Theoretical Analysis of Binaural Transfer Function MVDR Beamformers with Interference Cue Preservation Constraints

Elior Hadad, *Student Member, IEEE*, Daniel Marquardt, *Student Member, IEEE*, Simon Doclo, *Senior Member, IEEE*, and Sharon Gannot, *Senior Member, IEEE*

***Abstract*—The objective of binaural noise reduction algorithms is not only to selectively extract the desired speaker and to suppress interfering sources (e.g., competing speakers) and ambient background noise, but also to preserve the auditory impression of the complete acoustic scene. For directional sources this can be achieved by preserving the relative transfer function (RTF) which is defined as the ratio of the acoustical transfer functions relating the source and the two ears and corresponds to the binaural cues. In this paper, we theoretically analyze the performance of three algorithms that are based on the binaural minimum variance distortionless response (BMVDR) beamformer, and hence, process the desired source without distortion. The BMVDR beamformer preserves the binaural cues of the desired source but distorts the binaural cues of the interfering source. By adding an interference reduction (IR) constraint, the recently proposed BMVDR-IR beamformer is able to preserve the binaural cues of both the desired source and the interfering source. We further propose a novel algorithm for preserving the binaural cues of both the desired source and the interfering source by adding a constraint preserving the RTF of the interfering source, which will be referred to as the BMVDR-RTF beamformer. We analytically evaluate the performance in terms of binaural signal-to-interference-and-noise ratio (SINR), signal-to-interference ratio (SIR), and signal-to-noise ratio (SNR) of the three considered beamformers. It can be shown that the BMVDR-RTF beamformer outperforms the BMVDR-IR beamformer in terms of SINR and outperforms the BMVDR beamformer in terms of SIR. Among all beamformers which are distortionless with respect to the desired source and preserve the binaural cues of the interfering source, the newly proposed BMVDR-RTF beamformer is optimal in terms of SINR. Simulations using acoustic transfer functions measured on a binaural hearing aid validate our theoretical results.**

***Index Terms*—Binaural cues, hearing aids, linearly constrained minimum variance (LCMV) beamformer, minimum variance distortionless response (MVDR) beamformer, noise reduction, relative transfer function.**

Manuscript received May 04, 2015; revised August 21, 2015; accepted September 18, 2015. Date of publication October 02, 2015; date of current version October 14, 2015. This work was supported in part by a Grant from the German-Israeli Foundation for Scientific Research and Development, a joint Lower Saxony-Israeli Project financially supported by the State of Lower Saxony, Germany, and the Cluster of Excellence 1077 "Hearing4All," funded by the German Research Foundation (DFG). The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Thushara Abhayapala.

E. Hadad and S. Gannot are with the Faculty of Engineering, Bar-Ilan University, Ramat-Gan, 5290002, Israel (e-mail: elior.hadad@biu.ac.il; sharon.gannot@biu.ac.il).

D. Marquardt and S. Doclo are with the Department of Medical Physics and Acoustics and the Cluster of Excellence Hearing4All, University of Oldenburg, 26111 Oldenburg, Germany (e-mail: daniel.marquardt@uni-oldenburg.de; simon.doclo@uni-oldenburg.de).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TASLP.2015.2486381

I. Introduction

In a binaural system, the hearing-impaired person is fitted with two hearing aids, where the microphone signals of both hearing aids are shared (e.g. through a wireless link [1]), principally leading to an improved performance compared to a bilateral system, where both hearing aids operate independently of each other [2], [3]. The objective of a binaural noise reduction algorithm is not only to selectively extract the desired speaker and to suppress interfering sources and ambient background noise, but also to preserve the auditory impression for the hearing aid user. This can be achieved by preserving the binaural cues of the sound sources in the acoustic scene namely the interaural level difference (ILD) and the interaural time difference (ITD). For directional sources, preserving the binaural cues can be achieved by preserving the relative transfer function (RTF), which is defined as the ratio of the acoustical transfer functions relating the source and the two ears. In addition to monaural cues, these binaural cues play a major role in spatial awareness, i.e. for source localization and for determining the spaciousness of auditory objects [4], [5], and are very important for speech intelligibility due to so-called binaural unmasking [6]–[10].

There are many binaural noise reduction algorithms that aim to preserve the binaural cues of the sound sources in the acoustic scene, which can be divided into two main paradigms. The first paradigm utilizes two microphone signals, i.e. one on each hearing aid, where an identical (real-valued) spectral gain is applied to both microphone signals (e.g. [11]–[15]), hence, preserving the binaural cues of all sources. Within this paradigm several approaches have been developed, e.g. based on computational auditory scene analysis [11], [12] or super-directive beamforming [13]. These approaches, however, typically suffer from single-channel noise reduction artifacts, especially at low signal-to-noise ratios (SNRs). In the second paradigm all microphone signals from both hearing aids are centrally processed with two (different) complex-valued filters (e.g. [16]–[28]). Using this paradigm, a large noise reduction performance can be achieved, but the binaural cues of the residual interference and noise components are not guaranteed to be preserved. In [23], a binaural noise reduction algorithm based on multichannel Wiener filtering (MWF) has been introduced. It has been shown that this algorithm preserves the binaural cues of the desired source component but distorts the binaural cues of the noise component, since both desired source and noise components are perceived as coming from the desired source direction. Clearly, this is an undesired result, and in some situations (e.g. traffic) even dangerous for the hearing aid user. To optimally benefit from binaural unmasking and to preserve the spatial impression for the hearing aid user, several extensions of the binaural MWF have been proposed, which aim to also preserve the binaural cues of the residual noise component by including cue preservation terms in the binaural MWF cost function [20], [25], [27]. However, for all proposed binaural MWF extensions, a trade-off between noise reduction performance and binaural cue preservation exists and by design, MWF-based algorithms suffer from some distortion of the desired source. For the implementation of the MWF-based algorithms estimates of the spatial correlation matrices of the noisy and the noise-only microphone signals are required. The noise-only spatial correlation matrices can, e.g., be estimated during speech-absence segments, requiring voice activity detector (VAD). The noisy spatial correlation matrices can be estimated during speech-plus-noise segments.

If it is desired that the interference and background noise power is minimized while processing the desired source without distortion, the minimum variance distortionless response (MVDR) beamformer can be applied [23], [29], [30]. The MVDR beamformer can be extended into a binaural version producing two output signals [24]. In order to preserve the binaural cues of the desired source, it is sufficient to preserve the RTF of the desired source between the two reference microphone signals, i.e. one on each hearing aid [24]. These RTFs can be estimated, e.g. by exploiting the non-stationarity of speech signals [30]–[32], or using an eigenvalue decomposition (EVD) [33]. However, an important drawback of the BMVDR beamformer is the fact that the RTF of the (directional) interfering source is not preserved. To address this issue, an extension of the BMVDR beamformer, namely the binaural linearly constrained minimum variance (BLCMV) beamformer, has been proposed in [26] and examined in [34]. This beamformer aims to preserve the RTF of the interfering source by including an interference reduction (IR) constraint in the BMVDR cost function. For consistency in this paper, we will refer to this beamformer as the BMVDR-IR beamformer. In [26], it has been proven that the BMVDR-IR beamformer perfectly preserves the RTF of both the desired source and the interfering source at the expense of a degradation of noise reduction performance. For the implementation of the MVDR-based algorithms estimates of the spatial correlation matrices of the noisy and noise-only signals are required, as well as the acoustic transfer functions (ATFs) normalized by the ATFs of the reference microphones, which can be calculated e.g. from the noisy and noise-only spatial correlation matrices [30]–[33], [35], [36].

In this paper, we propose a novel binaural MVDR-based beamformer, which (similar to the BMVDR and BMVDR-IR beamformers) aims to extract a distortionless desired source, and to minimize the interference and background noise power,

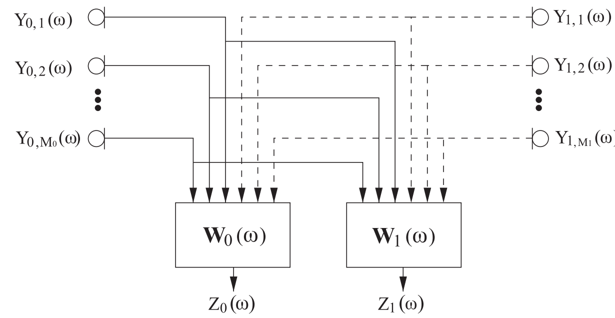

Fig. 1. General binaural processing scheme.

but aims to preserve the binaural cues of the interfering source by adding an RTF preservation constraint to the BMVDR cost function. We will refer to this beamformer as the BMVDR-RTF beamformer. A theoretical comparison between the BMVDR, the BMVDR-IR and the BMVDR-RTF beamformers is given in terms of binaural cue preservation and noise reduction performance. Analytical expressions of the performance in terms of binaural signal-to-interference-and-noise ratio (SINR), signal-to-interference ratio (SIR), and SNR of the three considered beamformers are derived. It can be shown that the BMVDR-RTF beamformer outperforms the BMVDR-IR beamformer in terms of binaural SINR and outperforms the BMVDR in terms of binaural SIR. Among all beamformers which are distortionless with respect to the desired source and preserve the RTF of the interfering source, the BMVDR-RTF beamformer is optimal in terms of binaural SINR. In addition, since the binaural input signal obviously satisfies the distortionless response constraint for the desired source and preserves the RTF of the interfering source, the binaural output SINR of the BMVDR-RTF beamformer is always larger than or equal to the binaural input SINR.

The paper is structured as follows: In Section II, the configuration and notation of the considered binaural hearing aid setup is introduced. In Section III, two recently proposed binaural noise reduction algorithms, namely the BMVDR and the BMVDR-IR beamformers, are briefly reviewed and we propose a novel binaural noise reduction algorithm, namely the BMVDR-RTF beamformer. In Section IV, instrumental performance measures are introduced and a theoretical performance comparison is provided for the considered binaural algorithms in terms of noise reduction. In Section V, the theoretical results are validated by experiments using measured acoustic transfer functions on a binaural hearing aid.

II. Configuration and Notation

In this section, we introduce the general binaural noise reduction and cue preservation problem. We consider a simplified cocktail party scenario consisting of two speakers, i.e. one desired speaker and one interfering speaker, in a noisy and reverberant environment.

A. Microphone and Output Signals

Consider the binaural hearing aid configuration in Fig. 1, consisting of a microphone array with $M = M_0 + M_1$ microphones on the left and the right hearing aid. The m-th microphone signal on the left hearing aid $Y_{0,m}(\omega)$ can be written in the frequency-domain as

$$Y_{0,m}(\omega) = X_{0,m}(\omega) + U_{0,m}(\omega) + N_{0,m}(\omega), m = 1 \dots M_0,$$

with $X_{0,m}(\omega)$ the desired source component, $U_{0,m}(\omega)$ the directional interfering source component, and $N_{0,m}(\omega)$ the ambient background noise component in the m-th microphone signal on the left hearing aid. The m-th microphone signal on the right hearing aid $Y_{1,m}(\omega)$ is defined similarly. In the remainder of the paper, the frequency variable ω will be omitted for the sake of brevity. All microphone signals from both hearing aids can be stacked in the M-dimensional vector $\mathbf{Y}$ as $\mathbf{Y} = [Y_{0,1} \dots Y_{0,M_0} Y_{1,1} \dots Y_{1,M_1}]^T$, which can be written as

$$\mathbf{Y} = \mathbf{X} + \mathbf{U} + \mathbf{N} = \mathbf{X} + \mathbf{V}, \tag{1}$$

where the vectors $\mathbf{X}$, $\mathbf{U}$, $\mathbf{V}$ and $\mathbf{N}$ are defined similarly as $\mathbf{Y}$. The vector $\mathbf{V} = \mathbf{U} + \mathbf{N}$ is defined as the overall noise component as received by the microphones, i.e. directional interfering source plus background noise. The desired source and the directional interfering source components $\mathbf{X}$ and $\mathbf{U}$ can be further written as

$$\mathbf{X} = S_{\mathrm{d}}\mathbf{A}, \quad \mathbf{U} = S_{\mathrm{u}}\mathbf{B}, \tag{2}$$

where S_{d} and S_{u} denote the desired source and interfering source signals and $\mathbf{A}$ and $\mathbf{B}$ denote the ATFs between the desired and interfering sources and the microphones, respectively. Fig. 2 depicts a schematic description of the considered scenario. Without loss of generality, the first microphone on the left hearing aid and the first microphone on the right hearing aid are chosen as the so-called reference microphones. For conciseness, the reference microphone signals $Y_{0,1}$ and $Y_{1,1}$ of the left and the right hearing aid are denoted as Y_0 and Y_1, and can be written as

$$Y_0 = \mathbf{e}_0^T\mathbf{Y}, \quad Y_1 = \mathbf{e}_1^T\mathbf{Y}, \tag{3}$$

where $\mathbf{e}_0$ and $\mathbf{e}_1$ are M-dimensional vectors with one element equal to 1 and all other elements equal to 0, i.e. $\mathbf{e}_0(1) = 1$ and $\mathbf{e}_1(M_0+1) = 1$. The reference microphone signals can then be written as

$$Y_0 = S_{\mathrm{d}}A_0 + S_{\mathrm{u}}B_0 + N_0, \quad Y_1 = S_{\mathrm{d}}A_1 + S_{\mathrm{u}}B_1 + N_1, \tag{4}$$

where $A_0 = \mathbf{e}_0^T\mathbf{A}$, $B_0 = \mathbf{e}_0^T\mathbf{B}$ and $N_0 = \mathbf{e}_0^T\mathbf{N}$. A_1, B_1 and N_1 are defined similarly. The spatial correlation matrices of the desired source, interfering source and noise components are defined as

$$\begin{aligned}\mathbf{R}_x &= \mathcal{E}\left\{\mathbf{X}\mathbf{X}^H\right\} = P_s\mathbf{A}\mathbf{A}^H,\\ \mathbf{R}_u &= \mathcal{E}\left\{\mathbf{U}\mathbf{U}^H\right\} = P_u\mathbf{B}\mathbf{B}^H,\\ \mathbf{R}_n &= \mathcal{E}\left\{\mathbf{N}\mathbf{N}^H\right\},\end{aligned} \tag{5}$$

where $\mathcal{E}\{\cdot\}$ 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 density (PSD) of the desired and interfering sources, respectively.

