MULTICHANNEL DEREVERBERATION FOR HEARING AIDS WITH INTERAURAL COHERENCE PRESERVATION

Sebastian Braun[1], *Matteo Torcoli*[1], *Daniel Marquardt*[2], *Emanuël A. P. Habets*[1] *and Simon Doclo*[2]

[1]International Audio Laboratories Erlangen*, Am Wolfsmantel 33, 91058 Erlangen, Germany
[2]University of Oldenburg, Department of Medical Physics and Acoustics, Oldenburg, Germany

ABSTRACT

Dereverberation for multichannel hearing aids is still a field of extensive research. If a single desired source is assumed, many binaural spatial filtering techniques such as the minimum variance distortionless response beamformer or the rank-one multichannel Wiener filter (MWF) distort the binaural cues of the residual undesired signal components. A recently proposed spatial filter minimizes the mean squared error plus an additional cost function to preserve the long-term interaural coherence of a diffuse noise field between the hearing aid signals. In this paper, we adapt this approach to binaural dereverberation by modeling the reverberation as a time-varying diffuse sound field. Using this approach, a considerable amount of reverberation and noise reduction can be achieved. Experimental results show that we can preserve the coherence at the output without significantly impairing the reverberation and noise reduction performance.

Index Terms— Binaural dereverberation, hearing aids, cue preservation

1. INTRODUCTION

Reverberation and background noise can severely decrease the sound quality and speech intelligibility for hearing aid users. There exist many algorithms for binaural hearing aids to reduce noise and reverberation. However, commonly used binaural algorithms are essentially single-channel algorithms that do not exploit the advantages of multichannel techniques in terms of superior performance in speech distortion and artifacts. In the approaches proposed in [1–3], multiple microphones are used to estimate parameters of the sound field, but a single microphone signal on each hearing aid is filtered by a simple real-valued gain. If the gain function is equal for the left and right hearing aid, the binaural cues are perfectly preserved.

Spatial filters that filter and sum all microphones usually alter the binaural cues of the undesired sound components such as interfering speakers, noise and diffuse sound. When a rank-one minimum variance distortionless response (MVDR) beamformer or multichannel Wiener filter (MWF) as in [4] is used, the undesired sound components will be perceived with the same binaural cues as the desired sound source. This means that also non-directional components such as the diffuse sound are perceived as a directional source from the same direction as the desired sound source, which is obiously undesired.

In [5], a method to preserve the interaural coherence (IC) of a diffuse noise field is presented, where a scenario with a static desired source and a stationary diffuse noise field is considered. The IC is preserved by adding an additional term to the cost function of the MWF. In [5], the filter aims at extracting the reverberant desired signal while suppressing a stationary diffuse noise field.

In this work, our aim is to estimate the direct component of a desired source at the left and right hearing aid and hence reduce both reverberation and noise. In addition, we aim at preserving the IC of the residual reverberation and the residual noise. In the following, we assume that the direction of arrival (DOA) of the desired source is known and that the reverberation can be modeled as a quickly time-varying diffuse sound field. The idea in [5] is adopted to preserve the IC of the residual interference, and the second order statistics of the reverberation are estimated using the maximum likelihood estimator proposed in [6].

The paper is structured as follows. In Section 2, the binaural setup is presented and the problem is formulated. In Section 3, the binaural MWF with coherence preservation is derived. Parameter estimation techniques required for the filter are presented in Section 4. The evaluation of the proposed approach is carried out in Section 5 and Section 6 concludes the paper.

2. SIGNAL MODEL

We consider a binaural hearing aid setup, where the sound field is captured by two hearing aids, each equipped with M microphones. We assume an ideal data link between the two hearing aid devices, i. e., no latency and transmission errors. The observed signals in the short-time Fourier transform (STFT) domain are stacked into the $2M \times 1$ vector

$$\begin{aligned} \boldsymbol{y}(k,n) = [&Y_{L,1}(k,n), \dots, Y_{L,M}(k,n), \\ &Y_{R,1}(k,n), \dots, Y_{R,M}(k,n)]^T, \end{aligned} \qquad (1)$$

