A Generalized Binaural MVDR Beamformer with Interferer Relative Transfer Function Preservation

Elior Hadad
Faculty of Engineering
Bar-Ilan University
Ramat-Gan, Israel
Email: elior.hadad@biu.ac.il

Simon Doclo
University of Oldenburg
Department of Medical Physics and
Acoustics and Cluster of Excellence
Hearing4All, Oldenburg, Germany
Email: simon.doclo@uni-oldenburg.de

Sharon Gannot
Faculty of Engineering
Bar-Ilan University
Ramat-Gan, Israel
Email: sharon.gannot@biu.ac.il

***Abstract*— In addition to interference and noise reduction, an important objective of binaural speech enhancement algorithms is the preservation of the binaural cues of both the target and the undesired sound sources. For directional sources, this can be achieved by preserving the relative transfer function (RTF). The recently proposed binaural minimum variance distortionless response (BMVDR) beamformer preserves the RTF of the target, but typically distorts the RTF of the interfering sources. Recently, two extensions of the BMVDR beamformer were proposed preserving the RTFs of both the target and the interferer, namely, the binaural linearly constrained minimum variance (BLCMV) and the BMVDR-RTF beamformers. In this paper, we generalize the BMVDR-RTF to trade off interference reduction and noise reduction. Three special cases of the proposed beamformer are examined, either maximizing the signal-to-interference-and-noise ratio (SINR), the signal-to-noise ratio (SNR), or the signal-to-interference ratio (SIR). Experimental validations in an office environment validate our theoretical results.**

I. Introduction

In the last decades, several binaural speech enhancement algorithms were developed, which in addition to reducing noise and limiting speech distortion aim to preserve the binaural cues of the sound sources [1]–[10]. By preserving the binaural cues, both an improved sound localisation as well as a better speech intelligibility in noisy environments can be achieved as a result of binaural unmasking [11]. For directional sources, preservation of the interaural level difference (ILD) and the interaural time difference (ITD) cues can be achieved by preserving the so-called relative transfer function (RTF), which is defined as the ratio of the acoustic transfer functions (ATFs) relating the source and the two ears.

In [3], the binaural multichannel Wiener filter (BMWF) was presented. The BMWF preserves the binaural cues of the target but distorts the binaural cues of the noise, such that both the target and the noise are perceived as arriving from the target direction [3]. To optimally exploit the benefits of binaural unmasking and to optimize the spatial awareness of the hearing aid user, several extensions of the BMWF have been proposed, which aim to preserve also the binaural cues of the residual noise [6]–[9].

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

If the target needs to be processed without distortion, the binaural minimum variance distortionless response (BMVDR) beamformer (BF) can be applied [3]. However, similarly to the BMWF, an important drawback of the BMVDR BF is the fact that the binaural cues of the noise are not preserved. To control both the suppression and the binaural cue preservation of a directional interferer, the binaural linearly constrained minimum variance (BLCMV) BF was proposed in [4], [10]. In the BLCMV criterion, a hard constraint controlling the amount of interference reduction was added to the BMVDR cost function. It was shown that the BLCMV BF is able to preserve the binaural cues of both the target and the interferer. In [5], another extension of the BMVDR BF was proposed, namely the BMVDR-RTF, which is also able to preserve the binaural cues of both the target and interferer by adding a hard constraint preserving the RTF of the interferer to the BMVDR cost function. It was analytically proven that the BMVDR-RTF BF outperforms the BLCMV BF in terms of SINR.

In this paper, we propose an extension of the BMVDR-RTF BF by generalizing the BMVDR-RTF cost function to trade off interference and noise reduction using a weighting parameter. The obtained BF, referred to as the G-BMVDR-RTF, is analyzed and specific settings of the weighting parameter are examined, in particular three special cases either maximizing the SINR, the SNR, or the SIR.