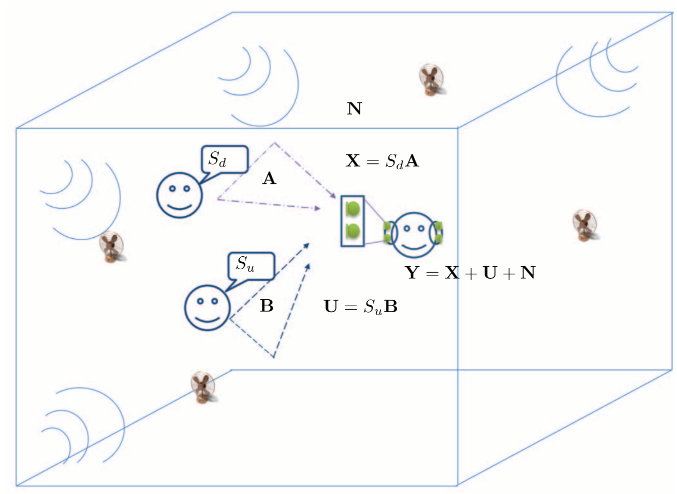

Fig. 2. A scheme of the scenario considered in the paper, depicting two speech sources and noise in reverberant room.

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

$$\mathbf{R}_y = \mathbf{R}_x + \underbrace{\mathbf{R}_u + \mathbf{R}_n}_{\mathbf{R}_v}, \tag{6}$$

with $\mathbf{R}_v$ the spatial correlation matrix of the overall noise component, i.e. interfering source plus background noise.

The output signals on the left and the right hearing aid Z_0 and Z_1 are obtained by applying the left and the right beamformers on all microphone signals from both hearing aids, i.e.

$$Z_0 = \mathbf{W}_0^H\mathbf{Y}, \quad Z_1 = \mathbf{W}_1^H\mathbf{Y}, \tag{7}$$

where $\mathbf{W}_0$ and $\mathbf{W}_1$ are M-dimensional complex-valued weight vectors for the left and the right hearing aid, respectively. Furthermore, we define the $2M$-dimensional complex-valued stacked weight vector $\mathbf{W}$ as

$$\mathbf{W} = \begin{bmatrix}\mathbf{W}_0\\ \mathbf{W}_1\end{bmatrix}. \tag{8}$$

B. Spatial Cues

The RTFs of the desired source 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.

$$\mathrm{RTF}_{\mathrm{x}}^{\mathrm{in}} = \frac{A_0}{A_1}, \quad \mathrm{RTF}_{\mathrm{u}}^{\mathrm{in}} = \frac{B_0}{B_1}. \tag{9}$$

The output RTFs of the desired source and the interfering source are defined as the ratio of the filtered components on the left and the right hearing aid, i.e.

$$\mathrm{RTF}_{\mathrm{x}}^{\mathrm{out}} = \frac{\mathbf{W}_0^H\mathbf{A}}{\mathbf{W}_1^H\mathbf{A}}, \quad \mathrm{RTF}_{\mathrm{u}}^{\mathrm{out}} = \frac{\mathbf{W}_0^H\mathbf{B}}{\mathbf{W}_1^H\mathbf{B}}. \tag{10}$$

Note that the RTF is a complex-valued frequency-dependent scalar from which the binaural ILD and ITD cues can be extracted [25], [37].

In the case of a diffuse noise field, the noise correlation matrix in (5) can be calculated as

$$\mathbf{R}_n = P_n \mathbf{\Gamma}, \tag{11}$$

with P_n the PSD of the noise component in all microphone signals and $\mathbf{\Gamma}$, the spatial coherence matrix of the diffuse noise field. For a binaural hearing aid setup, the input interaural coherence (IC) of a diffuse noise field is equal to [27]

$$\mathrm{IC}_{\mathrm{IN}} = \frac{\mathbf{e}_0^H \mathbf{R}_n \mathbf{e}_1}{\sqrt{\mathbf{e}_0^H \mathbf{R}_n \mathbf{e}_0}\sqrt{\mathbf{e}_1^H \mathbf{R}_n \mathbf{e}_1}}. \tag{12}$$

The output IC of the noise component is equal to

$$\mathrm{IC}_{\mathrm{OUT}} = \frac{\mathbf{W}_0^H \mathbf{R}_n \mathbf{W}_1}{\sqrt{\mathbf{W}_0^H \mathbf{R}_n \mathbf{W}_0}\sqrt{\mathbf{W}_1^H \mathbf{R}_n \mathbf{W}_1}}. \tag{13}$$

The (real-valued) magnitude squared coherence (MSC) is defined as $\mathrm{MSC} = |\mathrm{IC}|^2$. Note that coherent (directional) sources are characterized by an MSC equal to '1'.

C. Weighted Inner Products

In the next sections we will derive decompositions for the filters, for which the following definitions are useful. Assuming that $\mathbf{R}_v$, and hence, $\mathbf{R}_v^{-1}$ is a positive-definite Hermitian matrix, the following weighted inner products can be defined as:

$$\sigma_a = \mathbf{A}^H \mathbf{R}_v^{-1} \mathbf{A}, \quad \sigma_b = \mathbf{B}^H \mathbf{R}_v^{-1} \mathbf{B}, \quad \sigma_{ab} = \mathbf{A}^H \mathbf{R}_v^{-1} \mathbf{B}. \tag{14}$$

In the inner product vector space of M-dimensional complex vectors with the weighted inner product defined above, the squared cosine of the generalized angle between $\mathbf{A}$ and $\mathbf{B}$ can be written as

$$\Sigma = \frac{|\sigma_{ab}|^2}{\sigma_a \sigma_b}, \tag{15}$$

where using the Cauchy-Schwarz inequality it can be shown that

$$0 \leq \Sigma \leq 1. \tag{16}$$

Furthermore, by defining the following two vectors containing the ATFs of the reference microphone signals,

$$\tilde{\mathbf{A}} = \begin{bmatrix} A_0 \\ A_1 \end{bmatrix}, \quad \tilde{\mathbf{B}} = \begin{bmatrix} B_0 \\ B_1 \end{bmatrix}, \tag{17}$$

the squared cosine of the angle between $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ can be written as

$$\Upsilon = \frac{|\tilde{\mathbf{A}}^H \tilde{\mathbf{B}}|^2}{\|\tilde{\mathbf{A}}\|^2 \|\tilde{\mathbf{B}}\|^2}, \tag{18}$$

where $\|\tilde{\mathbf{A}}\|^2 = |A_0|^2 + |A_1|^2$ and $\|\tilde{\mathbf{B}}\|^2 = |B_0|^2 + |B_1|^2$. Again using the Cauchy-Schwartz inequality it can be shown that

$$0 \leq \Upsilon \leq 1. \tag{19}$$

III. Binaural Noise Reduction Techniques

In this section, we first briefly review the BMVDR and the BMVDR-IR beamformers. We then propose a new BMVDR-based beamformer, namely the BMVDR-RTF beamformer, by extending the BMVDR cost function with a constraint related to the RTF of the directional interfering source[1]. The objective of the proposed beamformers is to preserve the RTF of the directional sources. The preservation of the MSC of the background noise is not considered in this work. Combining RTF preservation for directional sources and MSC preservation of diffuse background noise is a topic for future work.

A. Binaural MVDR (BMVDR)

The BMVDR is a binaural extension of the well-known MVDR beamformer [23], [29], consisting of two beamformers, $\mathbf{W}_0$ and $\mathbf{W}_1$, designed to reproduce the desired source component of both reference microphone signals without distortion, while minimizing the overall noise power, i.e.

$$\boxed{\begin{aligned} &\min_{\mathbf{W}_0} \left\{\mathbf{W}_0^H \mathbf{R}_v \mathbf{W}_0\right\} \text{ subject to } \mathbf{W}_0^H \mathbf{A} = A_0, \\ &\min_{\mathbf{W}_1} \left\{\mathbf{W}_1^H \mathbf{R}_v \mathbf{W}_1\right\} \text{ subject to } \mathbf{W}_1^H \mathbf{A} = A_1. \end{aligned}} \tag{20}$$

Both constrained criteria can be combined as a general linearly constrained minimum variance (LCMV) criterion [29] with multiple constraints N, on the stacked vector $\mathbf{W}$, i.e.

$$\min_{\mathbf{W}} \left\{\mathbf{W}^H \mathbf{R} \mathbf{W}\right\} \quad \text{s.t.} \quad \mathbf{C}^H \mathbf{W} = \mathbf{b}, \tag{21}$$

where $\mathbf{C}$ denotes the constraint matrix, $\mathbf{b}$ denotes the desired vector. For the BMVDR beamformer, the constraint matrix $\mathbf{C}$ and the desired vector $\mathbf{b}$ are equal to

$$\mathbf{C}_{\mathrm{BMVDR}} = \begin{bmatrix} \mathbf{A} & \mathbf{0}_{M\times 1} \\ \mathbf{0}_{M\times 1} & \mathbf{A} \end{bmatrix}, \quad \mathbf{b}_{\mathrm{BMVDR}} = \begin{bmatrix} A_0^* \\ A_1^* \end{bmatrix}, \tag{22}$$

and

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_v & 0 \\ 0 & \mathbf{R}_v \end{bmatrix}. \tag{23}$$

Since the number of constraints imposed on the stacked vector $\mathbf{W}$ is equal to $\mathrm{N} = 2$, the number of degrees of freedom for the minimization is equal to $2\mathrm{M} - 2$. The solution for the general LCMV problem in (21) is given by [29]

$$\mathbf{W} = \mathbf{R}^{-1}\mathbf{C}\left[\mathbf{C}^H \mathbf{R}^{-1} \mathbf{C}\right]^{-1} \mathbf{b}. \tag{24}$$

By substituting (22) into (24), the filters of the left and the right hearing aid can be written as

$$\boxed{\begin{aligned} \mathbf{W}_0 &= \frac{A_0^* \mathbf{R}_v^{-1} \mathbf{A}}{\sigma_a}, \\ \mathbf{W}_1 &= \frac{A_1^* \mathbf{R}_v^{-1} \mathbf{A}}{\sigma_a}. \end{aligned}} \tag{25}$$

This implies that $\mathbf{W}_0$ and $\mathbf{W}_1$ are parallel, i.e. $\mathbf{W}_0 = (\mathrm{RTF}_{\mathrm{x}}^{\mathrm{in}})^* \mathbf{W}_1$. Hence, the RTF of the desired source at the

[1]Note that for constructing all considered beamformers it is sufficient to substitute the ATFs by the ATFs normalized by the ATFs of the reference microphones. Efficient procedures for estimating the normalized ATFs exist, e.g. in [30]–[33], [35], [36]. Nevertheless, the derivations will be given using ATFs for the sake of clarity.

output of the BMVDR beamformer is equal to the input RTF, i.e.

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

However, this also implies that *all* sound sources (including the interfering source) are perceived as coming from the desired source direction, i.e.

$$\mathrm{RTF}_{\mathrm{u}}^{\mathrm{out}} = \frac{A_0}{A_1} = \mathrm{RTF}_{\mathrm{x}}^{\mathrm{in}}. \tag{27}$$

Therefore, the RTF of the interfering source is typically distorted, which is not desired, since the spatial impression of the acoustic scene is altered. Since $\mathbf{W}_0$ and $\mathbf{W}_1$ are parallel, the residual noise at the output of the BMVDR beamformer is attributed with MSC equal to one, and hence fully coherent.

B. Binaural MVDR with Interference Reduction Constraints (BMVDR-IR)

The recently proposed BMVDR-IR beamformer [26] is an extension of the BMVDR beamformer, which 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. This is achieved by adding interference reduction (IR) constraints to the BMVDR cost function, i.e.

$$\boxed{\begin{aligned}&\min_{\mathbf{W}_0}\left\{\mathbf{W}_0^H\mathbf{R}_v\mathbf{W}_0\right\} \text{ s.t. } \mathbf{W}_0^H\mathbf{A}=A_0, \mathbf{W}_0^H\mathbf{B}=\eta B_0,\\&\min_{\mathbf{W}_1}\left\{\mathbf{W}_1^H\mathbf{R}_v\mathbf{W}_1\right\} \text{ s.t. } \mathbf{W}_1^H\mathbf{A}=A_1, \mathbf{W}_1^H\mathbf{B}=\eta B_1,\end{aligned}} \tag{28}$$

where the real-valued parameter η, with $0 \leq \eta \leq 1$, is defined as the *interference cue gain factor*, which allows the setting of the amount of interference reduction. Again, both constrained criteria can be combined as a general LCMV criterion with multiple constraints N on the stacked vector $\mathbf{W}$, i.e.

$$\min_{\mathbf{W}} \left\{\mathbf{W}^H\mathbf{R}\mathbf{W}\right\} \quad \text{s.t.} \quad \mathbf{C}_{\mathrm{IR}}^H\mathbf{W} = \mathbf{b}_{\mathrm{IR}}, \tag{29}$$

where the BMVDR-IR constraint set is given by

$$\mathbf{C}_{\mathrm{IR}} = \begin{bmatrix}\mathbf{A} & \mathbf{0}_{M\times1} & \mathbf{B} & \mathbf{0}_{M\times1}\\ \mathbf{0}_{M\times1} & \mathbf{A} & \mathbf{0}_{M\times1} & \mathbf{B}\end{bmatrix}, \mathbf{b}_{\mathrm{IR}} = \begin{bmatrix}A_0^*\\A_1^*\\ \eta B_0^*\\ \eta B_1^*\end{bmatrix}. \tag{30}$$

Since the number of constraints imposed on the stacked vector $\mathbf{W}$ is equal to N = 4, the number of degrees of freedom for the minimization is equal to 2M – 4. Substituting (30) into (24), the BMVDR-IR filters can be written as (see Appendix A-A)

$$\boxed{\begin{aligned}\mathbf{W}_0&=\left(\frac{A_0^*}{\sigma_a}-\frac{\eta B_0^*\sigma_{ab}}{\sigma_a\sigma_b}\right)\frac{\mathbf{R}_v^{-1}\mathbf{A}}{1-\Sigma}+\left(\frac{\eta B_0^*}{\sigma_b}-\frac{A_0^*\sigma_{ab}^*}{\sigma_a\sigma_b}\right)\frac{\mathbf{R}_v^{-1}\mathbf{B}}{1-\Sigma},\\\mathbf{W}_1&=\left(\frac{A_1^*}{\sigma_a}-\frac{\eta B_1^*\sigma_{ab}}{\sigma_a\sigma_b}\right)\frac{\mathbf{R}_v^{-1}\mathbf{A}}{1-\Sigma}+\left(\frac{\eta B_1^*}{\sigma_b}-\frac{A_1^*\sigma_{ab}^*}{\sigma_a\sigma_b}\right)\frac{\mathbf{R}_v^{-1}\mathbf{B}}{1-\Sigma}.\end{aligned}} \tag{31}$$

