COHERENCE PRESERVATION IN MULTI-CHANNEL WIENER FILTERING BASED NOISE REDUCTION FOR BINAURAL HEARING AIDS

Daniel Marquardt, Volker Hohmann, Simon Doclo

University of Oldenburg, Department of Medical Physics and Acoustics, 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, i.e. the Interaural Level Difference and the Interaural Time Difference of all sound sources. Recently, extensions of the binaural Multi-channel Wiener filter (MWF) have been presented, which aim to preserve the binaural cues of the residual noise component. However, since these algorithms aim to preserve the Interaural Transfer Function (ITF), they are well-suited only for directional noise sources but not for, e.g. spatially isotropic noise, which can not be fully described by the ITF. In this paper, we present an extension of the binaural MWF, aiming to preserve the Interaural Coherence of the residual noise component. Experimental results using spatially isotropic noise show that the proposed algorithm yields a good preservation of the Interaural Coherence without significantly degrading the output SNR compared to the binaural MWF and the binaural MWF with ITF preservation.

***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 confusions 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.

To achieve binaural cue preservation, two main concepts for binaural noise reduction have been developed. In the first concept, the multi-channel signals are used to calculate a real-valued gain, where the same gain is applied to the reference microphone in the left, respectively right hearing aid [2, 3, 4]. This processing strategy allows perfect preservation of the binaural cues of both the speech and the noise component, but typically suffers from limited noise reduction performance and single-channel noise reduction artifacts. The second concept is to apply a complex-valued filter to all available microphone signals on the left and the right hearing aid, combining spatial and spectral filtering. Using this processing strategy, a large noise reduction performance can be achieved, but the binaural cues of the residual noise component are not guaranteed to be preserved. Hence, algorithms have been proposed that aim to preserve the binaural cues of the residual noise component by adding a cue preservation term to the basic noise reduction cost function [5, 6, 7]. In [1] the binaural Speech Distortion Weighted Multi-channel Wiener Filter (MWF) has been presented. It has been theoretically proven in [5] that in case of a single speech source this technique preserves the Interaural Transfer Function (ITF), comprising the Interaural Level Difference (ILD) and the Interaural Time Difference (ITD) cues, of the speech component but typically distorts the ITF of the noise component. In addition, an extension of the MWF, namely the MWF-ITF, has been presented, by adding a term related to the preservation of the ITF of the noise component. It has been shown in [5] that a better preservation of the ITF of the noise component can be achieved, depending on the output SNR and a trade-off parameter. Hence, the MWF-ITF is well suited for directional noise sources since the spatial properties of directional noise sources are well described by the ILD and ITD cues.

In contrast to directional noise sources, the spatial characteristics of e.g. spatially isotropic noise however can not be properly described by the ITF, but rather by the Interaural Coherence (IC). Therefore, this paper proposes an extension of the MWF with a term related to the IC preservation of the noise component.

The binaural configuration and notation used throughout the paper are described in section 2. In section 3, we briefly review the MWF and the MWF-ITF and investigate the properties of these algorithms in preserving the IC of spatially isotropic noise. In section 4, we introduce a new cost function, extending the MWF with a term related to the IC of the noise component and in section 5 we provide simulation results in a reverberant environment.

This research was partly funded by the German-Israeli Foundation for Scientific Research and Development (GIF)

2. CONFIGURATION AND NOTATION

Consider the binaural hearing aid configuration in Figure 1, consisting of the left and the right microphone array with M microphones each.

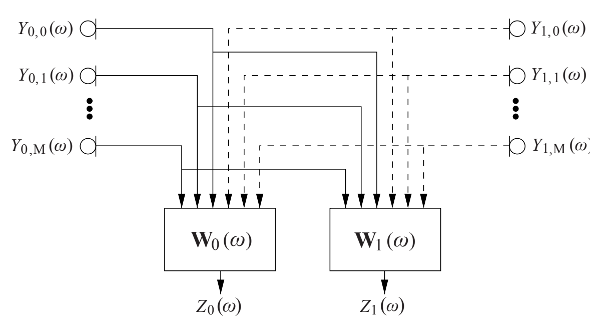

Fig. 1. Binaural hearing aid configuration

The frequency-domain representation of the m-th microphone signal in the left hearing aid $Y_{0,m}(\omega)$ can be written as

$$Y_{0,m}(\omega) = X_{0,m}(\omega) + V_{0,m}(\omega), \;\; m = 0 \ldots M-1,$$

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

$$\mathbf{Y} = \left[Y_{0,0} \ldots Y_{0,M-1} Y_{1,0} \ldots Y_{1,M-1}\right]^T. \tag{1}$$

