IMPROVED MULTI-MICROPHONE NOISE REDUCTION PRESERVING BINAURAL CUES

Andreas I. Koutrouvelis[†] *Richard C. Hendriks*[†] *Jesper Jensen*[⋆] *Richard Heusdens*[†]

[†]Circuits and Systems (CAS) Group, Delft University of Technology, the Netherlands
[⋆]Oticon A/S and Electronic Systems Department, Aalborg University, Denmark

ABSTRACT

We propose a new multi-microphone noise reduction technique for binaural cue preservation of the desired source and the interferers. This method is based on the linearly constrained minimum variance (LCMV) framework, where the constraints are used for the binaural cue preservation of the desired source and of multiple interferers. In this framework there is a trade-off between noise reduction and binaural cue preservation. The more constraints the LCMV uses for preserving binaural cues, the less degrees of freedom can be used for noise suppression. The recently presented binaural LCMV (BLCMV) method and the optimal BLCMV (OBLCMV) method require two constraints per interferer and introduce an additional interference rejection parameter. This unnecessarily reduces the degrees of freedom, available for noise reduction, and negatively influences the trade-off between noise reduction and binaural cue preservation. With the proposed method, binaural cue preservation is obtained using just a single constraint per interferer without the need of an interference rejection parameter. The proposed method can simultaneously achieve noise reduction and perfect binaural cue preservation of more than twice as many interferers as the BLCMV, while the OBLCMV can preserve the binaural cues of only one interferer.

Index Terms— Binaural cue preservation, LCMV, noise reduction.

1. INTRODUCTION

Multi-microphone noise reduction [1, 2] is of significant importance for hearing aids. The availability of multiple microphones per hearing aid results in potentially more noise suppression and a better speech intelligibility [3], than with single-channel algorithms, e.g. [4]. Apart from noise suppression, it is also important to preserve the binaural cues of both desired and noise sources. A binaural hearing-aid system consists of two collaborative hearing aids, usually with multiple microphones each. The collaboration between the hearing aids can be exploited to preserve the binaural cues and increase the amount of noise suppression.

A popular multi-microphone noise reduction technique is the multi-microphone extension of the Wiener filter [5], also known as the multi-channel Wiener filter (MWF) [6]. In [7], it was shown that the MWF perfectly preserves the binaural cues of the desired source, but distorts the ones of the interferers. Several multi-microphone noise reduction techniques that aim at simultaneous noise reduction and binaural cue preservation have been proposed in the literature [7–12]. Binaural cues are defined by the interaural level differences (ILDs) and the interaural time differences (ITDs) [8].

In [10], a variation of the MWF is proposed, which preserves the ITDs and ILDs of the desired source, but only partially preserves binaural cues of the interferers [8, 10]. More specifically, the method uses a trade-off parameter, which allows a portion of the noise to remain unprocessed in the final enhanced signal by mixing in a portion of the original noisy signal. The larger the portion of the unprocessed noise is, the better the ITD and ILD preservation, but the less noise reduction.

The binaural linearly constrained minimum variance (BLCMV) method [11] preserves the ITDs and ILDs of the desired source as well as multiple interferers. BLCMV consists of two beamformers, one for each hearing aid, that reproduce the desired signal as received by the reference microphones in each hearing aid by minimizing the noise power. Binaural cue preservation of interferers is obtained by placing constraints on the acoustic transfer functions (ATFs) of the interferers using a fixed interference rejection parameter which controls the amount of noise reduction. This implies that two constraints are needed for each interferer for which the binaural cues are to be preserved, i.e., a constraint for the right hearing aid and a constraint for the left hearing aid.

The optimal BLCMV (OBLCMV) [12] optimizes the interference rejection parameter, with respect to the binaural output SNR, leading to better noise reduction. However, in contrast to the BLCMV, this method can only preserve the ILDs and ITDs of the desired source and *one* interferer. Both the BLCMV and OBLCMV consist of two LCMV-based optimization problems using the recordings of all microphones of both hearing aids.

Both the BLCMV and OBLCMV depend on the interference rejection parameter and can simultaneously achieve noise reduction and binaural cue preservation of $M-2$ and 1 interferers, respectively, where M is the total number of microphones of both hearing aids. Both algorithms are characterized by the fact that for each interferer for which binaural cues are to be preserved, two constraints are introduced. As we show, this unnecessarily reduces the degrees of freedom, for noise reduction, in the optimization problem.