II. Problem Formulation

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

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

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

978-0-9928-6265-7/16/$31.00 ©2016 IEEE

The target and interferer components can be written as $\mathbf{x} = S_{\text{x}}\mathbf{a}$ and $\mathbf{u} = S_{\text{u}}\mathbf{b}$, where S_{x} and S_{u} denote the target and interferer signals and $\mathbf{a}$ and $\mathbf{b}$ denote the ATFs relating the microphones to the target and the interfering source, respectively. Assuming statistical independence between all components in (1), the spatial correlation matrix of the microphone signals $\mathbf{R}_y$ can be written as

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

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

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

where $P_s = \mathcal{E}\left\{|S_{\text{x}}|^2\right\}$ and $P_u = \mathcal{E}\left\{|S_{\text{u}}|^2\right\}$ denote the power spectral density (PSD) of the target and the interferer, respectively. The background noise correlation matrix $\mathbf{R}_n$ is assumed to be full-rank.

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

$$y_L = S_{\text{x}}a_L + S_{\text{u}}b_L + n_L,\ y_R = S_{\text{x}}a_R + S_{\text{u}}b_R + n_R. \qquad (4)$$

The RTFs of the target and the interferer between the reference microphones at both hearing devices are defined as the ratio of the respective ATFs, i.e.

$$\text{RTF}_{\text{x}}^{\text{in}} = \frac{a_L}{a_R}, \qquad \text{RTF}_{\text{u}}^{\text{in}} = \frac{b_L}{b_R}. \qquad (5)$$

The output signals at the left and the right hearing devices are given by $z_L = \mathbf{w}_L^H\mathbf{y}$ and $z_R = \mathbf{w}_R^H\mathbf{y}$, respectively, where $\mathbf{w}_L$ and $\mathbf{w}_R$ denote M-dimensional complex-valued weight vectors. Furthermore, we define the $2M$-dimensional stacked weight vector as $\mathbf{w} = \left[\begin{array}{cc}\mathbf{w}_L & \mathbf{w}_R\end{array}\right]^T$.

The binaural output SINR is defined as the ratio of the average output PSDs of the target component and the total noise component (interferer plus background noise) in the left and the right hearing aid, i.e.

$$\text{SINR}^{\text{out}} = \frac{\mathbf{w}_L^H\mathbf{R}_x\mathbf{w}_L + \mathbf{w}_R^H\mathbf{R}_x\mathbf{w}_R}{\mathbf{w}_L^H\mathbf{R}_v\mathbf{w}_L + \mathbf{w}_R^H\mathbf{R}_v\mathbf{w}_R}. \qquad (6)$$

Similarly, the binaural output SIR and SNR are defined as the ratio of the average output PSDs of the target component and the interferer and background noise components, respectively, in the left and the right hearing aid, i.e.

$$\text{SIR}^{\text{out}} = \frac{\mathbf{w}_L^H\mathbf{R}_x\mathbf{w}_L + \mathbf{w}_R^H\mathbf{R}_x\mathbf{w}_R}{\mathbf{w}_L^H\mathbf{R}_u\mathbf{w}_L + \mathbf{w}_R^H\mathbf{R}_u\mathbf{w}_R}, \qquad (7)$$

$$\text{SNR}^{\text{out}} = \frac{\mathbf{w}_L^H\mathbf{R}_x\mathbf{w}_L + \mathbf{w}_R^H\mathbf{R}_x\mathbf{w}_R}{\mathbf{w}_L^H\mathbf{R}_n\mathbf{w}_L + \mathbf{w}_R^H\mathbf{R}_n\mathbf{w}_R}. \qquad (8)$$

Using (6), (7), and (8), it can be easily shown that

$$\frac{1}{\text{SNR}^{\text{out}}} = \frac{1}{\text{SINR}^{\text{out}}} - \frac{1}{\text{SIR}^{\text{out}}}. \qquad (9)$$

III. Binaural noise reduction techniques

In section III-A, we briefly review the BMVDR, BLCMV and BMVDR-RTF BFs. In section III-B, we then propose a novel generalized MVDR-based BF.

