Open access • Journal Article • DOI:10.1109/TSP.2006.888897

Binaural Noise Reduction Algorithms for Hearing Aids That Preserve Interaural Time Delay Cues — Source link

Thomas Klasen, T. Van den Bogaert, Marc Moonen, Jan Wouters

Institutions: Katholieke Universiteit Leuven

Published on: 01 Apr 2007 - IEEE Transactions on Signal Processing (Institute of Electrical and Electronics Engineers)

Topics: Hearing aid, Noise, Binaural recording, Monaural and Speech processing

[15] J. Zhang, Y. Wu, S. Zhou, and J. Wang, "Joint linear transmitter and receiver design for the downlink of multiuser MIMO systems," *IEEE Commun. Lett.*, vol. 9, pp. 991–993, Nov. 2005.
[16] A. Mezghani, M. Joham, R. Hunger, and W. Utschick, "Transceiver design for multi-user MIMO systems," presented at the ITG/IEEE Workshop Smart Antennas (WSA'06), Ulm, Germany, Mar. 2006.
[17] A. M. Khachan, A. J. Tenenbaum, and R. S. Adve, "Linear processing for the downlink in multiuser MIMO systems with multiple data streams," presented at the IEEE Int. Conf. Communications (ICC'06), Istanbul, Turkey, Jun. 2006.
[18] M. Joham, W. Utschick, and J. A. Nossek, "Linear transmit processing in MIMO communications systems," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2700–2712, Aug. 2005.
[19] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commun.*, vol. 49, pp. 2198–2206, Dec. 2001.
[20] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Trans. Signal Processing*, vol. 51, no. 9, pp. 2381–2401, Sep. 2003.
[21] S. Serbetli and A. Yener, "Transceiver optimization for multiuser MIMO systems," *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 214–226, Jan. 2004.
[22] L. S. Lasdon, *Optimization Theory for Large Systems*. New York: Dover, 2002.
[23] N. Jindal, W. Rhee, S. Vishwanath, S. A. Jafar, and A. Goldsmith, "Sum power iterative water-filling for multi-antenna Gaussian broadcast channels," *IEEE Trans. Inf. Theory*, vol. 51, pp. 1570–1580, Apr. 2005.
[24] Q. H. Spencer and A. L. Swindlehurst, "A hybrid approach to spatial multiplexing in multi-user MIMO downlinks," *EURASIP J. Wireless Commun. Netw.*, vol. 2004, no. 2, pp. 236–247, 2004.

Binaural Noise Reduction Algorithms for Hearing Aids That Preserve Interaural Time Delay Cues

Thomas J. Klasen, *Student Member, IEEE*,
Tim Van den Bogaert, *Student Member, IEEE*,
Marc Moonen, *Member, IEEE*, and Jan Wouters

***Abstract*—Binaural hearing aids use microphone inputs from both the left and right hearing aid to generate an output for each ear. On the other hand, a monaural hearing aid generates an output by processing only its own microphone inputs. This correspondence presents a binaural extension of a monaural multichannel noise reduction algorithm for hearing aids based on Wiener filtering. In addition to significantly suppressing the noise interference, the algorithm preserves the interaural time delay (ITD) cues of the speech component, thus allowing the user to correctly localize the speech source. Unfortunately, binaural multichannel Wiener filtering distorts the ITD cues of the noise source. By adding a parameter to the cost function the amount of noise reduction performed by the algorithm can be controlled, and traded off for the preservation of the noise ITD cues.**

***Index Terms*—Binaural hearing, hearing aids, noise reduction, Wiener filtering.**

I. Introduction

Hearing impaired persons localize sounds better without their bilateral hearing aids than with them [1]. In addition, noise reduction algorithms currently used in hearing aids are not designed to preserve localization cues [2]. The inability to correctly localize sounds puts the hearing aid user at a disadvantage. The sooner the user can localize a speech signal, the sooner the user can begin to exploit visual cues. Generally, visual cues lead to large improvements in intelligibility for hearing impaired persons [3]. Furthermore, preserving the spatial separation between the target speech and the interfering signals leads to an improvement in speech understanding [4].

It is important to explain the difference between bilateral and binaural hearing aids. A hearing impaired person wearing a monaural hearing aid on each ear is said to be using bilateral hearing aids. Each monaural hearing aid uses its own microphone inputs to generate an output for its respective ear. No information is shared between the hearing aids. In contrast, binaural hearing aids use the microphone inputs from both the left and right hearing aid to generate an output for the left and right ear. Additional information regarding binaural hearing aids can be found in [5].