In this paper we present a different LCMV-based approach, where the degrees of freedom available for noise reduction is increased by spending only one constraint per interferer. As a result, the proposed method does not depend on the interference rejection parameter. The proposed method can simultaneously achieve noise reduction and binaural cue preservation of up to $2M-3$ interferers. The proposed algorithm has two advantages: a) more degrees of freedom, resulting in more sources for which binaural cues can be preserved and b) no need to predefine or optimize an additional interference rejection parameter.

2. SIGNAL MODEL AND NOTATION

The collaborating hearing aids are assumed to consist of $M = M_L + M_R$ microphones in total, where M_L and M_R are the number of microphones of the left and right hearing aid, respectively. Without loss of generality, we assume $M_L = M_R$. In this paper the

This work was supported by the Oticon Foundation and the Dutch Technology Foundation STW.

signals are processed on a frame-by-frame basis in the frequency domain. However, time-frame indices are neglected for notational convenience. Further, let ω denote the frequency variable. The M-element microphone array acquires the emitted sounds from one desired source, $S(\omega)$, and r interferers, $U_i(\omega)$, $i = 1, \cdots, r$, each placed at a potentially different location. Let $X_j(\omega) = A_j(\omega)S(\omega)$ and $N_{ij}(\omega) = B_{ij}(\omega)U_i(\omega)$ be the received desired source and received i-th interferer, respectively, at the j-th microphone with $A_j(\omega)$ and $B_{ij}(\omega)$ being the corresponding ATFs. The j-th microphone Fourier coefficient, $Y_j(\omega)$, is given by

$$Y_j(\omega) = X_j(\omega) + \sum_{i=1}^{r} N_{ij}(\omega) + V_j(\omega),\ j = 1, \cdots, M, \quad (1)$$

where $V_j(\omega)$ is additive uncorrelated noise. In the remainder of the paper, the frequency variable, ω, is omitted to simplify the notation. Using a stacked vector notation, i.e., $\mathbf{Y} = [Y_1, \cdots, Y_M]^T$, Eq. (1) can be written as

$$\mathbf{Y} = \mathbf{X} + \sum_{i=1}^{r} \mathbf{N}_i + \mathbf{V}, \quad (2)$$

where $\mathbf{X}$, $\mathbf{N}_i$ and $\mathbf{V}$ are the vector representations of the X_j, N_{ij}, V_j (for $j = 1, \cdots, M$) components, respectively. Note that $\mathbf{X} = \mathbf{A}S$ and $\mathbf{N}_i = \mathbf{B}_i U_i$, where $\mathbf{A} \in \mathbb{C}^{M\times 1}$ and $\mathbf{B}_i \in \mathbb{C}^{M\times 1}$ denote the vectors containing the ATFs of the desired source and the i-th interferer, respectively. Note also that the cross power spectral density (CPSD) matrix of $\mathbf{X}$ is given by

$$\mathbf{P_X} = E\big[\mathbf{X}\mathbf{X}^H\big] = P_S \mathbf{A}\mathbf{A}^H, \quad (3)$$

where $P_S = E\big[|S|^2\big]$ is the power spectral density (PSD) of S. Similarly, the CPSD matrix of $\mathbf{N}_i$ is given by

$$\mathbf{P}_{\mathbf{N}_i} = E\big[\mathbf{N}_i\mathbf{N}_i^H\big] = P_{U_i}\mathbf{B}_i\mathbf{B}_i^H, \quad (4)$$

where $P_{U_i} = E\big[|U_i|^2\big]$ is the PSD of U_i. Without loss of generality, we assume that $V_j(\omega)$ is spatially uncorrelated white Gaussian noise (WGN) with variance σ_j^2, and $\sigma_j = \sigma > 0$, for $j = 1, \cdots, M$. Therefore, the CPSD matrix of $\mathbf{V}$ is given by

$$\mathbf{P_V} = E\big[\mathbf{V}\mathbf{V}^H\big] = \sigma^2\mathbf{I}, \quad (5)$$

where $\mathbf{I} \in \mathbb{R}^{M\times M}$ is the identity matrix. Assuming that all sources and the additive uncorrelated noise are mutually uncorrelated, the CPSD matrix of $\mathbf{Y}$ is given by

$$\mathbf{P_Y} = \mathbf{P_X} + \underbrace{\sum_{i=1}^{r} \mathbf{P}_{\mathbf{N}_i} + \mathbf{P_V}}_{\mathbf{P}}, \quad (6)$$

where $\mathbf{P}$ is the CPSD matrix of all disturbances and is of full rank.