A. BMVDR, BLCMV, and BMVDR-RTF beamformers

The BMVDR BF is designed to reproduce the target component of both reference microphone signals without distortion, while minimizing the total noise power, i.e.

$$\begin{aligned}&\min_{\mathbf{w}_L}\left\{\mathbf{w}_L^H\mathbf{R}_v\mathbf{w}_L\right\} \text{ s.t. } \mathbf{w}_L^H\mathbf{a} = a_L,\\ &\min_{\mathbf{w}_R}\left\{\mathbf{w}_R^H\mathbf{R}_v\mathbf{w}_R\right\} \text{ s.t. } \mathbf{w}_R^H\mathbf{a} = a_R.\end{aligned} \qquad (10)$$

The well-known solution for the constrained optimization problem in (10) is given by [3]

$$\mathbf{w}_L = \frac{\mathbf{R}_v^{-1}\mathbf{a}}{\mathbf{a}^H\mathbf{R}_v^{-1}\mathbf{a}}a_L^*, \quad \mathbf{w}_R = \frac{\mathbf{R}_v^{-1}\mathbf{a}}{\mathbf{a}^H\mathbf{R}_v^{-1}\mathbf{a}}a_R^*. \qquad (11)$$

This implies that $\mathbf{w}_L$ and $\mathbf{w}_R$ are parallel, i.e. $\mathbf{w}_L = (\text{RTF}_{\text{x}}^{\text{in}})^*\mathbf{w}_R$. Hence, *all* sound sources (i.e. target, interferer, and background noise) are perceived as arriving from the target direction. Therefore, the RTF of the interferer is typically distorted, which is clearly an undesired phenomenon, since the spatial impression of the acoustic scene is altered.

In order to suppress the interferer, while preserving its RTF, the BLCMV BF was proposed in [4], [10], by adding constraints to the BMVDR cost function, i.e.

$$\begin{aligned}&\min_{\mathbf{w}_L}\left\{\mathbf{w}_L^H\mathbf{R}_v\mathbf{w}_L\right\} \text{ s.t. } \mathbf{w}_L^H\mathbf{a} = a_L\ , \mathbf{w}_L^H\mathbf{b} = \eta b_L,\\ &\min_{\mathbf{w}_R}\left\{\mathbf{w}_R^H\mathbf{R}_v\mathbf{w}_R\right\} \text{ s.t. } \mathbf{w}_R^H\mathbf{a} = a_R\ , \mathbf{w}_R^H\mathbf{b} = \eta b_R,\end{aligned} \qquad (12)$$

where the real-valued scaling parameter η, with $0 \leq \eta \leq 1$, sets the amount of interference reduction. Both constrained criteria in (12) can be combined as a general LCMV criterion with multiple constraints on the stacked vector $\mathbf{w}$, i.e.

$$\min_{\mathbf{w}}\left\{\mathbf{w}^H\mathbf{R}_V\mathbf{w}\right\} \quad \text{s.t.} \quad \mathbf{C}_{\text{BLCMV}}^H\mathbf{w} = \mathbf{b}_{\text{BLCMV}}, \qquad (13)$$

where the BLCMV constraint set is given by

$$\mathbf{C}_{\text{BLCMV}} = \begin{bmatrix}\mathbf{a} & \mathbf{0}_M & \mathbf{b} & \mathbf{0}_M\\ \mathbf{0}_M & \mathbf{a} & \mathbf{0}_M & \mathbf{b}\end{bmatrix}, \mathbf{b}_{\text{BLCMV}} = \begin{bmatrix}a_L^*\\ a_R^*\\ \eta b_L^*\\ \eta b_R^*\end{bmatrix}, \qquad (14)$$

with $\mathbf{0}_M$ the M-dimensional all-zero vector, and

$$\mathbf{R}_V = \left[\begin{array}{cc}\mathbf{R}_v & \mathbf{0}_{M\times M}\\ \mathbf{0}_{M\times M} & \mathbf{R}_v\end{array}\right], \qquad (15)$$

with $\mathbf{0}_{M\times M}$ the $M \times M$-dimensional all-zero matrix. The solution for the general LCMV problem in (13) is given by

$$\mathbf{w} = \mathbf{R}_V^{-1}\mathbf{C}_{\text{BLCMV}}\left[\mathbf{C}_{\text{BLCMV}}^H\mathbf{R}_V^{-1}\mathbf{C}_{\text{BLCMV}}\right]^{-1}\mathbf{b}_{\text{BLCMV}}. \qquad (16)$$

Since the BLCMV BF satisfies the distortionless response constraints in (12) for the target, the RTF of the target at the output of the BLCMV BF is equal to the input RTF, i.e.