What are the benefits of a binaural algorithm? First, noise reduction performance of the binaural algorithm will be better than that of the monaural algorithm. Double the number of microphones are now at the

Manuscript received December 7, 2005; revised June 1, 2006. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Jan C. de Munck. This research work was carried out at the ESAT Laboratory of the Katholieke Universiteit Leuven, in the framework of the Belgian Program on Inter-University Attraction Poles, initiated by the Belgian Federal Science Policy Office IUAP P5/22 (Dynamical Systems and Control: Computation, Identification and Modeling), the Concerted Research Action GOA-MEFISTO-666 (Mathematical Engineering for Information and Communication Systems Technology) of the Flemish Government, Research Project FWO nr.G.0233.01 (Signal processing and automatic patient fitting for advanced auditory prostheses), and IWT project 020540 (Innovative Speech Processing Algorithms for Improved Performance of Cochlear Implants). The scientific responsibility is assumed by its authors.

T. J. Klasen and M. Moonen are with ESAT-SCD Katholieke Universiteit Leuven, 2001 Leuven, Belgium.

T. Van den Bogaert and J. Wouters are with ExpORL Katholieke Universiteit Leuven, 2001 Leuven, Belgium (e-mail: tim.vandenbogaert@med.kuleuven.be)

Digital Object Identifier 10.1109/TSP.2006.888897

disposal of the algorithm and some of these microphones will have a better signal-to-noise ratio (SNR) than the others. Imagine the extreme scenario where the speech source is located at left side of the hearing aid user and the noise source at the right side of the user, the signals at the left ear have a much better SNR than the those at the right ear. The head blocks some of the noise coming from the right side. This binaural benefit is referred to as the best ear advantage [4]. Second, studies have shown that the spatial separation between the speech and noise sources contributes to an improvement in intelligibility [4], [13]. This is referred to as *spatial release from masking*. Therefore, the benefit of a noise reduction algorithm that preserves localization cues is twofold. First, noise reduction leads to an improvement in intelligibility. In addition, preserving localization cues preserves the spatial separation of the target speech and noise sources, resulting again in an improvement in intelligibility.

The goal of this correspondence is to develop binaural noise reduction algorithms that preserve speech localization cues without sacrificing noise reduction performance. We focus specifically on interaural time delay (ITD) cues, which help the listener localize sounds horizontally [6], [7]. ITD is the time delay in the arrival of the sound signal between the left and right ear. If the ITD cues of the processed signal are the same as the ITD cues of the unprocessed signal, we assume that a user will localize the processed signal and the unprocessed signal to the same source.

In [8], a binaural adaptive noise reduction algorithm is proposed. This algorithm takes a microphone signal from each ear as inputs. The inputs are filtered by a high-pass and a low-pass filter with the same cutoff frequency to create a high-frequency and a low-frequency portion. The high-frequency portion is adaptively processed and added to the delayed low-frequency portion. Since ITD cues are contained in the low-frequency regions, as the cutoff frequency increases more ITD information will arrive undistorted to the user [6]. The major drawback to this approach is that the low-frequency portion containing speech energy also contains noise energy. Consequently, noise, as well as speech energy, is passed from the input to the output unprocessed. Therefore, there is a tradeoff between noise reduction and speech ITD cue preservation.

This correspondence extends the monaural multichannel Wiener filtering algorithm discussed in [9]–[11] to a binaural algorithm first proposed in [12]. The monaural multichannel Wiener filtering algorithm is well suited for binaural extension, because it makes no assumption about the location of the speech source and is capable of estimating the speech components in all microphone channels and knowledge of the exact geometry of the microphone array is not necessary [10].

The binaural extension is shown to outperform two independent monaural multichannel Wiener filtering algorithms (bilateral algorithm) in terms of noise reduction performance. In addition, the binaural algorithm consistently preserves the ITD cues of the speech source without sacrificing noise reduction performance. In order to preserve the noise ITD cues, some of the noise signal is passed to the output of the algorithm unprocessed. Consequently, the binaural algorithm is modified so the emphasis on noise reduction can be controlled. As less emphasis is put on noise reduction, more noise arrives at the output of the algorithm unprocessed; accordingly more noise ITD cues will arrive undistorted to the user. Although a tradeoff between noise reduction performance and noise ITD cue preservation also exists for this algorithm, it differs from the one proposed in [8], since the current algorithm consistently preserves the speech ITD cues without sacrificing noise reduction.