3. PROBLEM STATEMENT

Let us denote the first and M-th microphone as the reference microphones for the left and right hearing aid, respectively. Hence, $Y_L = Y_1$, $X_L = X_1$, $A_L = A_1$, $B_{iL} = B_{i1}$, $N_L = N_1$, $V_L = V_1$ are the reference Fourier coefficients of the left hearing aid. Similarly, the corresponding reference Fourier coefficients of the right hearing aid are denoted with subscript R. Binaural beamforming can now be formulated by having two different spatial filters, $\hat{\mathbf{w}}_L$ and $\hat{\mathbf{w}}_R$, applied to the left and right hearing aid, respectively. Each filter produces a different output given by

$$\hat{X}_L = \hat{\mathbf{w}}_L^H\mathbf{Y} \quad \text{and} \quad \hat{X}_R = \hat{\mathbf{w}}_R^H\mathbf{Y}. \quad (7)$$

3.1. LCMV

As the proposed method is based on the LCMV framework, we briefly summarize its main aspects. The LCMV problem is given by [13, 14]

$$\hat{\mathbf{w}} = \arg\min_{\mathbf{w}} \mathbf{w}^H\mathbf{P}\mathbf{w} \text{ s.t. } \mathbf{\Lambda}^H\mathbf{w} = \mathbf{f}, \quad (8)$$

where the constraint matrix $\mathbf{\Lambda} \in \mathbb{C}^{M\times d}$ and d is the number of constraints. In Secs. 3.2, 3.3 and 4, $\hat{\mathbf{w}}$ denotes $\hat{\mathbf{w}}_L$ or $\hat{\mathbf{w}}_R$ or the concatenation of these two (the actual meaning of $\hat{\mathbf{w}}$ is clear from the context). Without loss of generality, we assume that $\mathbf{\Lambda}$ has full rank. There are then three interesting cases:

1. If $d < M$, the feasible set $\{\mathbf{w} : \mathbf{\Lambda}^H\mathbf{w} = \mathbf{f}\}$ has infinitely many solutions and the problem of Eq. (8) has a closed-form solution given by [14]

$$\hat{\mathbf{w}} = \mathbf{P}^{-1}\mathbf{\Lambda}\left(\mathbf{\Lambda}^H\mathbf{P}^{-1}\mathbf{\Lambda}\right)^{-1}\mathbf{f}. \quad (9)$$

   Note that in this case there are $M - d$ degrees of freedom left for noise reduction.

2. If $d = M$, the feasible set has one unique solution given by

$$\hat{\mathbf{w}} = (\mathbf{\Lambda}^H)^{-1}\mathbf{f}. \quad (10)$$

   In this case, there are no degrees of freedom left and, thus, $\hat{\mathbf{w}}$ is unable to control the suppression of noise.

3. If $d > M$, the feasible set is empty and the problem of Eq. (8) cannot be solved.

Therefore, in order to achieve noise reduction, the matrix $\mathbf{\Lambda}$ has to be "tall" (i.e., $d < M$). Generally, the larger $M - d$, the more degrees of freedom can be devoted to noise reduction.

3.2. BLCMV

BLCMV [11] aims to preserve the binaural cues of the desired source and b ($b \leq r$) interferers, where r is the number of all present interferers (see Sec. 2). It estimates $\mathbf{w}_L^H$ and $\mathbf{w}_R^H$ by solving two independent LCMV optimization problems; one for each hearing aid. The LCMV problem of the left hearing aid is given by

$$\begin{aligned} \hat{\mathbf{w}}_L^H \quad &= \quad \arg\min_{\mathbf{w}_L} \mathbf{w}_L^H\mathbf{P}\mathbf{w}_L \\ \text{s.t.} \quad & \quad \mathbf{w}_L^H\mathbf{A} = A_L \\ & \quad \mathbf{w}_L^H\mathbf{B}_1 = \eta_L B_{1L},\ \ldots, \mathbf{w}_L^H\mathbf{B}_b = \eta_L B_{bL}. \end{aligned} \quad (11)$$