$$\mathrm{RTF}_{\mathrm{x}}^{\mathrm{out}} = \frac{a_L}{a_R} = \mathrm{RTF}_{\mathrm{x}}^{\mathrm{in}}. \tag{17}$$

In addition, since the BLCMV BF satisfies the constraints in (12) for the interferer, the RTF of the interferer at the output of the BLCMV BF is equal to the input RTF, i.e.

$$\mathrm{RTF}_{\mathrm{u}}^{\mathrm{out}} = \frac{b_L}{b_R} = \mathrm{RTF}_{\mathrm{u}}^{\mathrm{in}}. \tag{18}$$

Hence, the BLCMV BF preserves the RTFs of both the target and the interferer.

In order to preserve the RTF of the interferer, another extension of the BMVDR BF was proposed in [5], referred to as the BMVDR-RTF BF, by adding an RTF constraint to the BMVDR cost function, i.e.

$$\min_{\mathbf{w}} \left\{ \mathbf{w}_L^H \mathbf{R}_v \mathbf{w}_L + \mathbf{w}_R^H \mathbf{R}_v \mathbf{w}_R \right\}$$
$$\text{s.t. } \mathbf{w}_L^H \mathbf{a} = a_L,\ \mathbf{w}_R^H \mathbf{a} = a_R,\ \frac{\mathbf{w}_L^H \mathbf{b}}{\mathbf{w}_R^H \mathbf{b}} = \frac{b_L}{b_R}. \tag{19}$$

Due to the constraints in (19), the BMVDR-RTF BF also preserves the RTFs of both the target and the interferer. Mathematical expressions for the BMVDR-RTF BF will be given in Section III-C.

B. G-BMVDR-RTF beamformer

In this section, we propose an extension of the BMVDR-RTF BF. The constrained optimization problem in (19) is generalized by substituting $\mathbf{R}_v$ with $\tilde{\mathbf{R}}_v = \xi \mathbf{R}_u + \mathbf{R}_n$, where $\xi \geq 0$ is a weighting parameter enabling to trade off interference reduction and background noise reduction. The resulting G-BMVDR-RTF criterion is hence given by

$$\min_{\mathbf{w}} \left\{ \mathbf{w}_L^H \tilde{\mathbf{R}}_v \mathbf{w}_L + \mathbf{w}_R^H \tilde{\mathbf{R}}_v \mathbf{w}_R \right\}$$
$$\text{s.t. } \mathbf{w}_L^H \mathbf{a} = a_L,\ \mathbf{w}_R^H \mathbf{a} = a_R,\ \frac{\mathbf{w}_L^H \mathbf{b}}{\mathbf{w}_R^H \mathbf{b}} = \frac{b_L}{b_R}. \tag{20}$$

Since the RTF constraint is equivalent to the linear constraint $\mathbf{w}_L^H \mathbf{b} - \mathrm{RTF}_{\mathrm{u}}^{\mathrm{in}} \mathbf{w}_R^H \mathbf{b} = 0$, the G-BMVDR-RTF criterion in (20) can be reformulated as

$$\min_{\mathbf{w}} \left\{ \mathbf{w}^H \tilde{\mathbf{R}}_V \mathbf{w} \right\} \ \text{s.t. } \mathbf{C}_{\mathrm{RTF}}^H \mathbf{w} = \mathbf{b}_{\mathrm{RTF}}, \tag{21}$$

where the G-BMVDR-RTF constraint set is given by

$$\mathbf{C}_{\mathrm{RTF}} = \begin{bmatrix} \mathbf{a} & \mathbf{0}_M & \mathbf{b} \\ \mathbf{0}_M & \mathbf{a} & -\mathrm{RTF}_{\mathrm{u}}^{\mathrm{in}} \mathbf{b} \end{bmatrix},\ \mathbf{b}_{\mathrm{RTF}} = \begin{bmatrix} a_L^* \\ a_R^* \\ 0 \end{bmatrix}, \tag{22}$$

and

$$\tilde{\mathbf{R}}_V = \begin{bmatrix} \tilde{\mathbf{R}}_v & \mathbf{0}_{M \times M} \\ \mathbf{0}_{M \times M} & \tilde{\mathbf{R}}_v \end{bmatrix}. \tag{23}$$

