Binaural LCMV Beamforming with Partial Noise Estimation

Nico Gößling, Student Member, IEEE, Elior Hadad, Member, IEEE, Sharon Gannot, Senior Member, IEEE and Simon Doclo, Senior Member, IEEE

Abstract—Besides suppressing all undesired signal components, another important task of binaural noise reduction algorithms is the preservation of the spatial impression of the acoustical scene, which can be achieved by preserving the binaural cues of all signal components. While the well-known binaural minimum variance distortionless response (BMVDR) beamformer at least can preserve the binaural cues of a single desired source, several extensions have been proposed to additionally preserve the binaural cues of interfering sources and diffuse noise. The binaural linearly-constrained minimum variance (BLCMV) beamformer uses additional constraints to preserve the binaural cues of interfering sources and enables a direct scaling using an interference scaling parameter. The BMVDR with partial noise estimation (BMVDR-N) is aimed at preserving a scaled version of the diffuse noise to partially preserve its interaural coherence and hence its perceived diffuseness. In this paper, we propose to combine both extensions of the BMVDR, leading to the BLCMV with partial noise estimation (BLCMV-N). It is shown that the BLCMV-N can be seen as a mixture of the noisy input signal and the output of a BLCMV that uses an adjusted interference scaling parameter. A theoretical analysis and comparison between the BMVDR, its extensions and the proposed BLCMV-N in terms of noise reduction and binaural cue preservation performance is provided. Experimental results and results of a subjective listening test show that the BLCMV-N is able to preserve the spatial impression of an interfering source, i.e., like the BLCMV, and yields a trade-off between noise reduction and binaural cue preservation of diffuse noise, i.e., like the BMVDR-N.

Index Terms—Binaural cues, MVDR, LCMV, partial noise estimation, binaural noise reduction

I. INTRODUCTION

ALGORITHMS for noise reduction in head-mounted assistive hearing devices (e.g., hearing aids, cochlear implants and headsets) are mandatory to improve speech intelligibility in noisy acoustical scenarios. In addition to the reduction of all undesired sources, another important objective of binaural noise reduction algorithms is the preservation of the so-called binaural cues of all present sources [1]. Binaural cues consist of differences in the signals arriving at the two ears of a listener and are crucial for the spatial impression of the acoustical scene [2]. Spatial hearing is not only important for the spatial awareness of the listener, but can also lead to improved understanding of a desired speaker as a result of binaural unmasking [3]. For coherent sources such as speakers, the most descriptive binaural cue is the interaural transfer function (ITF), from which the interaural level difference (ILD) and the interaural time difference (ITD) can be calculated [4]. For diffuse (incoherent) sound fields such as the noise produced by multiple talkers in a reverberant environment, the most descriptive binaural cue is the interaural coherence (IC).

While bilateral processing strategies use only the ipsilateral microphones per hearing device and can lead to a distortion of the binaural cues [5], [6], binaural processing strategies use the microphones on both sides of the head and are promising, because the spatial information on both sides of the head can be exploited. Binaural noise reduction algorithms can be split into two main processing paradigms. The first paradigm is based on binaural spectral post-filtering, where the same real-valued spectro-temporal gain is applied to two microphones (one on each side of the head), which intrinsically preserves the binaural cues of all sound sources [7]–[13]. The second paradigm, the focus of this paper, is based on binaural spatial filtering, where the microphones on the two sides of the head are processed by different complex-valued spatial filters, where the desired source is estimated in so-called reference microphones on each side of the head [11], [14]–[16]. The number of degrees-of-freedom available for noise reduction is larger in this second paradigm, but usually only the binaural cues of a desired source are preserved, whereas to also preserve the binaural cues of other sources, typically a trade-off between noise reduction and binaural cue preservation has to be applied.

A spatial filter that at least can preserve the binaural cues of one desired source is the binaural minimum variance distortionless response (BMVDR) beamformer [1], [4], [15], [17]. While the binaural cues of the desired source are preserved, the undesired sources are spatially shifted to the position of the desired source and hence are perceived as coming from the same direction. To allow in addition partial preservation of the binaural cues of diffuse noise, the BMVDR with partial noise estimation (BMVDR-N) was proposed [18], [19]. The BMVDR-N can be seen as a mixture of the BMVDR output and the noisy reference microphone signals and hence provides a trade-off between noise reduction and binaural cue preservation. For a diffuse noise, this trade-off can be psycho-acoustically motivated such that a listener should not perceive any spatial difference in the processed signal, compared to the input [18]. To preserve the binaural cues of coherent interfering sources, additional constraints can be used, resulting in the binaural linearly constrained minimum variance (BLCMV) beamformer [20]–[24]. The BLCMV preserves the binaural cues of the desired source and all additionally constrained interfering sources and enables a direct scaling of each interfering source by an interference scaling parameter. For every additionally constrained interfering source, there
is one degree-of-freedom less available for noise reduction, which results in a trade-off between noise reduction and the number of constrained interfering sources. Further, the number of possibly constrained interfering sources is limited by the number of microphones used for the spatial filtering.

In this paper, we combine the two approaches and introduce the BLCMV with partial noise estimation (BLCMV-N). It is shown that the solution can be considered a mixture of the BLCMV output and the noisy reference microphone signals, if an adjusted interfering scaling parameter is used in the BLCMV. We analytically derive the performance differences as compared to the BLCMV, as described in [22], and provide a general comparison to the BMVDR and BMVDR-N. The derivations are then validated using measurements of anechoic room impulse responses and quantified in a more realistic scenario using reverberant room impulse responses and recordings. Finally, we provide the results of a subjective listening test that prove the effectiveness of the proposed approach in terms of binaural cue preservation. The results show that the BLCMV-N is able to preserve the binaural cues and hence the spatial impression of an interfering source, i.e. like the BLCMV, and yields a trade-off between noise reduction and binaural cue preservation of diffuse noise, i.e. like the BMVDR-N.

The paper is structured as follows. In Section II the general configuration and notation is introduced that is used throughout the paper. In Section III the BMVDR, the BMVDR-N and the BLCMV are briefly reviewed. In Section IV the BLCMV-N is derived and compared to the BLCMV. Further, it is shown under which conditions the BLCMV-N is equal to the BMVDR-N. In Section V the performance of the BLCMV-N is theoretically investigated. In Section VI first a validation using anechoic measurements is shown, followed by results of experiments using reverberant measurements and recordings. Additionally, the results of a subjective listening test are shown to compare the binaural cue preservation performance and hence the preservation of the spatial impression for all the considered algorithms.

II. CONFIGURATION AND NOTATION

In this section, the notation and system configuration is introduced. In Section II-A the signal model is introduced, and interaural criteria and performance measures are introduced in Sections II-B and II-C respectively.

A. Signal Model

Consider the binaural hearing device configuration depicted in Figure 1 with \( M_L \) microphones on the left side, \( M_R \) microphones on the right side and \( M = M_L + M_R \) microphones in total. Consider a scenario that consists of a desired source and an interfering source in a noisy and reverberant environment, i.e. dual source scenario. The input signal of the \( m \)-th microphone \( Y_m(\omega) \) in the frequency domain is equal to

\[
Y_m(\omega) = X_m(\omega) + U_m(\omega) + N_m(\omega) \in \mathbb{C},
\]

with \( X_m(\omega) \) the desired source component, \( U_m(\omega) \) the interfering source component, \( N_m(\omega) \) the diffuse noise component and \( V_m(\omega) \) the undesired component. For the sake of conciseness, we omit the frequency index \( \omega \) in the remainder of the paper. When all the microphone signals are stacked in a vector, the input vector \( \mathbf{y} \) is equal to

\[
\mathbf{y} = [Y_1, \ldots, Y_{M_L}, Y_{M_L+1}, \ldots, Y_M]^T \in \mathbb{C}^{M \times 1},
\]

which further can be written as

\[
\mathbf{y} = \mathbf{x} + \mathbf{u} + \mathbf{n},
\]

where the vectors \( \mathbf{x}, \mathbf{u}, \mathbf{n} \) and \( \mathbf{v} \) are defined similarly to \( \mathbf{y} \) in (2). Without loss of generality, the first microphone on each side is defined as the so-called reference microphone. To simplify the notation, the reference microphone signals \( Y_1 \) and \( Y_{M_L+1} \) are denoted as \( Y_L \) and \( Y_R \), respectively, and are equal to

\[
Y_L = \mathbf{e}_L^T \mathbf{y}, \quad Y_R = \mathbf{e}_R^T \mathbf{y},
\]

where \( \mathbf{e}_L \) and \( \mathbf{e}_R \) are zero vectors with \( \mathbf{e}_L(1) = 1 \) and \( \mathbf{e}_R(M_L+1) = 1 \). Hence, the reference microphone signals are equal to

\[
Y_L = X_L + U_L + N_L, \quad Y_R = X_R + U_R + N_R.
\]

