

INTERAURAL COHERENCE PRESERVATION IN MWF-BASED BINAURAL NOISE REDUCTION ALGORITHMS USING PARTIAL NOISE ESTIMATION

Daniel Marquardt, Volker Hohmann, Simon Doclo

University of Oldenburg, Department of Medical Physics and Acoustics and Cluster of Excellence Hearing4All, Oldenburg, Germany

{daniel.marquardt, volker.hohmann, simon.doclo}@uni-oldenburg.de

ABSTRACT

Besides noise reduction an important objective of binaural speech enhancement algorithms is the preservation of the binaural cues of both desired and undesired sound sources. Recently an extension of the binaural Multi-channel Wiener filter (MWF), namely the MWF-IC, has been presented which aims to preserve the Interaural Coherence (IC) of the noise component. Since for the MWF-IC no closed-form solution exists, in this paper we propose to preserve the IC using the binaural MWF with partial noise estimation (MWF-N), for which a closed-form solution exists. Furthermore, we derive a closed-form expression for the trade-off parameter in the MWF-N yielding a predefined IC at the filter output. Experimental results in a diffuse noise scenario show that both the MWF-IC and the MWF-N preserve the IC of the output noise component. However, the MWF-IC yields a better noise reduction performance whereas the MWF-N introduces less speech distortion.

Index Terms— Hearing aids, binaural cues, noise reduction

1. INTRODUCTION

Noise reduction algorithms in hearing aids are crucial to improve speech understanding in background noise for hearing impaired persons. For binaural hearing aids, algorithms that exploit the microphone signals from both the left and the right hearing aid are considered to be promising techniques for noise reduction, because in addition to spectral information spatial information can be exploited [1]. In addition to reducing noise and limiting speech distortion, another important objective of binaural noise reduction algorithms is the preservation of the listener's impression of the acoustical scene, in order to exploit the binaural hearing advantage and to avoid confusion due to a mismatch between the acoustical and the visual information. This can be achieved by preserving the binaural cues of all sound sources in the acoustical scene.

In [1] the binaural Speech Distortion Weighted Multi-channel Wiener Filter (MWF) has been presented. It has been theoretically proven in [2] that in case of a single speech source the binaural MWF preserves the binaural cues of the speech component but distorts the binaural cues of the noise component such that both, speech and noise components, comprise the same binaural cues and hence are perceived as coming from the speech direction. Due to this perceptual disadvantage, several extensions of the binaural MWF in [2, 3] and the binaural LCMV [4] have been presented in order to also preserve the so-called Interaural Transfer Function (ITF) of

directional interferences. However, for diffuse noise fields whose characteristics can not be properly described by the ITF but rather by the Interaural Coherence (IC), these extensions are not able to preserve the spatial characteristics. Hence in [5] a MWF-based IC preservation filter, namely the MWF-IC, has been presented allowing to preserve the IC of the residual noise in diffuse noise fields. Since for the MWF-IC a trade-off between IC preservation and output SNR exists, in [6] it has been proposed to control the amount of IC preservation based on the IC discrimination abilities of the human auditory system. Unfortunately for the MWF-IC no closed form solution exists such that one has to rely on numerical optimization routines, which might be computationally intensive. In this paper we propose to preserve the IC of the noise component using the binaural MWF with partial noise estimation (MWF-N) [7] for which a closed form solution exists such that the computational complexity is rather low as compared to the MWF-IC. Furthermore, a closed form solution is also derived for the trade-off parameter in the MWF-N that yields a desired IC at the output of the MWF-N, whereas for the MWF-IC this trade-off parameter has to be determined in an exhaustive search. Experimental results illustrate that the MWF-IC and the MWF-N show the same capabilities in preserving the IC of the noise component whereas the intelligibility weighted output SNR of the MWF-IC is higher than for the MWF-N while the MWF-N introduces less speech distortion.