This correspondence is organized into six sections. In Section II, the listening scenario is discussed and several standard assumptions are made. The binaural extension of the multichannel Wiener filtering algorithm is derived in Section III. This algorithm is modified in Section IV, allowing one to control the amount of noise reduction and therefore the distortion of the noise ITD cues. Finally, Section V explores the performance of the binaural multichannel Wiener filtering algorithms and the binaural adaptive algorithm discussed in [8] and compares the performance of the binaural and bilateral multichannel Wiener filtering algorithms.

II. System Model

A. Listening Scenario

The speaker speaks intermittently in the continuous background noise caused by a noise source. There are M microphones on each hearing aid. We refer to the mth microphone of the left hearing aid and the mth microphone of the right hearing aid as the mth microphone pair. The received signals at time k at the mth microphone pair can be written as

$$y_{\mathrm{L}_m}[k] = x_{\mathrm{L}_m}[k] + v_{\mathrm{L}_m}[k] \tag{1}$$
$$y_{\mathrm{R}_m}[k] = x_{\mathrm{R}_m}[k] + v_{\mathrm{R}_m}[k]. \tag{2}$$

In (1) and (2), $x_{\mathrm{L}_m}[k]$ and $x_{\mathrm{R}_m}[k]$ represent the speech component in the mth microphone pair. Likewise, $v_{\mathrm{L}_m}[k]$ and $v_{\mathrm{R}_m}[k]$ represent the noise component at the mth microphone pair.

We make two standard assumptions that will be pertinent later. First, the speech signal is assumed to be statistically independent of the noise signal. Second, we assume that the noise is zero-mean and short-term stationary.

B. Voice Activity Detection

The signals received at the microphones of the left and right hearing aids contain either noise when speech is not present, or speech and noise. We assume that we have access to a perfect VAD algorithm. In other words, we can identify when there is only noise present, and when there is speech and noise present. For simplicity, let us call the time instants when there is only noise present k^n and when there is speech and noise present k^{sn}.

III. Binaural Multichannel Wiener Filtering

This algorithm is an extension of the multichannel Wiener filtering technique discussed in [9]–[11]. The goal of this algorithm is to estimate the speech components of the mth microphone pair $x_{\mathrm{L}_m}[k]$ and $x_{\mathrm{R}_m}[k]$ using all received microphone signals $y_{\mathrm{L}_{1:M}}[k]$ and $y_{\mathrm{R}_{1:M}}[k]$. In order to estimate the speech components of the mth microphone pair, we design two Wiener filters that estimate the noise components in the mth microphone pair. The noise estimates of the mth microphone pair and therefore the output of the two Wiener filters are $\tilde{v}_{\mathrm{L}_m}[k]$ and $\tilde{v}_{\mathrm{R}_m}[k]$. To obtain the estimates of the speech components of the mth microphone pair, the estimates of the noise components are subtracted from the original signals received at the two microphones. The speech and error estimates are defined below for the left microphone.

$$\tilde{x}_{\mathrm{L}_m}[k] = (x_{\mathrm{L}_m}[k] + v_{\mathrm{L}_m}[k]) - \tilde{v}_{\mathrm{L}_m}[k] \tag{3}$$
$$e_{\mathrm{L}_m}[k] = v_{\mathrm{L}_m}[k] - \tilde{v}_{\mathrm{L}_m}[k]. \tag{4}$$

The speech and error estimates for the right microphone can be defined similarly.