Consider an acoustic scenario with one desired source \( S_x \) and one interfering source \( S_u \). The desired source component vector \( \mathbf{x} \) and the interfering source component vector \( \mathbf{u} \) are equal to

\[
\mathbf{x} = S_x \mathbf{a}, \quad \mathbf{u} = S_u \mathbf{b},
\]

with \( \mathbf{a} \) and \( \mathbf{b} \) the acoustic transfer function (ATF) vectors, containing the acoustic transfer functions between the microphones and the desired source and the interfering source, respectively. The ATF vectors are equal to

\[
\mathbf{a} = [A_L, \ldots, A_{M_L}, A_R, \ldots, A_M]^T, \quad \mathbf{b} = [B_L, \ldots, B_{M_L}, B_R, \ldots, B_M]^T.
\]

The covariance matrix of the desired source \( \mathbf{R}_x \), the covariance matrix of the interfering source \( \mathbf{R}_u \) and the covariance matrix of the diffuse noise \( \mathbf{R}_n \) are equal to

\[
\mathbf{R}_x = \mathcal{E}\{\mathbf{x}\mathbf{x}^H\} = \Phi_x \mathbf{a} \mathbf{a}^H, \quad \mathbf{R}_u = \mathcal{E}\{\mathbf{u}\mathbf{u}^H\} = \Phi_u \mathbf{b} \mathbf{b}^H, \quad \mathbf{R}_n = \mathcal{E}\{\mathbf{n}\mathbf{n}^H\},
\]
with $\mathcal{E}\{\cdot\}$ the expectation operator, $\Phi_x = \mathcal{E}\{|S_x|^2\}$ the power spectral density (PSD) of the desired source and $\Phi_u = \mathcal{E}\{|S_u|^2\}$ the PSD of the interfering source. By assuming statistical independence between $x$, $u$ and $n$, the noisy input covariance matrix $R_y = \mathcal{E}\{yy^H\}$ can be written as
\[
R_y = R_x + R_u + R_n + R_w, \tag{12}
\]
with $R_x$ the covariance matrix of the undesired component. The left and right output signals $Z_L$ and $Z_R$ are obtained by filtering and summing all $M$ microphone signals using the (complex-valued) filter vectors $w_L$ and $w_R$, i.e.
\[
Z_L = w_L^H y, \quad Z_R = w_R^H y. \tag{13}
\]
To provide a concise notation, we define
\[
\gamma_a = a^H R_n^{-1} w, \quad \gamma_b = b^H R_n^{-1} b, \quad \gamma_{ab} = a^H R_n^{-1} b. \tag{14-16}
\]
Further, the squared cosine of the generalized angle between $a$ and $b$ is defined as
\[
\Psi = \frac{\gamma_{ab}^2}{\gamma_a \gamma_b}, \tag{17}
\]
where it can be shown that $0 \leq \Psi \leq 1$ using the Cauchy-Schwarz inequality.

**B. Interaural Criteria/Binaural Cues**

Binaural cues are used by the listener to localize sound sources and to perceive their width and diffuseness and hence provide the cues needed for spatial hearing [2]. They are described by interaural differences, i.e. differences of the signal arriving at the left and the right ear. For coherent sources such as speakers, the most descriptive binaural cue is the ITF. The input and output ITFs of the desired source can be calculated as [4]
\[
\text{ITF}_{x}^{\text{in}} = \frac{e_L^T R_x e_L}{e_L^T R_x e_R}, \quad \text{ITF}_{x}^{\text{out}} = \frac{w_L^H R_x w_L}{w_R^H R_x w_R}. \tag{18}
\]
The same definitions can be applied for the noisy input, the interfering source component, the diffuse noise component and the undesired component by substituting $R_y$, $R_n$, $R_u$ or $R_v$ for $R_x$, respectively. The ILD and ITD cues can then be calculated as [4]
\[
\text{ILD} = 20 \log_{10}(|\text{ITF}|), \quad \text{ITD} = \frac{\angle \text{ITF}}{\omega}, \tag{19}
\]
with $\angle$ the unwrapped phase. For diffuse (incoherent) sound fields, the most descriptive binaural cue is the IC. The input IC of the diffuse noise component is defined as [27]
\[
I_{c_{\text{in}}} = \frac{e_L^T R_x e_R}{\sqrt{(e_L^T R_x e_L)(e_R^T R_x e_R)}}, \tag{20}
\]
while the output IC of the diffuse noise component is defined as [27]
\[
I_{c_{\text{out}}} = \frac{w_L^H R_x w_R}{\sqrt{(w_L^H R_x w_L)(w_R^H R_x w_R)}}. \tag{21}
\]
Because the IC is complex-valued, the magnitude-squared coherence (MSC) is often used. The input and output MSCs can be calculated by
\[
\text{MSC}^{\text{in}} = |I^\text{in}_{\text{co}}|^2, \quad \text{MSC}^{\text{out}} = |I^\text{out}_{\text{co}}|^2. \tag{22}
\]
An MSC of 1 is attributed to a coherent source perceived as a distinct point source, while lower values lead to broader or even diffuse perception [2].

**C. Performance Measures**

The left and right input PSDs of the desired source component are given by
\[
\Phi_{x,L}^{\text{in}} = e_L^T R_x e_L, \quad \Phi_{x,R}^{\text{in}} = e_R^T R_x e_R. \tag{23}
\]
The left and right output PSDs of the desired source component are given by
\[
\Phi_{x,L}^{\text{out}} = w_L^H R_x w_L, \quad \Phi_{x,R}^{\text{out}} = w_R^H R_x w_R. \tag{24}
\]
The same definitions can be applied for the noisy input, the interfering source component, the diffuse noise component and the undesired component by substituting $R_y$, $R_n$, $R_u$ or $R_v$ for $R_x$, respectively.

The left and right input signal-to-noise ratios (SNRs) are defined as the ratio of the left and right input PSDs of the desired source and diffuse noise components, respectively, i.e.
\[
\text{SNR}_{x,L}^{\text{in}} = \frac{\Phi_{x,L}^{\text{in}}}{\Phi_{n,L}}, \quad \text{SNR}_{x,R}^{\text{in}} = \frac{\Phi_{x,R}^{\text{in}}}{\Phi_{n,R}}. \tag{25}
\]
The left and right input SNRs are defined as the ratio of the left and right input PSDs of the desired source and diffuse noise components, respectively, i.e.
\[
\text{SNR}_{x,L}^{\text{out}} = \frac{\Phi_{x,L}^{\text{out}}}{\Phi_{n,L}}, \quad \text{SNR}_{x,R}^{\text{out}} = \frac{\Phi_{x,R}^{\text{out}}}{\Phi_{n,R}}. \tag{26}
\]
The left and right output SNRs are defined as the ratio of the left and right output PSDs of the desired source and interfering source components, respectively, i.e.
\[
\text{SINR}_{x,L}^{\text{in}} = \frac{\Phi_{x,L}^{\text{in}}}{\Phi_{u,L}}, \quad \text{SINR}_{x,R}^{\text{in}} = \frac{\Phi_{x,R}^{\text{in}}}{\Phi_{u,R}}. \tag{27}
\]
The left and right output SNRs are defined as the ratio of the left and right output PSDs of the desired source and interfering source components, respectively, i.e.
\[
\text{SINR}_{x,L}^{\text{out}} = \frac{\Phi_{x,L}^{\text{out}}}{\Phi_{u,L}}, \quad \text{SINR}_{x,R}^{\text{out}} = \frac{\Phi_{x,R}^{\text{out}}}{\Phi_{u,R}}. \tag{28}
\]
The left and right output SINRs are defined as the ratio of the left and right output PSDs of the desired source and undesired components, respectively, i.e.
\[
\text{SINR}_{x,L}^{\text{in}} = \frac{\Phi_{x,L}^{\text{in}}}{\Phi_{v,L}}, \quad \text{SINR}_{x,R}^{\text{in}} = \frac{\Phi_{x,R}^{\text{in}}}{\Phi_{v,R}}. \tag{29}
\]
The left and right output SINRs are defined as the ratio of the left and right output PSDs of the desired source and undesired components, respectively, i.e.
\[
\text{SINR}_{x,L}^{\text{out}} = \frac{\Phi_{x,L}^{\text{out}}}{\Phi_{v,L}}, \quad \text{SINR}_{x,R}^{\text{out}} = \frac{\Phi_{x,R}^{\text{out}}}{\Phi_{v,R}}. \tag{30}
\]
III. BINAURAL NOISE REDUCTION

In this section, we briefly review three state-of-the-art binaural noise reduction algorithms: First, the BMVDR [1, 4, 15] using one constraint to preserve the desired source component in the reference microphones, is introduced in Section III-A; second, the BMVDR-N [18], additionally preserving a scaled version of the diffuse noise component in the reference microphones, is described in Section III-B; and third, the BLCMV [22] as a more general version of the BMVDR, using an additional constraint to preserve a scaled version of the interfering source component in the reference microphones, is described in Section III-C. For the sake of conciseness, we show only the equations for the left hearing device, as previously denoted by the subscript L. All the equations that follow can of course also be formulated for the right hearing device by changing the subscript to R.