Since the BMVDR-IR beamformer is satisfying the distortionless response constraints in (28) for the desired source, the RTF of the desired source at the output of the BMVDR-IR beamformer is equal to the input RTF, i.e.

$$\mathrm{RTF}_{\mathrm{x}}^{\mathrm{out}} = \frac{A_0}{A_1} = \mathrm{RTF}_{\mathrm{x}}^{\mathrm{in}}. \tag{32}$$

In addition, since the BMVDR-IR beamformer is satisfying the interference reduction constraints in (28) for the interfering source, the RTF of the directional interfering source at the output of the BMVDR-IR beamformer is equal to the input RTF, i.e.

$$\mathrm{RTF}_{\mathrm{u}}^{\mathrm{out}} = \frac{B_0}{B_1} = \mathrm{RTF}_{\mathrm{u}}^{\mathrm{in}}. \tag{33}$$

Hence, the BMVDR-IR beamformer perfectly preserves the RTFs of both the desired source and the interfering source [26]. Note that in general, $\mathbf{W}_0$ and $\mathbf{W}_1$ are not parallel. Hence, the MSC of the background noise at the output of the beamformer is not necessarily equal to one. The interference reduction performance of the BMVDR-IR beamformer obviously depends on the choice of the *interference cue gain factor*, i.e. the lower the η the more the interfering source will be reduced. When aiming at perfect interference reduction, it is reasonable to set $\eta \to 0$. Accordingly, the BMVDR-IR filters in (31) reduce to

$$\begin{aligned}\mathbf{W}_0 &= \frac{A_0^*}{\sigma_a(1-\Sigma)}\left[\mathbf{R}_v^{-1}\mathbf{A} - \frac{\sigma_{ab}^*}{\sigma_b}\mathbf{R}_v^{-1}\mathbf{B}\right],\\ \mathbf{W}_1 &= \frac{A_1^*}{\sigma_a(1-\Sigma)}\left[\mathbf{R}_v^{-1}\mathbf{A} - \frac{\sigma_{ab}^*}{\sigma_b}\mathbf{R}_v^{-1}\mathbf{B}\right].\end{aligned} \tag{34}$$

In this special case, the filters for the left and the right hearing aid are parallel and $\mathbf{W}_0^H\mathbf{B} = \mathbf{W}_1^H\mathbf{B} = 0$.

C. Binaural MVDR with RTF Constraint (BMVDR-RTF)

In order to preserve the RTF of the interfering source, we propose a novel extension of the BMVDR beamformer, namely the BMVDR-RTF, which instead of adding interference reduction constraints (cf. Section III-B), adds an RTF constraint to the BMVDR cost function in (20), i.e.

$$\boxed{\begin{aligned}&\min_{\mathbf{W}_0,\mathbf{W}_1}\left\{\mathbf{W}_0^H\mathbf{R}_v\mathbf{W}_0+\mathbf{W}_1^H\mathbf{R}_v\mathbf{W}_1\right\}\\&\text{s.t. } \mathbf{W}_0^H\mathbf{A}=A_0, \mathbf{W}_1^H\mathbf{A}=A_1, \frac{\mathbf{W}_0^H\mathbf{B}}{\mathbf{W}_1^H\mathbf{B}}=\frac{B_0}{B_1}.\end{aligned}} \tag{35}$$

The RTF constraint is equivalent to the linear constraint $\mathbf{W}_0^H\mathbf{B} - \mathrm{RTF}_{\mathrm{u}}^{\mathrm{in}}\mathbf{W}_1^H\mathbf{B} = 0$. Hence, similar to the BMVDR and the BMVDR-IR beamformers, the BMVDR-RTF beamformer is the solution of an LCMV criterion with multiple constraints N on the stacked vector $\mathbf{W}$, i.e.

$$\min_{\mathbf{W}} \left\{\mathbf{W}^H\mathbf{R}\mathbf{W}\right\} \quad \text{s.t.} \quad \mathbf{C}_{\mathrm{RTF}}^H\mathbf{W} = \mathbf{b}_{\mathrm{RTF}}, \tag{36}$$

where the BMVDR-RTF constraint set is given by

$$\mathbf{C}_{\mathrm{RTF}} = \begin{bmatrix}\mathbf{A} & \mathbf{0}_{M\times1} & \mathbf{B}\\ \mathbf{0}_{M\times1} & \mathbf{A} & -\mathrm{RTF}_{\mathrm{u}}^{\mathrm{in}}\mathbf{B}\end{bmatrix}, \mathbf{b}_{\mathrm{RTF}} = \begin{bmatrix}A_0^*\\A_1^*\\0\end{bmatrix}. \tag{37}$$

In this new criteria the number of constraints imposed on the stacked vector $\mathbf{W}$ is equal to N = 3. This implies that the number of degrees of freedom for the minimization is equal to 2M – 3. Note that, as opposed to the BMVDR in (20) and the

BMVDR-IR in (28), which could be formulated using two separate independent criteria (i.e. one for the left side and one for the right side), for the BMVDR-RTF both criteria are coupled.

The constrained optimization problem in (35) can be solved using the method of Lagrange multipliers (see Appendix B-A). Similarly to (25) and (31), the BMVDR-RTF filters of the left and the right hearing aid can be written as

$$\boxed{\begin{aligned}\mathbf{W}_0 &= \frac{1}{\sigma_a}\left[A_0^* + \frac{(A_0+\alpha A_1)^*}{(1+|\alpha|^2)}\frac{\Sigma}{(1-\Sigma)}\right]\mathbf{R}_v^{-1}\mathbf{A} \\ &\quad - \frac{(A_0+\alpha A_1)^*}{(1+|\alpha|^2)\,\sigma_{ab}}\frac{\Sigma}{(1-\Sigma)}\mathbf{R}_v^{-1}\mathbf{B}, \\ \mathbf{W}_1 &= \frac{1}{\sigma_a}\left[A_1^* + \alpha\frac{(A_0+\alpha A_1)^*}{(1+|\alpha|^2)}\frac{\Sigma}{(1-\Sigma)}\right]\mathbf{R}_v^{-1}\mathbf{A} \\ &\quad - \alpha\frac{(A_0+\alpha A_1)^*}{(1+|\alpha|^2)\,\sigma_{ab}}\frac{\Sigma}{(1-\Sigma)}\mathbf{R}_v^{-1}\mathbf{B},\end{aligned}} \tag{38}$$

where

$$\alpha = -\mathrm{RTF}_{\mathrm{u}}^{\mathrm{in}} = -\frac{B_0}{B_1}. \tag{39}$$

Since the BMVDR-RTF beamformer is satisfying the distortionless response constraints in (35) for the desired source, the RTF of the desired source at the output of the BMVDR-RTF beamformer is equal to the input RTF, i.e.

$$\mathrm{RTF}_{\mathrm{x}}^{\mathrm{out}} = \frac{A_0}{A_1} = \mathrm{RTF}_{\mathrm{x}}^{\mathrm{in}}. \tag{40}$$

In addition, since the BMVDR-RTF beamformer is satisfying the RTF constraint in (35) for the interfering source, the RTF of the interfering source at the output of the beamformer is equal to the input RTF, i.e.

$$\mathrm{RTF}_{\mathrm{u}}^{\mathrm{out}} = \frac{B_0}{B_1} = \mathrm{RTF}_{\mathrm{u}}^{\mathrm{in}}. \tag{41}$$

Hence, similarly to the BMVDR-IR beamformer, the BMVDR-RTF beamformer also perfectly preserves the RTFs of both the desired source and the interfering source. In general, $\mathbf{W}_0$ and $\mathbf{W}_1$ of the BMVDR-RTF beamformer are not parallel. Hence, the MSC of the background noise at the output of the beamformer is not necessarily equal to one.

IV. Performance Comparison

In this section, we compare the noise reduction performance of the BMVDR, BMVDR-IR and BMVDR-RTF beamformers. As was shown in Section III, in terms of cue preservation, all three binaural beamformers perfectly preserve the RTF of the desired source, whereas the RTF of the interfering source is only preserved by the BMVDR-IR and the BMVDR-RTF beamformers. For the BMVDR, the output RTF of the interfering source is equal to the RTF of the desired source, which is a drawback of the BMVDR beamformer.

All three binaural beamformers are designed to minimize the overall noise power under N linear constraints. In general, the number of degrees of freedom in the optimization is equal to 2M − N. While the BMVDR beamformer requires two linear constraints on the stacked filter (22), the BMVDR-RTF beamformer requires three linear constraints (37) and the BMVDR-IR beamformer requires four linear constraints (30). Therefore, in terms of noise reduction, intuitively it is expected that the performance of the BMVDR-RTF will outperform the BMVDR-IR but will be lower compared to the BMVDR. In this section, the theoretical relationship between the three considered binaural beamformers will be mathematically analyzed and analytical expressions for the binaural SINR, the binaural SIR and the binaural SNR are derived.

A. Performance Measures

The *binaural output SINR* is defined as the ratio of the average output PSDs of the desired source and the overall noise components (interfering source plus background noise) in the left and the right hearing aid, i.e.

$$\mathrm{SINR}^{\mathrm{out}} = \frac{\mathbf{W}_0^H\mathbf{R}_x\mathbf{W}_0 + \mathbf{W}_1^H\mathbf{R}_x\mathbf{W}_1}{\mathbf{W}_0^H\mathbf{R}_v\mathbf{W}_0 + \mathbf{W}_1^H\mathbf{R}_v\mathbf{W}_1}. \tag{42}$$

The *binaural output SIR* is defined as the ratio of the average output PSDs of the desired source and the interfering source components in the left and the right hearing aid, i.e.

$$\mathrm{SIR}^{\mathrm{out}} = \frac{\mathbf{W}_0^H\mathbf{R}_x\mathbf{W}_0 + \mathbf{W}_1^H\mathbf{R}_x\mathbf{W}_1}{\mathbf{W}_0^H\mathbf{R}_u\mathbf{W}_0 + \mathbf{W}_1^H\mathbf{R}_u\mathbf{W}_1}. \tag{43}$$

The *binaural output SNR* is defined as the ratio of the average output PSDs of the desired source and the background noise components in the left and the right hearing aid, i.e.

$$\mathrm{SNR}^{\mathrm{out}} = \frac{\mathbf{W}_0^H\mathbf{R}_x\mathbf{W}_0 + \mathbf{W}_1^H\mathbf{R}_x\mathbf{W}_1}{\mathbf{W}_0^H\mathbf{R}_n\mathbf{W}_0 + \mathbf{W}_1^H\mathbf{R}_n\mathbf{W}_1}. \tag{44}$$

B. Maximum Binaural Signal-to-Interference-and-Noise Ratio Criterion

In this section, we first relate the binaural MVDR-based beamformers to the maximum binaural SINR beamformers. The beamformer maximizing the binaural SINR under distortionless response constraints for the desired source is a binaural extension of the well-known maximum SINR beamformer in [38]. The constrained maximum binaural SINR criterion is given by

$$\begin{aligned} J_{\mathrm{SINR}}(\mathbf{W}) &= \max_{\mathbf{W}}\left\{\frac{\mathbf{W}_0^H\mathbf{R}_x\mathbf{W}_0 + \mathbf{W}_1^H\mathbf{R}_x\mathbf{W}_1}{\mathbf{W}_0^H\mathbf{R}_v\mathbf{W}_0 + \mathbf{W}_1^H\mathbf{R}_v\mathbf{W}_1}\right\} \\ &\quad \text{s.t.}\quad \mathbf{W}_0^H\mathbf{A} = A_0, \mathbf{W}_1^H\mathbf{A} = A_1. \end{aligned} \tag{45}$$

Due to the distortionless response constraints for the desired source and using (5), the numerator in J_{SINR} is equal to $P_s\|\tilde{\mathbf{A}}\|^2$. Therefore, the constrained optimization problem in (45) is equivalent to

$$\begin{aligned} &\min_{\mathbf{W}}\left\{\mathbf{W}_0^H\mathbf{R}_v\mathbf{W}_0 + \mathbf{W}_1^H\mathbf{R}_v\mathbf{W}_1\right\} \\ &\quad \text{s.t. } \mathbf{W}_0^H\mathbf{A} = A_0, \mathbf{W}_1^H\mathbf{A} = A_1, \end{aligned} \tag{46}$$

which is exactly equal to the BMVDR criterion in (20). Hence, the BMVDR beamformer maximizes the binaural SINR under distortionless response constraints for the desired source. Since the BMVDR-IR and the BMVDR-RTF beamformers are also satisfying the distortionless response constraints for the desired source, the binaural SINR for these beamformers is always lower than or equal to the binaural SINR for the BMVDR, i.e.

$$\boxed{\mathrm{SINR}^{\mathrm{out}}_{\mathrm{BMVDR}} \geq \mathrm{SINR}^{\mathrm{out}}_{\mathrm{BMVDR-RTF}},} \tag{47}$$

and in addition,

$$\boxed{\mathrm{SINR}^{\mathrm{out}}_{\mathrm{BMVDR}} \geq \mathrm{SINR}^{\mathrm{out}}_{\mathrm{BMVDR-IR}}.} \tag{48}$$

Obviously, since the reference microphone signals also satisfy the distortionless response constraints for the desired source, the binaural output SINR of the BMVDR is always larger than or equal to the binaural input SINR, i.e.

$$\mathrm{SINR}^{\mathrm{out}}_{\mathrm{BMVDR}} \geq \mathrm{SINR}^{\mathrm{in}}. \tag{49}$$

In addition, the constrained maximum binaural SINR criterion in (45) can be defined under additional constraints, e.g. a constraint for preserving the RTF of the interfering source, i.e.

$$J_{\mathrm{SINR}}(\mathbf{W}) = \max_{\mathbf{W}} \left\{ \frac{\mathbf{W}_0^H \mathbf{R}_x \mathbf{W}_0 + \mathbf{W}_1^H \mathbf{R}_x \mathbf{W}_1}{\mathbf{W}_0^H \mathbf{R}_v \mathbf{W}_0 + \mathbf{W}_1^H \mathbf{R}_v \mathbf{W}_1} \right\}$$
$$\text{s.t. } \mathbf{W}_0^H \mathbf{A} = A_0,\ \mathbf{W}_1^H \mathbf{A} = A_1,\ \frac{\mathbf{W}_0^H \mathbf{B}}{\mathbf{W}_1^H \mathbf{B}} = \frac{B_0}{B_1}. \tag{50}$$