2. CONFIGURATION AND NOTATION

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

$$Y_{0,m}(\omega) = X_{0,m}(\omega) + V_{0,m}(\omega), \quad m = 1, \dots, M,$$

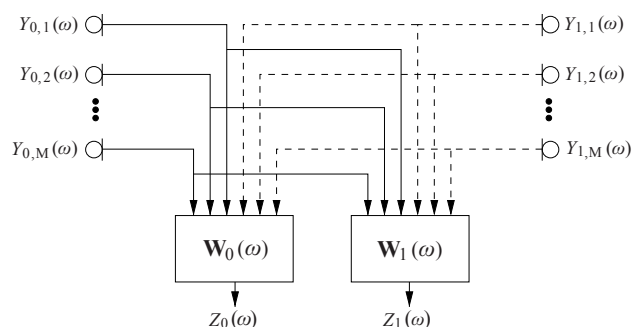


Fig. 1. Binaural hearing aid configuration

This work was supported in part by the Cluster of Excellence 1077 "Hearing4All", funded by the German Research Foundation (DFG) and a Grant from the GIF, the German-Israeli Foundation for Scientific Research and Development.

with $X_{0,m}(\omega)$ and $V_{0,m}(\omega)$ representing the speech and the noise component and ω denoting the frequency index. The m -th microphone signal in the right hearing aid $Y_{1,m}(\omega)$ is defined similarly. For conciseness we will omit the frequency variable ω in the remainder of the paper. We define the $2M$ -dimensional signal vector \mathbf{Y} as

$$\mathbf{Y} = [Y_{0,1} \dots Y_{0,M} \ Y_{1,1} \dots Y_{1,M}]^T, \quad (1)$$

which can be written as $\mathbf{Y} = \mathbf{X} + \mathbf{V}$, where \mathbf{X} and \mathbf{V} are defined similarly as \mathbf{Y} . Furthermore, we define the $4M$ -dimensional stacked weight vector \mathbf{W} as

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_0 \\ \mathbf{W}_1 \end{bmatrix}. \quad (2)$$

The output signal at the left hearing aid Z_0 is equal to

$$Z_0 = \mathbf{W}_0^H \mathbf{Y} = \mathbf{W}_0^H \mathbf{X} + \mathbf{W}_0^H \mathbf{V} = Z_{x,0} + Z_{v,0}, \quad (3)$$

where $Z_{x,0}$ represents the output speech component and $Z_{v,0}$ represents the output noise component. The output signal at the right hearing aid Z_1 can be defined similarly.

For the special case of a single speech source S , the received speech component can be written as $\mathbf{X} = S\mathbf{A}$, with \mathbf{A} the Acoustic Transfer Function (ATF) between the speech source and the microphones. The speech correlation matrix can then be calculated as

$$\mathbf{R}_x = \mathcal{E} \{ \mathbf{X}\mathbf{X}^H \} = P_s \mathbf{A}\mathbf{A}^H, \quad (4)$$

with $P_s = \mathcal{E} \{ |S|^2 \}$, the Power Spectral Density (PSD) of the speech source. Furthermore, in the case of a diffuse noise field, the noise correlation matrix can be calculated as

$$\mathbf{R}_v = \mathcal{E} \{ \mathbf{V}\mathbf{V}^H \} = P_v \mathbf{\Gamma}, \quad (5)$$

with $P_v = \mathcal{E} \{ |V|^2 \}$, the PSD of the noise component in all microphone signals and $\mathbf{\Gamma}$, the spatial coherence matrix of the diffuse noise field. The Relative Transfer Functions (RTF) of the input and output speech component are defined as

$$RTF_x^{in} = \frac{\mathbf{e}_0^T \mathbf{A}}{\mathbf{e}_1^T \mathbf{A}} = \frac{A_0}{A_1}, \quad RTF_x^{out} = \frac{\mathbf{W}_0^H \mathbf{A}}{\mathbf{W}_1^H \mathbf{A}}. \quad (6)$$