A. The Binaural MVDR Beamformer (BMVDR)

The BMVDR [1, 15, 19] minimizes the output PSD of the diffuse noise component while preserving the desired source component in the reference microphones. The constrained optimization problem for the left filter is given by

$$
\min_{w_L} w_L^H R_n w_L \quad \text{subject to} \quad w_L^H a = A_L .
$$

(31)

The filter solving the constrained optimization problem in (31) is equal to [1, 17]

$$
R_{BMVDR,L} = \frac{R_n^{-1} a}{\gamma_a} A_L^* .
$$

(32)

It has been shown that the BMVDR preserves the binaural cues of the desired source [4], i.e.

$$
\text{ITF}_{BMVDR,L} = \frac{A_L}{A_L} = \text{ITF}_{in} .
$$

(33)

The disadvantage is that the BMVDR distorts the binaural cues of the interfering source and the diffuse noise, such that

$$
\text{ITF}_{BMVDR,u} = \frac{A_L}{A_R} = \text{ITF}_{in} ,
$$

(34)

$$
\text{ITF}_{BMVDR,n} = \frac{A_L}{A_R} = \text{ITF}_{in} , \quad \text{MSC}_{BMVDR,n} = 1 .
$$

(35)

Hence, the interfering source and the diffuse noise are perceived as coming from the same position as the desired source at the output of the BMVDR.

The output SNR of the BMVDR is equal to [4, 15]

$$
\text{SNR}_{BMVDR,L} = \Phi_x a^H R_n^{-1} a = \Phi_x \gamma_a .
$$

(36)

Please note that the BMVDR can also be defined using the covariance matrix of the undesired component $R_v$, although $R_v$ is considerably more difficult to estimate in practice than $R_n$ and cannot easily be modelled. Further, using the noisy covariance matrix $R_y$ is sometimes referred to as the binaural minimum power distortionless response (MPDR) beamformer [28] and can lead to cancellation of the desired source, if estimation errors are present.

B. The Binaural MVDR-N Beamformer (BMVDR-N)

In addition to preserving the desired source component, the BMVDR-N [18, 19] preserves a scaled version of the diffuse noise component in the reference microphones. In literature, this approach is often referred to as partial noise estimation [15, 18, 29]. The constrained optimization problem for the left filter is given by

$$
\min_{w_L} E\{|w_L^H n - \eta N_L|^2\} \quad \text{subject to} \quad w_L^H a = A_L ,
$$

(37)

with $0 \leq \eta \leq 1$ the noise scaling parameter. The filter solving (37) is given by [18]

$$
w_{BMVDR-N,L} = \eta e_L + (1 - \eta) w_{BMVDR,L} .
$$

(38)

The BMVDR-N can be seen as a mix of the reference microphone signal (scaled with $\eta$) and the BMVDR output (scaled with $1 - \eta$). For $\eta = 0$, the BMVDR-N is equal to the BMVDR, whereas for $\eta = 1$, the BMVDR-N output is equal to the reference microphone signal.

It has been shown that the BMVDR-N preserves the binaural cues of the desired source [18], i.e.

$$
\text{ITF}_{BMVDR-N,x} = \frac{A_L}{A_R} = \text{ITF}_{in} .
$$

(39)

It can easily be shown that the binaural cues of the interfering source are preserved if $\eta = 1$, whereas for $\eta = 0$ the binaural cues of the interfering source are equal to the binaural cues of the desired source (as applied for the BMVDR), i.e.

$$
\text{ITF}_{BMVDR-N,u} = \begin{cases} 
\frac{A_L}{A_R} = \text{ITF}_{in} & \text{for } \eta = 1 \\
\frac{A_L}{A_R} = \text{ITF}_{in} & \text{for } \eta = 0 
\end{cases} .
$$

(40)

For the diffuse noise, it was shown in [18] that for $\eta = 1$ the output MSC is equal to the input MSC, whereas for $\eta = 0$ the output MSC is equal to 1, as for the BMVDR, i.e.

$$
\text{MSC}_{BMVDR-N,n} = \begin{cases} 
\text{MSC}_{in} & \text{for } \eta = 1 \\
1 & \text{for } \eta = 0 
\end{cases} .
$$

(41)

A larger noise scaling parameter $\eta$ hence leads to better binaural cue preservation of the diffuse noise but obviously leads to smaller noise reduction because the (noisy) reference microphone signals are increasingly mixed with the BMVDR output. The output SNR of the BMVDR-N is hence always smaller than or equal to the output SNR of the BMVDR [18], i.e.

$$
\text{SNR}_{BMVDR-N,L} \leq \text{SNR}_{BMVDR,L} .
$$

(42)

For the BMVDR-N, frequency-dependent and psycho-physically motivated noise scaling parameters were proposed in the literature [18, 27], such that a human listener should not perceive any spatial difference of diffuse noise at the input and output of the BMVDR-N. Another approach is to set $\eta = 0.2$ for all frequencies [4], which appears to be a good compromise between noise reduction performance and binaural cue preservation of the diffuse noise.
C. The Binaural LCMV Beamformer (BLCMV)

In addition to preserving the desired source component, the BLCMV uses an additional constraint to preserve a scaled version of the interfering source component. The BLCMV is aimed at minimizing the output PSD of the diffuse noise component subject to a constraint set \( \{ \eta, \delta \} \), i.e.,

\[
\min_{w_{L}} w_{L}^{H} R_{n} w_{L} \quad \text{subject to} \quad C^{H} w_{L} = g_{L}, \tag{43}
\]

where the response vector \( g_{L} \) is defined as

\[
g_{L} = \begin{bmatrix} A_{L}^{*} \\ \delta B_{L}^{*} \end{bmatrix}, \tag{44}
\]

with \( 0 \leq \delta \leq 1 \) the interference scaling parameter. In [22], also a scaling parameter for the desired source was used. Here, we assume that a distortionless response is wanted for the desired source. The constraint matrix \( C \) is defined as

\[
C = \begin{bmatrix} a \ b \end{bmatrix}, \tag{45}
\]

including the ATF vectors of the desired source and the interfering source as defined in (7) and (8), respectively. The filter solving (43) is equal to (22)

\[
w_{BLCMV,L} = R_{n}^{-1} C \left( C^{H} R_{n}^{-1} C \right)^{-1} g_{L}. \tag{46}
\]

It was shown in [22], that the BLCMV preserves the binaural cues of both the desired source and the interfering source, i.e.

\[
\text{ITF}_{BLCMV,x}^{\text{out}} = \frac{A_{L}}{A_{R}} = \text{ITF}_{x}^{\text{in}}, \tag{47}
\]

\[
\text{ITF}_{BLCMV,n}^{\text{out}} = \frac{B_{L}}{B_{R}} = \text{ITF}_{n}^{\text{in}}. \tag{48}
\]

It was also shown in [22], that the output IC of the diffuse noise is equal to

\[
\text{IC}_{BLCMV,n}^{\text{out}} = \frac{e_{L}^{T} \bar{R}_{sxu} e_{R}}{\sqrt{e_{L}^{T} \bar{R}_{sxu} e_{L}}}, \tag{49}
\]

with

\[
\bar{R}_{sxu} = \frac{1}{1 - \Psi} \begin{bmatrix} aa^{H} & \delta^{2} bb^{H} \\ \gamma_{a} & -2 \Psi \delta R \end{bmatrix}. \tag{50}
\]

The output MSC of the diffuse noise component for the BLCMV is equal to

\[
\text{MSC}_{BLCMV,n}^{\text{out}} = |\text{IC}_{BLCMV,n}^{\text{out}}|^{2}. \tag{51}
\]

Because \( \bar{R}_{sxu} \) is a rank-2 matrix, the diffuse noise component at the output is non-coherent and an output MSC smaller than one is attributed to it [22].