where $Y_{L,m}(k,n)$ with $m = \{1, \dots, M\}$ are the signals at the left hearing aid and $Y_{R,m}(k,n)$ the signals at the right hearing aid, respectively. The frequency and time frame indices are denoted by k and n. As proposed in [6], we assume a single plane wave originating from a single static sound source, propagating in a diffuse sound field and additive noise. The diffuse sound models the reverberation, which holds statistically for late reverberation above the Schroeder frequency [7]. The signal model is given by

$$\boldsymbol{y}(k,n) = \boldsymbol{a}_L(k) X_L(k,n) + \boldsymbol{d}(k,n) + \boldsymbol{v}(k,n), \qquad (2)$$

where $X_L(k,n)$ is the direct sound of the desired source at the reference microphone of the left hearing aid, $\boldsymbol{a}_L(k)$ is the relative transfer function (RTF) from the left reference microphone to all $2M$ microphones, $\boldsymbol{d}(k,n)$ is the diffuse sound and $\boldsymbol{v}(k,n)$ is additive

This research was partly supported by a Grant from the GIF, the German-Israeli Foundation for Scientific Research and Development.

*A joint institution of the Friedrich-Alexander-University Erlangen-Nürnberg (FAU) and Fraunhofer IIS, Germany.

stationary noise at each microphone. For a more compact notation, we define the undesired signal component by the interference vector

$$\boldsymbol{u}(k,n) = \boldsymbol{d}(k,n) + \boldsymbol{v}(k,n). \tag{3}$$

Alternatively, the direct sound component at the microphones can also be expressed in terms of the direct sound at the right reference microphone $X_R(k,n)$ and the corresponding RTF $\boldsymbol{a}_R(k)$, i. e.,

$$\boldsymbol{a}_L(k)X_L(k,n) = \boldsymbol{a}_R(k)X_R(k,n). \tag{4}$$

Note that the RTFs $\boldsymbol{a}_L(k)$ and $\boldsymbol{a}_R(k)$ are complex propagation vectors that are directly related to the head-related transfer functions (HRTFs) of the direct sound source to the hearing aids. By assuming a static scenario, the RTFs $\boldsymbol{a}_L(k)$ and $\boldsymbol{a}_R(k)$ are independent of time.

The power spectral density (PSD) matrix of the observed signal is given by

$$\boldsymbol{\Phi}_{\boldsymbol{y}}(k,n) = E\left\{\boldsymbol{y}(k,n)\boldsymbol{y}^H(k,n)\right\}, \tag{5}$$

where $E\{\cdot\}$ denotes the expectation operator. The PSD matrices $\boldsymbol{\Phi}_{\boldsymbol{u}}(k,n)$ and $\boldsymbol{\Phi}_{\boldsymbol{v}}(k,n)$ are defined similarly. As in commonly used related sound field models [6, 8], we assume all three sound components in (2) to be mutually uncorrelated. The diffuse sound field is described by a homogenous and isotropic sound field with the coherence matrix $\boldsymbol{\Gamma}_{\text{d}}(k)$ that also takes the head shadowing into account. Therefore, the diffuse sound can be expressed in terms of the diffuse coherence matrix scaled by the time-varying diffuse sound PSD $\phi_{\text{d}}(k,n)$. As a consequence, the interference PSD matrix is given by

$$\boldsymbol{\Phi}_{\boldsymbol{u}}(k,n) = \phi_{\text{d}}(k,n)\boldsymbol{\Gamma}_{\text{d}}(k) + \boldsymbol{\Phi}_{\boldsymbol{v}}(k,n). \tag{6}$$

Our aim is to obtain an estimate of the direct sound at the left and right reference microphones $X_L(k,n)$ and $X_R(k,n)$ by applying the complex filter weights $\boldsymbol{h}_L(k,n)$ and $\boldsymbol{h}_R(k,n)$ to the input signals, i. e.,

$$\begin{bmatrix}\hat{X}_L(k,n)\\ \hat{X}_R(k,n)\end{bmatrix} = \left[\boldsymbol{h}_L(k,n)\ \boldsymbol{h}_R(k,n)\right]^H \boldsymbol{y}(k,n), \tag{7}$$

where the $2M \times 1$ filter coefficients $\boldsymbol{h}_L(k,n)$ and $\boldsymbol{h}_R(k,n)$ are derived in Section 3.