Again, since the numerator in J_{SINR} is equal to $P_s\|\tilde{\mathbf{A}}\|^2$, the constrained optimization problem in (50) is equivalent to

$$\min_{\mathbf{W}} \left\{ \mathbf{W}_0^H \mathbf{R}_v \mathbf{W}_0 + \mathbf{W}_1^H \mathbf{R}_v \mathbf{W}_1 \right\}$$
$$\text{s.t. } \mathbf{W}_0^H \mathbf{A} = A_0,\ \mathbf{W}_1^H \mathbf{A} = A_1,\ \frac{\mathbf{W}_0^H \mathbf{B}}{\mathbf{W}_1^H \mathbf{B}} = \frac{B_0}{B_1}, \tag{51}$$

which is exactly equal to the BMVDR-RTF criterion in (35). Therefore, the BMVDR-RTF beamformer maximizes the binaural SINR under distortionless response constraints for the desired source and an RTF preservation constraint for the interfering source. Since the BMVDR-IR beamformer is also satisfying the same constraints as in (50), by imposing two interference reduction constraints for preserving the RTF of the interfering source (hence, reducing the degrees of freedom for the filter $\mathbf{W}$), the binaural SINR of the BMVDR-IR beamformer is always lower than or equal to the binaural SINR of the BMVDR-RTF beamformer, i.e.

$$\boxed{\mathrm{SINR}^{\mathrm{out}}_{\mathrm{BMVDR-RTF}} \geq \mathrm{SINR}^{\mathrm{out}}_{\mathrm{BMVDR-IR}}.} \tag{52}$$

Again, since the reference microphone signals also satisfy the distortionless response constraints for the desired source and the RTF preservation constraint for the interfering source, the binaural output SINR of the BMVDR-RTF is always larger than or equal to the binaural input SINR, i.e.

$$\mathrm{SINR}^{\mathrm{out}}_{\mathrm{BMVDR-RTF}} \geq \mathrm{SINR}^{\mathrm{in}}. \tag{53}$$

Similarly, the BMVDR-IR beamformer can be interpreted as the maximum binaural SINR beamformer under the constraints in (28). However, it is important to note that only for the special case of $\eta = 1$, the reference microphone signals satisfy the constraints for the BMVDR-IR beamformer. For $\eta < 1$ the reference microphone signals do not satisfy the constraints. Hence, the binaural output SINR of the BMVDR-IR is larger than or equal to the binaural input SINR for $\eta = 1$, whereas in general, the binaural output SINR of the BMVDR-IR is not necessarily larger than or equal to the binaural input SINR. E.g. if we set η to a low value for a scenario in which the interfering source is close to the desired source, a contradiction between the constraints may lead to noise amplification.

C. Analytical Expressions for Binaural SINR

In this section, analytical expressions for the binaural output SINR of the BMVDR, BMVDR-IR and BMVDR-RTF beamformers are derived.

Due to the distortionless response constraints for the desired source, the average output PSD of the desired source for all beamformers is equal and is given by

$$\frac{\mathbf{W}_0^H \mathbf{R}_x \mathbf{W}_0 + \mathbf{W}_1^H \mathbf{R}_x \mathbf{W}_1}{2} = \frac{\|\tilde{\mathbf{A}}\|^2}{2} P_s, \tag{54}$$

with $\tilde{\mathbf{A}}$ defined in (17).

1) For the BMVDR Beamformer: the average output PSD of the overall noise component can be computed, using (25), as

$$\frac{\mathbf{W}_0^H \mathbf{R}_v \mathbf{W}_0 + \mathbf{W}_1^H \mathbf{R}_v \mathbf{W}_1}{2} = \frac{\|\tilde{\mathbf{A}}\|^2}{2\sigma_a}, \tag{55}$$

with σ_a defined in (14). The binaural output SINR of the BMVDR is hence given by

$$\boxed{\mathrm{SINR}^{\mathrm{out}}_{\mathrm{BMVDR}} = P_s \sigma_a.} \tag{56}$$

2) For the BMVDR-IR Beamformer: the average output PSD of the overall noise component is given by (see Appendix A-B)

$$\frac{\mathbf{W}_0^H \mathbf{R}_v \mathbf{W}_0 + \mathbf{W}_1^H \mathbf{R}_v \mathbf{W}_1}{2} = \frac{(\|\tilde{\mathbf{A}}\|^2}{2\sigma_a(1-\Sigma)} \left(\zeta\eta^2 - 2\beta\eta + 1\right), \tag{57}$$

with

$$\zeta = \frac{\sigma_a \|\tilde{\mathbf{B}}\|^2}{\sigma_b \|\tilde{\mathbf{A}}\|^2}, \quad \beta = \frac{Re\left\{\sigma_{ab} \tilde{\mathbf{A}}^H \tilde{\mathbf{B}}\right\}}{\sigma_b \|\tilde{\mathbf{A}}\|^2}, \tag{58}$$

with σ_a, σ_b and σ_{ab} defined in (14), Σ defined in (15) and $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ defined in (17). In Appendix A-B, it is shown that the average output PSD of the overall noise component for the BMVDR-IR beamformer can be minimized for $\eta_{opt} = \frac{\beta}{\zeta}$ and is given by

$$\frac{\mathbf{W}_0^H \mathbf{R}_v \mathbf{W}_0 + \mathbf{W}_1^H \mathbf{R}_v \mathbf{W}_1}{2} = \frac{\|\tilde{\mathbf{A}}\|^2}{2\sigma_a(1-\Sigma)} \left(1 - \frac{\beta^2}{\zeta}\right). \tag{59}$$

For the special case of $\eta = 0$, the average output PSD of the overall noise component for the BMVDR-IR beamformer is equal to

$$\frac{\mathbf{W}_0^H \mathbf{R}_v \mathbf{W}_0 + \mathbf{W}_1^H \mathbf{R}_v \mathbf{W}_1}{2} = \frac{\|\tilde{\mathbf{A}}\|^2}{2\sigma_a(1-\Sigma)}. \tag{60}$$

Using (54) and (57), the binaural output SINR for the BMVDR-IR beamformer is equal to

$$\boxed{\mathrm{SINR}^{\mathrm{out}}_{\mathrm{BMVDR-IR}} = P_s\sigma_a\frac{1-\Sigma}{\zeta\eta^2-2\beta\eta+1}.} \tag{61}$$

Using (54) and (59), the binaural output SINR for η_{opt} is equal to

$$\mathrm{SINR}^{\mathrm{out}}_{\mathrm{BMVDR-IR}} = P_s\sigma_a\frac{1-\Sigma}{1-\frac{\beta^2}{\zeta}}, \tag{62}$$

and, using (54) and (60), the binaural output SINR for $\eta = 0$ is equal to

$$\mathrm{SINR}^{\mathrm{out}}_{\mathrm{BMVDR-IR}} = P_s\sigma_a(1-\Sigma). \tag{63}$$

3) For the BMVDR-RTF Beamformer: the average output PSD of the overall noise component is given by (see Appendix B-B)

$$\frac{\mathbf{W}_0^H\mathbf{R}_v\mathbf{W}_0+\mathbf{W}_1^H\mathbf{R}_v\mathbf{W}_1}{2} = \frac{\|\tilde{\mathbf{A}}\|^2}{2}\frac{1-\Sigma\Upsilon}{\sigma_a(1-\Sigma)}, \tag{64}$$

with Υ defined in (18). Using (54) and (64), the binaural output SINR of the BMVDR-RTF beamformer is equal to

$$\boxed{\mathrm{SINR}^{\mathrm{out}}_{\mathrm{BMVDR-RTF}} = P_s\sigma_a\frac{1-\Sigma}{1-\Sigma\Upsilon}.} \tag{65}$$

In Appendix C-A it is analytically proven that for any value of η, the average output PSD of the overall noise of the BMVDR-IR beamformer is larger than or equal to the average output PSD of the overall noise of the BMVDR-RTF beamformer. Since the output PSD of the desired source in both beamformers is equal, the binaural SINR of the BMVDR-RTF beamformer is always larger than or equal to the binaural SINR of the BMVDR-IR beamformer, as was already shown in (52) without analytical expressions. Furthermore, by comparing (56) and (65) and noting that $0 \leq \Sigma \leq 1$ and $0 \leq \Upsilon \leq 1$, it is also clear that the binaural SINR of the BMVDR beamformer is larger than or equal to the binaural SINR of the BMVDR-RTF beamformer, as was already shown in (47) without analytical expressions. Hence, in conclusion:

$$\begin{aligned}\underbrace{P_s\sigma_a}_{\mathrm{SINR}^{\mathrm{out}}_{\mathrm{BMVDR}}} &\geq \underbrace{\frac{P_s\sigma_a(1-\Sigma)}{1-\Sigma\Upsilon}}_{\mathrm{SINR}^{\mathrm{out}}_{\mathrm{BMVDR-RTF}}} \geq \underbrace{\frac{P_s\sigma_a(1-\Sigma)}{1-\frac{\beta^2}{\zeta}}}_{\mathrm{SINR}^{\mathrm{out}}_{\mathrm{BMVDR-IR},\eta_{opt}}}\\ &\geq \underbrace{\frac{P_s\sigma_a(1-\Sigma)}{\zeta\eta^2-2\beta\eta+1}}_{\mathrm{SINR}^{\mathrm{out}}_{\mathrm{BMVDR-IR}}}\end{aligned} \tag{66}$$

When examining the binaural SINR of the BMVDR-RTF beamformer in (65) for the extremum values of Υ, i.e. the squared cosine of the angle between $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$, it can be noted that for $\Upsilon = 1$ the binaural SINR is equal to the binaural SINR of the BMVDR beamformer, whereas for $\Upsilon = 0$ the binaural SINR is equal to the binaural SINR of the BMVDR-IR beamformer for $\eta_{opt} = 0$.

D. Analytical Expressions for Binaural SIR

In this section, analytical expressions for the binaural output SIR of the BMVDR, BMVDR-IR and BMVDR-RTF beamformers are derived.

1) For the BMVDR Beamformer: the average output PSD of the interfering source component can be computed, using (25), as

$$\frac{\mathbf{W}_0^H\mathbf{R}_u\mathbf{W}_0+\mathbf{W}_1^H\mathbf{R}_u\mathbf{W}_1}{2} = \frac{\|\tilde{\mathbf{A}}\|^2}{2}\frac{P_u|\sigma_{ab}|^2}{\sigma_a^2}, \tag{67}$$

with σ_a and σ_{ab} defined in (14), such that, using (54) and (67), the binaural output SIR of the BMVDR beamformer is equal to

$$\boxed{\mathrm{SIR}^{\mathrm{out}}_{\mathrm{BMVDR}} = \frac{P_s}{P_u}\frac{\sigma_a^2}{|\sigma_{ab}|^2}.} \tag{68}$$

2) For the BMVDR-IR Beamformer: the average output PSD of the interfering source component can be computed, using (31), as

$$\frac{\mathbf{W}_0^H\mathbf{R}_u\mathbf{W}_0+\mathbf{W}_1^H\mathbf{R}_u\mathbf{W}_1}{2} = P_u\eta^2\frac{\|\tilde{\mathbf{B}}\|^2}{2}, \tag{69}$$

such that, using (54) and (69), the binaural output SIR of the BMVDR-IR beamformer is equal to

$$\boxed{\mathrm{SIR}^{\mathrm{out}}_{\mathrm{BMVDR-IR}} = \frac{P_s}{P_u}\frac{\|\tilde{\mathbf{A}}\|^2}{\|\tilde{\mathbf{B}}\|^2\eta^2}.} \tag{70}$$

Depending on the value of η, the binaural output SIR of the BMVDR-IR beamformer can hence be any value between the binaural input SIR ($\eta = 1$) and ∞ ($\eta = 0$), i.e.

$$\begin{aligned}\mathrm{SIR}^{\mathrm{in}} &= \mathrm{SIR}^{\mathrm{out}}_{\mathrm{BMVDR-IR}_{\eta=1}}\\ &\leq \mathrm{SIR}^{\mathrm{out}}_{\mathrm{BMVDR-IR}} \leq \mathrm{SIR}^{\mathrm{out}}_{\mathrm{BMVDR-IR}_{\eta=0}}.\end{aligned} \tag{71}$$

3) For the BMVDR-RTF Beamformer: the average output PSD of the interfering source component is given by (see Appendix B-C)

$$\frac{\mathbf{W}_0^H\mathbf{R}_u\mathbf{W}_0+\mathbf{W}_1^H\mathbf{R}_u\mathbf{W}_1}{2} = \frac{\|\tilde{\mathbf{A}}\|^2}{2}\frac{P_u|\sigma_{ab}|^2}{\sigma_a^2}\Upsilon, \tag{72}$$

such that, using (54) and (72), the binaural output SIR of the BMVDR-RTF beamformer is equal to

$$\boxed{\mathrm{SIR}^{\mathrm{out}}_{\mathrm{BMVDR-RTF}} = \frac{P_s}{P_u}\frac{\sigma_a^2}{|\sigma_{ab}|^2\Upsilon}.} \tag{73}$$

First, by comparing the SIR expressions in (68) and (73), and noting that $0 \leq \Upsilon \leq 1$, the binaural output SIR of the BMVDR beamformer is always smaller than or equal to the binaural output SIR of the BMVDR-RTF beamformer, i.e.

$$\boxed{\mathrm{SIR}^{\mathrm{out}}_{\mathrm{BMVDR}} \leq \mathrm{SIR}^{\mathrm{out}}_{\mathrm{BMVDR-RTF}}.} \tag{74}$$

Secondly, by comparing the SIR expressions in (70) and (73), the binaural output SIR of the BMVDR-RTF and BMVDR-IR beamformers are equal for (see Appendix C-B)

$$\eta_{\mathrm{SIR}} = \frac{|\sigma_{ab}|}{\sigma_a} \frac{|\tilde{\mathbf{A}}^H \tilde{\mathbf{B}}|}{\|\tilde{\mathbf{B}}\|^2}. \tag{75}$$

Hence, the binaural output SIR of the BMVDR-IR beamformer is larger than or equal to the binaural output SIR of the BMVDR-RTF beamformer for $\eta \leq \eta_{SIR}$ and vice versa.