The output SIR of the BLCMV is equal to [22]

\[
\text{SIR}_{BLCMV,L}^{\text{out}} = \text{SIR}_{x}^{\text{in}} \frac{1}{\delta}, \tag{52}
\]

and is hence directly controlled by the interference scaling parameter \( \delta \). Practical insights in how to set \( \delta \) when estimation errors are present were provided in [26].

The output SNR of the BLCMV is equal to [22]

\[
\text{SNR}_{BLCMV,L}^{\text{out}} = \frac{\Phi_{x} |A_{L}|^{2}}{e_{L}^{T} \bar{R}_{sxu} e_{L}}. \tag{53}
\]

IV. THE BINAURAL LCMV-N BEAMFORMER (BLCMV-N)

In this section, we combine the approaches of the BMVDR-N and the BLCMV so that the binaural cue preservation of both the interfering source and the diffuse noise can be controlled. This novel approach is called the BLCMV with partial noise estimation (BLCMV-N). In Section [IV-A] the BLCMV-N is derived and interpreted. As was shown for the BLCMV in [22], in Section [IV-B] and [IV-C] two decompositions of the BLCMV-N are shown to provide a more intuitive view of the solution. The output components of the BLCMV-N are calculated in Section [IV-D] and a value for the interference scaling parameter is derived in Section [IV-E] that leads to equivalence between the BLCMV-N and the BMVDR-N.

A. Derivation and Interpretation

Compared to the BMVDR, the BLCMV-N uses an additional constraint to preserve a scaled version of the interfering source component in the reference microphones, i.e. like the BLCMV, and is aimed at preserving a scaled version of the diffuse noise component in the reference microphones, i.e. like the BMVDR-N. The cost function of the BLCMV-N is given by

\[
J_{L} = \mathcal{E} \left\{ |w_{L}^{H} n - \eta N_{L}|^{2} \right\} \quad \text{subject to} \quad C^{H} w_{L} = g_{L}, \tag{54}
\]

with the constraint matrix \( C \) defined in (45) and the response vector \( g_{L} \) defined in (44). The filter solving (54) is equal to (see Appendix for derivation)

\[
w_{BLCMV-N,L} = \begin{bmatrix} \eta & 1 - \eta \end{bmatrix} R_{n}^{-1} C \left( C^{H} R_{n}^{-1} C \right)^{-1} \begin{bmatrix} A_{L}^{*} \\ \delta B_{L}^{*} \end{bmatrix}. \tag{55}
\]

with

\[
\tilde{\delta} = \frac{\delta - \eta}{1 - \eta} \tag{56}
\]

the adjusted interference scaling parameter. The BLCMV-N can hence be seen as a mixture of the reference microphone signal (scaled with \( \eta \)) and a BLCMV (scaled with \( 1 - \eta \)) using the adjusted interference scaling parameter \( \delta \) instead of the interference scaling parameter \( \delta \). For \( \eta = 0 \) the BLCMV-N is equal to the BLCMV in (46), whereas for \( \eta = 1 \) the BLCMV-N is equal to the reference microphone signal in (5).

The adjusted interference scaling parameter \( \tilde{\delta} \) is related to both, \( \delta \) and \( \eta \), because both the additional constraint and the partial noise estimation affect the interfering source component. Figure 2 depicts \( \tilde{\delta} \) as a function of \( \eta \) for different values of \( \delta \). It can be seen that

\[
\tilde{\delta}(\eta, \delta) = \begin{cases} > 0 & \text{for } \delta > \eta \\ < 0 & \text{for } \delta < \eta \\ 0 & \text{for } \delta = \eta \end{cases} \tag{57}
\]
and the interfering source. Due to the first term in (58), the BLCMV-N can hence preserve the binaural cues of both the desired source component and steers a null towards the direction of the interfering source, i.e. is aimed at interference rejection. The sub-BLCMV $w_{u,L}$ denotes a BLCMV that preserves the interfering source component in the reference microphone signal and steers a null towards the direction of the desired source. It can therefore be shown that [22]

$$\begin{align*}
  w^H_{x,L}a &= A_L, \\
  w^H_{u,L}a &= 0, \\
  w^H_{u,L}b &= B_L.
\end{align*}$$

Hence, the expression in (58) can be interpreted as a mixture of the reference microphone signal (scaled with $\eta$), a BLCMV with interference rejection (scaled with $1-\eta$) and a BLCMV preserving the interfering source and rejecting the desired source (scaled with $\delta-\eta$). The sub-BLCMV $w_{x,L}$ lies in the nullspace of the constraint subspace of the interfering source and binaural cues of the desired source component can be controlled such that the interfering source component is not affected [22]. Additionally, the sub-BLCMV beamformer $w_{u,L}$ lies in the nullspace of the constraint subspace of the desired source and the binaural cues of the interfering source component can be controlled such that the desired source component is not affected. A combination of $w_{x,L}$ and $w_{u,L}$ can hence preserve the binaural cues of both the desired and the interfering source. Due to the first term in (58), the BLCMV-N can additionally (partially) preserve the binaural cues of the diffuse noise (cf. Section V-B).

### C. Filter Decomposition using Binauralization Postfilters

For a decomposition that intuitively combines both the left and the right BLCMVs, it was shown in [22], that (59) and (60) can be written as BLCMVs with binauralization postfilters. The BLCMV steering towards the desired source and rejecting the interfering source $w_{x,L}$ can be written as

$$w_{x,L} = w_x A_L^\ast,$$

with

$$w_x = \frac{1}{1-\Psi} \left( R^{-1}_n a - \Psi R^{-1}_n b \right).$$

the desired BLCMV (D-BLCMV) and $A_L$ the ATF between the desired source and the left reference microphone, used as the binauralization postfilter for the desired source. The BLCMV steering towards the interfering source and rejecting the desired source $w_{u,L}$ can be written as

$$w_{u,L} = w_u B_L^\ast,$$

with

$$w_u = \frac{1}{1-\Psi} \left( R^{-1}_n b - \Psi R^{-1}_n a \right).$$

the undesired BLCMV (U-BLCMV) and $B_L$ the ATF between the interfering source and the left reference microphone, used as the binauralization postfilter for the interfering source.

Using (63) and (65) in (58), the BLCMV-N can hence be written as

$$w_{BLCMV-N,L} = \eta e_L + (1-\eta)w_x + (\delta-\eta)B_L^\ast w_u.$$

Figure 3 depicts the scheme of the BLCMV-N decomposition using binauralization postfilters. It can be interpreted as a mixture of the reference microphone signals (scaled with $\eta$), the binauralized D-BLCMV beamformer output (scaled with $1-\eta$) and the binauralized U-BLCMV beamformer output (scaled with $\delta-\eta$).

### D. Output Components of the BLCMV-N

The desired source component at the output of the BLCMV-N is given by

$$w_{BLCMV-N,L}^H x = A_L S_x,$$
i.e. the BLCMV-N perfectly preserves the desired source. The interfering source component at the output of the BLCMV-N is given by

\[ \mathbf{w}_{BLCMV-N,L}^H \mathbf{u} = \delta B_L S_u, \]  

(69)

which is equal to the \( \delta \)-scaled interfering source component in the reference microphone. The diffuse noise component at the output of the BLCMV-N is given by

\[ \mathbf{w}_{BLCMV-N,L}^H \mathbf{n} = \eta N_L + (1 - \eta) N_x A_L + (\delta - \eta) N_u B_L, \]  

(70)

with \( N_x = \mathbf{w}_{BLCMV-N,L}^H \mathbf{n} \) and \( N_u = \mathbf{w}_{BLCMV-N,L}^H \mathbf{n} \) the diffuse noise component at the output of the D-BLCMV beamformer and the U-BLCMV beamformer, respectively. The expression in (70) can be considered a mixture of the noise component in the reference microphone signal (scaled with \( \eta \)) and two (coherent) noise sources.