The time and frequency dependent IC of the undesired component at the input is given by

$$\text{IC}_u^{\text{in}}(k,n) = \frac{\boldsymbol{e}_L^T\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{e}_R}{\sqrt{\boldsymbol{e}_L^T\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{e}_L\boldsymbol{e}_R^T\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{e}_R}}, \tag{8}$$

where $\boldsymbol{e}_L$ and $\boldsymbol{e}_R$ are zero-column vectors with a one at the corresponding position selecting the reference microphone, such that $\boldsymbol{e}_L^T\,\boldsymbol{a}_L(k) = 1$ and $\boldsymbol{e}_R^T\,\boldsymbol{a}_R(k) = 1$. The frequency and time indices k and n are partly omitted here and in following equations wherever necessary for space constraints.

The IC of the undesired component at the output is given by

$$\text{IC}_u^{\text{out}}(k,n) = \frac{\boldsymbol{h}_L^H\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{h}_R}{\sqrt{\boldsymbol{h}_L^H\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{h}_L\boldsymbol{h}_R^H\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{h}_R}}. \tag{9}$$

3. MWF WITH BINAURAL CUE PRESERVATION

In this section, the binaural MWF is derived and an additional term for IC preservation is added.

3.1. Binaural multichannel Wiener filter

A typical approach well suited for joint dereverberation and noise reduction is the MWF. The binaural MWF is obtained by minimizing the cost function

$$J_{\text{MWF}}(\boldsymbol{h}) = E\left\{\left\|\begin{bmatrix}X_L - \boldsymbol{h}_L^H\boldsymbol{y}\\ X_R - \boldsymbol{h}_R^H\boldsymbol{y}\end{bmatrix}\right\|^2\right\}, \tag{10}$$

where $\boldsymbol{h}$ is the stacked $4M \times 1$ vector containing the filter coefficients for the left and right output signal. Using the assumption that only a single desired sound source is active, the filter minimizing (10) is given by [4]

$$\boldsymbol{h}_{\text{MWF,L}}(k,n) = \frac{\phi_{X_L}\,\boldsymbol{\Phi}_{\boldsymbol{u}}^{-1}\,\boldsymbol{a}_L}{1\ +\ \phi_{X_L}\,\boldsymbol{a}_L^H\,\boldsymbol{\Phi}_{\boldsymbol{u}}^{-1}\,\boldsymbol{a}_L}$$
$$\boldsymbol{h}_{\text{MWF,R}}(k,n) = \frac{\phi_{X_R}\,\boldsymbol{\Phi}_{\boldsymbol{u}}^{-1}\,\boldsymbol{a}_R}{1\ +\ \phi_{X_R}\,\boldsymbol{a}_R^H\,\boldsymbol{\Phi}_{\boldsymbol{u}}^{-1}\,\boldsymbol{a}_R}, \tag{11}$$

where $\phi_{X_L}(k,n)$ and $\phi_{X_R}(k,n)$ denote the PSDs of the direct sound at the left and right reference microphones.

3.2. Additional constraint for coherence preservation

In [5, 9] it has been shown that the obtained filter $\boldsymbol{h}_{\text{MWF}}(k,n)$ perfectly preserves the binaural cues of the desired source, i. e., the interaural transfer function (ITF) and the IC. In contrast, the binaural cues of the interference are severely distorted since the residual interference is perceived with the same ITF and IC as the desired sound source. For a single desired source, this is theoretically proven in [9].

In our signal model, the interference consists of a quickly time-varying diffuse field and potentially slowly time-varying noise. The IC describes how a non-directional sound field is perceived. In [5], a method to preserve the IC has been presented in the context of noise reduction of a stationary diffuse noise field. The authors of [5] add a term to the cost function (10), such that a trade-off between interference reduction and coherence preservation is achieved. The additional coherence preservation term is given by

$$J_{\text{IC}}(\boldsymbol{h}) = \left|\frac{\boldsymbol{h}_L^H\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{h}_R}{\sqrt{\boldsymbol{h}_L^H\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{h}_L\boldsymbol{h}_R^H\boldsymbol{\Phi}_{\boldsymbol{u}}\boldsymbol{h}_R}} - \text{IC}_u^{\text{des}}(k,n)\right|^2, \tag{12}$$

where $\text{IC}_u^{\text{des}}(k,n)$ denotes the desired coherence of the interference component at the output.

