Katholieke Universiteit Leuven

Departement Elektrotechniek ESAT-SISTA/TR 10-120

Binaural cue preservation in binaural hearing aids with reduced-bandwidth multichannel Wiener filter based noise reduction[1]

Bram Cornelis[2] , Marc Moonen[2], Jan Wouters[3]

Published in the proceedings of the
12th International Workshop on Acoustic Echo and Noise Control
(IWAENC) ,
Tel Aviv, Israel, Sep. 2010

[1]This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/sista/bcorneli/reports/IWAENC10.pdf*

[2]K.U.Leuven, Dept. of Electrical Engineering (ESAT), Kasteelpark Arenberg 10, 3001 Leuven, Belgium. Tel. +32 16 321797, Fax +32 16 321970, WWW: *http://www.esat.kuleuven.ac.be/sista*, E-mail: *bram.cornelis@esat.kuleuven.ac.be*. Bram Cornelis is funded by a Ph.D. grant of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen). This research work was carried out at the ESAT Laboratory of Katholieke Universiteit Leuven in the frame of the Belgian Programme on Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Office IUAP P6/04 (DYSCO, 'Dynamical systems, control and optimization', 2007-2011), Concerted Research Action GOA-MaNet and research project FWO nr. G.0600.08 ('Signal processing and network design for wireless acoustic sensor networks'). The scientific responsibility is assumed by its authors.

[3]K.U.Leuven, Dept. of Neurosciences, ExpORL, Herestraat 49/721, 3000 Leuven, Belgium.

Abstract

The Speech Distortion Weighted Multichannel Wiener Filter (SDW-MWF) is a promising noise reduction algorithm, especially for binaural hearing aid applications where microphone signals can be exchanged between the left and right hearing aid. In a binaural hearing aid application, it is also desirable to preserve the binaural cues of both the target speech signal and the residual noise, in particular the interaural level differences (ILD's) and interaural time differences (ITD's), which are used to localize sound sources. It was previously shown that the SDW-MWF, when all the microphone signals are exchanged, allows for a significant noise reduction as well as the preservation of the target speech binaural cues. The MWF cost function can be extended so that the residual noise binaural cues are also preserved. As in practice not all microphone signals can be exchanged, reduced-bandwidth MWF algorithms will be used. In this paper a framework is proposed that allows the study of the binaural cue preservation of these reduced-bandwidth MWF algorithms. It is shown that the speech ITD cues are always preserved, but that the speech ILD cues may be distorted. A reduced-bandwidth binaural MWF reduces the speech ILD errors compared to a bilateral MWF, especially so when a good estimate of the speech signal is transmitted.

Binaural cue preservation in binaural hearing aids with reduced-bandwidth multichannel Wiener filter based noise reduction

Bram Cornelis*, Marc Moonen* and Jan Wouters†
* Katholieke Universiteit Leuven, ESAT–SCD, Kasteelpark Arenberg 10, B-3001 Leuven, Belgium
† Katholieke Universiteit Leuven, Dept. of Neurosciences, ExpORL, Herestraat 49/721, B-3000 Leuven, Belgium

***Abstract*—The Speech Distortion Weighted Multichannel Wiener Filter (SDW-MWF) is a promising noise reduction algorithm, especially for binaural hearing aids where microphone signals can be exchanged between left and right hearing aid. It is also desirable to preserve the binaural cues of both the speech signal and the residual noise, in particular the interaural level differences (ILD's) and interaural time differences (ITD's), which are used to localize sound. It was previously shown that the SDW-MWF, when all microphone signals are exchanged, allows for a significant noise reduction as well as the preservation of the speech binaural cues. The MWF cost function can be extended so that the residual noise binaural cues are also preserved. As in practice not all microphone signals can be exchanged, reduced-bandwidth MWF algorithms will be used. In this paper a framework is proposed that allows the study of the binaural cue preservation of these reduced-bandwidth MWF algorithms. It is shown that the speech ITD cues are always preserved, but that the speech ILD cues may be distorted. A reduced-bandwidth binaural MWF reduces the speech ILD errors compared to a bilateral MWF, especially so when a good estimate of the speech signal is transmitted.**