E. Equivalence Between the BMVDR-N and the BLCMV-N

By using (58), it can be shown that the interfering source component at the output of the BMVDR-N is equal to

\[ \mathbf{w}_{BMVDR-N,L}^H \mathbf{u} = \left( \eta B + (1 - \eta) \gamma_{ab} \frac{A_L}{r_a} \right) S_u. \]  

(71)

A comparison of (69) and (71) shows that, to allow the BLCMV-N to have the same influence on the interfering source, the interference scaling parameter \( \delta \) has to be equal to

\[ \delta_{BMVDR-N,L} = \eta + (1 - \eta) \gamma_{ab} \frac{A_L}{r_a} B_L. \]  

(72)

Similar values can be obtained for the right BLCMV-N, i.e. \( \delta_{BMVDR-N,R} = \eta + (1 - \eta) \gamma_{ab} \frac{A_L}{r_a} B_L \). Equivalently, the adjusted interference scaling parameter has to be equal to

\[ \bar{\delta}_{BMVDR-N,L} = \gamma_{ab} \frac{A_L}{r_a} B_L. \]  

(73)

Please note that if considering complex-valued interfering scaling parameters, the response vector in (44) would be defined as \( \mathbf{g}_L = [A_L, \delta B_L]^H \). It is clear that equivalence between the BLCMV-N and the BMVDR-N can be achieved only by using (different) complex-valued interference scaling parameters in the left and right filter. This can also be explained intuitively, because by using the same parameter for both sides, the binaural cues of the interfering source are preserved by the BLCMV-N. To obtain equivalence, the binaural cues of the interfering source need to be distorted exactly as they are distorted by the BMVDR-N. Obviously, the BLCMV-N no longer preserves the binaural cues of the interfering source in this case.

For example, by substituting (72) for \( \delta \) in (55) and comparing to (59) the equivalence between the BLCMV-N and the BMVDR-N can easily be proven, i.e.

\[ \mathbf{w}_{BLCMV-N,L} = \mathbf{w}_{BMVDR-N,L} \quad \text{if} \quad \delta = \delta_{BMVDR-N,L}. \]  

(74)

The same can of course be proven for the right BLCMV-N, by using \( \delta_{BMVDR-N,R} \) in \( \mathbf{w}_{BLCMV-N,R} \) and comparing to \( \mathbf{w}_{BMVDR-N,R} \).

V. PERFORMANCE OF THE BLCMV-N

In this section, the performance of the BLCMV-N is examined. First, the output PSDs of the signal components are calculated in Section [V.-A]. Then, the binaural cue preservation performance and the interference and noise reduction performance are analyzed in Section [V.-B] and [V.-C] respectively. Finally, in Section [V.-D] the setting of the two scaling parameters \( \delta \) and \( \eta \) is explained.

A. Power Spectral Densities

The output PSD of the desired source component for the BLCMV-N is equal to

\[ \Phi_{BLCMV-N,x,L}^{\text{out}} = \Phi_x |A_L|^2. \]  

(75)

The output PSD of the interfering source component for the BLCMV-N is equal to

\[ \Phi_{BLCMV-N,u,L}^{\text{out}} = \delta^2 \Phi_u |B_L|^2. \]  

(76)

By substituting (67) in (24), it can be shown that the output PSD of the diffuse noise component for the BLCMV-N is equal to

\[ \Phi_{BLCMV-N,n,L}^{\text{out}} = \mathbf{e}_L^T \left( \eta^2 \mathbf{R}_n + \mathbf{R}_{xu} \right) \mathbf{e}_L, \]  

(77)

with

\[ \mathbf{R}_{xu} = \frac{1}{1 - \psi} \left[ (1 - \eta) \frac{\mathbf{a}_n \mathbf{a}_f^H}{\gamma_a} + (\delta^2 - \eta^2) \frac{\mathbf{b}_L \mathbf{b}_L^H}{\gamma_b} - 2 \psi (\delta - \eta^2)^2 \mathbb{R} \left\{ \frac{\mathbf{ab}_L^H}{\gamma_{ab}} \right\} \right]. \]  

(78)

It can be seen that (77) is a quadratic function in both scaling parameters \( \delta \) and \( \eta \). Please note that \( \mathbf{R}_{xu} = \mathbf{R}_{xu} \) for \( \eta = 0 \), cf. (50) and (78).

B. Binaural Cue Preservation

Like the BLCMV, due to the hard constraints, the BLCMV-N preserves the ITF of both the desired source and the interfering source, i.e.

\[ \text{ITF}_{BLCMV-N,x}^{\text{out}} = \frac{A_L}{A_R} = \text{ITF}_{x}^{\text{in}}, \]  

(79)

\[ \text{ITF}_{BLCMV-N,u}^{\text{out}} = \frac{B_L}{B_R} = \text{ITF}_{u}^{\text{in}}. \]  

(80)
TABLE I
Binaural Cue Preservation of Binaural Beamformers

| Algorithm   | Desired source | Interfering source | Diffuse noise |
|-------------|----------------|--------------------|---------------|
| BMVDR       | Yes            | No                 | No            |
| BMVDR-N     | Yes            | Trade-off (η)      | Trade-off (η) |
| BLCMV       | Yes            | Yes                | No            |
| BLCMV-N     | Yes            | Yes                | Trade-off (η) |

Substituting (67) in (21), the output IC of the diffuse noise component is equal to

\[
\text{IC}^{\text{out}}_{\text{BLCMV}-N,n} = \frac{e^T_L (\eta^2 R_n + R_{xxu}) e_R}{\sqrt{e^T_L (\eta^2 R_n + R_{xxu}) e_L e^*_R (\eta^2 R_n + R_{xxu}) e^*_L}}, \tag{81}
\]

with \( R_{xxu} \) defined in (78). The output MSC of the diffuse noise component is then equal to

\[
\text{MSC}^{\text{out}}_{\text{BLCMV}-N,n} = |\text{IC}^{\text{out}}_{\text{BLCMV}-N,n}|^2. \tag{82}
\]

It can be shown that for \( \eta = 0 \) the expressions in (81) and (49) are equal and consequently

\[
\text{MSC}^{\text{out}}_{\text{BLCMV}-N,n} = \text{MSC}^{\text{out}}_{\text{BLCMV},n} \quad \text{for} \quad \eta = 0, \tag{83}
\]

leading to an output MSC smaller than 1. For \( \delta = 0 \) and \( \eta = 0 \) the matrix \( R_{xxu} \) in (78) becomes rank-1 and \( R_{xxu} \) disappears from the expression in (81). Hence, the MSC would be equal to 1, which is again attributed to a coherent source, in this case perceived as coming from the direction of the desired source, i.e.

\[
\text{MSC}^{\text{out}}_{\text{BLCMV}-N,n} = 1 \quad \text{for} \quad \eta = 0 \quad \text{and} \quad \delta = 0. \tag{84}
\]

In contrast to the BMVDR-N described in Section III-B for \( \eta = 1 \) the BLCMV-N does not always preserve the MSC of the diffuse noise component. Only if \( \eta = 1 \) and \( \delta = 1 \) can the MSC of the diffuse noise be preserved, i.e.

\[
\text{MSC}^{\text{out}}_{\text{BLCMV}-N,n} = \text{MSC}^{\text{in}}_n \quad \text{for} \quad \eta = 1 \quad \text{and} \quad \delta = 1, \tag{85}
\]

which would lead to using only the reference microphone signals and hence no noise reduction would be applied. The BLCMV-N hence provides a trade-off between noise reduction and the IC preservation of the diffuse noise. Table I shows a summary of the binaural cue preservation capabilities of all the considered binaural beamformers. We further elaborate on the output MSC of the diffuse noise component in Section VI.

C. Interference and Noise Reduction Performance

The output SIR of the BLCMV-N is equal to the output SNR of the BLCMV in (52), i.e.

\[
\text{SIR}^{\text{out}}_{\text{BLCMV}-N,L} = \text{SIR}^{\text{in}}_L \frac{1}{\delta^2} = \text{SIR}^{\text{out}}_{\text{BLCMV},L}, \tag{86}
\]

which is hence also solely controlled by the interference scaling parameter \( \delta \). The output SNR of the BLCMV-N is equal to

\[
\text{SNR}^{\text{out}}_{\text{BLCMV}-N,L} = \frac{\Phi_n |A_L|^2}{\alpha^2_L (\eta^2 R_n + R_{xxu}) e^*_L}, \tag{87}
\]

which like (77) is a quadratic function in both scaling parameters \( \delta \) and \( \eta \). For \( \eta = 1 \) and \( \delta = 1 \), the output SNR is equal to the input SNR, because no noise reduction is applied, i.e.

\[
\text{SNR}^{\text{out}}_{\text{BLCMV}-N,L} = \text{SNR}^{\text{in}}_L \quad \text{for} \quad \eta = 1 \quad \text{and} \quad \delta = 1. \tag{88}
\]

For \( \eta = 0 \), the output SNR of the BLCMV-N is equal to the output SNR of the BLCMV in (53), i.e.

\[
\text{SNR}^{\text{out}}_{\text{BLCMV}-N,L} = \text{SNR}^{\text{out}}_{\text{BLCMV},L} \quad \text{for} \quad \eta = 0. \tag{89}
\]

D. Scaling Parameter Settings

As described above, the noise scaling parameter \( \eta \) provides a trade-off between noise reduction and the binaural cue preservation, i.e. IC preservation, of the diffuse noise. It is easy to show that, to maximize the output SNR (or to minimize the output PSD of the diffuse noise in (77)), \( \eta \) should be set equal to 1, although in this case no additional binaural cue preservation of the diffuse noise is applied, cf. (83). Raising the noise scaling parameter \( \eta \) leads to a lower noise reduction performance but increased binaural cue preservation of the diffuse noise. As can be seen in (86), the output SIR is not affected by \( \eta \).