The vectors \mathbf{e}_0 and \mathbf{e}_1 are zero column vectors with $\mathbf{e}_0(1) = 1$ and $\mathbf{e}_1(M+1) = 1$, defining the reference microphones of the left and the right hearing aid. The input Interaural Coherence (IC) of the noise component is defined as

$$IC_v^{in} = \frac{\mathbf{e}_0^T \mathbf{R}_v \mathbf{e}_1}{\sqrt{\mathbf{e}_0^T \mathbf{R}_v \mathbf{e}_0 \mathbf{e}_1^T \mathbf{R}_v \mathbf{e}_1}} = \mathbf{e}_0^T \mathbf{\Gamma} \mathbf{e}_1. \quad (7)$$

For a binaural hearing aid setup, the input IC of a diffuse noise field is real-valued and can be calculated e.g. based on physical models of the head [8] or a modified sinc-function [9]. The output IC of the noise component is equal to

$$IC_v^{out} = \frac{\mathbf{W}_0^H \mathbf{R}_v \mathbf{W}_1}{\sqrt{\mathbf{W}_0^H \mathbf{R}_v \mathbf{W}_0 \mathbf{W}_1^H \mathbf{R}_v \mathbf{W}_1}} \quad (8)$$

The (real-valued) Magnitude Squared Coherence (MSC) is defined as $MSC = |IC|^2$. The IC of the input and output speech component can be defined similarly as for the noise component.

3. BINAURAL NOISE REDUCTION ALGORITHMS

In this section we introduce the binaural MWF with partial noise estimation (MWF-N) and show its relation to the binaural MWF. Furthermore, we review the MWF with IC preservation (MWF-IC).

3.1. Binaural multi-channel Wiener filter with partial noise estimation (MWF-N)

The MWF-N is an extension of the MWF which in addition to preserving the binaural cues of the speech component also aims to partially preserve the binaural cues of the noise component [7, 2]. The MWF-N produces a minimum mean-square error (MMSE) estimate of the speech component and a portion of the noise component in the reference microphone signal for both hearing aids. The MWF-N cost function can be written as

$$J_{\text{MWF-N}}(\mathbf{W}) = \mathcal{E} \left\{ \left\| \begin{bmatrix} X_0 - \mathbf{W}_0^H \mathbf{X} \\ X_1 - \mathbf{W}_1^H \mathbf{X} \end{bmatrix} \right\|^2 + \mu \left\| \begin{bmatrix} \eta V_0 - \mathbf{W}_0^H \mathbf{V} \\ \eta V_1 - \mathbf{W}_1^H \mathbf{V} \end{bmatrix} \right\|^2 \right\}, \quad (9)$$

where the parameter $\mu \in \mathbb{R}$ provides a trade-off between noise reduction and speech distortion and the parameter $\eta \in \mathbb{R}$ provides a trade-off between noise reduction and the preservation of the binaural cues of the noise component. If $\eta = 0$, the MWF-N cost function reduces to the cost function of the MWF. The solution to the optimization problem in (9) is equal to

$$\mathbf{W}_{\text{MWF-N},0} = (1 - \eta) \mathbf{W}_{\text{MWF},0} + \eta \mathbf{e}_0, \quad (10)$$

$$\mathbf{W}_{\text{MWF-N},1} = (1 - \eta) \mathbf{W}_{\text{MWF},1} + \eta \mathbf{e}_1, \quad (11)$$

with

$$\mathbf{W}_{\text{MWF},0} = (\mathbf{R}_x + \mu \mathbf{R}_v)^{-1} \mathbf{R}_x \mathbf{e}_0, \quad (12)$$

$$\mathbf{W}_{\text{MWF},1} = (\mathbf{R}_x + \mu \mathbf{R}_v)^{-1} \mathbf{R}_x \mathbf{e}_1. \quad (13)$$