Similarly to (16), the solution for the LCMV problem in (21) is given by

$$\mathbf{w} = \tilde{\mathbf{R}}_V \mathbf{C}_{\mathrm{RTF}} \left[\mathbf{C}_{\mathrm{RTF}}^H \tilde{\mathbf{R}}_V^{-1} \mathbf{C}_{\mathrm{RTF}} \right]^{-1} \mathbf{b}_{\mathrm{RTF}}. \tag{24}$$

Since the G-BMVDR-RTF BF satisfies the distortionless response constraints in (20) for the target, the RTF of the target at the output of the G-BMVDR-RTF BF is equal to the input RTF, i.e.

$$\mathrm{RTF}_{\mathrm{x}}^{\mathrm{out}} = \frac{a_L}{a_R} = \mathrm{RTF}_{\mathrm{x}}^{\mathrm{in}}. \tag{25}$$

In addition, since the G-BMVDR-RTF BF satisfies the RTF constraint in (20) for the interferer, the RTF of the interferer at the output of the G-BMVDR-RTF BF is equal to the input RTF, i.e.

$$\mathrm{RTF}_{\mathrm{u}}^{\mathrm{out}} = \frac{b_L}{b_R} = \mathrm{RTF}_{\mathrm{u}}^{\mathrm{in}}. \tag{26}$$

Hence, similarly to the BLCMV BF, the G-BMVDR-RTF BF preserves the RTFs of both the target and the interferer.

As a result of the distortionless response constraints for the target in (20) and the rank-1 structure of $\mathbf{R}_x$ in (3), the output PSD of the target is equal to

$$\mathbf{w}_L^H \mathbf{R}_x \mathbf{w}_L + \mathbf{w}_R^H \mathbf{R}_x \mathbf{w}_R = P_s(|a_L|^2 + |a_R|^2). \tag{27}$$

Therefore, the criterion in (20) is equivalent to

$$\min_{\mathbf{w}} \left\{ \frac{\mathbf{w}_L^H \tilde{\mathbf{R}}_v \mathbf{w}_L + \mathbf{w}_R^H \tilde{\mathbf{R}}_v \mathbf{w}_R}{\mathbf{w}_L^H \mathbf{R}_x \mathbf{w}_L + \mathbf{w}_R^H \mathbf{R}_x \mathbf{w}_R} \right\} \text{s.t. } \mathbf{C}_{\mathrm{RTF}}^H \mathbf{w} = \mathbf{b}_{\mathrm{RTF}}. \tag{28}$$

Using (7) and (8), the criterion in (28) can be interpreted as

$$\min_{\mathbf{w}} \left\{ \frac{\xi}{\mathrm{SIR}^{\mathrm{out}}} + \frac{1}{\mathrm{SNR}^{\mathrm{out}}} \right\} \text{s.t. } \mathbf{C}_{\mathrm{RTF}}^H \mathbf{w} = \mathbf{b}_{\mathrm{RTF}}. \tag{29}$$

Hence, by substituting (9) into (29) and rearranging terms, (29) can be rewritten as

$$\min_{\mathbf{w}} \left\{ \frac{\xi}{\mathrm{SINR}^{\mathrm{out}}} + \frac{1-\xi}{\mathrm{SNR}^{\mathrm{out}}} \right\} \text{s.t. } \mathbf{C}_{\mathrm{RTF}}^H \mathbf{w} = \mathbf{b}_{\mathrm{RTF}}, \tag{30}$$

i.e. the weighting parameter ξ trades off output SINR and output SNR.

C. Special cases of the G-BMVDR-RTF beamformer

Three special cases of the G-BMVDR-RTF BF deserve attention, namely $\xi = 1$, $\xi = 0$, and $\xi \to \infty$.