The LCMV problem of the right hearing aid has a similar form. In both LCMV problems, the constraints $\mathbf{w}_L^H\mathbf{A} = A_L$ and $\mathbf{w}_R^H\mathbf{A} = A_R$ preserve the binaural cues of the desired source, while the constraints $\mathbf{w}_L^H\mathbf{B}_i = \eta_L B_{iL}$ and $\mathbf{w}_R^H\mathbf{B}_i = \eta_R B_{iR}$, for $i = 1, \ldots, b$, preserve the binaural cues of the b interferers. η_L and η_R are predefined (η_L, $\eta_R < 1$) real-valued interference rejection parameters, which control the noise reduction of the interferers. A necessary condition for binaural cue preservation of the interferers is $\eta = \eta_L = \eta_R$ [12]. Eq. (11) can be reformulated compactly as

$$\hat{\mathbf{w}}_L = \arg\min_{\mathbf{w}_L} \mathbf{w}_L^H\mathbf{P}\mathbf{w}_L \text{ s.t. } \mathbf{\Lambda}^H\mathbf{w}_L = \mathbf{f}_L, \quad (12)$$

where

$$\mathbf{\Lambda} = \begin{bmatrix}\mathbf{A} & \mathbf{B}_1 & \cdots & \mathbf{B}_b\end{bmatrix} \in \mathbb{C}^{M\times(b+1)}, \quad (13)$$

$$\mathbf{f}_L^H = \begin{bmatrix}A_L & \eta_L B_{1L} & \cdots & \eta_L B_{bL}\end{bmatrix} \in \mathbb{C}^{1\times(b+1)}. \quad (14)$$

As explained in Sec. 3.1, $\mathbf{\Lambda}$ should be "tall" ($M > b+1$) to allow noise reduction. The least "tall" $\mathbf{\Lambda}$ has dimensions $M \times (M-1)$, where one of the columns is dedicated to the desired source. Therefore, the maximum number of columns that can be dedicated to binaural cue preservation of interferers is $M-2$. Consequently, BLCMV can simultaneously achieve noise suppression and binaural cue preservation of at most $b_{\text{max}} = M-2$ interferers.

3.3. OBLCMV

OBLCMV can preserve the binaural cues of the desired source and only one interferer (i.e., $b_{\text{max}} = 1$), say, the k-th interferer, where $k \in \{1, \ldots, r\}$. Consequently, OBLCMV solves the two LCMV problems of BLCMV where $\mathbf{\Lambda}$ and $\mathbf{f}_L^H$, $\mathbf{f}_R^H$ are given by [12]

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{A} & \mathbf{B}_k \end{bmatrix} \in \mathbb{C}^{M \times 2}, \tag{15}$$

$$\mathbf{f}_L^H = \begin{bmatrix} A_L & \eta_L B_{kL} \end{bmatrix} \in \mathbb{C}^{1 \times 2} \tag{16}$$

and

$$\mathbf{f}_R^H = \begin{bmatrix} A_R & \eta_R B_{kR} \end{bmatrix} \in \mathbb{C}^{1 \times 2}. \tag{17}$$

Unlike BLCMV, in OBLCMV η_L and η_R are complex-valued. OBLCMV estimates $\eta = \eta_L = \eta_R$ such that the binaural output SNR (defined in Sec. 5.2.2) is maximized. Note also that the matrix $\mathbf{\Lambda}$ has dimensions $M \times 2$ and, therefore, there are $M-2$ degrees of freedom left that can be devoted to noise reduction.

4. PROPOSED METHOD

Instead of solving the problem of Eq. (11) and the corresponding problem for the right hearing aid separately, the proposed method has the advantage of solving them jointly, without the need to introduce η_L and η_R. Preserving binaural cues of the i-th interferer implies that the following constraint has to be satisfied

$$\frac{\mathbf{w}_L^H \mathbf{B}_i}{\mathbf{w}_R^H \mathbf{B}_i} = \frac{B_{iL}}{B_{iR}}, \tag{18}$$

which can be reformulated as:

$$\mathbf{w}_L^H \mathbf{B}_i B_{iR} - \mathbf{w}_R^H \mathbf{B}_i B_{iL} = 0. \tag{19}$$

By using this unified constraint, the total number of constraints dedicated to binaural cue preservation of interferers is reduced by a factor of 2 (compared to (O)BLCMV). Therefore, for a given number of interferers, more degrees of freedom can be devoted to noise reduction. The proposed method thus solves the following problem

$$\begin{aligned} \hat{\mathbf{w}}_L^H, \hat{\mathbf{w}}_R^H \quad &= \quad \underset{\mathbf{w}_L, \mathbf{w}_R}{\arg\min}\ \mathbf{w}_L^H \mathbf{P} \mathbf{w}_L + \mathbf{w}_R^H \mathbf{P} \mathbf{w}_R \\ \text{s.t.} \quad & \quad \mathbf{w}_L^H \mathbf{A} = A_L \\ & \quad \mathbf{w}_R^H \mathbf{A} = A_R \\ & \quad \mathbf{w}_L^H \mathbf{B}_1 B_{1R} - \mathbf{w}_R^H \mathbf{B}_1 B_{1L} = 0 \\ & \quad \vdots \\ & \quad \mathbf{w}_L^H \mathbf{B}_b B_{bR} - \mathbf{w}_R^H \mathbf{B}_b B_{bL} = 0. \end{aligned} \tag{20}$$