I. Introduction

In future binaural hearing aids microphone signals will be exchanged between the left and right hearing aid to generate an output signal for each device. The microphone signals can be processed by the Speech Distortion Weighted Multichannel Wiener Filter (SDW-MWF) to achieve significant noise reduction in a speech + noise scenario [1], [2]. In addition to providing noise reduction, it is also desirable to preserve the so-called binaural cues of both the target speech signal and the residual noise [3], [4], in particular the interaural level differences (ILD's) and interaural time differences (ITD's), which are used to localize sound sources. It was previously shown that the SDW-MWF, when all the microphone signals are exchanged, allows for a significant noise reduction as well as the preservation of the target speech binaural cues [5]. The MWF cost function can be extended so that the residual noise binaural cues are also preserved. In the analysis in [5], it is however assumed that there are no bandwidth restrictions on the binaural link, so that all microphone signals can be exchanged between the left and right device.

Due to power and bandwidth limitations of the binaural link, the number of microphone signals that can be transmitted will be restricted. Several reduced-bandwidth MWF algorithms have thefore been presented in [2], where each device transmits a filtered combination of its microphone signals. A distributed MWF approach, which for certain scenarios converges to the optimal binaural solution, has also been proposed. In [6], [7], the problem was studied from an information-theoretic point of view. Using rate-distortion theory, a theoretically optimal (but not practically feasible) solution has been obtained in [6]. The reduced-bandwidth MWF approaches of [2] were also studied in this framework and compared to the theoretical upper bound [7]. Although these previous works indeed show the potential SNR improvement of the reduced-bandwidth MWF algorithms, the binaural cue preservation of these algorithms has not yet been studied.

In this paper, a general theoretical framework is proposed that allows the study of the binaural cue preservation of these reduced-bandwidth MWF algorithms, as well as of the bilateral case where no signals are exchanged. The binaural cue preservation of the target speech signal is then studied in this framework. It is shown that the bilateral MWF and the reduced-bandwidth MWF algorithms preserve the speech ITD cues, but distort the speech ILD cues. This is supported by simulations on a binaural hearing aid setup. The simulations also illustrate that in general, a reduced-bandwidth binaural MWF reduces the speech ILD errors compared to a bilateral MWF, especially so when a good estimate of the speech signal is transmitted.

The paper is organized as follows. In section II, the notation and the reduced-bandwidth framework are introduced. The SDW-MWF is briefly reviewed in section III. In section IV, the binaural cue preservation of the target speech signal is analyzed in the proposed framework. Simulations in section V then illustrate the noise reduction performance and cue preservation of reduced-bandwidth MWF algorithms. Finally, conclusions are drawn in section VI.

II. Configuration and notation

A. Microphone signals and output signals

We consider a binaural hearing aid configuration, where both hearing aids have a microphone array consisting of M

Bram Cornelis is funded by a Ph.D. grant of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen). This research work was carried out at the ESAT Laboratory of Katholieke Universiteit Leuven in the frame of the Belgian Programme on Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Office IUAP P6/04 (DYSCO, 'Dynamical systems, control and optimization', 2007-2011), Concerted Research Action GOA-MaNet and research project FWO nr. G.0600.08 ('Signal processing and network design for wireless acoustic sensor networks'). The scientific responsibility is assumed by its authors.

microphones. The mth microphone signal in the left hearing aid $Y_{0,m}(\omega)$ can be specified in the frequency-domain as[1]

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

where $X_{0,m}(\omega)$ represents the speech component and $V_{0,m}(\omega)$ represents the noise component. Similarly, the mth microphone signal in the right hearing aid can be specified as $Y_{1,m}(\omega) = X_{1,m}(\omega) + V_{1,m}(\omega)$. For conciseness, we will omit the frequency-domain variable ω from now on.

We define the M-dimensional stacked vectors $\mathbf{Y}_0$ and $\mathbf{Y}_1$ and the $2M$-dimensional signal vector $\mathbf{Y}$ as

$$\mathbf{Y}_0 = \begin{bmatrix} Y_{0,0} \\ \vdots \\ Y_{0,M-1} \end{bmatrix}, \quad \mathbf{Y}_1 = \begin{bmatrix} Y_{1,0} \\ \vdots \\ Y_{1,M-1} \end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix} \mathbf{Y}_0 \\ \mathbf{Y}_1 \end{bmatrix}. \quad (2)$$

The signal vector can be written as $\mathbf{Y} = \mathbf{X} + \mathbf{V}$, where $\mathbf{X}$ and $\mathbf{V}$ are defined similarly as $\mathbf{Y}$.