Case 1: By setting $\xi = 1$, the G-BMVDR-RTF BF reduces to the BMVDR-RTF BF in (19). In this case the constrained criterion in (30) is equivalent to

$$\max_{\mathbf{w}} \left\{ \mathrm{SINR}^{\mathrm{out}} \right\} \text{s.t. } \mathbf{C}_{\mathrm{RTF}}^H \mathbf{w} = \mathbf{b}_{\mathrm{RTF}}, \tag{31}$$

hence resulting in the constrained maximum binaural SINR criterion subject to a distortionless response constraint for the target and an RTF preservation constraint for the interferer. Among all BFs that are distortionless with respect to the target and preserve the binaural cues of the interferer, the proposed G-BMVDR-RTF BF with $\xi = 1$ is optimal in terms of SINR, and is hence referred to as the maxSINR-RTF BF.

Since the BLCMV BF satisfies the same constraints as in (31), the output SINR of the BLCMV BF is always lower than or equal to the output SINR of the maxSINR-RTF BF, i.e.

$$\text{SINR}^{\text{out}}_{\text{SINR-RTF}} \geq \text{SINR}^{\text{out}}_{\text{BLCMV}}. \tag{32}$$

In addition, since the reference microphone signals at the input to the binaural BF also satisfy the same constraints as in (31), the output SINR of the maxSINR-RTF BF is always larger than or equal to the input SINR, i.e.

$$\text{SINR}^{\text{out}}_{\text{SINR-RTF}} \geq \text{SINR}^{\text{in}}. \tag{33}$$

Case 2: By setting $\xi = 0$, the constrained criterion in (30) is equivalent to

$$\max_{\mathbf{w}} \left\{ \text{SNR}^{\text{out}} \right\} \text{s.t. } \mathbf{C}^H_{\text{RTF}} \mathbf{w} = \mathbf{b}_{\text{RTF}}, \tag{34}$$

hence resulting in the constrained maximum binaural SNR criterion subject to a distortionless response constraint for the target and an RTF preservation constraint for the interferer. Therefore, among all BFs that are distortionless with respect to the target and preserve the binaural cues of the interferer, the proposed G-BMVDR-RTF BF with $\xi = 0$ is optimal in terms of SNR, and is hence referred to as the maxSNR-RTF BF.

Since the BLCMV BF satisfies the same constraints as in (34), the output SNR of the BLCMV BF is always lower than or equal to the output SNR of the maxSNR-RTF BF, i.e.

$$\text{SNR}^{\text{out}}_{\text{SNR-RTF}} \geq \text{SNR}^{\text{out}}_{\text{BLCMV}}. \tag{35}$$

In addition, since the reference microphone signals at the input to the binaural BF also satisfy the same constraints as in (34), the output SNR of the maxSNR-RTF BF is always larger than or equal to the input SNR, i.e.

$$\text{SNR}^{\text{out}}_{\text{SNR-RTF}} \geq \text{SNR}^{\text{in}}. \tag{36}$$

Case 3: By setting $\xi \to \infty$, minimizing the criterion in (29) can only be obtained if SIR $\to \infty$, corresponding to a perfect null directed toward the interferer, i.e. $\mathbf{w}^H_L \mathbf{b} = 0$ and $\mathbf{w}^H_R \mathbf{b} = 0$. Hence, the constrained criterion reduces to

$$\begin{aligned} &\min_{\mathbf{w}} \left\{ \mathbf{w}^H_L \mathbf{R}_n \mathbf{w}_L + \mathbf{w}^H_R \mathbf{R}_n \mathbf{w}_R \right\} \\ &\text{s.t. } \mathbf{w}^H_L \mathbf{a} = a_L, \mathbf{w}^H_R \mathbf{a} = a_R, \mathbf{w}^H_L \mathbf{b} = 0, \mathbf{w}^H_R \mathbf{b} = 0. \end{aligned} \tag{37}$$

The G-BMVDR-RTF BF with $\xi = \infty$ hence maximizes the SIR and is equivalent to the BLCMV BF in (12) with $\eta = 0$.

Note that implementing all considered BFs requires knowledge of the RTF vectors, defined as the ATF vectors **a** and **b** normalized by the respective ATFs of the reference microphones, rather than the ATF vectors. Estimation procedures for all components of the BFs exist, but are beyond the scope of the current contribution. Practical aspects concerning these estimation procedures, as well as a generalization to multiple interferers, are described in [10].