Hence, the solution to the MWF-N corresponds to a mixing of the output signals of the MWF weighted with the factor $(1 - \eta)$ and the reference microphone signals weighted with η . Again, setting $\eta = 0$ gives the MWF solution. Applying the matrix inversion lemma to (12) and (13), the MWF can be decomposed into a MVDR beamformer $\bar{\mathbf{W}}$ and a single-channel Wiener postfilter applied to the output of the MVDR beamformer [10], i.e.,

$$\mathbf{W}_0 = \bar{\mathbf{W}}_0 \frac{\rho}{\mu + \rho}, \quad \bar{\mathbf{W}}_0 = \frac{\mathbf{\Gamma}^{-1} \mathbf{H}_0^H}{\mathbf{H}_0^H \mathbf{\Gamma}^{-1} \mathbf{H}_0}, \quad (14)$$

$$\mathbf{W}_1 = \bar{\mathbf{W}}_1 \frac{\rho}{\mu + \rho}, \quad \bar{\mathbf{W}}_1 = \frac{\mathbf{\Gamma}^{-1} \mathbf{H}_1^H}{\mathbf{H}_1^H \mathbf{\Gamma}^{-1} \mathbf{H}_1}, \quad (15)$$

with

$$\rho = \frac{P_s |A_0|^2}{P_v} \mathbf{H}_0^H \mathbf{\Gamma}^{-1} \mathbf{H}_0 = \frac{P_s |A_1|^2}{P_v} \mathbf{H}_1^H \mathbf{\Gamma}^{-1} \mathbf{H}_1 \quad (16)$$

the narrowband output SNR of the left and the right MVDR beamformer. Furthermore,

$$SNR_0^{in} = \frac{P_s |A_0|^2}{P_v}, \quad SNR_1^{in} = \frac{P_s |A_1|^2}{P_v}, \quad (17)$$

is the narrowband input SNR in the left and the right hearing aid and

$$\mathbf{H}_0 = \frac{\mathbf{A}}{A_0}, \quad \mathbf{H}_1 = \frac{\mathbf{A}}{A_1} \quad (18)$$

is the RTF vector of the left and the right hearing aid. Due to the mixing of the MWF output with the input signal, the output SNR of the MWF-N will be lower than for the MWF which has been analytically proven in [2].

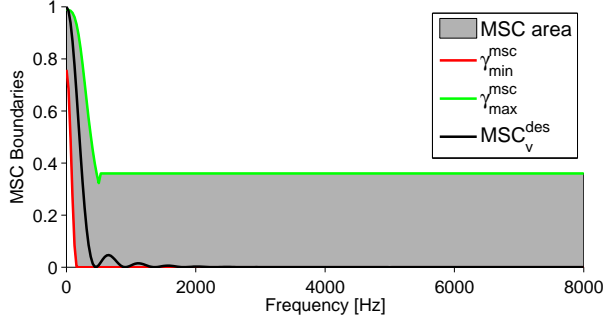


Fig. 2. MSC constraint boundaries for the MWF-IC

3.2. Binaural MWF with psychoacoustically optimized Interaural Coherence preservation (MWF-IC)

Aiming at preserving the Interaural Coherence of diffuse noise fields, a coherence preservation term has been defined in [5] as

$$J_{IC}(\mathbf{W}) = \left| \frac{\mathbf{W}_0^H \mathbf{R}_v \mathbf{W}_1}{\sqrt{\mathbf{W}_0^H \mathbf{R}_v \mathbf{W}_0 \mathbf{W}_1^H \mathbf{R}_v \mathbf{W}_1}} - IC_v^{des} \right|^2, \quad (19)$$

where IC_v^{des} represents the desired output IC. When adding this term to the MWF cost function, i.e.,

$$J_{MWF-IC}(\mathbf{W}) = J_{MWF}(\mathbf{W}) + \lambda J_{IC}(\mathbf{W}), \quad (20)$$

a trade-off between noise reduction and IC preservation arises which can be controlled by the trade-off parameter λ . Since no closed-form expression is available for the filter $\mathbf{W}(\lambda)$ minimizing the cost function J_{MWF-IC} , an iterative numerical optimization method has been used. To control the amount of MSC preservation in the MWF-IC it has been proposed in [6] to limit the possible solutions of the optimization problem in (20) by imposing a constraint on the MSC of the output noise component, i.e.,

$$\gamma_{min}^{msc} \leq MSC_v^{out}(\mathbf{W}(\lambda)) \leq \gamma_{max}^{msc} \quad (21)$$