In the binaural processing scheme, signals are exchanged between the two devices. We assume that, due to bandwidth limitations, not every microphone signal can be transmitted, so that each device does not have access to the full signal vector $\mathbf{Y}$. We assume that each device transmits N linear combinations of its microphone signals, i.e.

$$\mathbf{Y}_{10} = \mathbf{F}_{10}^H \mathbf{Y}_0 \ , \quad \mathbf{Y}_{01} = \mathbf{F}_{01}^H \mathbf{Y}_1 \ , \quad (3)$$

where $\mathbf{F}_{10}$ and $\mathbf{F}_{01}$ are $M \times N$ dimensional complex matrices with $1 \leq N < M$. The available signals in the left and right device[2] can then be written as:

$$\mathbf{Y}_L = \begin{bmatrix} \mathbf{Y}_0 \\ \mathbf{Y}_{01} \end{bmatrix} = \left[\begin{array}{c:c} \mathbf{I}_M & \mathbf{0}_{M\times M} \\ \hdashline \mathbf{0}_{N\times M} & \mathbf{F}_{01}^H \end{array}\right] \mathbf{Y} = \mathbf{Q}_L^H \mathbf{Y} \ , \quad (4)$$

$$\mathbf{Y}_R = \begin{bmatrix} \mathbf{Y}_{10} \\ \mathbf{Y}_1 \end{bmatrix} = \left[\begin{array}{c:c} \mathbf{F}_{10}^H & \mathbf{0}_{N\times M} \\ \hdashline \mathbf{0}_{M\times M} & \mathbf{I}_M \end{array}\right] \mathbf{Y} = \mathbf{Q}_R^H \mathbf{Y} \ , \quad (5)$$

where $\mathbf{Q}_L$ and $\mathbf{Q}_R$ are $2M \times (M+N)$ matrices which compress the signal vector $\mathbf{Y}$ into the lower-dimensional $\mathbf{Y}_L$ and $\mathbf{Y}_R$. The speech components $\mathbf{X}_L$ and $\mathbf{X}_R$ and the noise components $\mathbf{V}_L$ and $\mathbf{V}_R$ are similarly defined. For the special case of a bilateral setup (where no signals are transmitted, i.e. $N = 0$), the same framework is used, but $\mathbf{Q}_L$ and $\mathbf{Q}_R$ are then $2M \times M$ matrices, defined as:

$$\mathbf{Q}_L^H = \left[\begin{array}{c:c} \mathbf{I}_M & \mathbf{0}_{M\times M} \end{array}\right] , \quad (6)$$

$$\mathbf{Q}_R^H = \left[\begin{array}{c:c} \mathbf{0}_{M\times M} & \mathbf{I}_M \end{array}\right] . \quad (7)$$

The correlation matrix $\mathbf{R}_y$, the speech correlation matrix $\mathbf{R}_x$ and the noise correlation matrix $\mathbf{R}_v$ are defined as

$$\mathbf{R}_y = \mathcal{E}\{\mathbf{Y}\mathbf{Y}^H\}, \ \mathbf{R}_x = \mathcal{E}\{\mathbf{X}\mathbf{X}^H\}, \ \mathbf{R}_v = \mathcal{E}\{\mathbf{V}\mathbf{V}^H\} \ , \quad (8)$$

where $\mathcal{E}$ denotes the expected value operator. Assuming that the speech and the noise components are uncorrelated, $\mathbf{R}_y = \mathbf{R}_x + \mathbf{R}_v$. By using definitions (4) and (5), the left and right correlation matrices (i.e. the correlation matrices estimated in the left and right devices) can be written as

$$\mathbf{R}_{yL} = \mathbf{Q}_L^H \mathbf{R}_y \mathbf{Q}_L, \quad \mathbf{R}_{yR} = \mathbf{Q}_R^H \mathbf{R}_y \mathbf{Q}_R \ . \quad (9)$$

The speech correlation matrices $\mathbf{R}_{xL}$ and $\mathbf{R}_{xR}$ and the noise correlation matrices $\mathbf{R}_{vL}$ and $\mathbf{R}_{vR}$ are similarly defined.