Let $\hat{\mathbf{w}}^H = \begin{bmatrix} \hat{\mathbf{w}}_L^H, \hat{\mathbf{w}}_R^H \end{bmatrix}$. The above problem can then be written as

$$\hat{\mathbf{w}} = \underset{\mathbf{w}}{\arg\min}\ \mathbf{w}^H \tilde{\mathbf{P}} \mathbf{w} \text{ s.t. } \mathbf{\Lambda}^H \mathbf{w} = \mathbf{f}, \tag{21}$$

where

$$\tilde{\mathbf{P}} = \begin{bmatrix} \mathbf{P} & 0 \\ 0 & \mathbf{P} \end{bmatrix} \in \mathbb{C}^{2M \times 2M}, \tag{22}$$

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{A} & \mathbf{0} & \mathbf{B}_1 B_{1R} & \mathbf{B}_2 B_{2R} & \cdots & \mathbf{B}_b B_{bR} \\ \mathbf{0} & \mathbf{A} & -\mathbf{B}_1 B_{1L} & -\mathbf{B}_2 B_{2L} & \cdots & -\mathbf{B}_b B_{bL} \end{bmatrix} \tag{23}$$

and

$$\mathbf{f}^H = \begin{bmatrix} A_L & A_R & 0 & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{C}^{1 \times (2+b)}, \tag{24}$$

where $\mathbf{\Lambda} \in \mathbb{C}^{2M \times (2+b)}$. As explained in Sec. 3.1, if matrix $\mathbf{\Lambda}$ is "tall" (i.e., $2M > 2+b$), the proposed method can, simultaneously, achieve noise reduction and binaural cue preservation of up to $b_{\text{max}} = 2M-3$ interferers. For instance, if $M=4$, the proposed method can achieve noise reduction and preserve the binaural cues of $b_{\text{max}} = 2M-3 = 5$ interferers, while BLCMV can preserve the binaural cues of only $b_{\text{max}} = M-2 = 2$ interferers, and OBLCMV can preserve the binaural cues of $b_{\text{max}} = 1$ interferer.

5. EXPERIMENTS

We compare the proposed method with BLCMV and OBLCMV in simulation experiments using a target speech signal degraded by several additive point noise sources. We study performance in terms of noise reduction and binaural cue preservation as a function of simultaneously present interferers.

5.1. Experiment Setup

Fig. 1(a) shows the top-view of the experimental setup. Two virtual hearing aids ('+') are used. The center of the head is at the origin, $(0, 0)$. Each hearing aid consists of a linear array (in the direction of the y-axis) of two omnidirectional microphones (i.e., $M_L = M_R = 2$) having a distance of 1.2 cm. The distance between the two hearing aids is 20 cm. There is one desired talker at 135 degrees, denoted by 'o'. Experiments are performed as a function of the number, r ($1 \leq r \leq 5$), of simultaneously present interferers. Two of the interferers are speech shaped white noise (at 15 and 105 degrees) denoted by $\star$ and three are WGN sources (at 45, 75, 165 degrees) denoted by 'x' markers. In Fig. 1(a), next to each interferer, a set of r-values is indicated. For instance, the interferer with the set $\{4, 5\}$ is present for $r = 4, 5$. All sources have a duration of 24 seconds and are located 1 m from the origin. In this initial work, we ignore the presence of the head of the user, i.e., we consider a free-field and near-field acoustic situation.