The signal vector 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

$$\mathbf{W} = \begin{bmatrix}\mathbf{W}_0 \\ \mathbf{W}_1\end{bmatrix}. \tag{2}$$

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

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

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

$$\mathbf{R}_y = \mathcal{E}\left\{\mathbf{Y}\mathbf{Y}^H\right\}, \mathbf{R}_v = \mathcal{E}\left\{\mathbf{V}\mathbf{V}^H\right\}, \mathbf{R}_x = \mathcal{E}\left\{\mathbf{X}\mathbf{X}^H\right\}. \tag{4}$$

The input ITF of the noise component is defined as

$$ITF_v^{in} = \frac{\mathcal{E}\left\{V_0V_1^*\right\}}{\mathcal{E}\left\{V_1V_1^*\right\}} = \frac{\mathbf{e}_0^T\mathbf{R}_v\mathbf{e}_1}{\mathbf{e}_1^T\mathbf{R}_v\mathbf{e}_1}. \tag{5}$$

The vectors $\mathbf{e}_0$ and $\mathbf{e}_1$ are zero column vectors with $\mathbf{e}_0(1) = 1$ and $\mathbf{e}_1(M+1) = 1$ such that $V_0 = \mathbf{e}_0^T\mathbf{V}$ and $V_1 = \mathbf{e}_1^T\mathbf{V}$ are the noise components in the reference microphones.
The input Interaural Coherence of the noise component is defined as

$$IC_v^{in} = \frac{\mathcal{E}\left\{V_0V_1^*\right\}}{\sqrt{\mathcal{E}\left\{V_0V_0^*\right\}\mathcal{E}\left\{V_1V_1^*\right\}}} = \frac{\mathbf{e}_0^T\mathbf{R}_v\mathbf{e}_1}{\sqrt{\mathbf{e}_0^T\mathbf{R}_v\mathbf{e}_0\,\mathbf{e}_1^T\mathbf{R}_v\mathbf{e}_1}}. \tag{6}$$

The real-valued Magnitude Squared Coherence (MSC) is defined as $MSC = |IC|^2$. In case of a directional noise source, the input ITF is equal to the Relative Transfer Function [8] and the input IC is equal to the normalized ITF, i.e.

$$ITF_v^{dir} = \frac{A_0}{A_1}, \quad IC_v^{dir} = \frac{ITF_v^{dir}}{|ITF_v^{dir}|} = e^{j\angle\frac{A_0}{A_1}}, \tag{7}$$

where A_0 and A_1 are the acoustic transfer functions from the speech source to the reference microphone in the left, respectively right hearing aid and $\angle$ defines the phase. The output ITF and output IC of the noise component are defined as

$$ITF_v^{out} = \frac{\mathbf{W}_0^H\mathbf{R}_v\mathbf{W}_1}{\mathbf{W}_1^H\mathbf{R}_v\mathbf{W}_1}, \quad IC_v^{out} = \frac{\mathbf{W}_0^H\mathbf{R}_v\mathbf{W}_1}{\sqrt{\mathbf{W}_0^H\mathbf{R}_v\mathbf{W}_0\mathbf{W}_1^H\mathbf{R}_v\mathbf{W}_1}}.$$

The ITF and 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 briefly review the cost functions for the binaural MWF [1] and the MWF-ITF [5] and investigate the properties of these two algorithms in preserving the IC of spatially isotropic noise.

3.1. Binaural multi-channel Wiener filter (MWF)

The binaural MWF produces a minimum mean-square error (MMSE) estimate of the speech component in the reference microphone signal for both hearing aids. The MWF cost function estimating the speech components $X_{0,0}$ and $X_{1,0}$ in the left and the right hearing aid can be written as

$$J_{MWF}(\mathbf{W}) = \mathcal{E}\left\{\left\|\begin{bmatrix}X_{0,0} - \mathbf{W}_0^H\mathbf{X} \\ X_{1,0} - \mathbf{W}_1^H\mathbf{X}\end{bmatrix}\right\|^2 + \mu\left\|\begin{bmatrix}\mathbf{W}_0^H\mathbf{V} \\ \mathbf{W}_1^H\mathbf{V}\end{bmatrix}\right\|^2\right\}, \tag{8}$$