We will use the r_L-th microphone on the left device and the r_R-th microphone on the right device as the so-called reference microphones for the noise reduction algorithms. Typically, the front microphones are used as reference microphones (i.e. $r_L = r_R = 0$). The reference microphone signals at the left and the right device are denoted as $Y_{L,\mathrm{ref}}$ and $Y_{R,\mathrm{ref}}$, which are then equal to

$$Y_{L,\mathrm{ref}} = \mathbf{e}_L^H \mathbf{Y}_L, \quad Y_{R,\mathrm{ref}} = \mathbf{e}_R^H \mathbf{Y}_R \ , \quad (10)$$

where $\mathbf{e}_L$ and $\mathbf{e}_R$ are $(M+N)$-dimensional vectors with only one non-zero element, namely $\mathbf{e}_L(r_L + 1) = 1$ and $\mathbf{e}_R(N + r_R + 1) = 1$.

The output signals Z_L and Z_R for the left and the right device are obtained by filtering and summing the left and right signal vectors, i.e.

$$Z_L = \mathbf{W}_L^H \mathbf{Y}_L, \quad Z_R = \mathbf{W}_R^H \mathbf{Y}_R \ , \quad (11)$$

where $\mathbf{W}_L$ and $\mathbf{W}_R$ are $(M+N)$-dimensional complex weight vectors.

B. Special case: single target speech source

In the case of a single target speech source, the target speech signal vector can be modelled as

$$\mathbf{X} = \mathbf{A}S \ , \quad (12)$$

where the $2M$-dimensional steering vector $\mathbf{A}$ contains the acoustic transfer functions from the speech source to the microphones (including room acoustics, microphone characteristics and headshadow effect) and S denotes the target speech signal. As in (4) and (5), the vectors $\mathbf{A}_L$ and $\mathbf{A}_R$ can be defined as $\mathbf{A}_L = \mathbf{Q}_L^H \mathbf{A}$ and $\mathbf{A}_R = \mathbf{Q}_R^H \mathbf{A}$. Similarly to (10), we have $A_{L,\mathrm{ref}} = \mathbf{e}_L^H \mathbf{A}_L$ and $A_{R,\mathrm{ref}} = \mathbf{e}_R^H \mathbf{A}_R$.

With assumption (12), the speech correlation matrix is a rank-1 matrix, i.e.

$$\mathbf{R}_x = P_s \mathbf{A}\mathbf{A}^H \ , \quad (13)$$

with $P_s = \mathcal{E}\{|S|^2\}$ the power of the target speech signal. The speech correlation matrices $\mathbf{R}_{xL}$ and $\mathbf{R}_{xR}$ (i. e. the correlation matrix estimates at the left and right devices) can be similarly written as $\mathbf{R}_{xL} = P_s \mathbf{A}_L \mathbf{A}_L^H$ and $\mathbf{R}_{xR} = P_s \mathbf{A}_R \mathbf{A}_R^H$.

C. Performance measures

The *input and output SNR's* are defined as

$$SNR_L^{in} = \frac{\mathbf{e}_L^H \mathbf{R}_{xL} \mathbf{e}_L}{\mathbf{e}_L^H \mathbf{R}_{vL} \mathbf{e}_L}, \ SNR_L^{out} = \frac{\mathbf{W}_L^H \mathbf{R}_{xL} \mathbf{W}_L}{\mathbf{W}_L^H \mathbf{R}_{vL} \mathbf{W}_L}, \quad (14)$$

for the left device. SNR_R^{in} and SNR_R^{out} are similarly defined. The input and output *interaural transfer functions (ITF)* of

[1]The notational conventions of [2] will be followed.

[2]In contrast to [2], [5], a distinction is made here between the microphone signal vectors $\mathbf{Y}_0$ and $\mathbf{Y}_1$, and the signal vectors $\mathbf{Y}_L$ and $\mathbf{Y}_R$ which represent the available signals in the left and right devices.

the target speech source are defined as the ratio of the target speech signals at the left and right devices, i.e.

$$ITF_x^{in} = \frac{\mathbf{e}_L^H \mathbf{X}_L}{\mathbf{e}_R^H \mathbf{X}_R}, \quad ITF_x^{out} = \frac{\mathbf{W}_L^H \mathbf{X}_L}{\mathbf{W}_R^H \mathbf{X}_R} \; . \tag{15}$$

