N perfectly preserve the ILD and the ITD of the target. On the other hand, for the interferer, the binaural MWF introduces a large ILD error (up to 13 dB). The ITD error of the binaural MWF varies around 0.2 ms for all interferer positions. The MWF-N introduces a lower ILD error (around 5 dB) and an ITD error that varies around 0.1 ms for all interferer positions. The MWF-IR perfectly preserves the binaural cues of the interferer.

The global output SINR is depicted in Fig. 1(e). It is observed that the global output SINR is lower for both the MWF-N and the MWF-IR as compared to the global output SINR for the binaural MWF.

V. Discussion and conclusions

In this paper, two extensions of the binaural MWF algorithm, namely, the MWF-N and the MWF-IR, were explored. Both the MWF-N and the MWF-IR incorporate binaural cue preservation of the interferer in the binaural MWF cost function. However, their properties are different.

In the MWF-N, the modification of the binaural MWF cost function is for the overall noise signal. The MWF-N filter vectors can be decomposed to the weighted sum of the binaural MWF filter vectors (multiplied by $(1-\eta_{\mathrm{N}})$) and the noisy reference microphone indicator input vectors (multiplied by η_{N}). The binaural cues of the target are preserved. However, the parameter η_{N} trades-off the noise reduction and the interferer binaural cue preservation.

In the MWF-IR, a hard constraint is added to the binaural MWF cost function regarding the ATF of the interferer rather than the signal itself. The MWF-IR filter vectors can be decomposed to the sum of the MWF-IR$_0$ filter vectors and MVDR-U filter vectors directed toward the interferer direction (multiplied by η_{IR}). The MWF-IR$_0$ filter vectors direct a null toward the interferer direction, and the binaural cues of the target are imposed at the output signal, such that the binaural cues of the target are preserved. In order to mask residual interference due to estimation errors and to control the amount of interference reduction, η_{IR} should be set to a value greater than zero. Since the MVDR-U preserves the binaural cues of the interferer, the binaural cues of the interferer at the output of the MWF-IR are preserved. The parameter η_{IR} trades-off the interference reduction, estimation error, and interferer binaural cue preservation. However, since the the binaural cues of the target at the output of the MVDR-U are equal to the binaural cues of the interferer, for η_{IR} greater than zero, the binaural cues of the target at the output of the MWF-IR are slightly distorted.

To implement the MWF-N, it is necessary to estimate the spatial correlation matrices of the noisy and the noise-only microphone signals (similar to the implementation requirement for the binaural MWF). To implement the MWF-IR, in addition, it is necessary to estimate the RTF vectors of the interferer (ATF vectors normalized by the ATFs of the reference microphones).

In general, there is a degradation in the performance in terms of SINR for both algorithms as compared to the binaural MWF, as a result of the additional requirement of the binaural cue preservation of the interferer.

Acknowledgment

This work was supported by a joint Lower Saxony-Israeli Project, funded by the State of Lower Saxony in Germany.

References

[1] S. Doclo, S. Gannot, M. Moonen, and A. Spriet, "Acoustic beamforming for hearing aid applications," in *Handbook on Array Processing and Sensor Networks*, S. Haykin and K. J. R. Liu, Eds. Wiley, 2010, pp. 269–302.

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

[3] R. Beutelmann and T. Brand, "Prediction of speech intelligibility in spatial noise and reverberation for normal-hearing and hearing-impaired listeners," *Journal 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," *Journal of the Acoustical Society of America*, vol. 127, pp. 387–399, 2010.

[5] T. Lotter and P. Vary, "Dual-channel speech enhancement by superdirective beamforming," *EURASIP Journal on Applied Signal Processing*, vol. 2006, no. 1, pp. 175–175, Jan 2006.

[6] A. H. Kamkar-Parsi and M. Bouchard, "Instantaneous Binaural Target PSD Estimation for Hearing Aid Noise Reduction in Complex Acoustic Environments," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 4, pp. 1141–1154, Apr. 2011.

[7] Y. Suzuki, S. Tsukui, F. Asano, R. Nishimura, and T. Sone, "New design method of a binaural microphone array using multiple constraints," *IEICE Trans. Fundamentals*, vol. E82-A, no. 4, pp. 588–596, apr 1999.

[8] B. Cornelis, S. Doclo, T. van den Bogaert, J. Wouters, and M. Moonen, "Theoretical analysis of binaural multi-microphone noise reduction techniques," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 2, pp. 342–355, Feb 2010.

[9] 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 Transactions on Signal Processing*, vol. 55, no. 4, pp. 1579–1585, Apr 2007.

[10] E. Hadad, S. Gannot, and S. Doclo, "Binaural linearly constrained minimum variance beamformer for hearing aid applications," in *Proceedings of International Workshop on Acoustic Signal Enhancement*, Aachen, Germany, Sep. 2012, pp. 117–120.

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

[12] E. Hadad, S. Doclo, and S. Gannot, "The binaural LCMV beamformer and its performance analysis," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 3, pp. 543–558, Mar 2016.

[13] 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 Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 12, pp. 2384–2397, Dec 2015.

[14] E. Hadad, D. Marquardt, S. Doclo, and S. Gannot, "Binaural multichannel Wiener filter with directional interference rejection," in *Proceedings of IEEE International Conference on Audio, Speech and Signal Processing*, Brisbane, Australia, Apr. 2015, pp. 644–648.

[15] E. Hadad, D. Marquardt, S. Doclo, and S. Gannot, "Extensions of the binaural MWF with interference reduction preserving the binaural cues of the interfering source," in *Proceedings of IEEE International Conference on Audio, Speech and Signal Processing*, Shanghai, China, March 2016.

[16] W. C. Liao, Z. Q. Luo, I. Merks, and T. Zhang, "An effective low complexity binaural beamforming algorithm for hearing aids," in *Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, Oct 2015, pp. 1–5.

[17] H. Kayser, S. Ewert, J. Annemüller, 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*, vol. 2009, p. 10 pages, 2009.