To model microphone self noise, WGN is added to each microphone at an SNR of 50 dB with respect to the desired source measured at the microphones. The total average binaural input SNR (defined in Sec. 5.2.2) for $r = 1, \ldots, 5$ is -9.1, -18.9, -21.4, -21.9, -23.3 dB, respectively. The sampling frequency is 16 kHz. The enhancement is performed using the overlap-and-add analysis/synthesis method [15] with 50% overlap and using a square-root-Hann window for analysis and synthesis. Matrix $\mathbf{P}$ is calculated from the true ATFs of the r interferers and the estimated PSDs, using Welch's method, of all disturbances. Note that each of the three compared methods has a different b_{max}. All methods are tested using $b = b_{\text{max}}$, if $b_{\text{max}} \leq r$, otherwise, they are tested using $b = r$. The BLCMV algorithm is tested for $\eta = 0.1$ and 0.001. The former value was used in [11] in order to mitigate possible artifacts from estimation inaccuracies. In the present paper, we also tested the latter value which enables BLCMV to achieve even higher noise reduction and better ILD preservation than with $\eta = 0.1$.

5.2. Performance Evaluation

In this section we specify the used performance measures, which are based on [8, 12]. These measures are averaged over all frequency bins and frames. Moreover, ITD and ILD errors are averaged over all present interferers.

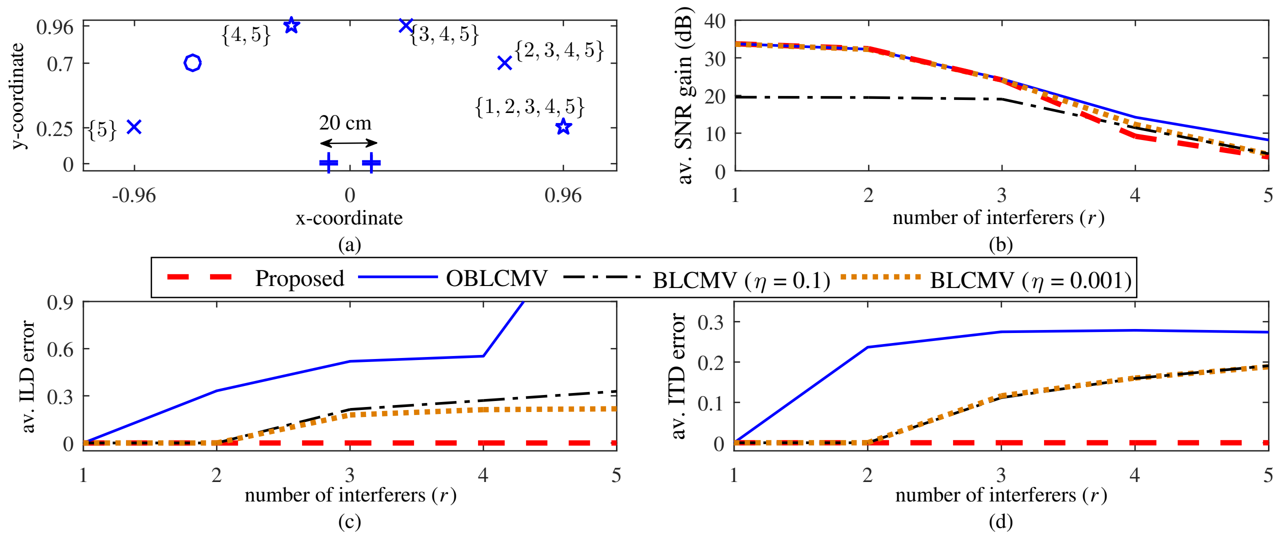

Fig. 1. Locations of sources and microphones (a), average SNR gain over all frequencies and frames (b), average ILD error over all frequencies and all interferers that are present (c), average ITD error over all frequencies and all interferers that are present (d). In (a), $\star$ and 'x' denote speech shaped white noise and WGN interferers, respectively, 'o' denotes the desired talker source and '+' denote the microphones.

5.2.1. ITFs, ITDs & ILDs

The input and output interaural transfer functions (ITFs) of the desired source are given by [8]

$$\mathrm{ITF}_{\mathbf{X}}^{\mathrm{in}} = \frac{X_L}{X_R} = \frac{A_L}{A_R} \quad \text{and} \quad \mathrm{ITF}_{\mathbf{X}}^{\mathrm{out}} = \frac{\mathbf{w}_L^H \mathbf{X}}{\mathbf{w}_R^H \mathbf{X}} = \frac{\mathbf{w}_L^H \mathbf{A}}{\mathbf{w}_R^H \mathbf{A}}. \qquad (25)$$