where μ provides a trade-off between noise reduction and speech distortion and the first microphone has been used as a reference microphone. The filter minimizing $J_{MWF}(\mathbf{W})$ is equal to

$$\mathbf{W}_{MWF} = \mathbf{R}^{-1}\mathbf{r}_x, \tag{9}$$

with

$$\mathbf{R} = \begin{bmatrix}\mathbf{R}_x + \mu\mathbf{R}_v & \mathbf{0}_{2M} \\ \mathbf{0}_{2M} & \mathbf{R}_x + \mu\mathbf{R}_v\end{bmatrix}, \quad \mathbf{r}_x = \begin{bmatrix}\mathbf{R}_x\mathbf{e}_0 \\ \mathbf{R}_x\mathbf{e}_1\end{bmatrix}. \tag{10}$$

Based on the theoretical analysis in [5] we can show that in case of a single speech source the output IC of the speech and the noise component is equal to the input IC of the speech component, i.e.

$$IC_x^{out} = IC_v^{out} = IC_x^{in} = e^{j\angle ITF_x^{in}}. \tag{11}$$

Equation (11) also implies that in the case of a spatially isotropic noise field the residual noise component would be perceived as a point source coming from the speech direction, what is obviously undesired.

3.2. MWF with ITF preservation (MWF-ITF)

To allow for the preservation of the noise ITF, an extension of the MWF cost function with a term related to the ITF of the noise component has been proposed and analyzed in [5] as

$$J_{ITF}(\mathbf{W}) = \mathcal{E}\left\{\left|\mathbf{W}_0^H\mathbf{V} - ITF_v^{in}\mathbf{W}_1^H\mathbf{V}\right|^2\right\}. \tag{12}$$

This term is added to the binaural MWF cost function, i.e.

$$J_{MWF-ITF}(\mathbf{W}) = J_{MWF}(\mathbf{W}) + \delta J_{ITF}(\mathbf{W}). \tag{13}$$

The filter minimizing $J_{MWF-ITF}(\mathbf{W})$ is equal to

$$\mathbf{W}_{MWF-ITF} = \left(\mathbf{R} + \delta\mathbf{R}_{vt}\right)^{-1}\mathbf{r}_x, \tag{14}$$

with

$$\mathbf{R}_{vt} = \begin{bmatrix}\mathbf{R}_v & -ITF_v^{in,*}\mathbf{R}_v \\ -ITF_v^{in}\mathbf{R}_v & \left|ITF_v^{in}\right|^2\mathbf{R}_v\end{bmatrix}. \tag{15}$$

The parameter δ controls the emphasis on the noise ITF preservation term. Based on the theoretical analysis in [5] we can show that for a single speech source the output IC of the speech and the noise component are still the same and are equal to

$$IC_x^{out} = IC_v^{out} = e^{j\angle ITF^{out}}, \tag{16}$$

where ITF^{out} can be calculated analytically depending on the parameter δ and the output SNR. If $\delta = 0$ then $ITF^{out} = ITF_x^{in}$ and if $\delta \rightarrow \infty$ then $ITF^{out} = ITF_v^{in}$ such that the solution is always a trade-off between preserving the ITF of the speech component and preserving the ITF of the noise component. Using the MWF-ITF, the IC of a directional noise source can be preserved on the cost of distorting the IC of the speech component. The IC of the output components in (16) is complex-valued and the MSC is still equal to 1 for all frequencies and independent of the parameter δ. Hence, the real-valued IC of a spatially isotropic noise field, which is dependent on the frequency and the microphone distance [9, 10, 11], can not be preserved using the MWF-ITF.

4. MWF WITH INTERAURAL COHERENCE PRESERVATION (MWF-IC)

To allow for the preservation of the Interaural Coherence of spatially isotropic noise fields, we define the following coherence preservation term

$$J_{IC}(\mathbf{W}) = \left| \frac{\mathbf{W}_0^H \mathbf{R}_v \mathbf{W}_1}{\sqrt{\mathbf{W}_0^H \mathbf{R}_v \mathbf{W}_0 \mathbf{W}_1^H \mathbf{R}_v \mathbf{W}_1}} - IC_v^{des} \right|^2, \quad (17)$$

where IC_v^{des} represents the desired output IC, which can be equal to the estimated IC_v^{in} as in (6) or can be defined using HRTF measurements or head models [10, 11]. Similarly as for the MWF-ITF, we add this additional term to the MWF cost function i.e,