The target speech binaural cues, i.e. the interaural level differences (*ILD's*) and interaural time differences (*ITD's*), are defined as

$$\begin{aligned} ILD_x^{in} &= \frac{\mathbf{e}_L^H \mathbf{R}_{xL} \mathbf{e}_L}{\mathbf{e}_R^H \mathbf{R}_{xR} \mathbf{e}_R} \; , ITD_x^{in} = \angle\Big(\mathbf{e}_L^H \mathbf{R}_{xLR} \mathbf{e}_R\Big), \\ ILD_x^{out} &= \frac{\mathbf{W}_L^H \mathbf{R}_{xL} \mathbf{W}_L}{\mathbf{W}_R^H \mathbf{R}_{xR} \mathbf{W}_R} \; , ITD_x^{out} = \angle\Big(\mathbf{W}_L^H \mathbf{R}_{xLR} \mathbf{W}_R\Big), \end{aligned} \tag{16}$$

with $\mathbf{R}_{xLR} = \mathbf{Q}_L^H \mathbf{R}_x \mathbf{Q}_R$.

III. Review of Multichannel Wiener Filter

A. General case

The Speech Distortion Weighted Multichannel Wiener Filter (SDW-MWF) minimizes a weighted sum of the residual noise energy and the speech distortion energy [1]. It is straightforward to extend the SDW-MWF for a binaural hearing aid application, where a left and a right output are generated [2]. The optimal SDW-MWF's for the left and right devices are then equal to

$$\boxed{\begin{aligned} \mathbf{W}_{MWF,L} &= (\mathbf{R}_{xL} + \mu \mathbf{R}_{vL})^{-1} \mathbf{R}_{xL} \mathbf{e}_L \; , \\ \mathbf{W}_{MWF,R} &= (\mathbf{R}_{xR} + \mu \mathbf{R}_{vR})^{-1} \mathbf{R}_{xR} \mathbf{e}_R \; , \end{aligned}} \tag{17}$$

where we make use of the correlation matrices defined in (9), and where μ is a trade-off parameter which allows to weight noise reduction against speech distortion.

B. Single target speech source

By plugging (13) into (17), it can be seen that in the single target speech source case, the optimal filters reduce to:

$$\boxed{\begin{aligned} \mathbf{W}_{MWF,L} &= \mathbf{R}_{vL}^{-1} \mathbf{A}_L \; \tfrac{P_s}{\mu+\rho_L} A_{L,\mathrm{ref}}^* \; , \\ \mathbf{W}_{MWF,R} &= \mathbf{R}_{vR}^{-1} \mathbf{A}_R \; \tfrac{P_s}{\mu+\rho_R} A_{R,\mathrm{ref}}^* \; , \end{aligned}} \tag{18}$$

with

$$\rho_L = P_s \mathbf{A}_L^H \mathbf{R}_{vL}^{-1} \mathbf{A}_L \; , \tag{19}$$

$$\rho_R = P_s \mathbf{A}_R^H \mathbf{R}_{vR}^{-1} \mathbf{A}_R \; . \tag{20}$$

Using definition (14), it can be seen that the output SNR's obtained at the left and right device are equal to these values ρ_L and ρ_R, i.e.

$$\boxed{SNR_L^{out} = \rho_L \; , \qquad SNR_R^{out} = \rho_R \; .} \tag{21}$$

IV. Target speech binaural cue preservation

By using the framework of section II-A, and by assuming a single speech signal as in section II-B, the output ITF of the target speech source can now be calculated for any transmission scheme (3). By plugging the optimal SDW-MWF filters (18) into the ITF definitions (15), we obtain:

$$\boxed{ITF_x^{out} = \frac{1 + \frac{\mu}{\rho_R}}{1 + \frac{\mu}{\rho_L}} \; ITF_x^{in} \; .} \tag{22}$$

Formula (22) shows that the output speech ITF is a scaled version of the input speech ITF, where the scaling factor is a real number. As the ITD is equal to the phase of the ITF [5], this implies that the speech ITD cues are not distorted by the SDW-MWF, regardless of the transmission scheme. This also holds for the bilateral case where no signals are transmitted. As the output speech ITF is scaled compared to the input speech ITF, the speech ILD cues will however be distorted.

Using (22), the relative speech ITF error can be derived as

$$\Delta ITF_x = \frac{|ITF_x^{out} - ITF_x^{in}|}{ITF_x^{in}} = \frac{\mu \; |\rho_L - \rho_R|}{(\mu + \rho_L) \; \rho_R} \; . \tag{23}$$

For the special case where $\rho_L = \rho_R$, which means the obtained left and right output SNR's are equal, it can thus be seen that the ILD cues are also undistorted. This is e.g. the case when all signals are transmitted over the wireless link, i.e. $\mathbf{Q}_L = \mathbf{Q}_R = \mathbf{I}_{2M}$. This is in accordance with the results in [5], where it was indeed assumed that all microphone signals are available in the left and right device.