E. Analytical Expressions for Binaural SNR

Using (42), (43) and (44), the following relation holds for any beamformer:

$$\frac{1}{\mathrm{SNR}^{\mathrm{out}}} = \frac{1}{\mathrm{SINR}^{\mathrm{out}}} - \frac{1}{\mathrm{SIR}^{\mathrm{out}}}. \tag{76}$$

Hence, using (47) and (74), the binaural output SNR of the BMVDR beamformer is always larger than or equal to the binaural output SNR of the BMVDR-RTF beamformer, i.e.

$$\boxed{\mathrm{SNR}^{\mathrm{out}}_{\mathrm{BMVDR}} \geq \mathrm{SNR}^{\mathrm{out}}_{\mathrm{BMVDR-RTF}}.} \tag{77}$$

For $\eta \leq \eta_{SIR}$, since the binaural output SIR of the BMVDR-IR is larger than or equal to the binaural output SIR of the BMVDR-RTF, using (52) and (76), the binaural output SNR of the BMVDR-IR is smaller than or equal to the binaural output SNR of the BMVDR-RTF. For $\eta > \eta_{SIR}$, the relation of the binaural output SNR performance between the BMVDR-IR and the BMVDR-RTF beamformers is dependant on the scenario. For $\eta = 1$, since the binaural output SIR is equal to one, the binaural output SINR is larger than or equal to the binaural input SINR, and using (76), the binaural output SNR of the BMVDR-IR is always larger than or equal to the binaural input SNR.

V. Experimental Validation

In this section, the analytical expressions for the BMVDR, BMVDR-RTF and BMVDR-IR beamformers derived in Section IV will be validated using acoustic transfer functions measured on Behind-The-Ear hearing aids [37]. First, in Section V-A, the simulation setup and the algorithm parameters will be discussed. In Section V-B, the beampattern of the three considered beamformers is examined for an anechoic environment. In Section V-C, the noise reduction performance is compared, and the MSC of the diffuse noise component is examined for a reverberant office environment. In order to analyze the full potential of the considered beamformers, we assume that a perfect estimate of the normalized ATFs of the desired and interfering sources, and the spatial correlation matrix of the overall noise is available. Since the experiments are aiming at the validation of the theoretical analysis we do not include estimation errors. By construction, the RTFs of the constrained sources are therefore perfectly preserved.

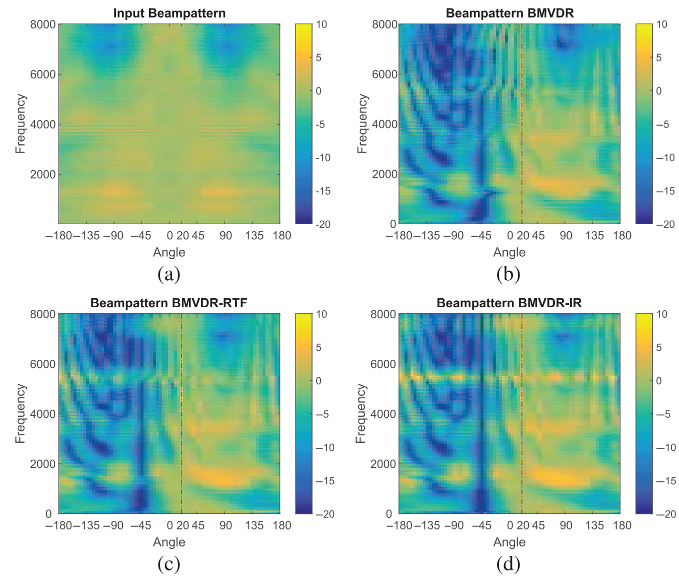

Fig. 3. Beampattern of the BMVDR, BMVDR-RTF and BMVDR-IR ($\eta = 0$) beamformers, together with the input beampattern (anechoic environment, desired source at $20°$, interfering source at $-45°$, $M = 3$). (a) Input (b) BMVDR (c) BMVDR-RTF (d) BMVDR-IR.

A. Simulation Setup and Algorithm Parameters

The performance of the three considered beamformers was evaluated using measured binaural Behind-the-Ear Impulse Responses (BTE-IRs) from [37] at a sampling frequency of 16 kHz. Each hearing aid was equipped with 2 microphones and was mounted on an artificial head. The BTE-IRs were measured both in an anechoic environment (angles ranging from $-180°$ to $180°$ in steps on $5°$, with the source at 3 m from the artificial head) and in an office environment with a reverberation time of approximately 300 ms (angles ranging from $-90°$ to $90°$ in steps on $5°$, with the source at 1 m from the artificial head). The angle $\theta = 0^o$ denotes the frontal direction of the head, and angle $\theta = 90^o$ denotes the right side of the head. The ATFs $\mathbf{A}$ and $\mathbf{B}$ of the desired source and the interfering source were calculated from the BTE-IRs. All experiments in this section were carried out using $M = 3$ microphones, i.e. two microphones on the left hearing aid and one microphone on the right hearing aid.

The PSDs of the desired and interfering sources P_s and P_u were calculated from two different speech signals [39] (Welch method using FFT size of 512 and Hann window). The ratio of the PSDs of the desired source and the interfering source, averaged over frequency, was set to 0 dB. For the background noise, a cylindrically isotropic noise field was assumed. The (i, j)-th element of the noise correlation matrix $\mathbf{R}_n^{i,j}$ was calculated using the ATFs of the anechoic BTE-IRs as

$$\mathbf{R}_n^{i,j} = P_n \frac{\sum_{k=1}^{K} H_i(\theta_k) H_j^*(\theta_k)}{\sqrt{\sum_{k=1}^{K} |H_i(\theta_k)|^2 \sum_{k=1}^{K} |H_j(\theta_k)|^2}}, \tag{78}$$

with $H(\theta_k)$ denoting the measured ATF at angle θ_k and K the total number of angles. The PSD of the background noise P_n is equal to the PSD of speech-shaped noise calculated by averaging multiple speech PSDs taken from [39]. The MSC of the noise component at the reference microphones is depicted in

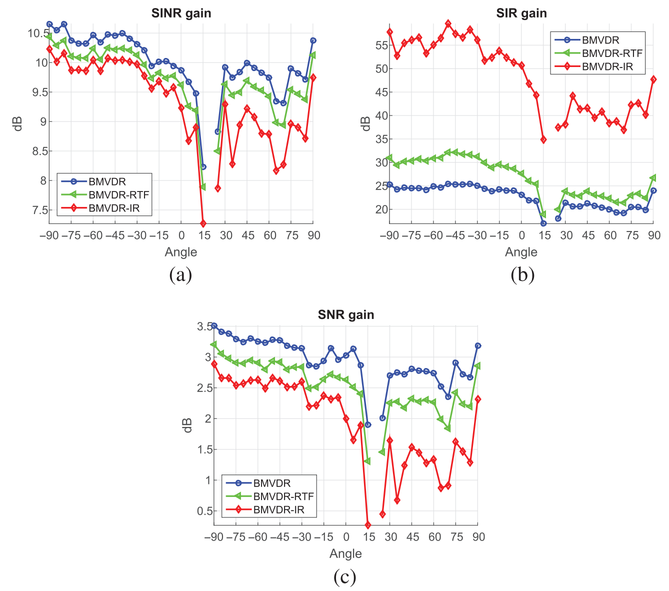

Fig. 4. Global binaural SINR, SIR and SNR gains for BMVDR, BMVDR-RTF and BMVDR-IR ($\eta = \eta_{opt}$) beamformers for different angles of the interfering source (Office environment, desired source at 20°, $M = 3$). (a) SINR gain (b) SIR gain (c) SNR gain.

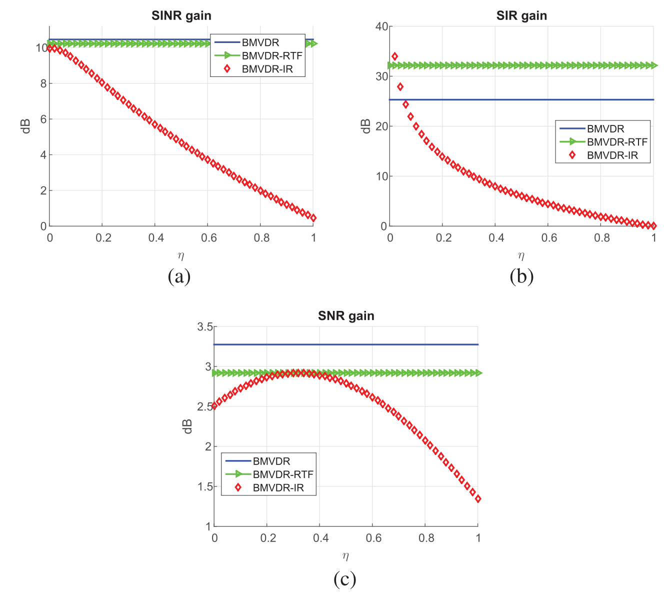

Fig. 5. Global binaural SINR, SIR and SNR gains for BMVDR, BMVDR-RTF and BMVDR-IR beamformers as a function of η (Office environment, desired source at 20°, interfering source at −45°, $M = 3$). (a) SINR gain (b) SIR gain (c) SNR gain.

Fig. 6. The global binaural input SNR averaged over frequency was equal to 0 dB.

B. Beampattern

The beampattern is an effective way to visualize the spatial filtering behavior of beamformers. Fig. 3 depicts the beampattern computed using anechoic ATFs for the BMVDR, the BMVDR-RTF and the BMVDR-IR ($\eta = 0$) beamformers, together with the input beampattern for a scenario comprising a desired source at $\theta_x = 20°$, an interfering source at $\theta_u = -45°$, and diffuse background noise. The directivity of the input microphone reflects the head shadowing effects. It is clear that the directivity of the considered beamformers is much higher than the input directivity. It can be observed that for all three beamformers, the desired source at 20° is processed without distortion (red line). In the beampattern of the BMVDR-IR beamformer, a null is obtained for all frequencies for the direction of the interfering source at −45°, which is not the case for the BMVDR and for the BMVDR-RTF beampatterns. In addition, it should be noted that the BMVDR and BMVDR-RTF beampatterns are very similar, i.e. leading to a similar noise reduction performance, although the BMVDR-RTF beamformer preserves the binaural cues of the interfering source, while the BMVDR beamformer changes the binaural cues of the interfering source to the binaural cues of the desired source.

C. Performance in Reverberant Environment

The noise reduction performance of the considered beamformers in a reverberant office environment is evaluated in terms of global binaural SINR, global binaural SIR and global binaural SNR, averaged in dB over all frequencies.

We first examine a scenario comprised of a desired source at $\theta_x = 20°$ and diffuse background noise for different angles of the interfering source, where the interfering source at 20° has not been evaluated.

Fig. 4(a) depicts the global binaural SINR gain for the BMVDR, BMVDR-RTF and BMVDR-IR ($\eta = \eta_{opt}$) beamformers for different angles of the interfering source. As shown in the theoretical analysis in Section IV, and as can be observed from the experimental results in Fig. 4(a), the BMVDR beamformer yields the largest binaural output SINR, followed by the BMVDR-RTF beamformer and the BMVDR-IR beamformer. The global binaural output SINR of the BMVDR-RTF beamformer is quite close to the global binaural output SINR of the BMVDR beamformer for all angles of the interfering source, whereas especially for angles between 20° and 80° the global binaural output SINR of the BMVDR-IR beamformer is substantially lower.

Fig. 4(b) depicts the global binaural SIR gain for all beamformers for different angles of the interfering source. The global binaural SIR gain of the BMVDR-IR is substantially higher. The global binaural SIR gain for the BMVDR-RTF beamformer is approximately 2-4 dB higher than for the BMVDR beamformer for all interfering source angles as predicted by the theoretical analysis. Fig. 4(c) depicts the global binaural SNR gain for all beamformers for different angles of the interfering source. As predicted by the theoretical analysis, the BMVDR-IR beamformer achieves the lowest SNR gain compared to the BMVDR and BMVDR-RTF beamformers, while the SNR gain of the BMVDR beamformer is the largest. All results correspond with the theoretical analysis in Section IV.

In Fig. 5 we examine the global binaural SINR, SIR and SNR gains for the BMVDR-IR beamformer for different values of η compared to the BMVDR and the BMVDR-RTF beamformers for desired source at 20° and interfering source at −45°. As predicted by the theoretical analysis, the global binaural SINR, SIR and SNR of the BMVDR-IR beamformer depend on η. The BMVDR and BMVDR-RTF beamformers outperform the BMVDR-IR beamformer for any η in terms of binaural SINR and binaural SNR for the examined scenario.

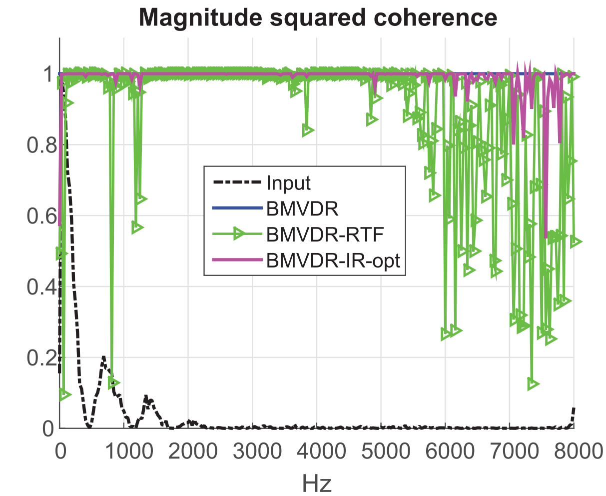

Fig. 6. The MSC for diffuse noise field at the input and at the output of the BMVDR, BMVDR-RTF and BMVDR-IR ($\eta = \eta_{opt}$) beamformers (Office, desired source at 20°, interfering source at $-45°$, $M = 3$).

Fig. 6 depicts the frequency-dependent MSC of the background noise component for the BMVDR, BMVDR-RTF and BMVDR-IR ($\eta = \eta_{opt}$) beamformers, together with the input MSC. As expected from the theoretical analysis is Section III, it is evident that the output MSC of the noise component of the BMVDR beamformer is equal to one for all frequencies. The output MSC of the noise component of the BMVDR-RTF and the BMVDR-IR beamformers is not equal but quite close to one for most frequencies, such that the input MSC of the noise component is not preserved for all three beamformers.

VI. Discussion and Conclusion

In this paper, a class of binaural noise reduction algorithms based on the MVDR criterion has been discussed, consisting of two recently proposed beamformers, namely the BMVDR and the BMVDR-IR beamformers, as well as a novel beamformer, namely the BMVDR-RTF beamformer. All three binaural beamformers process the desired source without distortion, hence, preserving the binaural cues of the desired source. Among all beamformers processing the desired source without distortion, the BMVDR beamformer achieves the maximum binaural output SINR. The drawback of the BMVDR beamformer however, is that all sources will be perceived as arriving from the direction of the desired source. The BMVDR-IR beamformer extends the BMVDR cost function with a constraint on the interfering source in order to control the amount of interference reduction. The BMVDR-IR beamformer preserves the binaural cues of the interfering source, however, trading off noise and interference reduction compared with the BMVDR beamformer.