Note that in contrast to the problem presented in [5], in our signal model the interference PSD matrix, input and output IC are time-varying. How to choose the desired IC is a critical task. If the aim is to preserve the spatial sound impression as close as possible to the true sound scene while just reducing the diffuse sound and noise, the desired interference output IC has to be chosen equal to the corresponding input IC as

$$\text{IC}_u^{\text{des}}(k,n) = \text{IC}_u^{\text{in}}(k,n). \tag{13}$$

Therefore, also the desired coherence $\text{IC}_u^{\text{des}}(k,n)$ can be highly time-varying, depending on the diffuse-to-noise ratio.

The MWF with coherence preservation, denoted as MWF-IC, is obtained by minimizing the total cost function

$$J_{\text{MWF-IC}}(\boldsymbol{h}) = J_{\text{MWF}}(\boldsymbol{h}) + \lambda\, J_{\text{IC}}(\boldsymbol{h}), \tag{14}$$

where the parameter λ allows for a trade-off between interference reduction and coherence preservation. Since there is no closed-form

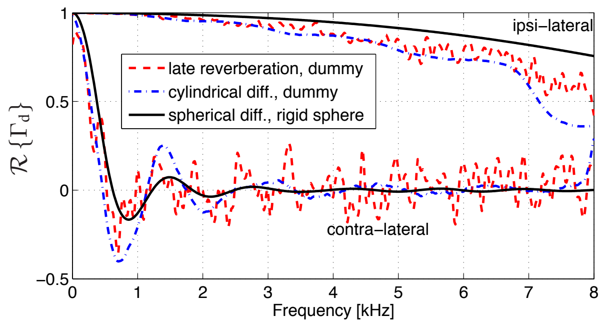

Fig. 1. Coherence of various diffuse fields for ipsi-lateral and contra-lateral hearing aid microphone pairs.

solution available for the filter minimizing (14), we use an iterative optimization technique based on a trust-region method using analytical expressions of the gradient and the Hessian of $J_{\text{MWF-IC}}(\boldsymbol{h})$. A more detailed description can be found in [4, 5].

4. PARAMETER ESTIMATION

To compute the MWF and MWF-IC filters discussed in Sections 3.1 and 3.2, the RTFs $\boldsymbol{a}_L(k), \boldsymbol{a}_R(k)$, the desired signal PSDs $\phi_{X_L}(k,n), \phi_{X_R}(k,n)$ and the interference PSD matrix $\boldsymbol{\Phi_u}(k,n)$ are required. Of paramount importance in the context of dereverberation is the diffuse sound PSD $\phi_{\text{d}}(k,n)$. In addition, the binaural diffuse coherence matrix is not as easy to obtain as in a free-field condition due to head shadowing effects. The diffuse coherence matrix $\boldsymbol{\Gamma}_{\text{d}}(k)$ can be obtained by measuring an ideal diffuse field with hearing aids worn by the human user or an artificial dummy head. This can be approximated by a full spherical set of anechoic impulse responses measured with a sufficiently high angular resolution. Another possibility is to use a head model, where analytical solutions are available, for example a rigid sphere head model [10]. For a spherical diffuse field, a solution for the coherence between two points on a rigid sphere is given in [11] using a spherical harmonics approximation.

Figure 1 shows the real part of the coherence between two ipsi-lateral (high coherence) and contra-lateral microphones (lower coherence) using different methods. The red dashed line is the coherence obtained from room impulse response measurements of hearing aids installed on a dummy head as provided by [12]. The coherence is computed from the late part of the impulse responses and averaged over multiple measurement positions and rooms. The blue dash-dotted line shows the diffuse coherence computed from a circular set of measurement points around the azimuth, taken from the anechoic impulse responses from the same database. This is the coherence of a cylindrical diffuse field, since no measurement points for a full sphere are available. The black solid line shows the coherence of an ideal spherical diffuse field obtained using the rigid sphere model proposed in [11]. The diameter of the sphere and the microphone positions were chosen accordingly to the parameters of the dummy head wearing hearing aids. This theoretical coherence is a quite good approximation of the measured coherence of the late reverberant sound field. The late reverberation of the measured rooms is neither a fully spherical nor cylindrical diffuse field, but both assumptions seem to be a reasonable choice for these rooms.