The input and output interaural level differences (ILDs) are the squared magnitudes of the input and output ITFs, respectively. That is,

$$\mathrm{ILD}_{\mathbf{X}}^{\mathrm{in}} = |\mathrm{ITF}_{\mathbf{X}}^{\mathrm{in}}|^2 \quad \text{and} \quad \mathrm{ILD}_{\mathbf{X}}^{\mathrm{out}} = |\mathrm{ITF}_{\mathbf{X}}^{\mathrm{out}}|^2. \qquad (26)$$

The input and output interaural time differences (ITDs) are defined as the phases of the input and output ITFs, respectively. That is,

$$\mathrm{ITD}_{\mathbf{X}}^{\mathrm{in}} = \angle \mathrm{ITF}_{\mathbf{X}}^{\mathrm{in}} \quad \text{and} \quad \mathrm{ITD}_{\mathbf{X}}^{\mathrm{out}} = \angle \mathrm{ITF}_{\mathbf{X}}^{\mathrm{out}}. \qquad (27)$$

Preservation of binaural cues implies

$$\mathrm{ITF}_{\mathbf{X}}^{\mathrm{in}} = \mathrm{ITF}_{\mathbf{X}}^{\mathrm{out}}, \quad \mathrm{ILD}_{\mathbf{X}}^{\mathrm{in}} = \mathrm{ILD}_{\mathbf{X}}^{\mathrm{out}}, \quad \mathrm{ITD}_{\mathbf{X}}^{\mathrm{in}} = \mathrm{ITD}_{\mathbf{X}}^{\mathrm{out}}. \qquad (28)$$

Note that $\mathrm{ITF}_{\mathbf{X}}^{\mathrm{in}} = \mathrm{ITF}_{\mathbf{X}}^{\mathrm{out}}$, implies preservation of the ILDs and ITDs. The preservation errors of the ILDs and ITDs are given by

$$\mathrm{ER}_{\mathrm{ILD}_{\mathbf{X}}} = |\mathrm{ILD}_{\mathbf{X}}^{\mathrm{out}} - \mathrm{ILD}_{\mathbf{X}}^{\mathrm{in}}|, \quad \mathrm{ER}_{\mathrm{ITD}_{\mathbf{X}}} = \frac{|\mathrm{ITD}_{\mathbf{X}}^{\mathrm{out}} - \mathrm{ITD}_{\mathbf{X}}^{\mathrm{in}}|}{\pi}, \qquad (29)$$

where $\mathrm{ER}_{\mathrm{ITD}_{\mathbf{X}}} \in [0, 1]$. Similar expressions can be defined for each interferer by replacing $\mathbf{X}$ with $\mathbf{N}_i$, S with U_i and $\mathbf{A}$ with $\mathbf{B}_i$.

5.2.2. SNR measures

The binaural input SNR is defined as [12]

$$\mathrm{SNR}^{\mathrm{in}} = 10\log_{10}\left(\frac{\mathbf{e}_L^T \mathbf{P}_{\mathbf{X}} \mathbf{e}_L + \mathbf{e}_R^T \mathbf{P}_{\mathbf{X}} \mathbf{e}_R}{\mathbf{e}_L^T \mathbf{P} \mathbf{e}_L + \mathbf{e}_R^T \mathbf{P} \mathbf{e}_R}\right) \text{ dB}, \qquad (30)$$

where $\mathbf{e}_L^T = [1, 0, \cdots, 0]$ and $\mathbf{e}_R^T = [0, \cdots, 0, 1]$. The binaural output SNR is defined as

$$\mathrm{SNR}^{\mathrm{out}} = 10\log_{10}\left(\frac{\mathbf{w}_L^H \mathbf{P}_{\mathbf{X}} \mathbf{w}_L + \mathbf{w}_R^H \mathbf{P}_{\mathbf{X}} \mathbf{w}_R}{\mathbf{w}_L^H \mathbf{P} \mathbf{w}_L + \mathbf{w}_R^H \mathbf{P} \mathbf{w}_R}\right) \text{ dB}. \qquad (31)$$

The binaural SNR gain is then defined as

$$\mathrm{SNR}^{\mathrm{gain}} = \mathrm{SNR}^{\mathrm{out}} - \mathrm{SNR}^{\mathrm{in}} \text{ dB}. \qquad (32)$$