$$J_{MWF-IC}(\mathbf{W}) = J_{MWF}(\mathbf{W}) + \lambda\, J_{IC}(\mathbf{W}), \quad (18)$$

where the parameter λ allows for a trade-off between noise reduction and coherence preservation. The resulting filter will be denoted as the MWF-IC. Since no closed-form expression is available for the filter minimizing J_{MWF-IC}, we will resort to iterative optimization techniques. We have used a *trust-region method* where analytical expressions for the gradient and the Hessian of the cost function $J_{MWF-IC}(\mathbf{W})$ have been provided in order to improve the numerical robustness and the convergence speed. The analytical expression of the gradient and the Hessian are omitted due to space constraints.

5. EXPERIMENTAL RESULTS

In this section we perform simulations for a cafeteria scenario with different desired speaker locations to investigate the performance of the MWF, MWF-ITF and MWF-IC with respect to the intelligibility weighted output SNR and the IC preservation for the speech and the noise component.

5.1. Setup

Binaural Behind-The-Ear Impulse Responses (BTE-IR) measured in a cafeteria from [12] have been used to generate the speech signals. Each hearing aid was equipped with 2 microphones, therefore in total 4 microphone signals are available. The speech source was located on different positions in the cafeteria (cf. Table 1). The noise components, corresponding to a spatially isotropic noise field were generated using the method described in [13], where the power spectral density (PSD) of the noise components was chosen to be the PSD of speech-shaped noise and the coherence matrix of the binaural setup in a cylindrical isotropic noise field was estimated using the anechoic BTE-IR from [12]. The signals were processed at $f_s = 16$ kHz using an weighted overlap-add framework with a block size of $N = 256$ samples and an overlap of 75% between successive blocks. The speech + noise signal had a length of 10 s and was preceded by a noise-only signal of 10 s length. The noise-only part was not taken into account during evaluation. The correlation matrices of the signal components are estimated as

$$\mathbf{R}_y(k) = \frac{1}{L_y} \sum_{i=0}^{L_y-1} \mathbf{Y}(k,i)\mathbf{Y}^H(k,i) \quad \text{speech present}, \quad (19)$$

$$\mathbf{R}_v(k) = \frac{1}{L_v} \sum_{i=0}^{L_v-1} \mathbf{V}(k,i)\mathbf{V}^H(k,i) \quad \text{speech absent}, \quad (20)$$

with $k = 0, 1, \ldots, N-1$ denoting the frequency index, i denoting the block index, L_y denoting the available signal vectors when speech is present and L_v denoting the available signal vectors when speech is absent, using a perfect Voice Activity Detector (VAD). The correlation matrix of the speech component is estimated as

$$\mathbf{R}_x(k) = \mathbf{R}_y(k) - \mathbf{R}_v(k). \quad (21)$$

Position	**A**	**B**	**C**	**D**	**E**
Azimuth [°]	0	45	90	270	225
Distance [cm]	102	118	52	162	129

Table 1. Speech source positions and distances relative to the listener. 0° defines the position in front of the listener. The azimuth angle is defined counter-clockwise.

The desired IC in the MWF-IC was calculated from the estimate of the noise correlation matrix $\mathbf{R}_v$ as in (6). The performance was evaluated for an average intelligibility weighted input SNR in the reference microphones of 0 dB. The parameter μ was set to 1. The trade-off parameter δ in the MWF-ITF was set to 3 [**MWF-ITF (3)**] and to 10^4 [**MWF-ITF (pow4)**]. The frequency-dependent trade-off parameter λ in the MWF-IC was chosen such that the MSC error of the noise component in each frequency bin was kept below 0.1 [**MWF-IC (0.1)**], respectively below 0.01 [**MWF-IC (0.01)**].