The diffuse PSD $\phi_{\text{d}}(k,n)$ can be estimated using the maximum likelihood estimator using reference signals from a blocking matrix as proposed in [6]. This estimator only requires knowledge of the RTFs $\boldsymbol{a}_L(k)$ or $\boldsymbol{a}_R(k)$ that are also required to compute the filter. The RTFs can be obtained by computing them from measured HRTFs and selecting the corresponding direction from an estimate of the DOA.

The left direct signal PSD was obtained by

$$\phi_{X_L}(k,n) = \frac{\boldsymbol{i}_L^T \operatorname{diag}\{\boldsymbol{\Phi_y} - \boldsymbol{\Phi_u}\}}{M}, \tag{15}$$

where the elements $1 \ldots M$ of the vector $\boldsymbol{i}_L$ are one and the elements $M+1 \ldots 2M$ are zero. The right direct signal PSD was computed similarly. Alternatively, we can compute $\phi_{X_R}(k,n)$ using (4).

As an estimate of the interference matrix $\boldsymbol{\Phi_u}(k,n)$ is available, the desired coherence $\text{IC}_u^{\text{des}}(k,n)$ can be calculated using (8) and (13). In this case, $\text{IC}_u^{\text{des}}(k,n)$ varies depending on the diffuse-to-noise ratio $\zeta(k,n)$. The two extreme cases for either a completely diffuse or a completely uncorrelated noise field are given by

$$\text{IC}_u^{\text{des}}(k,n) = \begin{cases} \boldsymbol{e}_L^T \boldsymbol{\Gamma}_{\text{d}}(k) \boldsymbol{e}_R & \text{for } \zeta(k,n) \to \infty \\ 0 & \text{for } \zeta(k,n) \to 0. \end{cases} \tag{16}$$

5. EVALUATION

In this section, the proposed system is evaluated in terms of its coherence preservation and speech enhancement performance.

5.1. Setup

We evaluated the proposed algorithm using the database of measurements with a dummy head equipped with hearing aids published in [12]. Each hearing aid is equipped with $M = 3$ microphones. We used measured reverberant impulse resposnes from a cafeteria scenario. The source is 1.6 m slightly elevated in front of the dummy head. The reverberation time is about 800 ms. A speech signal was convolved with the impulse response and stationary white Gaussian noise was added with an SNR of 50 dB. This rather high SNR was chosen to focus on the reverberation. The sampling rate was 16 kHz, the STFT length was 512 samples with a Hann window of length of 32 ms and 75 % overlap. The input PSD matrix was computed by recursive averaging with a time constant of 40 ms. The noise PSD matrix was assumed to be known and computed by recursive averaging with a time constant of 150 ms. The RTFs were also assumed to be known and were calculated using the direct-path extracted from the reverberant impulse responses. The diffuse coherence matrix was computed with the rigid sphere model as discussed in Section 4. The trade-off parameter for coherence preservation λ was in [5] computed adaptively so that the magnitude-squared coherence (MSC) error was kept below a certain value. In contrast to [5], we now need to compute the filter for each time-frequency bin. To reduce the computation time, a fixed $\lambda = 100$ for all frequencies was used. The performance can be possibly improved by computing λ adaptively, since it provides a trade-off between coherence preservation and interference reduction.

5.2. Preservation of the coherence

Since the actual time-frequency dependent IC can be complex, we evaluate the MSC of the interference, given by

$$\text{MSC}_u(k,n) = |\text{IC}_u(k,n)|^2. \tag{17}$$

The MSC at the output of the filter is calculated from the time-domain signal. This means that there is an additional inverse STFT

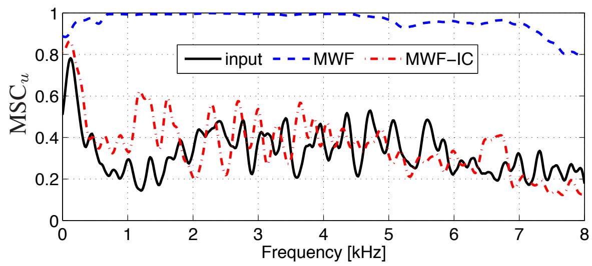

Fig. 2. Temporally averaged MSC at the input and at the output of the conventional MWF and the MWF-IC.

and STFT between the domain where the coherence error (12) is minimized and the evaluation domain of the MSC. This should give results closer to the perceptual impression. The STFT parameters for this analysis are the same as for the processing as described in Section 5.1. The expectation of the short-term MSC was carried out by averaging over a rectangular sliding window with a length of 100 ms.