Setting the derivative of (87) with respect to \( \delta \) to zero and solving for \( \delta \) yields the interference scaling parameter \( \delta \) that is optimal in terms of SNR, i.e.

\[
\delta_{\text{opt},L} = \frac{\alpha_L}{\beta_L}, \tag{90}
\]

with

\[
\alpha_L = \Psi \Re \left\{ A_L B^2_{\gamma_{ab}} \right\}, \quad \beta_L = \frac{|B_L|^2}{\gamma_{ab}}. \tag{91}
\]

By setting \( \delta = 0 \), obviously the output SIR would be maximized, but problems can arise in a practical implementation when estimation errors are present [26].

VI. SIMULATIONS

In this section, we first show a validation of the analytical expressions derived in the previous sections by using measured anechoic impulse responses. Second, we show experimental results and results of a subjective listening test using measured reverberant impulse responses and recorded signals to quantify the filter performances in a more realistic scenario with estimation errors.

A. Validation with Measured Anechoic ATFs

Here, we examine the theoretical performance of the considered binaural noise reduction algorithms using measured anechoic impulse responses of two behind-the-ear (BTE) hearing aids mounted on a head-and-torso-simulator (HATS) [50]. Two microphones on each side of the head were used, i.e. \( M = 4 \), and the impulse responses were resampled to a sampling frequency of 16 kHz. We considered an acoustic scenario with one desired source, one interfering source and diffuse noise, as described in Section II-A. The ATFs were calculated from the BTE impulse responses using a 512 point FFT. To simulate diffuse (cylindrically isotropic) noise and
spatially white sensor noise, the noise covariance matrix was generated as

\[ R_n = \Phi_{n,w} I_M + \Phi_{n,diff} \Gamma, \]  

(92)

with \( \Phi_{n,w} \) the PSD of sensor noise, \( I_M \) the \( M \times M \) identity matrix, \( \Phi_{n,diff} \) the PSD of the diffuse noise and \( \Gamma \) the spatial coherence matrix of the diffuse noise. The \((i,j)\)-th element of \( \Gamma \) can be calculated using the set of anechoic A TFs as [18], [22],

\[ \Gamma_{i,j} = \frac{\sum_{k=1}^{K} H_i(\theta) H_j^*(\theta)}{\sqrt{\sum_{k=1}^{K} |H_i(\theta)|^2 \sum_{k=1}^{K} |H_j(\theta)|^2}}, \]  

(93)

with \( H(\theta) \) the anechoic A TF at angle \( \theta \) and \( K \) the total number of angles in the database (\( K = 72 \) for [22]).

The desired source was placed at 30° and the interfering source was placed at -70°, both at a distance of 3 m (0° indicates the look direction and 90° indicates the right hand side). The covariance matrices of the desired and interfering sources \( R_x \) and \( R_n \) were constructed using the respective A TFs according to (9) and (10), respectively. \( \Phi_x \) and \( \Phi_n \) were both set to 1. The PSD of the spatially white sensor noise \( \Phi_{n,w} \) was set to -55 dB.

1) Binaural Cue Preservation: Without any estimation errors, all of the binaural noise reduction algorithms considered in this paper perfectly preserve the binaural cues of the desired source (cf. Table I). The BMVDR does not constrain the interfering source and hence the interfering source is collocated with the desired source at the output (not shown here for conciseness, cf. Section III). The BMVDR-N can trade off noise reduction performance for the binaural cue preservation of the diffuse noise and the interfering source. Whereas for a diffuse noise this approach showed promising results and was explained in detail in [18], the effect of the partial noise estimation approach on a coherent interfering source strongly depends on the position of the interfering source relative to the desired source. To encounter this, the BLCMV and BLCMV-N hence use additional constraints such that the binaural cues of the interference source are perfectly preserved.

Next, the MSC of the diffuse noise is examined. Figure 5 depicts the MSC of the diffuse noise at the input of the reference microphones and the output of the binaural MVDR, MVDR-N and LCMV beamformers. As expected, at the output of the BMVDR the diffuse noise is a coherent source (co-located with the desired source, cf. [35]). The BMVDR-N can use the trade-off to preserve the MSC of the diffuse noise, such that if \( \eta = 1 \) the MSC is perfectly preserved (but no noise reduction is achieved). The BLCMV is not designed to preserve the MSC of diffuse noise, although if \( \delta = 1 \) an output MSC smaller than 1 is attributed to the diffuse noise [22]. This effect depends on the relative position of the interfering source to the desired source and cannot easily be controlled.

Figure 6 depicts the output MSC of the diffuse noise for the BLCMV-N for different values of \( \eta \) and \( \delta \). It can be seen that both \( \eta \) and \( \delta \) influence the output MSC of the diffuse noise as expected from the previous observations. Perfect preservation of the MSC of the diffuse noise is possible only if \( \delta = \eta = 1 \).
The BLCMV hence not only preserves the binaural cues of the interfering source but also allows the output SIR to be calculated using (90). In combination with Figure 7, the trade-off between MSC preservation of the diffuse noise and noise reduction in the BLCMV-N can be assessed. The output SIR for the BLCMV-N is equal to the output SIR of the BLCMV and therefore is not depicted.

B. Experimental Results in Reverberant Environment

In this section, we evaluate and compare the performances of the considered binaural noise reduction algorithms in a reverberant environment. The same two BTE hearing aids have again been mounted to the HATS and have been used in [30] to measure reverberant room impulse responses with a reverberation time of about 1.25 s in a university cafeteria. The HATS was placed at a cafeteria table corresponding to a sitting position. Again, two microphones on each side of the head were used, such that $M = 4$, and all room impulse responses and signals were resampled to a sampling frequency of 16 kHz. Two directional speakers, i.e. the desired source and the interfering source, were generated by convolving clean speech recordings with room impulse responses corresponding to specific positions. The desired source was placed at 0°, i.e. in the look direction of the HATS, at a distance of about 102 cm, corresponding to a person sitting on the other side of the table. The interfering source was placed at about $-35^\circ$ at a distance of about 118 cm, corresponding to a person sitting to the left of the desired source. The desired source signal consisted of a male German speaker speaking eight sentences with a pause of 1 s between the sentences. The interfering source signal consisted of a male Dutch speaker speaking seven sentences with a pause of 0.25 s between the sentences. As diffuse noise, we used recordings of real multi-talker babble noise, recorded using the same setup in the same cafeteria [30]. The diffuse noise consisted of babble noise, clacking plates and temporary dominant interfering speakers around the setup. The entire signal had a length of about 28 s. The desired source and the diffuse noise were active the entire time, whereas the interfering source started to speak after about 14 s.

The processing was done in the STFT domain with a frame length of 4096 samples and a square-root Hann window with 50% overlap. We used a perfect voice activity detector in the time domain to estimate the covariance matrices $R_n$, $R_v$ and $R_{vn} = R_v + R_n$ (desired source plus diffuse noise), averaged over the entire signal. All the algorithms were implemented using the estimated noise covariance matrix $R_n$ and relative transfer functions (RTFs) [31], relating the ATF vectors in (7) and (8) to the reference microphones. The RTF vectors were estimated based on a generalised eigenvalue decomposition (GEVD) of $R_{vn}$ and $R_n$ or of $R_v$ and $R_n$. Please refer to [22], [26], [32], [33] for further details. The algorithm parameters were set as $\delta = 0.3$ and $\eta = 0.3$.

1) Noise Reduction Performance: As the performance measure for noise and interference reduction the SNR, SIR and SINR improvements (relative to the reference microphone signals) were calculated using the signal components at the output in the time domain. Table II presents the results for all the considered algorithms.