5.3. Results of Experiments

The proposed algorithm and reference algorithms are compared with respect to the average binaural SNR gain (Fig. 1 (b)), average ILD errors (Fig. 1(c)) and average ITD errors (Fig. 1(d)) as a function of the number of the present interferers, r. Note that all methods preserve the ILD and ITD of the desired source, since the two constraints, $\mathbf{w}_R^H \mathbf{A} = A_R$ and $\mathbf{w}_L^H \mathbf{A} = A_L$, guarantee $\mathrm{ITF}_{\mathbf{X}}^{\mathrm{in}} = \mathrm{ITF}_{\mathbf{X}}^{\mathrm{out}}$. The figure curves showing this are left out due to space limitations. As was stated in Secs. 3 and 4, the proposed method can simultaneously perform noise suppression and preserve the binaural cues of a maximum $2M - 3$ interferers, while the BLCMV and OBLCMV algorithms preserve the binaural cues of up to only $M-2$ and 1 interferer, respectively. It is clear from the results in Figs. 1(b), (c) and (d) that with $M = 4$ microphones, the proposed method is capable of preserving the binaural cues of $2M - 3 = 5$ interferers and achieve noise reduction, while the BLCMV and OBLCMV can preserve the binaural cues of only $M - 2 = 2$ and 1 interferers, respectively. The noise reduction performance (see Fig. 1(b)) of the proposed method and BLCMV for $\eta = 0.001$ is similar, while OBLCMV achieves slightly better noise reduction for $r > 3$.

6. CONCLUSION

A new multi-microphone LCMV-based noise reduction technique is proposed, which jointly estimates the left and right beamformers of the two hearing aids. We showed that the proposed approach can simultaneously perform noise suppression and preserve the binaural cues of $2M-3$ interferers. This means that, unlike existing methods, the proposed method can preserve the binaural cues of more interferers than the number of microphones, while still achieving some noise reduction. Notice that a similar strategy to preserve the binaural cues in combination with a MWF has been proposed in [16] upon submission of the current paper.

7. REFERENCES

[1] J. Benesty, M. M. Sondhi, and Y. Huang (Eds), *Springer handbook of speech processing*, Springer, 2008.

[2] M. Brandstein and D. Ward (Eds.), *Microphone arrays: signal processing techniques and applications*, Springer, 2001.

[3] K. Eneman et al., "Evaluation of signal enhancement algorithms for hearing instruments," in *EURASIP Europ. Signal Process. Conf. (EUSIPCO)*, Lausanne, Switzerland, Aug. 2008.

[4] R. C. Hendriks, T. Gerkmann, and J. Jensen, *DFT-Domain Based Single-Microphone Noise Reduction for Speech Enhancement: A Survey of the State of the Art*, Morgan & Claypool, 2013.

[5] N. Wiener, *Extrapolation, Interpolation and Smoothing of Stationary Time Series: With Engineering Applications*, MIT Press, principles of electrical engineering series edition, 1949.

[6] P. Vary and R. Martin, *Digital speech transmission: Enhancement, coding and error concealment*, John Wiley & Sons, 2006.

[7] S. Doclo, T. J. Klasen, T. Van den Bogaert, J. Wouters, and M. Moonen, "Theoretical analysis of binaural cue preservation using multi-channel Wiener filtering and interaural transfer functions," in *Int. Workshop Acoustic Echo, Noise Control (IWAENC)*, Paris, France, Sep. 2006.

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

[9] T. J. Klasen, T. Van den Bogaert, M. Moonen, and J. Wouters, "Preservation of interaural time delay for binaural hearing aids through multi-channel Wiener filtering based noise reduction," in *IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Philadelphia PA, USA, Mar. 2005, pp. 29–32.

[10] 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 Process.*, vol. 55, no. 4, pp. 1579–1585, Apr. 2007.

[11] E. Hadad, S. Gannot, and S. Doclo, "Binaural linearly constrained minimum variance beamformer for hearing aid applications," in *Int. Workshop Acoustic Signal Enhancement (IWAENC)*, Sep. 2012, pp. 1–4.

[12] D. Marquardt, E. Hadad, S. Gannot, and S. Doclo, "Optimal binaural lcmv beamformers for combined noise reduction and binaural cue preservation," in *Int. Workshop Acoustic Signal Enhancement (IWAENC)*, Sep. 2014, pp. 288–292.

[13] O. L. Frost III, "An algorithm for linearly constrained adaptive array processing," *Proceedings of the IEEE*, vol. 60, no. 8, pp. 926–935, Aug. 1972.

[14] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Mag.*, vol. 5, no. 5, pp. 4–24, Apr. 1988.

[15] T. F. Quatieri, *Discrete-Time Speech Signal Processing: Principles and Practice*, Prentice Hall, Upper Saddle River, NJ, 2002.

[16] 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, Language Process.*, pre-published.