IV. Experimental Validation

In this section, the performance of the BMVDR, G-BMVDR-RTF(ξ), and BLCMV(η) BFs is evaluated for various weighting and scaling parameters ξ and η. To verify the theoretical analysis, we used measured Behind-The-Ear impulse responses (BTE-IRs) from [12] and artificial sources. All experiments were carried out using $M = 2$ microphones, i.e. one microphone on each hearing aid, at a sampling frequency of 16 kHz. The acoustic scenario comprised one target at $\theta_x = -30°$ and 1 m from the artificial head, one interferer at different angles[1] (also 1 m from the artificial head) and diffuse background noise. The angle $\theta = 0^o$ corresponds to the frontal direction and $\theta = 90^o$ corresponds to the right side of the head. The reverberation time was approximately 400 ms. The PSDs of the target and the interferer P_s and P_u were calculated from two different speech signals (Welch method using FFT size of 512 and Hann window). For the background noise a cylindrically isotropic noise field was assumed, where the spatial coherence matrix was calculated using anechoic ATFs of the same database and the PSD of the background noise was equal to the PSD of speech-shaped noise. The wide-band input SIR and SNR were both set to 0 dB. The G-BMVDR-RTF(ξ) was evaluated for weighting parameters ξ equal to 0, 0.1, 1, 100, and 10^6. The BLCMV(η) was evaluated for scaling parameters η equal to 0, 0.1, and 0.5.

Fig. 1 depicts the performance of the considered BFs in terms of wide-band SINR, SIR, and SNR, which are defined as the ratio of the average PSDs of the target and the total noise, interferer and background noise, respectively, over all frequency bands. Please note that the y-axis scaling of the left and the right sub-figures is different as a result of a huge performance difference for different values of the weighting parameter ξ and for the BLCMV BF. It is observed that the BMVDR BF outperforms all other considered BFs in terms of SNR and SINR. However, since the BMVDR BF does not preserve the binaural cues of the interferer, we will mainly focus on the comparison between all other cue preserving BFs. Fig. 1(a) and Fig. 1(b) show that the G-BMVDR-RTF(1) BF, i.e. the maxSINR-RTF BF, outperforms all other BFs in terms of SINR. Fig. 1(c) and Fig. 1(d) show that the G-BMVDR-RTF(0) BF, i.e. the maxSNR-RTF BF, outperforms all other BFs in terms of SNR. While for ξ equal to 0, 0.1 and 1, the SINR and SNR gains are higher than zero, for larger values of ξ, the relative importance of interference reduction increases (cf. (29)), resulting in SINR and SNR gains that are possibly even lower than zero. Fig. 1(e) and Fig. 1(f) show that for ξ equal to or larger than 1, the G-BMVDR-RTF BF outperforms the BMVDR BF in terms of SIR. As ξ increases, the SIR becomes higher. For very large ξ, e.g., $\xi = 10^6$, the G-BMVDR-RTF BF achieves the same SINR and SNR gains as the BLCMV BF for $\eta = 0$.

Note that the dimension of the stacked vector $\mathbf{w}$ in the examined scenario is equal to $2M = 4$. In order to maintain degrees of freedom for the minimization in (12) and (20), it is

[1] Note that the interferer at $-30°$ was not evaluated.