In a reduced-bandwidth scheme, with $N < M$, in general $\rho_L \neq \rho_R$, which leads to ILD distortions. If more signals are transmitted or if a better estimate of the target speech signal is transmitted, the obtained output SNR's ρ_L and ρ_R will generally increase while the mismatch $|\rho_L - \rho_R|$ decreases [2]. As can be seen from (23), the ITF error (or ILD distortion) will then also decrease.

V. Simulations

A. Experimental setup

We consider a binaural hearing aid configuration where the left and right devices each have 2 microphones ($M = 2$). Head-related transfer functions (HRTF's) are measured in a low-reverberant room ($RT_{60} = 210$ ms [8]) on a dummy-head, so that the headshadow effect is taken into account. Twelve different speech-noise configurations are tested, where the azimuthal angles (defined clockwise with 0° as frontal direction) and the number of noise sources are varied. The target speech source is always placed at 0°, except for the last scenario. The first eight scenarios have a single noise source at an angle between 30° and 330°. Scenario N2 has 2 noise sources at 120° and 240°, scenario N3 has 3 noise sources at 90°, 180° and 270°, and scenario N4 has 4 noise sources at 60°, 120°, 180° and 210°. Finally, for scenario S90N270 the target speech source is at 90° and the noise source is at 270°. The HRTF data is used to construct steering vectors so that the speech correlation matrix can be constructed by the rank one model (13), and multitalker babble [9] is used as noise signal(s). The signals are scaled so that the average input SNR at the right front microphone is equal to 0 dB.

The filters are calculated in a batch procedure, with μ set to 5. A filter length of 128 taps, at a sampling frequency of 20480 Hz, is used. For the performance comparison, the 15-th frequency bin (corresponding to 2240 Hz) is selected, which is in the frequency range where ILD cues are the dominant binaural cues.

As performance measures, we use the input and output SNR's defined in (14), as well as the speech ILD error, i.e.

$$\Delta ILD_x = |10\log_{10}(ILD_x^{out}) - 10\log_{10}(ILD_x^{in})|\,, \quad (24)$$

where ILD_x^{out} and ILD_x^{in} are defined in (16).
The considered algorithms are:

- *MWF-all*, i.e. all microphone signals are transmitted, so that $\mathbf{Q}_L = \mathbf{Q}_R = \mathbf{I}$.
- *MWF-front*, i.e. the front signals are transmitted, so that $\mathbf{F}_{10}^H = \mathbf{F}_{01}^H = [1\ 0]$.
- *MWF-contra*, i.e. the transmitted signals are the outputs of a 2-microphone monaural MWF, as in [2].
- *MWF-bilateral*, i.e. no signals are transmitted, so that $\mathbf{Q}_L$ and $\mathbf{Q}_R$ are given by (6) and (7).

B. Results

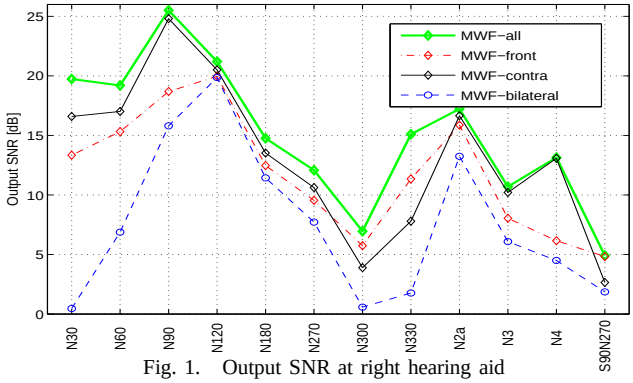

Fig. 1. Output SNR at right hearing aid

In figure 1, the output SNR obtained in the right device is shown for the twelve different spatial scenarios. It is apparent that the bilateral MWF obtains a significantly lower SNR than the binaural approaches, especially so for scenarios where the angle between the target speech source and the noise source is small ($\leq 60°$) and for scenarios with multiple noise sources, as was also shown in [8]. The results also illustrate that MWF-contra provides a better performance than MWF-front in many cases, sometimes even close to the performance of MWF-all. The exceptions are scenarios N300 and N330, which can be explained as follows: as the monaural MWF at the left device cannot estimate a high SNR signal when the noise source is located at these angles, the transmitted signal will be degraded so that a worse performance is obtained in the right device. This is also observed in [2], [7]. For the output SNR in the left device similar results are obtained, but these are omitted here because of space restrictions.