where γ_{min}^{msc} and γ_{max}^{msc} are lower and upper bounds for the MSC of the output noise component and $MSC_v^{out} = |IC_v^{out}|^2$. This inequality constraint limits the range of the trade-off parameter λ such that the output MSC lies within the boundaries γ_{min}^{msc} and γ_{max}^{msc} . The constraint boundaries γ_{min}^{msc} and γ_{max}^{msc} have been defined based on subjective listening experiments evaluating the IC discrimination abilities of the human auditory system in a diffuse noise field [11, 12] and are depicted in Fig. 2. Based on the subjective listening tests in [11] and [12] it is assumed that if the output MSC lies within the gray area in Fig. 2 the spatial impression of the output noise component is perceptually not discriminable from the spatial impression of a diffuse noise field. The trade-off parameter λ is then determined in an exhaustive search such that the inequality constraint in (21) is satisfied.

4. IC PRESERVATION USING THE BINAURAL MWF-N

To define an alternative solution to the rather computational complex MWF-IC for the preservation of the IC of the noise component, in this section we derive a closed form expression for the trade-off parameter η in the MWF-N that yields a predefined IC at the output of the filter to obtain the highest possible noise reduction performance while still preserving the spatial impression. For the MWF-N it has

been shown in [2] that the binaural cues of the speech component are preserved independent of the trade-off parameter η , i.e.,

$$RTF_x^{out} = \frac{(1-\eta) \frac{\rho}{(\mu+\rho)} A_0 + \eta A_0}{(1-\eta) \frac{\rho}{(\mu+\rho)} A_1 + \eta A_1} = \frac{A_0}{A_1}. \quad (22)$$

Plugging (10) and (11) in (8), the noise component output IC of the MWF-N is equal to

$$IC_v^{out} = \frac{\psi \mathbf{e}_0^T \mathbf{R}_x \mathbf{e}_1 + \eta^2 P_v IC_v^{in}}{\sqrt{(\psi \mathbf{e}_0^T \mathbf{R}_x \mathbf{e}_0 + \eta^2 P_v) (\psi \mathbf{e}_1^T \mathbf{R}_x \mathbf{e}_1 + \eta^2 P_v)}}, \quad (23)$$

with

$$\psi = (1-\eta)^2 \frac{\rho}{(\mu+\rho)^2} + 2\eta(1-\eta) \frac{1}{(\mu+\rho)}. \quad (24)$$

Equation (23) implies that for the MWF (i.e. setting $\eta = 0$), the output noise component IC is equal to the input IC of the speech component. In order to achieve a desired IC at the output of the MWF-N, the parameter η should be computed such that $IC_v^{out} = IC_v^{des}$. Similarly as for the MWF-IC in [6], IC_v^{des} can then be defined based on the IC discrimination ability of the auditory system. Since the trade-off parameter η is defined to be real-valued, the subset of IC_v^{des} that can be imposed is limited to the values that can be obtained using a real-valued η . Nevertheless, assuming the common hearing aid scenario of a speech source in front of the listener and assuming symmetry of the head, i.e. $A_0 = A_1$, the expression in (23) simplifies to

$$IC_v^{des} = \frac{\psi P_x + \eta^2 P_v IC_v^{in}}{\psi P_x + \eta^2 P_v}, \quad (25)$$

with

$$P_x = P_s |A_0|^2 = P_s |A_1|^2, \quad (26)$$

such that $SNR^{in} = SNR_0^{in} = SNR_1^{in}$ and $\mathbf{H}_0 = \mathbf{H}_1$. Now for any real-valued IC_v^{des} a real-valued trade-off parameter η can be analytically determined. Rearranging terms in (25) gives

$$\eta^2 - \eta \frac{2\mu}{2\mu + \rho + a} - \frac{\rho}{2\mu + \rho + a} = 0, \quad (27)$$

with

$$a = \frac{(IC_v^{in} - IC_v^{des}) (\mu + \rho)^2}{(IC_v^{des} - 1) SNR^{in}}. \quad (28)$$