The proposed BMVDR-RTF beamformer extends the BMVDR cost function with an RTF preservation constraint for the interfering source. It has been proved that this beamformer achieves the maximum binaural output SINR under the additional constraint of preserving the binaural cues of the interfering source, hence, outperforming the BMVDR-IR beamformer in terms of binaural SINR.

Analytical expressions were derived for the binaural output SINR, SIR and SNR of the three considered beamformers. It is analytically shown that in terms of binaural output SINR, the BMVDR beamformer outperforms the BMVDR-RTF beamformer, and that the BMVDR-RTF beamformer outperforms the BMVDR-IR beamformer for all values of η. Furthermore, it is shown that in terms of binaural output SIR, the BMVDR-RTF beamformer outperforms the BMVDR beamformer.

The estimation of the ATFs is important. In practice estimation errors may distort the binaural cues of the constrained sources. Blind identification of the normalized ATFs is usually regarded as significantly easier task than the task of blind identification of ATFs. Normalized ATFs estimation procedures can be found in [30]–[33], [35], [36]. The advantages and disadvantages of the different binaural MVDR-based beamformers are summarized in Table I.

The three considered beamformers can be generalized by taking into account multiple sources, i.e. multiple desired sources and multiple interfering sources. In general, we require that the dimension of the stacked weight vector $\mathbf{W}$ (which is equal to $2M$) will be larger than the number of constraints N, where the degrees of freedom for the minimization of $\mathbf{W}^H\mathbf{R}\mathbf{W}$ is equal to $2M - N$. Each directional source can be either constrained by the optimization criterion or not constrained. For a scenario with N_x constrained desired sources and N_u constrained interfering sources, the BMVDR beamformer requires two linear constraints on the stacked weight vector $\mathbf{W}$ for each constrained desired source, hence if $N = 2N_x < 2M$, at least one degree of freedom for the minimization is maintained. The BMVDR-RTF beamformer requires two linear constraints for each constrained desired source and one linear constraint for each constrained interfering source, hence $N = 2N_x + N_u < 2M$. The number of interferers that can be included in the constraints is therefore, $N_u < 2(M - N_x)$. The BMVDR-IR beamformer requires two linear constraints for each constrained desired source and two linear constraints for each constrained interfering source, hence $N = 2(N_x + N_u) < 2M$. The number of interferers that can be included in the constraints for the BMVDR-IR is therefore $N_u < M - N_x$. This implies that the number of constrained interferers for the BMVDR-RTF can be up to twice larger than the number of constrained interferers for the BMVDR-IR. We note that the power of the unconstrained directional interfering sources is reduced by applying the minimization to $\mathbf{W}^H\mathbf{R}\mathbf{W}$, however their RTF at the output of the beamformer is not constrained.

Appendix A
BMVDR-IR Beamformer

A. Filter Decomposition of BMVDR-IR

In this appendix, the BMVDR-IR filters for the left and right hearing aid are derived. Since the constrained BMVDR-IR criterion in (28) consists of two separate, independent criteria, the left and right filters can be calculated separately. The left BMVDR-IR criterion is given by

$$\arg\min_{\mathbf{W}_0}\{\mathbf{W}_0^H\mathbf{R}_v\mathbf{W}_0\} \quad \text{s.t.} \quad \mathbf{C}_{\text{IR},0}^H\mathbf{W}_0 = \mathbf{b}_{\text{IR},0}, \tag{A.1}$$

TABLE I
COMPARISON OF BMVDR, BMVDR-RTF AND BMVDR-IR BEAMFORMERS

	BMVDR-IR	BMVDR-RTF	BMVDR
SINR	+ Low	++ Medium	+++ **High**
SIR	+++ High for $\eta \leq \eta_{\text{SIR}}$	++ Medium	+ Low
SNR	+ Low for $\eta \leq \eta_{\text{SIR}}$	++ Medium	+++ **High**
Distortion	+ No	+ No	+ No
RTF$_\mathbf{x}$	+ Preserved	+ Preserved	+ Preserved
RTF$_\mathbf{u}$	+ **Preserved**	+ **Preserved**	- Not Preserved
Const. $N_x = 1\ N_u = 1$	$2N_x + 2N_u$ 4	$2N_x + N_u$ 3	$2N_x$ 2
Deg. Freedom $N_x = 1\ N_u = 1$	$2M - 2N_x - 2N_u$ $2M - 4$	$2M - 2N_x - N_u$ $2M - 3$	$2M - 2N_x$ $2M - 2$

where the left BMVDR-IR constraint set is given by

$$\mathbf{C}_{\text{IR},0} = [\,\mathbf{A} \quad \mathbf{B}\,], \quad \mathbf{b}_{\text{IR},0} = \begin{bmatrix} A_0^* \\ \eta B_0^* \end{bmatrix}. \tag{A.2}$$

Substituting (A.2) into (24), the solution for the LCMV problem in (A.1) is given by

$$\begin{aligned} \mathbf{W}_0 &= [\,\mathbf{R}_v^{-1}\mathbf{A} \quad \mathbf{R}_v^{-1}\mathbf{B}\,][\,\mathbf{C}^H\mathbf{R}_v^{-1}\mathbf{C}\,]^{-1} \begin{bmatrix} A_0^* \\ \eta B_0^* \end{bmatrix} \\ &= [\,\mathbf{R}_v^{-1}\mathbf{A} \quad \mathbf{R}_v^{-1}\mathbf{B}\,] \begin{bmatrix} \sigma_a & \sigma_{ab} \\ \sigma_{ab}^* & \sigma_b \end{bmatrix}^{-1} \begin{bmatrix} A_0^* \\ \eta B_0^* \end{bmatrix} \\ &= [\,\mathbf{R}_v^{-1}\mathbf{A} \quad \mathbf{R}_v^{-1}\mathbf{B}\,] \frac{1}{\sigma_a\sigma_b(1-\Sigma)} \\ &\quad \times \begin{bmatrix} \sigma_b & -\sigma_{ab} \\ -\sigma_{ab}^* & \sigma_a \end{bmatrix} \begin{bmatrix} A_0^* \\ \eta B_0^* \end{bmatrix}, \end{aligned} \tag{A.3}$$

with σ_a, σ_b and σ_{ab} defined in (14) and Σ defined in (15). After rearranging terms, the left BMVDR-IR filer is obtained, i.e.

$$\begin{aligned} \mathbf{W}_0 = &\left(\frac{A_0^*}{\sigma_a} - \frac{\eta B_0^*\sigma_{ab}}{\sigma_a\sigma_b}\right)\frac{\mathbf{R}_v^{-1}\mathbf{A}}{1-\Sigma} \\ &+ \left(\frac{\eta B_0^*}{\sigma_b} - \frac{A_0^*\sigma_{ab}^*}{\sigma_a\sigma_b}\right)\frac{\mathbf{R}_v^{-1}\mathbf{B}}{1-\Sigma}. \end{aligned} \tag{A.4}$$

The right BMVDR-IR filter can be obtained in a similar way, i.e.

$$\begin{aligned} \mathbf{W}_1 = &\left(\frac{A_1^*}{\sigma_a} - \frac{\eta B_1^*\sigma_{ab}}{\sigma_a\sigma_b}\right)\frac{\mathbf{R}_v^{-1}\mathbf{A}}{1-\Sigma} \\ &+ \left(\frac{\eta B_1^*}{\sigma_b} - \frac{A_1^*\sigma_{ab}^*}{\sigma_a\sigma_b}\right)\frac{\mathbf{R}_v^{-1}\mathbf{B}}{1-\Sigma}. \end{aligned} \tag{A.5}$$

B. Overall Noise Output PSD of BMVDR-IR

For any LCMV beamformer (24), it can be shown that the overall noise output PSD is equal to

$$\mathbf{W}^H\mathbf{R}\mathbf{W} = \mathbf{b}^H[\mathbf{C}^H\mathbf{R}^{-1}\mathbf{C}]^{-1}\mathbf{b}. \tag{A.6}$$

Since the constrained BMVDR-IR criterion in (28) consists of two separate independent criteria, we first evaluate the output PSD of the overall noise component for the left BMVDR-IR filter, i.e.

$$\mathbf{W}_0^H\mathbf{R}_v\mathbf{W}_0 = \mathbf{b}_{\text{IR},0}^H\left[\mathbf{C}_{\text{IR},0}^H\mathbf{R}_v^{-1}\mathbf{C}_{\text{IR},0}\right]^{-1}\mathbf{b}_{\text{IR},0}, \tag{A.7}$$

with $\mathbf{b}_{\text{IR},0}$ and $\mathbf{C}_{\text{IR},0}$ defined in (A.2). Substituting (A.2) into (A.7) yields

$$\begin{aligned} &\mathbf{W}_0^H\mathbf{R}_v\mathbf{W}_0 \\ &= [\,A_0 \quad \eta B_0\,] \frac{1}{\sigma_a\sigma_b(1-\Sigma)} \begin{bmatrix} \sigma_b & -\sigma_{ab} \\ -\sigma_{ab}^* & \sigma_a \end{bmatrix} \begin{bmatrix} A_0^* \\ \eta B_0^* \end{bmatrix} \\ &= \frac{\sigma_b|A_0|^2 - \eta(\sigma_{ab}A_0B_0^* + \sigma_{ab}^*B_0A_0^*) + \eta^2\sigma_a|B_0|^2}{\sigma_a\sigma_b(1-\Sigma)}. \end{aligned} \tag{A.8}$$

The output PSD of the overall noise component for the right BMVDR-IR filter is given similarly. Hence, the average output PSD is equal to

$$\begin{aligned} &\frac{\mathbf{W}_0^H\mathbf{R}_v\mathbf{W}_0 + \mathbf{W}_1^H\mathbf{R}_v\mathbf{W}_1}{2} \\ &= \frac{\sigma_b(|A_0|^2+|A_1|^2) + \eta^2\sigma_a(|B_0|^2+|B_1|^2)}{2\sigma_a\sigma_b(1-\Sigma)} \\ &\quad - \frac{\eta(\sigma_{ab}(A_0B_0^* + A_1B_1^*) + \sigma_{ab}^*(B_0A_0^* + B_1A_1^*))}{2\sigma_a\sigma_b(1-\Sigma)}, \end{aligned} \tag{A.9}$$

which can be written as

$$\begin{aligned} &\frac{\mathbf{W}_0^H\mathbf{R}_v\mathbf{W}_0 + \mathbf{W}_1^H\mathbf{R}_v\mathbf{W}_1}{2} \\ &= \frac{(|A_0|^2+|A_1|^2)}{2\sigma_a(1-\Sigma)}\left(\zeta\eta^2 - 2\beta\eta + 1\right), \end{aligned} \tag{A.10}$$

with

$$\begin{aligned} \zeta &= \frac{\sigma_a(|B_0|^2+|B_1|^2)}{\sigma_b(|A_0|^2+|A_1|^2)} = \frac{\sigma_a\|\tilde{\mathbf{B}}\|^2}{\sigma_b\|\tilde{\mathbf{A}}\|^2}, \\ \beta &= \frac{[\sigma_{ab}(A_0B_0^* + A_1B_1^*) + \sigma_{ab}^*(A_0^*B_0 + A_1^*B_1)]}{(|A_0|^2+|A_1|^2)\sigma_b} \\ &= \frac{Re\left\{\sigma_{ab}\tilde{\mathbf{A}}^H\tilde{\mathbf{B}}\right\}}{\sigma_b\|\tilde{\mathbf{A}}\|^2}, \end{aligned} \tag{A.11}$$

with $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ defined in (17).

C. Optimum Interference Cue Gain Factor η of BMVDR-IR

The average output PSD of the overall noise component for the BMVDR-IR beamformer in (57) is a function of η and can be written as

$$\begin{aligned} &\frac{\mathbf{W}_0^H\mathbf{R}_v\mathbf{W}_0 + \mathbf{W}_1^H\mathbf{R}_v\mathbf{W}_1}{2} \\ &= \frac{\|\tilde{\mathbf{A}}\|^2}{2\sigma_a(1-\Sigma)}\left(\zeta\eta^2 - 2\beta\eta + 1\right), \end{aligned} \tag{A.12}$$

with ζ and β defined in (A.11). Setting the derivative of (A.12) with respect to η to zero yields the optimum value $\eta_{opt} = \frac{\beta}{\zeta}$. Since the second-order derivative with respect to η is positive, the minimum is achieved. Substituting η_{opt} into (A.12) yields

$$\frac{\|\tilde{\mathbf{A}}\|^2}{2\sigma_a(1-\Sigma)}\left(1-\frac{\beta^2}{\zeta}\right). \tag{A.13}$$

APPENDIX B
BMVDR-RTF BEAMFORMER

A. Filter Decomposition of BMVDR-RTF

In this appendix, the BMVDR-RTF filters for the left and right hearing aid are derived. The BMVDR-RTF criterion is given in (36) and (37), i.e.

$$\begin{aligned}&\arg\min_{\mathbf{W}_0,\mathbf{W}_1}\{\mathbf{W}_0^H\mathbf{R}_v\mathbf{W}_0+\mathbf{W}_1^H\mathbf{R}_v\mathbf{W}_1\};\\&\text{s.t. } \mathbf{W}_0^H\mathbf{A}=A_0,\ \mathbf{W}_1^H\mathbf{A}=A_1,\\&\mathbf{W}_0^H\mathbf{B}-\text{RTF}_{\text{u}}^{\text{in}}\mathbf{W}_1^H\mathbf{B}=0.\end{aligned} \tag{B.1}$$