Figure 2 shows the MSC of the interference $\boldsymbol{u}(k,n)$ between the reference microphones at the input in black averaged over time. Note that the actual input coherence that we want to preserve can differ from the estimated desired IC due to the required short-term estimation procedure of $\mathrm{IC}_u^{\mathrm{in}}(k,n)$. We can clearly observe that the conventional binaural MWF distorts the IC: as expected, the blue dashed line is close to one as the MSC of the desired direct source. The output MSC of the MWF-IC (red dashed-dotted line) is similar to the input MSC. The output MSC of the MWF-IC deviates slightly from the input MSC due to the overlap-add effects of the inverse STFT and model errors between the input IC and the estimate of the desired IC.

The binaural cues are mainly characterized in terms of the interaural time difference (ITD) and interaural level difference (ILD). Figure 3 shows the joint probability density function (PDF) for ITD and ILD of the interference component. The ITD and ILD are computed of short-term segments per critical band of a Gammatone filterbank [13], without using the coherence thresholding for computing reliable cues. We show the results for two bands with center frequencies at 0.5 kHz in the left column and 2 kHz in the right column. Since the input interference is quite diffuse, the ILD and ITD are kind of Gaussian distributed. This is not the case for the MWF: ITD and ILD are very narrowly distributed around the cues of the direct sound. By using the MWF-IC, the cues of the interference become more spread and are very similar to the cues at the input.

5.3. Dereverberation performance

Table 1 shows some objective measures for the dereverberation and noise reduction performance of the MWF-IC compared to the MWF. We used the measures PESQ [14], cepstral distance (CD) [15], signal-to-reverberation-modulation ratio (SRMR) [16] and the segmental signal-to-interference ratio (segSIR). The table shows averaged results between the left and right signal. It can be observed that both approaches (i. e., the MWF and the MWF-IC) contain less reverberation and noise compared to the unprocessed signal. By preserving the coherence, the performance slightly degrades as expected. Informal listening tests revealed that the perceptual difference of the interference reduction between the MWF and the MWF-IC is quite small[1]. On the other hand by using the MWF-IC, the diffuse sound is also perceived as a diffuse sound and not as a very narrow directional source as with the MWF. Formal studies about the benefits of this effect on the cognitive human hearing system are still topics of future work.

[1] Sound examples are available at http://www.audiolabs-erlangen.de/resources/2014-IWAENC-BDICP

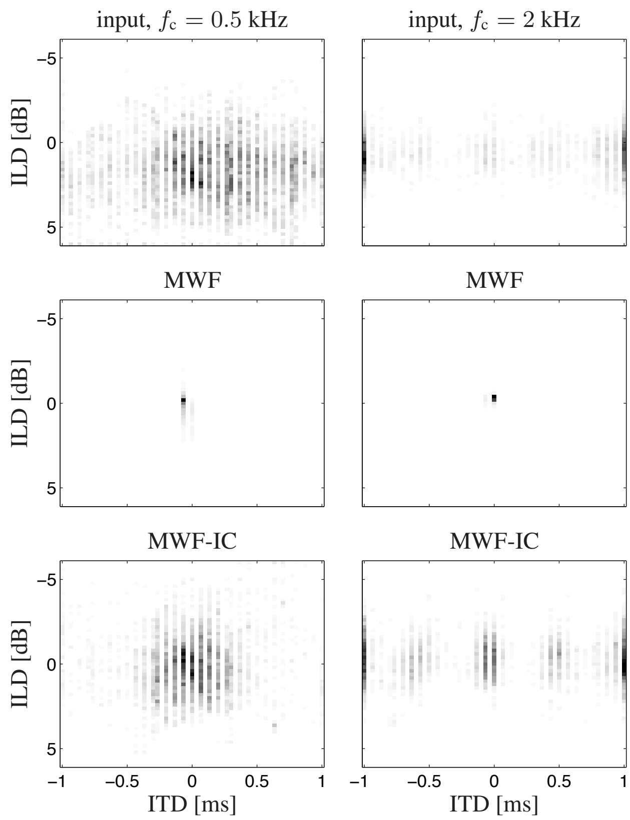

Fig. 3. PDF of the ITD and ILD of the undesired signal component for critical bands centered at 0.5 kHz in the left column and 2 kHz in the right column.