For the left and right SNR improvement ($\Delta \text{SNR}_L$ and $\Delta \text{SNR}_R$), it can be seen that the BMVDR yields the best...
results. This is expected, because the BMVDR-N trades off noise reduction performance for binaural cue preservation and the BLCMV uses an additional constraint for the interfering source and hence has one degree-of-freedom less available for noise reduction. The BLCMV-N does both and hence leads to the lowest noise reduction.

For the left and right SIR improvement (ΔSIR_L and ΔSIR_R), it can be seen that the BMVDR and BMVDR-N cannot reduce the interfering speaker, because R_u is used in the filters. By setting δ = 0.3 for the BLCMV and BLCMV-N, an SIR improvement of about 10 dB is expected. This performance cannot be completely achieved due to estimation errors, which is in line with the results in [22]. It is important to note, however, that the SIR improvement of both the BLCMV and BLCMV-N is roughly the same and controlled by δ. Hence, the partial noise estimation introduced in the BLCMV-N does not affect the reduction of the interfering source, i.e. for any η value.

For the left and right SINR improvement (ΔSINR_L and ΔSINR_R), i.e. the overall reduction of all undesired signal components, it can be seen that the BLCMV and BLCMV-N outperform the BMVDR and BMVDR-N in this scenario. This is of course specifically because the BMVDR and BMVDR-N do not take into account explicitly the reduction of the interfering source signal in the optimization. A comparison of the BLCMV with the BLCMV-N shows that the introduced partial noise estimation reduced the SINR improvement by roughly 1 dB.

2) Subjective Listening Test: To investigate the binaural cue preservation performance of the four considered filters a procedure similar to the Multi-Stimulus Test with Hidden Reference and Anchor (MUSHRA) was conducted. The signal processing framework and the acoustic scenario were the same as described in [13], except that the desired source was placed at −35° and the interfering source was placed at 90°. The input SIR and SNR were both set to 0 dB, measured in the right reference microphone.

Thirteen self-reported normal hearing subjects participated in the listening test. None of the authors participated in the listening test. All subjects gave informed consent and ethical approval was obtained by the ethics committee of the University of Oldenburg.

The listening test was conducted in a sound proof listening booth using a MATLAB implementation of the MUSHRA and an RME Fireface UCX with Sennheiser HD 580 headphones.

The task was to rate the perceived spatial difference to a reference signal. In the case of coherent sources, spatial similarity corresponds to the same localization as for the reference signal, and, in the case of a diffuse noise field, to the same amount of perceived diffuseness. A score of 0 is associated with an extreme difference, whereas a score of 100 is associated with no perceived difference. The reference signal was chosen to be the left and right unprocessed reference microphone signals, i.e. the hearing aid input signals. An anchor signal was produced by using only the left reference microphone signal and playing it back via both loudspeakers. The anchor was hence a monaural signal with no binaural cues. Additionally, the anchor was lowpass filtered using a 1st order Butterworth filter with a cutoff frequency of 1 kHz. In a MUSHRA, the anchor should be associated with a very low score and encourage the subjects to use the entire rating scale.

The subjects had time to get comfortable with the sound material in a training round. In the first evaluation round, only the desired source and the interfering source were active. In the second evaluation round, only the desired source and the diffuse noise were active. In the third evaluation round, all signal components were active.

The results are depicted in Figure 10. A one-way repeated-measures ANOVA was performed using IBM SPSS Statistics. The analysis revealed a significant within-subjects effect for all three evaluation rounds. Hence, post-hoc comparison t-tests with Bonferroni adjustment were performed.

a) Interfering source: The within-subjects effect was significant [F(2.098, 25.176) = 219.2, p < .001, Greenhouse-Geisser correction]. As expected, both BLCMVs preserved the spatial perception of the interfering source significantly better than both BMVDRs (p < .001). The BMVDR-N performed significantly better than the BMVDR (p < .001). No significant difference was found between the BLCMV and BLCMV-N (p = 1).

b) Diffuse noise: The within-subjects effect was significant [F(3.072, 36.869) = 332.066, p < .001, Greenhouse-Geisser correction]. The BMVDR-N and the BLCMV-N, both using partial noise estimation, significantly outperformed the BMVDR and the BLCMV (p < .001). No significant difference was found between the BMVDR-N and BLCMV-N (p = 1). No significant difference was found between the BMVDR and BLCMV (p = .614).

c) Complete signal: The within-subjects effect was significant [F(2.905, 34.858) = 171.783, p < .001, Greenhouse-Geisser correction]. The BMVDR-N scored significantly higher than the BMVDR (p < .001). The BLCMV significantly outperformed the BMVDR-N (p = .014). Further, the BLCMV-N could outperform the BLCMV significantly (p = .025) and hence also significantly outperformed the BMVDR and BMVDR-N (p < .001).

To summarize the subjective listening test, the BLCMV-N is capable of preserving the binaural cues of an interfering source

---

**TABLE II**

| Algorithm     | η   | δ   | ΔSINR_L | ΔSINR_R | ΔSIR_L | ΔSIR_R | ΔSINR_L | ΔSINR_R |
|---------------|-----|-----|---------|---------|--------|--------|---------|---------|
| BMVDR        | -   | -   | 12.3    | 11.9    | -0.1   | -2.1   | 2.5     | 1.5     |
| BMVDR-N      | 0.3 | -   | 8.2     | 7.7     | 0.9    | -0.13  | 3.0     | 2.72    |
| BLCMV        | -   | 0.3 | 9.9     | 8.9     | 8.8    | 9.1    | 9.4     | 9.2     |
| BLCMV-N      | 0.3 | 0.3 | 7.6     | 6.7     | 8.9    | 9.6    | 8.4     | 7.8     |
Fig. 10. Boxplot of the subjective listening scores. The plot depicts the median score (red line), the mean score (red dot), the first and third quartiles (blue boxes) and the interquartile ranges (whiskers). Outliers are indicated by the red + markers.

and diffuse noise in a realistic acoustic scenario, whereas all the other considered binaural noise reduction algorithms lack in preserving the binaural cues of at least one signal component.

VII. CONCLUSIONS

In this paper, we combined two extensions of the well-known BMVDR, namely the BLCMV and the BMVDR-N, to propose the BLCMV with partial noise estimation (BLCMV-N). It was shown, that the proposed BLCMV-N can be seen as a mixture of the noisy reference microphone signals and a BLCMV output, where an adjusted interference scaling parameter is used. Theoretical expressions for the noise reduction and binaural cue preservation performance of the BLCMV-N were derived and equality between the BLCMV-N and BMVDR-N was shown for a specific (complex-valued) interference scaling parameter. An anechoic validation and experimental results using reverberant signals showed that the BLCMV-N is capable of preserving the binaural cues of an interfering source, like the BLCMV, and yields a trade-off between noise reduction and binaural cue preservation of diffuse noise, like the MVDR-N. Results of a subjective listening test showed that the BLCMV-N can preserve the spatial impression of an interfering source and diffuse noise in a realistic acoustic scenario, whereas the BMVDR, the BMVDR-N and the BLCMV lack in preserving the spatial impression of at least one signal component.

APPENDIX

DERIVATION OF THE BLCMV-N

The Lagrangian of the cost function in (54) is given by

$$\mathcal{L}(w_L) = w_L^H R_n w_L - \eta e^T_n R_n w_L - \eta w_L^H R_n e_L + \eta^2 |N| \eta L + \lambda^H (C^H w_L - g_L) + (w_L^H C - g^H) \lambda_L,$$

with the Lagrangian multiplier \(\lambda_L\). The gradient with respect to \(w_L\) is equal to

$$\nabla \mathcal{L}(w_L) = 2R_n w_L - 2\eta R_n e_L + 2C \lambda_L.$$

Setting (96) to 0, the filter minimizing (54) is equal to

$$w_L = \eta e_L - \frac{1}{2} R_n^{-1} C \lambda_L.$$

Substituting (97) into (54), the Lagrangian multiplier is equal to

$$\lambda_L = (C^H R_n^{-1} C)^{-1} (\eta C^H e_L - g_L).$$

Substituting (98) into (97), the solution to (54) is equal to

$$w_{\text{BLCMV-N,L}} = \eta e_L + R_n^{-1} C (C^H R_n^{-1} C)^{-1} (g_L - \eta C^H e_L),$$

which can be written as

$$w_{\text{BLCMV-N,L}} = \eta e_L + (1 - \eta) R_n^{-1} C (C^H R_n^{-1} C)^{-1} \frac{A^*_L}{\delta B^*_L},$$

with

$$\delta = \frac{\delta - \eta}{1 - \eta}.$$

REFERENCES

[1] S. Doclo, S. Gannot, D. Marquardt, and E. Hadad, “Binaural speech processing with application to hearing devices,” in Audio Source Separation and Speech Enhancement. Wiley, 2018, ch. 18, pp. 413–442.