The closed form expression for η is then equal to

$$\eta = \frac{\mu + \sqrt{(\mu + \rho)^2 + a\rho}}{2\mu + \rho + a}. \quad (29)$$

Besides the components of the MVDR beamformer, for the calculation of η an estimate of the input SNR is required due to the spectral Wiener postfilter in the MWF-N. We propose to estimate the input SNR based on the overall gain of the MVDR beamformer which is defined as

$$g = \frac{\overline{\mathbf{W}_0^H \mathbf{R}_y \mathbf{W}_0}}{\mathbf{e}_0^H \mathbf{R}_y \mathbf{e}_0} = \frac{P_x + P_v (\mathbf{H}_0^H \mathbf{\Gamma}^{-1} \mathbf{H}_0)^{-1}}{P_x + P_v}. \quad (30)$$

Rearranging the terms in (30), the input SNR is then equal to

$$SNR^{in} = \frac{(\mathbf{H}_0^H \mathbf{\Gamma}^{-1} \mathbf{H}_0)^{-1} - g}{g - 1}. \quad (31)$$

Based on the boundaries in Fig. 2, the desired IC for the MWF-N is calculated as

$$IC_v^{des} = \sqrt{\gamma_{max}^{msc}}. \quad (32)$$

5. EXPERIMENTAL RESULTS

In this section we present simulation results for a cafeteria scenario to compare the performance of the MWF, MWF-IC and the MWF-N with respect to objective evaluation measures.

5.1. Setup

Binaural Behind-The-Ear Impulse Responses (BTE-IR) measured in a cafeteria from [13] have been used to generate the speech component in the signals. Each hearing aid was equipped with 2 microphones, therefore in total 4 microphone signals are available. The speaker was located in front of the listener at a distance of 1m. A time-varying diffuse babble noise signal was generated using the method described in [14] and added to the speech component at an intelligibility weighted input SNR of 0 dB. The signals had a length of 10 s and were processed at $f_s = 16$ kHz using an overlap-add framework with a block size of $N = 512$ samples and an overlap of 75% between successive blocks. Cepstral smoothing [15, 16] has been applied to the SNR estimate and the minimum gain of the postfilter in (14) and (15) has been set to -15 dB.

5.2. Objective measures

For the evaluation of the noise reduction capabilities of the proposed filters, the intelligibility weighted output SNR (iSNR) [17] has been used. The speech quality was evaluated using PESQ [18], where the input speech component has been compared to the processed speech component. For a combined objective comparison of speech distortion and noise reduction, the frequency-weighted segmental SNR (FwSegSnr) [19] has been used. To evaluate the IC preservation capabilities of the algorithms, the broadband MSC error has been calculated by averaging the frequency-dependent MSC errors. The ILD and ITD errors (Δ ILD and Δ ITD) of the speech component have been evaluated using a model of binaural auditory processing described in [20]. The objective measures iSNR, FwSegSnr and PESQ for the left and the right hearing aid have been averaged to avoid a separated analysis for the left and the right hearing aid.

5.3. Experiment 1

In the first experiment we compare the MWF, the MWF-IC and the MWF-N using batch processing and disregarding estimation errors. The correlation matrices \mathbf{R}_x and \mathbf{R}_y have been calculated offline from the underlying components. Hence, the time-varying spectral properties of the speech and noise PSDs are not taken into account avoiding single-channel noise reduction artifacts but also limiting the noise reduction capabilities of the Wiener postfilter.

The results are depicted in Table 1. The MWF achieves an average output iSNR of 5.1 dB and the output iSNR decreases to 4.5 for the MWF-IC and to 3.6 for the MWF-N. On the contrary as expected, the MWF shows a very large MSC error compared to the MWF-IC and the MWF-N. Hence, using the MWF-IC and the MWF-N a similar MSC error can be obtained whereas the output iSNR of the MWF-N is 0.9 dB worse compared to the MWF-IC. However, because of less speech distortion introduced by the MWF-N (reflected in the higher PESQ score) the FwSegSnr is 0.1 dB better than for the MWF-IC. The PESQ score for the MWF and the MWF-IC is 3.7 and increases to 3.9 for the MWF-N. This can be explained by the mixing of the MWF output signal with the input signal, partially compensating speech distortion introduced by the MWF. The ILD and ITD errors are very low for all algorithms.