First, we define $\alpha = -\text{RTF}_{\text{u}}^{\text{in}}$. To solve the constrained optimization in (B.1), we define the complex Lagrangian,

$$\begin{aligned}\mathcal{L}(\mathbf{W}_0,\mathbf{W}_1)&=\mathbf{W}_0^H\mathbf{R}_v\mathbf{W}_0+\mathbf{W}_1^H\mathbf{R}_v\mathbf{W}_1\\&+\lambda\left[\mathbf{W}_0^H\mathbf{B}+\alpha\mathbf{W}_1^H\mathbf{B}\right]+\lambda^*\left[\mathbf{B}^H\mathbf{W}_0+\alpha^*\mathbf{B}^H\mathbf{W}_1\right]\\&+\lambda_0\left[A_0-\mathbf{W}_0^H\mathbf{A}\right]+\lambda_0^*\left[A_0^*-\mathbf{A}^H\mathbf{W}_0\right]\\&+\lambda_1\left[A_1-\mathbf{W}_1^H\mathbf{A}\right]+\lambda_1^*\left[A_1^*-\mathbf{A}^H\mathbf{W}_1\right],\end{aligned} \tag{B.2}$$

where λ, λ_0 and λ_1 are Lagrange multipliers. Setting the gradient with respect to $\mathbf{W}_0^H$ and $\mathbf{W}_1^H$ to $\mathbf{0}$ yields

$$\begin{aligned}\nabla_{\mathbf{W}_0^H}\mathcal{L}(\mathbf{W}_0,\mathbf{W}_1)&=\mathbf{R}_v\mathbf{W}_0+\lambda\mathbf{B}-\lambda_0\mathbf{A}=\mathbf{0},\\\nabla_{\mathbf{W}_1^H}\mathcal{L}(\mathbf{W}_0,\mathbf{W}_1)&=\mathbf{R}_v\mathbf{W}_1+\lambda\alpha\mathbf{B}-\lambda_1\mathbf{A}=\mathbf{0},\end{aligned} \tag{B.3}$$

which can be formulated in matrix notation as

$$\begin{bmatrix}\mathbf{W}_0\\\mathbf{W}_1\end{bmatrix}=\begin{bmatrix}\mathbf{R}_v^{-1}(\lambda_0\mathbf{A}-\lambda\mathbf{B})\\\mathbf{R}_v^{-1}(\lambda_1\mathbf{A}-\lambda\alpha\mathbf{B})\end{bmatrix}. \tag{B.4}$$

Multiplying each row in (B.4) with $\mathbf{A}^H$ and using the first two constraints, we obtain

$$\begin{bmatrix}\mathbf{A}^H\mathbf{W}_0\\\mathbf{A}^H\mathbf{W}_1\end{bmatrix}=\begin{bmatrix}A_0^*\\A_1^*\end{bmatrix}=\begin{bmatrix}\lambda_0\sigma_a-\lambda\sigma_{ab}\\\lambda_1\sigma_a-\lambda\alpha\sigma_{ab}\end{bmatrix}, \tag{B.5}$$

with σ_a and σ_{ab} defined in (14). Then, by substituting (B.4) into the third constraint $\mathbf{B}^H\mathbf{W}_0+\alpha^*\mathbf{B}^H\mathbf{W}_1=0$, we obtain

$$\sigma_{ab}^*\lambda_0-\sigma_b\lambda+\alpha^*\sigma_{ab}^*\lambda_1-\sigma_b|\alpha|^2\lambda=0, \tag{B.6}$$

such that

$$\lambda=\frac{\sigma_{ab}^*(\lambda_0+\alpha^*\lambda_1)}{\sigma_b(1+|\alpha|^2)}=\frac{\Sigma\sigma_a(\lambda_0+\alpha^*\lambda_1)}{\sigma_{ab}(1+|\alpha|^2)}. \tag{B.7}$$

Substituting (B.7) into (B.5) yields

$$\begin{aligned}\begin{bmatrix}A_0^*\\A_1^*\end{bmatrix}&=\begin{bmatrix}\lambda_0\sigma_a-\frac{\Sigma\sigma_a(\lambda_0+\alpha^*\lambda_1)}{(1+|\alpha|^2)}\\\lambda_1\sigma_a-\alpha\frac{\Sigma\sigma_a(\lambda_0+\alpha^*\lambda_1)}{(1+|\alpha|^2)}\end{bmatrix}\\&=\sigma_a\begin{bmatrix}1-\tau&-\alpha^*\tau\\1-\alpha\tau&-|\alpha|^2\tau\end{bmatrix}\begin{bmatrix}\lambda_0\\\lambda_1\end{bmatrix},\end{aligned} \tag{B.8}$$

with $\tau=\frac{\Sigma}{1+|\alpha|^2}$. Solving (B.8) for the Lagrange multipliers λ_0 and λ_1 yields

$$\begin{bmatrix}\lambda_0\\\lambda_1\end{bmatrix}=\frac{1}{\sigma_a\alpha^*\tau(1-\alpha)}\begin{bmatrix}-|\alpha|^2\tau&\alpha^*\tau\\-(1-\alpha\tau)&1-\tau\end{bmatrix}\begin{bmatrix}A_0^*\\A_1^*\end{bmatrix}. \tag{B.9}$$

Substituting (B.7) and (B.9) into (B.4), the BMVDR-RTF filters in (38) are obtained.

B. Overall Noise Output PSD of BMVDR-RTF

For the BMVDR-RTF beamformer, the overall noise output PSD is derived by substituting the BMVDR-RTF constraint set (37) into (A.6). First, using (23) and (37),

$$\begin{aligned}\mathbf{Q}&=[\mathbf{C}^H\mathbf{R}^{-1}\mathbf{C}]^{-1}\\&=\begin{bmatrix}\mathbf{A}^H\mathbf{R}_v^{-1}\mathbf{A}&0&\mathbf{A}^H\mathbf{R}_v^{-1}\mathbf{B}\\0&\mathbf{A}^H\mathbf{R}_v^{-1}\mathbf{A}&\alpha\mathbf{A}^H\mathbf{R}_v^{-1}\mathbf{B}\\\mathbf{B}^H\mathbf{R}_v^{-1}\mathbf{A}&\alpha^*\mathbf{B}^H\mathbf{R}_v^{-1}\mathbf{A}&\mathbf{B}^H\mathbf{R}_v^{-1}\mathbf{B}(1+|\alpha|^2)\end{bmatrix}^{-1}\\&=\begin{bmatrix}\sigma_a&0&\sigma_{ab}\\0&\sigma_a&\alpha\sigma_{ab}\\\sigma_{ab}^*&\alpha^*\sigma_{ab}^*&\sigma_b(1+|\alpha|^2)\end{bmatrix}^{-1}.\end{aligned} \tag{B.10}$$

Substituting (37) and (B.10) into (A.6), we obtain

$$\begin{aligned}\mathbf{W}^H\mathbf{R}\mathbf{W}&=[\,A_0\quad A_1\quad 0\,]\\&\quad\times\begin{bmatrix}\sigma_a&0&\sigma_{ab}\\0&\sigma_a&\alpha\sigma_{ab}\\\sigma_{ab}^*&\alpha^*\sigma_{ab}^*&\sigma_b(1+|\alpha|^2)\end{bmatrix}^{-1}\times\begin{bmatrix}{A_0}^*\\{A_1}^*\\0\end{bmatrix}\\&=|A_0|^2Q_{11}+A_0^*A_1Q_{21}+A_0A_1^*Q_{12}+|A_1|^2Q_{22},\end{aligned} \tag{B.11}$$

with

$$Q_{ik}=\frac{C_{ki}}{\det\left[\mathbf{C}^H\mathbf{R}^{-1}\mathbf{C}\right]}, \tag{B.12}$$

and the cofactor $C_{ki}=(-1)^{k+i}\det[\mathbf{M}_{ki}]$ with $\mathbf{M}_{ki}$ the matrix of minors. These cofactors are equal to

$$\begin{aligned}C_{11}&=\sigma_a\sigma_b\left(1+|\alpha|^2\right)-|\alpha|^2|\sigma_{ab}|^2\\&=\sigma_a\sigma_b\left[\left(1+|\alpha|^2\right)-|\alpha|^2\Sigma\right],\\C_{12}&=\alpha|\sigma_{ab}|^2=\sigma_a\sigma_b\alpha\Sigma,\\C_{21}&=\alpha^*|\sigma_{ab}|^2=\sigma_a\sigma_b\alpha^*\Sigma,\\C_{22}&=\sigma_a\sigma_b\left(1+|\alpha|^2\right)-|\sigma_{ab}|^2=\sigma_a\sigma_b\left[\left(1+|\alpha|^2\right)-\Sigma\right],\end{aligned} \tag{B.13}$$

and the determinant is equal to

$$\begin{aligned}&det[\mathbf{C}^H\mathbf{R}^{-1}\mathbf{C}]\\&=\sigma_a\left[\sigma_a\sigma_b\left(1+|\alpha|^2\right)-|\alpha|^2|\sigma_{ab}|^2\right]-\sigma_a|\sigma_{ab}|^2\\&=\sigma_a\left(1+|\alpha|^2\right)\left[\sigma_a\sigma_b-|\sigma_{ab}|^2\right]=\sigma_a^2\sigma_b\left(1+|\alpha|^2\right)(1-\Sigma).\end{aligned} \tag{B.14}$$

Substituting (B.13) and (B.14) into (B.12), we obtain

$$\begin{aligned}Q_{11}&=\sigma_a\sigma_b\left(1+|\alpha|^2\right)-|\alpha|^2|\sigma_{ab}|^2\\&=\sigma_a\sigma_b\left[1+|\alpha|^2-|\alpha|^2\Sigma\right],\\Q_{12}&=\alpha|\sigma_{ab}|^2=\sigma_a\sigma_b\alpha\Sigma,\\Q_{21}&=\alpha^*|\sigma_{ab}|^2=\sigma_a\sigma_b\alpha^*\Sigma,\\Q_{22}&=\sigma_a\sigma_b(1+|\alpha|^2)-|\sigma_{ab}|^2=\sigma_a\sigma_b\left[1+|\alpha|^2-\Sigma\right].\end{aligned} \tag{B.15}$$

Finally, substituting (B.15) into (B.11) yields

$$\mathbf{W}^H\mathbf{R}\mathbf{W} = \frac{A_0A_1^*\alpha^*\Sigma + |A_0|^2\left(1+|\alpha|^2-|\alpha|^2\Sigma\right)}{\sigma_a\left(1+|\alpha|^2\right)(1-\Sigma)} + \frac{A_0^*A_1\alpha\Sigma + |A_1|^2\left(1+|\alpha|^2-\Sigma\right)}{\sigma_a\left(1+|\alpha|^2\right)(1-\Sigma)}, \quad \text{(B.16)}$$

which can be simplified to

$$\mathbf{W}^H\mathbf{R}\mathbf{W} = \frac{|A_0|^2+|A_1|^2}{\sigma_a(1-\Sigma)}\left[1-\Sigma\frac{|A_1^*-\alpha A_0^*|^2}{|A_0|^2+|A_1|^2}\right]. \quad \text{(B.17)}$$

Noting that

$$\frac{|A_1^*-\alpha A_0^*|^2}{|A_0|^2+|A_1|^2} = \frac{|A_0^*B_0+A_1^*B_1|^2}{(|A_0|^2+|A_1|^2)(|B_0|^2+|B_1|^2)} = \Upsilon, \quad \text{(B.18)}$$

where Υ is defined in (18), the average output PSD of the overall noise component for the BMVDR-RTF beamformer can be simplified to

$$\frac{\mathbf{W}_0^H\mathbf{R}_v\mathbf{W}_0 + \mathbf{W}_1^H\mathbf{R}_v\mathbf{W}_1}{2} = \frac{\|\tilde{\mathbf{A}}\|^2}{2}\frac{1-\Sigma\Upsilon}{\sigma_a(1-\Sigma)}. \quad \text{(B.19)}$$

C. Interfering Source Output PSD of BMVDR-RTF

In order to compute the average output PSD of the interfering source component, we first calculate the right BMVDR-RTF filter response for the ATF of the interfering source using (38), i.e.

$$\mathbf{B}^H\mathbf{W}_1 = \frac{1}{\sigma_a}\left[A_1^* + \alpha\frac{(A_0+\alpha A_1)^*}{(1+|\alpha|^2)}\frac{\Sigma}{(1-\Sigma)}\right]\sigma_{ab}^* - \alpha\frac{(A_0+\alpha A_1)^*}{(1+|\alpha|^2)\,\sigma_{ab}}\frac{\Sigma}{(1-\Sigma)}\sigma_b. \quad \text{(B.20)}$$

By using $\frac{\Sigma}{1-\Sigma} = \left(\frac{1}{1-\Sigma}-1\right)$ and some simplifications, it can be shown that

$$\mathbf{B}^H\mathbf{W}_1 = \frac{\sigma_{ab}^*}{\sigma_a(1+|\alpha|^2)}(A_1^*-\alpha A_0^*). \quad \text{(B.21)}$$

Using (5), we then obtain

$$\mathbf{W}_1^H\mathbf{R}_u\mathbf{W}_1 = \frac{P_u|\sigma_{ab}|^2}{\sigma_a^2(1+|\alpha|^2)^2}|A_1^*-\alpha A_0^*|^2. \quad \text{(B.22)}$$

Hence, using $\mathbf{W}_0^H\mathbf{B}\mathbf{B}^H\mathbf{W}_0 = |\alpha|^2\mathbf{W}_1^H\mathbf{B}\mathbf{B}^H\mathbf{W}_1$, the average output PSD of the interfering source component for the BMVDR-RTF beamformer is equal to

$$\frac{\mathbf{W}_0^H\mathbf{R}_u\mathbf{W}_0 + \mathbf{W}_1^H\mathbf{R}_u\mathbf{W}_1}{2} = \frac{P_u|\sigma_{ab}|^2(1+|\alpha|^2)}{2\sigma_a^2(1+|\alpha|^2)^2}|A_1^*-\alpha A_0^*|^2, \quad \text{(B.23)}$$

which can be simplified to

$$\frac{\mathbf{W}_0^H\mathbf{R}_u\mathbf{W}_0 + \mathbf{W}_1^H\mathbf{R}_u\mathbf{W}_1}{2} = \frac{(|A_0|^2+|A_1|^2)}{2}\frac{P_u|\sigma_{ab}|^2\Upsilon}{\sigma_a^2}, \quad \text{(B.24)}$$

with Υ defined in (18).