5.2. Performance measures

For comparing the performance of the algorithms we have used 3 objective performance measures. The *intelligibility weighted SNR* [14] is defined as

$$iSNR = \sum_k I(k) 10 \log_{10} \left(\frac{P_x(k)}{P_v(k)} \right), \quad (22)$$

where $P_x(k)$ and $P_v(k)$ are the PSDs of the speech component, respectively noise component of the input signal (input $iSNR$), respectively the output signal (output $iSNR$). $I(k)$ is a weighting function that takes the importance of different frequency bands for the speech intelligibility into account.
To avoid a separated analysis for the real and imaginary part of the complex-valued IC we evaluate the performance using the real-valued MSC. The MSC of the input noise component was calculated during the 10 s speech + noise period, i.e.

$$MSC_v^{in}(k) = \left| \frac{\sum_{i=0}^{L_y-1} V_0(k,i) V_1^*(k,i)}{\sqrt{\sum_{i=0}^{L_y-1} |V_0(k,i)|^2 \sum_{i=0}^{L_y-1} |V_1(k,i)|^2}} \right|^2 \quad (23)$$

The MSC of the output noise component (MSC_V^{out}) was calculated by replacing V with Z_v in (23). The broadband MSC error is calculated by averaging the frequency-dependent MSC errors, i.e.

$$MSC_v^{err} = \frac{1}{N-1} \sum_{k=1}^{N-1} \left| MSC_v^{in}(k) - MSC_v^{out}(k) \right|. \quad (24)$$

For the directional speech component the MSC error is however not an appropriate objective measure. The MSC contains information about the amount of correlation of a signal in the microphones but does not contain information about the perceived direction of a directional source. Hence, for the evaluation of the binaural cue preservation of the speech component we use an objective measure which is based on a model of binaural auditory processing [15] and has been already applied for binaural cue preservation evaluation in [16].

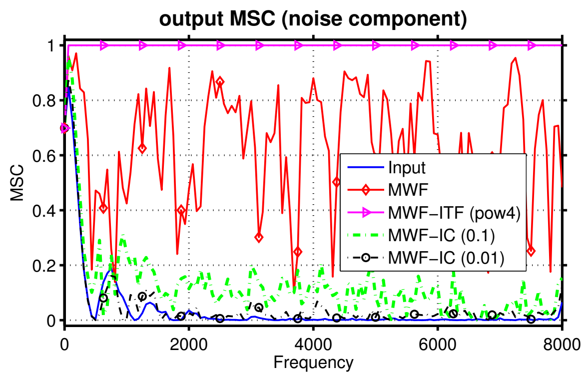

Fig. 2. Output noise component MSC for the MWF, MWF-ITF and MWF-IC algorithms for the speech source position A

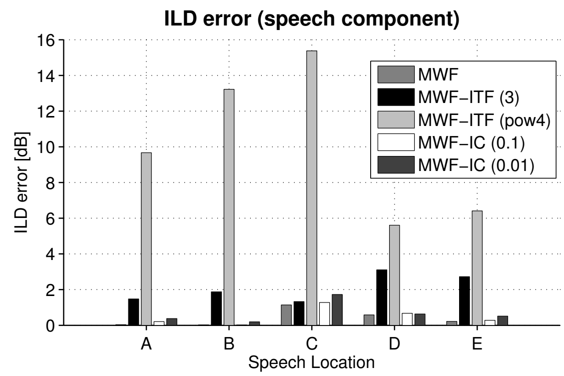

Fig. 4. Speech component ILD error for the MWF, MWF-ITF and MWF-IC algorithms

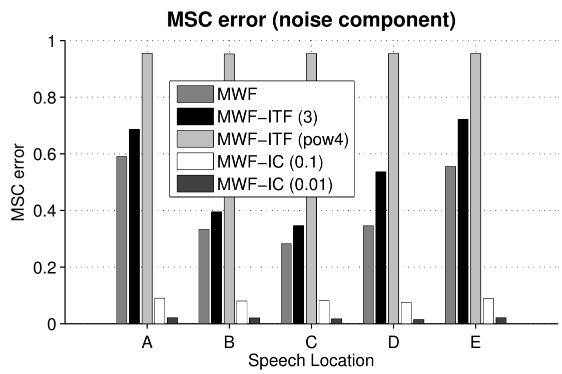

Fig. 3. Noise component broadband MSC error for the MWF, MWF-ITF and MWF-IC algorithms