In figure 2, the speech ILD error is shown for the twelve different spatial scenarios. As in [5], MWF-all does not introduce a speech ILD error, while the reduced-bandwidth and bilateral algorithms introduce ILD errors in several spatial scenarios. MWF-bilateral introduces the largest error for most scenarios, especially so for scenarios where only a low SNR is obtained at either the left and/or right device. Exceptions (such as N300) are due to the fact that the obtained output SNR's at the left and right devices are almost equal so that $|\rho_L - \rho_R|$ in the numerator of (23) is small. MWF-contra

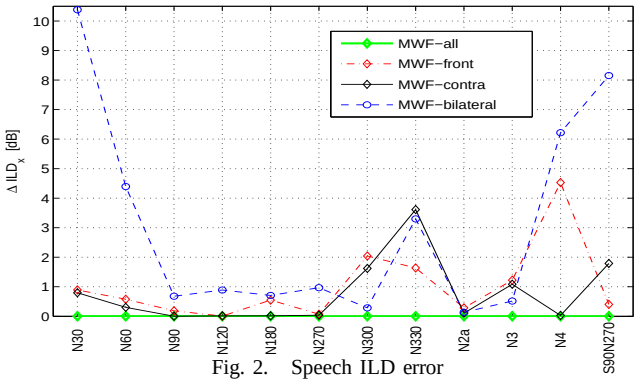

Fig. 2. Speech ILD error

generally outperforms MWF-front except for N330, where MWF-front obtains a higher output SNR as was explained before. Overall, it can be stated that a speech ILD error is introduced when the obtained output SNR's at the left and right devices are very different.

VI. Conclusion

In this paper, a general framework was proposed for the analysis of target speech binaural cue preservation by reduced-bandwidth and bilateral MWF-based noise reduction algorithms. It was proven theoretically and illustrated through experiments that these algorithms preserve the speech ITD cues, but distort the speech ILD cues. The experiments also show that, in addition to providing a better noise reduction performance, using a binaural link also reduces the speech ILD errors compared to the bilateral case, especially so when a good estimate of the speech signal is transmitted.

References

[1] S. Doclo, A. Spriet, J. Wouters, and M. Moonen, "Frequency-Domain Criterion for Speech Distortion Weighted Multichannel Wiener Filter for Robust Noise Reduction," *Speech Communication, special issue on Speech Enhancement*, vol. 49, no. 7–8, pp. 636–656, Jul.-Aug. 2007.

[2] S. Doclo, T. Van den Bogaert, J. Wouters, and M. Moonen, "Reduced-bandwidth and distributed MWF-based noise reduction algorithms for binaural hearing aids," *IEEE Trans. Audio, Speech and Language Processing*, vol. 17, no. 1, pp. 38–51, Jan. 2009.

[3] J. Desloge, W. Rabinowitz, and P. Zurek, "Microphone-array hearing aids with binaural output–Part I: Fixed-processing systems," *IEEE Trans. Speech and Audio Processing*, vol. 5, no. 6, pp. 529–542, Nov. 1997.

[4] T. Lotter and P. Vary, "Dual-channel speech enhancement by superdirective beamforming," *EURASIP Journal on Applied Signal Processing*, vol. 2006, Article ID 63297, 2006.

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

[6] O. Roy and M. Vetterli, "Rate-constrained beamforming for collaborating hearing aids," in *Proc. International Symposium on Information Theory (ISIT)*, Seattle WA, USA, July 2006, pp. 2809–2813.

[7] S. Srinivasan and A. den Brinker, "Rate-constrained beamforming in binaural hearing aids," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, pp. 1–9, 2009.

[8] T. Van den Bogaert, S. Doclo, M. Moonen, and J. Wouters, "Speech enhancement with multichannel Wiener filter techniques in multi-microphone binaural hearing aids," *Journal of the Acoustical Society of America*, vol. 125, no. 1, pp. 360–371, 2009.

[9] Auditec, "Auditory tests (revised), Compact Disc, Auditec," St.Louis, 1997.