Appendix C
Performance Comparison between BMVDR-IR and BMVDR-RTF

A. Comparison of Overall Noise Output PSD

Using (58) it can be shown that

$$\frac{\beta^2}{\zeta} = \frac{\frac{\left(Re\{\sigma_{ab}\tilde{\mathbf{A}}^H\tilde{\mathbf{B}}\}\right)^2}{\sigma_b^2\|\tilde{\mathbf{A}}\|^4}}{\frac{\sigma_a\|\tilde{\mathbf{B}}\|^2}{\sigma_b\|\tilde{\mathbf{A}}\|^2}} = \frac{\left(Re\left\{\sigma_{ab}\tilde{\mathbf{A}}^H\tilde{\mathbf{B}}\right\}\right)^2}{\sigma_a\sigma_b\|\tilde{\mathbf{A}}\|^2\|\tilde{\mathbf{B}}\|^2} = \Upsilon\frac{\|\tilde{\mathbf{A}}\|^2\|\tilde{\mathbf{B}}\|^2}{|\tilde{\mathbf{A}}^H\tilde{\mathbf{B}}|^2}\frac{|\sigma_{ab}|^2\left(Re\left\{\sigma_{ab}\tilde{\mathbf{A}}^H\tilde{\mathbf{B}}\right\}\right)^2}{\sigma_a\sigma_b|\sigma_{ab}|^2\|\tilde{\mathbf{A}}\|^2\|\tilde{\mathbf{B}}\|^2} = \Sigma\Upsilon\frac{\left(Re\left\{\sigma_{ab}\tilde{\mathbf{A}}^H\tilde{\mathbf{B}}\right\}\right)^2}{|\sigma_{ab}|^2|\tilde{\mathbf{A}}^H\tilde{\mathbf{B}}|^2}, \quad \text{(C.1)}$$

with Υ defined in (18). Hence, $\frac{\beta^2}{\zeta}$ is always smaller than or equal to $\Sigma\Upsilon$, such that, using (59) and (64), the average output PSD on the overall noise of the BMVDR-IR beamformer is larger than or equal to the output PSD of the overall noise of the BMVDR-RTF beamformer.

B. Comparison of Binaural Output SIR

By comparing the SIR expressions in (70) and (73), the binaural output SIR of the BMVDR-IR beamformer is equal to the binaural output SIR of the BMVDR-RTF beamformer if

$$\frac{P_s}{P_u}\frac{\|\tilde{\mathbf{A}}\|^2}{\|\tilde{\mathbf{B}}\|^2\eta^2} = \frac{P_s}{P_u}\frac{\sigma_a^2}{|\sigma_{ab}|^2\Upsilon}, \quad \text{(C.2)}$$

with $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ defined in (17) and Υ defined in (18). Substituting (18) into (C.2) and cancelling common elements, we obtain

$$\eta_{\text{SIR}} = \frac{|\sigma_{ab}|}{\sigma_a}\frac{|\tilde{\mathbf{A}}^H\tilde{\mathbf{B}}|}{\|\tilde{\mathbf{B}}\|^2}. \quad \text{(C.3)}$$

References

[1] A. Boothroyd, K. Fitz, J. Kindred, S. Kochkin, H. Levitt, B. Moore, and J. Yanz, "Hearing aids and wireless technology," *Hear. Rev.*, vol. 14, no. 6, p. 44, 2007.

[2] T. Van den Bogaert, S. Doclo, M. Moonen, and J. Wouters, "The effect of multimicrophone noise reduction systems on sound source localization by users of binaural hearing aids," *J. Acoust. Soc. Amer.*, vol. 124, pp. 484–497, 2008.

[3] T. Van den Bogaert, S. Doclo, J. Wouters, and M. Moonen, "Speech enhancement with multichannel Wiener filter techniques in multimicrophone binaural hearing aids," *J. Acoust. Soc. Amer.*, vol. 125, pp. 360–379, 2009.

[4] F. Baumgarte and C. Faller, "Binaural cue coding-part I: Psychoacoustic fundamentals and design principles," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 11, no. 6, pp. 509–519, Nov. 2003.

[5] J. Käsbach, M. Marschall, B. Epp, and T. Dau, "The relation between perceived apparent source width and interaural cross-correlation in sound reproduction spaces with low reverberation," in *Proc. 39th German Annu. Conf. Acoust.*, 2013.

[6] A. Bronkhorst and R. Plomp, "The effect of head-induced interaural time and level differences on speech intelligibility in noise," *J. Acoust. Soc. Amer.*, vol. 83, no. 4, pp. 1508–1516, 1988.

[7] J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization*. Cambridge, MA, USA: MIT Press, 1997.

[8] A. W. Bronkhorst, "The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions," *Acta Acust. united with Acust.*, vol. 86, no. 1, pp. 117–128, 2000.

[9] M. L. Hawley, R. Y. Litovsky, and J. F. Culling, "The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer," *J. Acoust. Soc. Amer.*, vol. 115, no. 2, pp. 833–843, 2004.

[10] R. Beutelmann and T. Brand, "Prediction of speech intelligibility in spatial noise and reverberation for normal-hearing and hearing-impaired listeners," *J. Acoust. Soc. Amer.*, vol. 120, no. 1, pp. 331–342, 2006.

[11] B. Kollmeier, J. Peissig, and V. Hohmann, "Real-time multiband dynamic compression and noise reduction for binaural hearing aids," *J. Rehabil. Res. and Develop.*, vol. 30, pp. 82–94, 1993.

[12] T. Wittkop and V. Hohmann, "Strategy-selective noise reduction for binaural digital hearing aids," *Speech Commun.*, vol. 39, no. 1, pp. 111–138, 2003.

[13] T. Lotter and P. Vary, "Dual-channel speech enhancement by superdirective beamforming," *EURASIP J. Adv. Signal Process.*, vol. 2006, pp. 1–14, Jan. 2006.

[14] A. Tsilfidis, E. Georganti, and J. Mourjopoulos, "Binaural extension and performance of single-channel spectral subtraction dereverberation algorithms," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Prague, Czech Republic, May 2011, pp. 1737–1740.

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

[16] J. Desloge, W. Rabinowitz, and P. Zurek, "Microphone-array hearing aids with binaural output-Part I: Fixed-processing systems," *IEEE Trans. Speech Audio Process.*, vol. 5, no. 6, pp. 529–542, Nov. 1997.

[17] D. Welker, J. Greenberg, J. Desloge, and P. Zurek, "Microphone-array hearing aids with binaural output-part II: A two-microphone adaptive system," *IEEE Trans. Speech Audio Process.*, vol. 5, no. 6, pp. 543–551, Nov. 1997.

[18] I. Merks, M. Boone, and A. Berkhout, "Design of a broadside array for a binaural hearing aid," in *Proc. IEEE Workshop Applicat. Signal Process. Audio Acoust. (WASPAA)*, New Paltz, NY, USA, Oct. 1997, pp. 1–4.

[19] Y. Suzuki, S. Tsukui, F. Asano, and R. Nishimura, "New design method of a binaural microphone array using multiple constraints," *IEICE Trans. Fundam. Elect., Commun. Comput. Sci.*, vol. 82, no. 4, pp. 588–596, 1999.

[20] S. Doclo, R. Dong, T. Klasen, J. Wouters, S. Haykin, and M. Moonen, "Extension of the multi-channel Wiener filter with ITD cues for noise reduction in binaural hearing aids," in *Proc. IEEE Workshop Applicat. Signal Process. Audio Acoust. (WASPAA)*, 2005, pp. 70–73.

[21] S. Wehr, M. Zourub, R. Aichner, and W. Kellermann, "Post-processing for BSS algorithms to recover spatial cues," in *Proc. Int. Workshop Acoust. Signal Enhance. (IWAENC)*, Paris, France, Sep. 2006.

[22] R. Aichner, H. Buchner, M. Zourub, and W. Kellermann, "Multichannel source separation preserving spatial information," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Honolulu, HI, USA, Apr. 2007, pp. 5–8.

[23] S. Doclo, S. Gannot, M. Moonen, and A. Spriet, "Acoustic Beamforming for Hearing Aid Applications," in *Handbook on Array Processing and Sensor Networks*. New York, NY, USA: Wiley, 2008, pp. 269–302.

[24] S. Markovich-Golan, S. Gannot, and I. Cohen, "A reduced bandwidth binaural MVDR beamformer," in *Proc. Int. Workshop Acoust. Signal Enhance. (IWAENC)*, Tel-Aviv, Israel, Sep. 2010.

[25] 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, Lang. Process.*, vol. 18, no. 2, pp. 342–355, Feb. 2010.

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

[27] D. Marquardt, V. Hohmann, and S. Doclo, "Coherence preservation in multi-channel Wiener filtering based noise reduction for binaural hearing aids," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Vancouver, BC, Canada, May 2013, pp. 8648–8652.

[28] E. Hadad, D. Marquardt, S. Doclo, and S. Gannot, "Binaural multichannel Wiener filter with directional interference rejection," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Brisbane, Australia, Apr. 2015, pp. 644–648.

[29] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 5, no. 2, pp. 4–24, Feb. 1988.

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

[31] R. Talmon, I. Cohen, and S. Gannot, "Relative transfer function identification using convolutive transfer function approximation," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 17, no. 4, pp. 546–555, May 2009.

[32] I. Cohen, "Relative transfer function identification using speech signals," *IEEE Trans. Speech Audio Process.*, vol. 12, no. 5, pp. 451–459, Sep. 2004.

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

[34] D. Marquardt, E. Hadad, S. Gannot, and S. Doclo, "Optimal binaural LCMV beamformers for combined noise reduction and binaural cue preservation," in *Proc. Int. Workshop Acoust. Signal Enhance. (IWAENC)*, Antibes-Juan les Pins, France, Sep. 2014.

[35] 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 Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Brisbane, Australia, Apr. 2015, pp. 544–548.

[36] 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 Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Brisbane, Australia, Apr. 2015, pp. 320–324.

[37] 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 J. Adv. Signal Process.*, vol. 2009, p. 10, 2009.

[38] H. L. Van Trees, "Detection, estimation, and modulation theory," in *Optimum Array Processing*. New York, NY, USA: Wiley, 2004.

[39] N. J. Versfeld, L. Daalder, J. M. Festen, and T. Houtgast, "Method for the selection of sentence materials for efficient measurement of the speech reception threshold," *J. Acoust. Soc. Amer.*, vol. 107, no. 3, pp. 1671–1684, 2000.

Elior Hadad (S'13) received the B.Sc. (summa cum laude) and the M.Sc. (cum laude) degrees in electrical engineering from Ben-Gurion University of the Negev (BGU), Beer-Sheva, Israel, in 2001 and 2007, respectively. He is currently pursuing his Ph.D. degree at the Engineering Faculty, Bar-Ilan University, Israel. His research interests include statistical signal processing and in particular binaural noise reduction algorithms using microphone arrays.

Daniel Marquardt (S'12) received the Dipl.-Ing. degree in Media Technology with focus on Audio-Visual Technology from Ilmenau University of Technology, Germany, in 2010. From 2009 to 2010, he was with Nuance Communications, Inc. and Fraunhofer IDMT where he worked in the field of voice activity detection and binaural acoustics. Since 2010, he has been a Ph.D. student in the Signal Processing Group at the University of Oldenburg, Germany. His research interests are in the area of signal processing for binaural hearing aids.

Simon Doclo (S'95–M'03–SM'13) received the M.Sc. degree in electrical engineering and the Ph.D. degree in applied sciences from the Katholieke Universiteit Leuven, Belgium, in 1997 and 2003. From 2003 to 2007, he was a Postdoctoral Fellow with the Research Foundation Flanders at the Electrical Engineering Department (Katholieke Universiteit Leuven) and the Adaptive Systems Laboratory (McMaster University, Canada). From 2007 to 2009, he was a Principal Scientist with NXP Semiconductors at the Sound and Acoustics Group in Leuven, Belgium. Since 2009, he has been a Full Professor at the University of Oldenburg, Germany, and Scientific Advisor for the project group Hearing, Speech and Audio Technology of the Fraunhofer Institute for Digital Media Technology. His research activities center around signal processing for acoustical applications, more specifically microphone array processing, active noise control, acoustic sensor networks and hearing aid processing. Prof. Doclo received the Master Thesis Award of the Royal Flemish Society of Engineers in 1997 (with Erik De Clippel), the Best Student Paper Award at the International Workshop on Acoustic Echo and Noise Control in 2001, the EURASIP Signal Processing Best Paper Award in 2003 (with Marc Moonen) and the IEEE Signal Processing Society 2008 Best Paper Award (with Jingdong Chen, Jacob Benesty, Arden Huang). He was member of the IEEE Signal Processing Society Technical Committee on Audio and Acoustic Signal Processing (2008–2013) and Technical Program Chair for the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA) in 2013. Prof. Doclo has served as guest editor for several special issues (*IEEE Signal Processing Magazine*, *Elsevier Signal Processing*) and is associate editor for IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH AND LANGUAGE PROCESSING and *EURASIP Journal on Advances in Signal Processing*.

Sharon Gannot (S'92–M'01–SM'06) received his B.Sc. degree (summa cum laude) from the Technion Israel Institute of Technology, Haifa, Israel, in 1986 and the M.Sc. (cum laude) and Ph.D. degrees from Tel-Aviv University, Israel, in 1995 and 2000, respectively, all in electrical engineering. In 2001, he held a post-doctoral position at the department of Electrical Engineering (ESAT-SISTA) at K.U.Leuven, Belgium. From 2002 to 2003, he held a research and teaching position at the Faculty of Electrical Engineering, Technion-Israel Institute of Technology, Haifa, Israel. Currently, he is a Full Professor at the Faculty of Engineering, Bar-Ilan University, Israel, where he is heading the Speech and Signal Processing laboratory and the Signal Processing Track. Prof. Gannot is the recipient of Bar-Ilan University outstanding lecturer award for 2010 and 2014. Prof. Gannot has served as an Associate Editor of the *EURASIP Journal of Advances in Signal Processing* in 2003–2012, and as an Editor of several special issues on Multi-microphone Speech Processing of the same journal. He has also served as a guest editor of *ELSEVIER Speech Communication and Signal Processing* journals. Prof. Gannot has served as an Associate Editor of IEEE TRANSACTIONS ON SPEECH, AUDIO AND LANGUAGE PROCESSING in 2009–2013. Currently, he is a Senior Area Chair of the same journal. He also serves as a reviewer of many IEEE journals and conferences. Prof. Gannot is a member of the Audio and Acoustic Signal Processing (AASP) technical committee of the IEEE since Jan., 2010. He is also a member of the Technical and Steering committee of the International Workshop on Acoustic Signal Enhancement (IWAENC) since 2005 and was the general co-chair of IWAENC held at Tel-Aviv, Israel, in August 2010. Prof. Gannot has served as the general co-chair of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA) in October 2013. Prof. Gannot is the recipient of IEEE Signal Processing Letters best paper award for 2015. Prof. Gannot was selected (with colleagues) to present a tutorial sessions in ICASSP 2012, EUSIPCO 2012, ICASSP 2013 and EUSIPCO 2013. Prof. Gannot research interests include multi-microphone speech processing and specifically distributed algorithms for ad hoc microphone arrays for noise reduction and speaker separation; dereverberation; single microphone speech enhancement and speaker localization and tracking.