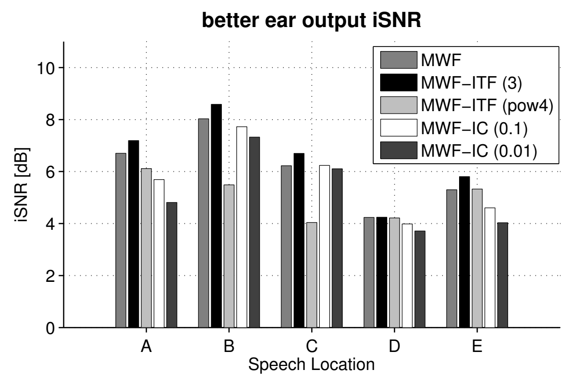

Fig. 5. Better ear iSNR for the MWF, MWF-ITF and MWF-IC algorithms

5.3. Performance Results

The output noise component MSC is depicted in Fig. 2. From the theoretical analysis (cf. section 3.1 and 3.2), the output noise component MSC is expected to be 1 for all frequencies, what is the case for the MWF-ITF (pow4) but not for the MWF due to estimation errors in the speech correlation matrix $\mathbf{R}_x$. The input noise component MSC is well preserved using the MWF-IC and a better preservation can be achieved by increasing the trade-off parameter λ. To avoid overcrowded plots, the output noise component MSC for the MWF-ITF (3) is omitted since it is comparable to the result for the MWF.

The noise component broadband MSC error is depicted in Fig. 3. From the theoretical analysis (cf. section 3.1 and 3.2), the noise component MSC error is expected to be the same for the MWF and the MWF-ITF, independent of the trade-off parameter δ, what is again not exactly the case due to estimation errors. For the MWF-ITF, the MSC error of the noise component increases with increasing trade-off parameter δ. Compared to the MWF and the MWF-ITF, the broadband MSC error of the noise component is significantly decreased using the MWF-IC, especially for speaker locations A and E.

The speech component ILD error is depicted in Fig. 4. The ILD error is slightly increased using the MWF-ITF (3) and dramatically increased using the MWF-ITF (pow 4) compared to the MWF. The MWF-IC does not increase the ILD error of the speech component compared to the MWF. The results for the ITD error of the speech component are similar but omitted due to space constraints. Hence, a preservation of the noise component IC is possible using the MWF-IC without distorting the speech component cues.

Since the iSNR of the better ear mainly affects speech intelligibility, the iSNR of the better ear for all algorithms is depicted in Fig. 5. The MWF-ITF (3) shows the best performance for all speaker positions compared to the other considered algorithms. The output iSNR is slightly decreased for the MWF-IC, since a better preservation of the IC leads to less noise reduction due to the higher impact of the coherence preservation term in (18). Hence, the MWF-IC allows for a controllable trade-off between better IC preservation and output SNR.

6. RELATION TO PRIOR WORK

Contrary to prior work in [5, 6, 7] we define an additional cost term for the binaural MWF that is not related to the ITF, ILD or ITD of the noise component but to the IC of the noise component. Due to the IC preservation the MWF-IC is also well-suited for spatially isotropic noise fields.

7. CONCLUSION

In this paper we have shown that for spatially isotropic noise fields the MWF-IC yields a better preservation of the IC of the residual noise component compared to the MWF and the MWF-ITF without distorting the speech component cues. The influence of the trade-off between better IC preservation and output SNR on spatial awareness and speech intelligibility needs to be examined in future work.

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

[3] T. Rohdenburg, V. Hohmann, and B. Kollmeier, "Robustness analysis of binaural hearing aid beamformer algorithms by means of objective perceptual quality measures," in *Proc. WASPAA*, New Paltz, NY, Oct. 2007, pp. 315–318.

[4] 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, pp. 1141–1154, Apr. 2011.

[5] 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.

[6] S. Doclo, R. Dong, T. J. Klasen, J. Wouters, S. Haykin, and M. Moonen, "Extension of the multi-channel Wiener filter with localisation cues for noise reduction in binaural hearing aids," in *Proc. IWAENC*, Eindhoven, The Netherlands, Sep. 2005.

[7] 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.

[8] S. Gannot, D. Burshtein, and E. Weinstein, "Signal enhancement using beamforming and non-stationarity with applications to speech," *IEEE Trans. Signal Processing*, vol. 49, no. 8, pp. 1614–1626, Aug. 2001.

[9] H. Cox, "Spatial correlation in arbitrary noise fields with application to ambient sea noise," *Journal of the Acoustical Society of America*, vol. 54, no. 5, pp. 1289–1301, 1973.

[10] 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, 1986.

[11] 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.

[12] H. Kayser, S. Ewert, J. Annemueller, 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.

[13] 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.

[14] 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.

[15] 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.

[16] D. Marquardt, L. Wang, V. Hohmann, and S. Doclo, "Improved noise estimation for the binaural MWF with instantaneous ITF preservation," in *Proc. European Signal Processing Conference (EUSIPCO)*, Bucharest, Romania, Aug. 2012, pp. 300–304.