5.4. Experiment 2

In the second experiment we compare the MWF and the MWF-N using a-priori assumptions about the acoustic scenario and short term estimates of the input SNR utilized in the Wiener postfilter. The spatial coherence matrix Γ is calculated based on a modified sinc-function [9] given the relative distance between the microphones. For the estimation of the RTF vectors \mathbf{H}_0 and \mathbf{H}_1 , the anechoic BTE-IR's of the same database [13] have been used and the input SNR is then estimated according to (31) where the input correlation matrix \mathbf{R}_y is estimated using recursive averaging with a time constant of 40 ms.

The results are depicted in Table 2. The output iSNR of the MVDR beamformer is 4.9 dB and can be further increased to 7.1 dB using the spectral Wiener postfilter while the PESQ score decreases from 3.6 to 3.3. The output iSNR of the MWF-N is decreased by 0.7 dB and the output FwSegSnr is increased by 0.5 dB compared to the MWF while the MSC error is significantly reduced. For the MWF-N, the PESQ score is the same as for the MVDR despite the applied spectral postfilter. Due to the mixing, in addition to partially compensating speech distortion also dereverberation effects of the MVDR beamformer are partially compensated, leading to an increased PESQ score. Again, the ILD and ITD errors are very low for both algorithms.

	MWF	MWF-IC	MWF-N
iSNR[dB]	5.1	4.5	3.6
FwSegSnr[dB]	12	11.6	11.7
MSC_v^{err}	0.97	0.37	0.36
PESQ score	3.7	3.7	3.9
Δ ILD[dB]	0.03	0.29	0.12
Δ ITD[ms]	0.005	0.003	0.001

Table 1. Objective evaluation results for Experiment 1. The PESQ score ranges from -0.5 to 4.5.

	MVDR	MWF	MWF-N
iSNR[dB]	4.9	7.1	6.4
FwSegSnr[dB]	11.6	12.2	12.7
MSC_v^{err}	0.97	0.97	0.38
PESQ score	3.6	3.3	3.6
Δ ILD[dB]	0.15	0.03	0.04
Δ ITD[ms]	0.004	0.005	0.005

Table 2. Objective evaluation results for Experiment 2. The PESQ score ranges from -0.5 to 4.5.

6. CONCLUSION

In this paper we have shown that both the MWF-IC and the MWF-N preserve the IC of the output noise component in diffuse noise fields whereas the MWF-IC shows a better noise reduction performance and the MWF-N achieves a higher PESQ score. Contrary to the MWF-IC, the MWF-N filter coefficients and the trade-off parameter η , that trades off between noise reduction and IC preservation, can be calculated using a closed form expression, making the MWF-N more appealing for real-time implementations. The determination of the (complex-valued) trade-off parameter η in the MWF-N for a more general scenario remains a topic for further research.