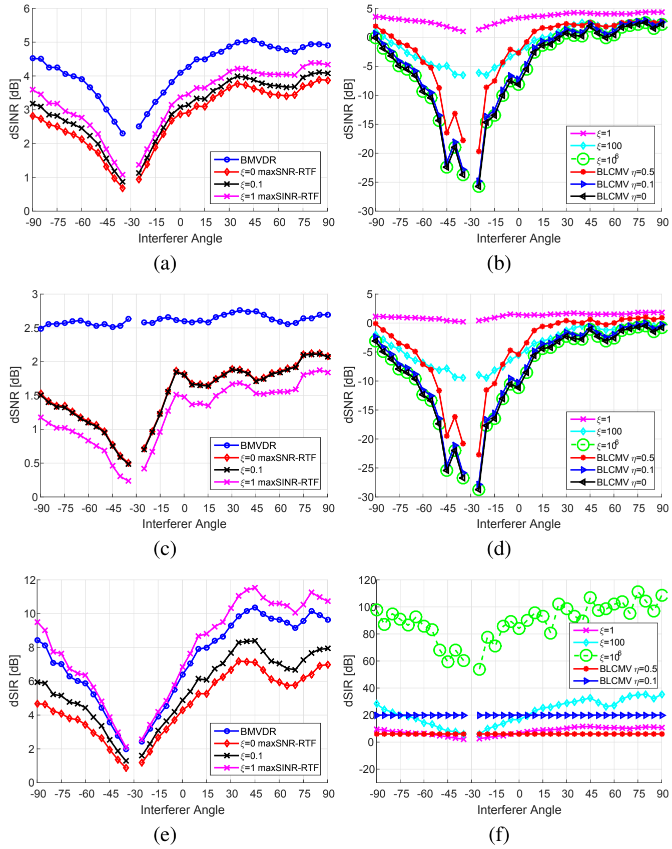

Fig. 1. Wide-band binaural SINR, SIR and SNR gains for the BMVDR, G-BMVDR-RTF and BLCMV BFs for various weighting and scaling parameters as a function of interferer angle (office environment, target at $-30°$, $M = 2$).

required that the number of constraints is lower than $2M = 4$. While the number of the constraints in the G-BLCMV-RTF constraint set in (22) is equal to 3, there are 4 constraints in the BLCMV constraint set in (14). Hence, in the BLCMV BF there are no available degrees of freedom, resulting in SINR and SNR gains that are possibly lower than zero (cf. Fig. 1(b) and Fig. 1(d)). Fig. 1(f) show that the SIR for the BLCMV BF is inversely proportional to η.

V. Discussion and Conclusion

In this paper, a novel binaural noise reduction algorithm based on the binaural MVDR criterion was introduced, designed to trade off the amount of interference reduction and noise reduction while preserving the binaural cues of the interferer. The obtained G-BMVDR-RTF BF is related to the BLCMV BF proposed in [4], [10]. Both BFs preserve the binaural cues of the target and the interferer and are controlled by trade off parameters. However, the role of these parameters is different.

For the BLCMV BF, the scaling parameter η controls the amount of interference reduction, whereas the amount of noise reduction cannot be controlled. E.g., setting η to a low value for a scenario in which the interferer is close to the target, may lead to noise amplification.

For the G-BMVDR-RTF BF, the weighting parameter ξ defines the relative importance of the interference reduction and the noise reduction in the generalized cost function. As ξ increases, the interference reduction is higher at the expense of lower noise reduction. However, the exact amount of interference reduction cannot be controlled. When ξ is set to one, the obtained BMVDR-RTF BF is optimal in terms of SINR. Moreover, the SINR at the output of the BMVDR-RTF BF is always larger than or equal to the input SINR (which cannot be guaranteed for the BLCMV BF). Similarly, when ξ is set to zero, the obtained G-BMVDR-RTF BF is optimal in terms of SNR. Moreover, the SNR at the output of the BMVDR-RTF BF is always larger than or equal to the input SNR (which again cannot be guaranteed for the BLCMV BF). There is a close relation between the weighting parameter of the G-BMVDR-RTF BF and the scaling parameter of the BLCMV BF, such that in the extreme case obtained by setting $\xi = \infty$ and $\eta = 0$, the G-BMVDR-RTF and the BLCMV BFs coincide. The relation for other values of ξ and η is left to future research.

References

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

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

[3] S. Doclo, S. Gannot, M. Moonen, and A. Spriet, "Acoustic beamforming for hearing aid applications," *Handbook on Array Processing and Sensor Networks*, pp. 269–302, 2010.

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

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

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

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

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

[9] E. Hadad, D. Marquardt, S. Doclo, and S. Gannot, "Extensions of the binaural MWF with interference reduction preserving the binaural cues of the interfering source," in *Proc. IEEE Intl. Conf. on Acous., Sp. and Sig. Proc. (ICASSP)*, Shanghai, China, March 2016.

[10] E. Hadad, S. Doclo, and S. Gannot, "The binaural LCMV beamformer and its performance analysis," *IEEE Trans. Audio, Speech and Language Proc.*, vol. 24, no. 3, pp. 543–558, March 2016.

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

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