Method	PESQ	CD	SRMR	segSIR [dB]
unprocessed	2.46	3.41	4.65	7.64
MWF	3.04	2.14	6.22	15.64
MWF-IC	2.98	2.41	6.13	14.69

Table 1. Objective measures for dereverberation and noise reduction performance. Average between left and right signal.

6. CONCLUSION

We proposed a method for dereverberation in the presence of sensor noise for binaural hearing aids while the interaural coherence of the diffuse sound is preserved. By assuming a parametric sound field model consisting of a single plane wave per time-frequency bin, time-varying diffuse sound and stationary noise, we can obtain a dereverberated binaural signal using a modified MWF. The modified MWF mitigates the distortion of the binaural cues of the residual diffuse sound and stationary noise components. It is shown with measured signals that the interaural coherence can be sufficiently preserved while the interference reduction performance does not deteriorate severely. In this way, the spatial impression of the sound scene can be preserved while reverberation and noise is reduced.

7. REFERENCES

[1] T. Lotter and P. Vary, "Dual-channel speech enhancement by superdirective beamforming," *EURASIP Journal on Advances in Signal Processing*, vol. 2006, no. 1, pp. 063297, 2006.

[2] T. Rohdenburg, V. Hohmann, and B. Kollmeier, "Robustness analysis of binaural hearing aid beamformer algorithms by means of objective perceptual quality measures," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, Oct. 2007, pp. 315–318.

[3] A.H. Kamkar-Parsi and M. Bouchard, "Instantaneous binaural target psd estimation for hearing aid noise reduction in complex acoustic environments," *IEEE Trans. on Instrumentation and Measurement*, vol. 60, no. 4, pp. 1141–1154, Apr. 2011.

[4] S. Doclo, S. Gannot, M. Moonen, and A. Spriet, "Acoustic beamforming for hearing aid applications," in *Handbook on Array Processing and Sensor Networks*, S. Haykin and K.J. Ray Liu, Eds., chapter 9. Wiley, 2010.

[5] D. Marquardt, V. Hohmann, and S. Doclo, "Coherence preservation in multi-channel Wiener filtering based noise reduction for binaural hearing aids," in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, 2013, pp. 8648–8652.

[6] S. Braun and E. A. P. Habets, "Dereverberation in noisy environments using reference signals and a maximum likelihood estimator," in *Proc. European Signal Processing Conf. (EUSIPCO)*, Marrakesh, Morocco, Sept. 2013.

[7] M. R. Schroeder, "Statistical parameters of the frequency response curves of large rooms," *Journal Audio Eng. Soc.*, vol. 35, pp. 299–306, 1954.

[8] O. Thiergart and E. A. P. Habets, "An informed LCMV filter based on multiple instantaneous direction-of-arrival estimates," in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, 2013.

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

[10] R. O. Duda and W. L. Martens, "Range dependence of the response of a spherical head model," *J. Acoust. Soc. Am.*, vol. 104, no. 5, pp. 3048–3058, Jan. 1998.

[11] D. P. Jarrett, E. A. P. Habets, M. R. P. Thomas, N. D. Gaubitch, and P. A. Naylor, "Dereverberation performance of rigid and open spherical microphone arrays: Theory & simulation," in *Proc. Joint Workshop on Hands-Free Speech Communication and Microphone Arrays (HSCMA)*, Edinburgh, UK, June 2011, pp. 145–150.

[12] H. Kayser, S. Ewert, J. Anemuller, 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, no. 1, pp. 298605, 2009.

[13] C. Faller and J. Merimaa, "Sound localization in complex listening situations: Selection of binaural cues based on interaural coherence," *J. Acoust. Soc. Am.*, vol. 116, pp. 3075–3089, 2004.

[14] ITU-T, "Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs," Feb. 2001.

[15] N. Kitawaki, H. Nagabuchi, and K. Itoh, "Objective quality evaluation for low bit-rate speech coding systems," *IEEE J. Sel. Areas Commun.*, vol. 6, no. 2, pp. 262–273, 1988.

[16] T. Falk, C. Zheng, and W.-Y. Chan, "A non-intrusive quality and intelligibility measure of reverberant and dereverberated speech," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 18, no. 7, pp. 1766–1774, Sept. 2010.