[2] J. Blauert, Spatial hearing: the psychophysics of human sound localization. Cambridge, Mass. MIT Press, 1997.

[3] A. W. Bronkhorst and R. Plomp, “The effect of head-induced interaural time and level differences on speech intelligibility in noise,” The Journal of the Acoustical Society of America, vol. 89, no. 4, pp. 1508–1516, Apr. 1988.

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

[5] T. Van den Bogaert, T. Klasen, M. Moonen, L. Van Den, and J. Wouters, “Horizontal localisation with bilateral hearing aids: Without is better than with,” The Journal of the Acoustical Society of America, vol. 119, no. 1, pp. 515–526, Jan. 2006.

[6] G. Keidser, A. O’Brien, J. U. Hain, M. McLelland, and I. Yeend, “The effect of frequency-dependent microphone directionality on horizontal localization performance in hearing-aid users,” International Journal of Audiology, vol. 48, no. 11, pp. 789–803, Dec. 2009.

[7] T. Wittkop and V. Hohmann, “Strategy-selective noise reduction for binaural digital hearing aids,” Speech Communication, vol. 39, no. 1-2, pp. 111–138, Jan. 2003.

[8] T. Lotter and P. Vary, “Dual-channel speech enhancement by superdirective beamforming,” EURASIP Journal on Applied Signal Processing, vol. 2006, pp. 1–14, Dec. 2006.

[9] G. Grimm, V. Hohmann, and B. Kollmeier, “Increase and subjective evaluation of feedback stability in hearing aids by a binaural coherence-based noise reduction scheme,” IEEE Transactions on Audio, Speech, and Language Processing, vol. 17, no. 7, pp. 1408–1419, Sep. 2009.
T. Klasen, T. Van den Bogaert, M. Moonen, and J. Wouters, “Binaural interaural coherence preservation for binaural hearing aids operating in a diffuse noise field environment,” IEEE Transactions on Audio, Speech, and Language Processing, vol. 17, no. 4, pp. 521–533, May 2009.

A. Kamkar-Parsi and M. Bouchard, “Improved noise power spectrum density estimation for binaural hearing aids operating in a diffuse noise field environment,” IEEE Transactions on Audio, Speech, and Language Processing, vol. 17, no. 4, pp. 521–533, May 2009.

A. Kamkar-Parsi and M. Bouchard, “Instantaneous binaural target PSD estimation for hearing aid noise reduction in complex acoustic environments,” IEEE Transactions on Instrumentation and Measurement, vol. 60, no. 4, pp. 1141–1154, Apr. 2011.

M. Jeub, M. Dörrecker, and P. Vary, “A semi-analytical model for the binaural coherence of noise fields,” IEEE Signal Processing Letters, vol. 18, no. 3, pp. 197–200, Mar. 2011.

R. Baumgärtel, M. Krawczyk-Becker, D. Marquardt, C. Völker, H. Hu, T. Herzke, G. Coleman, K. Addiloglu, S. M. A. Ernst, T. Gerkmann, S. Doclo, B. Kollmeier, V. Hohmann, and M. Dietz, “Comparing binaural signal processing strategies: Instrumental evaluation,” Trends in Hearing, vol. 19, pp. 1–16, 2015.

V. Hamacher, U. Kornagel, T. Lotter, and H. Puder, “Binaural signal processing in hearing aids: Technologies and algorithms,” in Advances in Digital Speech Transmission. New York, NY, USA: Wiley, 2008, pp. 401–429.

S. Doclo, S. Gannot, M. Moonen, and A. Spriet, “Acoustic beamforming for hearing aid applications,” in Handbook on Array Processing and Sensor Networks. Wiley, 2010, pp. 269–302.

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

S. Doclo, W. Kellermann, S. Makino, and S. E. Nordholm, “Multi-channel signal enhancement algorithms for assisted listening devices: Exploiting spatial diversity using multiple microphones,” IEEE Signal Processing Magazine, vol. 32, no. 2, pp. 18–30, Mar. 2015.

D. Marquardt and S. Doclo, “Interaural coherence preservation for binaural noise reduction using partial noise estimation and spectral postfiltering,” IEEE/ACM Transactions on Audio, Speech and Language Processing, vol. 26, no. 7, pp. 1257–1270, 2018.

T. Klasen, T. Van den Bogaert, M. Moonen, and J. Wouters, “Binaural noise reduction algorithms for hearing aids that preserve interaural time delay cues,” IEEE Transactions on Signal Processing, vol. 55, no. 4, pp. 1579–1585, Apr. 2007.

E. Hadad, S. Gannot, and S. Doclo, “Binaural linearly constrained minimum variance beamformer for hearing aid applications,” in International Workshop on Acoustic Signal Enhancement (IWAENC), Aachen, Germany, Sep. 2012, pp. 4–6.

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

E. Hadad, S. Doclo, and S. Gannot, “The binaural LCMV beamformer and its performance analysis,” IEEE/ACM Transactions on Audio, Speech, and Lang. Proc., vol. 24, no. 3, pp. 543–558, Mar. 2016.

D. Marquardt, E. Hadad, S. Gannot, and S. Doclo, “Optimal binaural LCMV beamformers for combined noise reduction and binaural cue preservation,” in Proc. International Workshop on Acoustic Signal Enhancement (IWAENC), Juan-les-Pins, France, Sep. 2014, pp. 288–292.

A. I. Koutrouvelis, R. C. Hendriks, R. Heusdens, and J. Jensen, “Relaxed binaural lcmv beamforming,” IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 25, no. 1, pp. 137–152, Jan. 2017.

A. I. Koutrouvelis, R. C. Hendriks, R. Heusdens, J. Jensen, and M. Guo, “Binaural beamforming using pre-determined relative acoustic transfer functions,” in Signal Processing Conference (EUSIPCO), 2017 25th European, Kos, Greece, Aug. 2017, pp. 1–5.

N. Gölling, D. Marquardt, I. Merks, T. Zhang, and S. Doclo, “Optimal binaural LCMV beamforming in complex acoustic scenarios: Theoretical and practical insights,” in Proc. International Workshop on Acoustic Signal Enhancement (IWAENC), Tokyo, Japan, Sep. 2018, pp. 381–385.

D. Marquardt, V. Hohmann, and S. Doclo, “Interaural coherence preservation in multi-channel Wiener filtering based noise reduction for binaural hearing aids,” IEEE/ACM Transactions on Audio, Speech and Language Processing, vol. 23, no. 12, pp. 2162–2176, Dec. 2015.

L. Ehrenberg, S. Gannot, A. Leshem, and E. Zehavi, “Sensitivity analysis of MVDR and MPDR beamformers,” in Proc. IEEE Convention of Electrical and Electronics Engineers in Israel, Eilat, Israel, Nov. 2010, pp. 416–420.

N. Gölling, D. Marquardt, and S. Doclo, “Performance analysis of the extended binaural MVDR beamformer with partial noise estimation in a homogeneous noise field,” in Proc. Joint Workshop on Hands-Free Speech Communication and Microphone Arrays (HSCMA), San Francisco, USA, Mar. 2017, pp. 1–5.

H. Kayser, S. D. Ewert, J. Anemüller, T. Rohdenburg, V. Hohmann, and B. Kollmeier, “Database of multichannel In-Ear and Behind-The-Ear head-related and binaural room impulse responses,” Eurasip Journal on Advances in Signal Processing, vol. 2009, p. 10 pages, 2009.

S. Gannot, D. Burshtein, and E. Weinstein, “Signal enhancement using beamforming and non-stationarity with applications to speech,” IEEE Transactions on Signal Processing, vol. 49, no. 8, pp. 1614–1626, Aug. 2001.

S. Markovich, S. Gannot, and I. Cohen, “Multichannel eigenspace beamforming in a reverberant noisy environment with multiple interfering speech signals,” IEEE Transactions on Audio, Speech, and Language Processing, vol. 17, no. 6, pp. 1071–1086, Aug. 2009.

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

ITU-R BS.1534-1, Method for the subjective assessment of intermediate quality level of coding systems., International Telecommunications Union (ITU-T) Recommendation, Jan. 2003.