7. REFERENCES

- [1] S. Doclo, S. Gannot, M. Moonen, and A. Spriet, "Acoustic beamforming for hearing aid applications," in *Handbook on Array Processing and Sensor Networks*, pp. 269–302. Wiley, 2010.
- [2] B. Cornelis, S. Doclo, T. Van den Bogaert, J. Wouters, and M. Moonen, "Theoretical analysis of binaural multi-microphone noise reduction techniques," *IEEE Trans. on Audio, Speech and Language Proc.*, vol. 18, no. 2, pp. 342–355, Feb. 2010.
- [3] D. Marquardt, V. Hohmann, and S. Doclo, "Binaural cue preservation for hearing aids using multi-channel Wiener filter with instantaneous ITF preservation," in *Proc. ICASSP*, Kyoto, Japan, Mar. 2012, pp. 21–24.
- [4] E. Hadad, S. Gannot, and S. Doclo, "Binaural linearly constrained minimum variance beamformer for hearing aid applications," in *Proc. International Workshop on Acoustic Signal Enhancement (IWAENC)*, Aachen, Germany, Sep. 2012, pp. 117–120.
- [5] D. Marquardt, V. Hohmann, and S. Doclo, "Coherence preservation in multi-channel Wiener filtering based noise reduction for binaural hearing aids," in *Proc. ICASSP*, Vancouver, Canada, May 2013, pp. 8648–8652.
- [6] D. Marquardt, V. Hohmann, and S. Doclo, "Perceptually motivated coherence preservation in multi-channel wiener filtering based noise reduction for binaural hearing aids," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, Florence, Italy, May 2014, pp. 3660–3664.
- [7] T.J. Klasen, T.V. den Bogaert, M. Moonen, and J. Wouters, "Binaural noise reduction algorithms for hearing aids that preserve interaural time delay cues," *Signal Processing, IEEE Transactions on*, vol. 55, no. 4, pp. 1579–1585, April 2007.
- [8] M. Jeub, M. Dorbecker, and P. Vary, "A semi-analytical model for the binaural coherence of noise fields," *Signal Processing Letters, IEEE*, vol. 18, no. 3, pp. 197–200, Mar. 2011.
- [9] I. Lindevald and A. Benade, "Two-ear correlation in the statistical sound fields of rooms," *Journal of the Acoustical Society of America*, vol. 80, no. 2, pp. 661–664, Aug. 1986.
- [10] K. U. Simmer, J. Bitzer, and C. Marro, "Post-filtering techniques," in *Microphone Arrays*, M. Brandstein and D. Ward, Eds., pp. 39–60. Springer Berlin Heidelberg, 2001.
- [11] A. Walther and C Faller, "Interaural correlation discrimination from diffuse field reference correlations," *Journal of the Acoustical Society of America*, vol. 133, no. 3, pp. 1496–1502, Mar. 2013.
- [12] M. Tohyama and A. Suzuki, "Interaural cross-correlation coefficients in stereo-reproduced sound fields," *Journal of the Acoustical Society of America*, vol. 85, no. 2, pp. 780–786, Feb. 1989.
- [13] H. Kayser, S. Ewert, J. Annemüller, T. Rohdenburg, V. Hohmann, and B. Kollmeier, "Database of multichannel In-Ear and Behind-The-Ear Head-Related and Binaural Room Impulse Responses," *Eurasip Journal on Advances in Signal Processing*, vol. 2009, pp. 10 pages, 2009.
- [14] E. A. P. Habets, I. Cohen, and S. Gannot, "Generating non-stationary multisensor signals under a spatial coherence constraint," *Journal of the Acoustical Society of America*, vol. 124, no. 5, pp. 2911–2917, Nov. 2008.
- [15] C. Breithaupt, T. Gerkmann, and R. Martin, "A novel a priori SNR estimation approach based on selective cepstro-temporal smoothing," in *Proc. ICASSP*, Las Vegas, NV, USA, Mar. 2008, pp. 4897–4900.
- [16] T. Gerkmann and R. Martin, "On the statistics of spectral amplitudes after variance reduction by temporal cepstrum smoothing and cepstral nulling," *Signal Processing, IEEE Transactions on*, vol. 57, no. 11, pp. 4165–4174, Nov 2009.
- [17] J. E. Greenberg, P. M. Peterson, and P. M. Zurek, "Intelligibility-weighted measures of speech-to-interference ratio and speech system performance," *Journal of the Acoustical Society of America*, vol. 94, no. 5, pp. 3009–3010, Nov. 1993.
- [18] 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 P.862*, International Telecommunications Union (ITU-T) Recommendation, Feb. 2001.
- [19] Yi Hu and P.C. Loizou, "Evaluation of objective quality measures for speech enhancement," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 16, no. 1, pp. 229–238, Jan 2008.
- [20] M. Dietz, S. D. Ewert, and V. Hohmann, "Auditory model based direction estimation of concurrent speakers from binaural signals," *Speech Communication*, vol. 53, pp. 592–605, 2011.