Before going any further, a few definitions are necessary. We will define only the filter $\mathbf{w}_{\mathrm{Left}}[k]$, which operates on all microphone signals to create an estimate of the noise signal in the mth left microphone. A similar definition is made for $\mathbf{w}_{\mathrm{Right}}$. We choose the filters $\mathbf{w}_{\mathrm{L}_m}[k]$

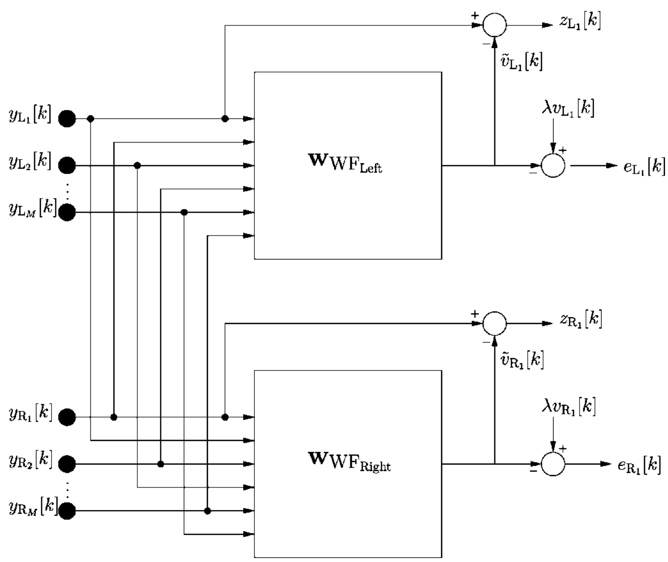

Fig. 1. Binaural ($\lambda = 1$) and controlled binaural multichannel Wiener filtering.

and $\mathbf{w}_{R_m}[k]$ to be of length N. The filter $\mathbf{w}_{L_m}[k]$ is expressed in the following equation:

$$\mathbf{w}_{L_m}[k] = \left[w^0_{L_m} w^1_{L_m} \dots w^{N-1}_{L_m}\right]^T \tag{5}$$

The filter $\mathbf{w}_{R_m}[k]$ is defined similarly. Next, we create a stacked vector, length $2MN$, of the individual left and right microphone filters, as follows:

$$\mathbf{w}_{\text{Left}}[k] = \left[\mathbf{w}^T_{L_1}[k] \dots \mathbf{w}^T_{L_M}[k] \quad \mathbf{w}^T_{R_1}[k] \dots \mathbf{w}^T_{R_M}[k]\right]^T \tag{6}$$

The received microphone signal at mth left microphone is defined as follows;

$$\mathbf{y}_{L_m}[k] = [y_{L_m}[k] y_{L_m}[k-1] \ \dots \ y_{L_m}[k-N+1]]^T. \tag{7}$$

The mth right microphone signal can be written in a similar fashion. Again, we create a stacked input vector, length $2MN$, of the individual left and right microphone inputs, as follows:

$$\mathbf{y}[k] = \left[\mathbf{y}^T_{L_1}[k] \dots \mathbf{y}^T_{L_M}[k] \quad \mathbf{y}^T_{R_1}[k] \dots \mathbf{y}^T_{R_M}[k]\right]^T. \tag{8}$$

In this section, we derive the left and right multichannel Wiener filters in a statistical setting. The goal is to develop left and right multichannel Wiener filters that minimize the error signals $e_{L_m}[k]$ and $e_{R_m}[k]$. The filtering process is illustrated in Fig. 1, with $\lambda = 1$. Minimizing the following cost function:

$$E\left\{|\mathbf{y}^T[k][\mathbf{w}_{\text{Left}}[k]\mathbf{w}_{\text{Right}}[k]] - [v_{L_m}[k] v_{R_m}[k]]|^2\right\} \tag{9}$$

minimizes the error signals. In (9), $E\{\cdot\}$ is the expectation operator. The filters achieving the minimum of the cost function are the well-known Wiener filters expressed as

$$[\mathbf{w}_{WF_{\text{Left}}}[k]\mathbf{w}_{WF_{\text{Right}}}[k]] = E\{\mathbf{y}[k]\mathbf{y}^T[k]\}^{-1} E\{\mathbf{y}[k][v_{L_m}[k] v_{R_m}[k]]\}. \tag{10}$$

Owing to (1) and (2), we can define $\mathbf{x}[k]$ and $\mathbf{v}[k]$, where $\mathbf{y}[k] = \mathbf{x}[k] + \mathbf{v}[k]$. In Section II-A, the first assumption asserts that the speech signal and the noise signal are statistically independent. More specifically, the following equation must hold:

$$E\{\mathbf{x}[k]\mathbf{v}^T[k]\} = 0. \tag{11}$$

Using this assumption, we can rewrite (10) by making the following substitution:

$$E\{\mathbf{y}[k][v_{L_m}[k] v_{R_m}[k]]\} = E\{\mathbf{v}[k][v_{L_m}[k] v_{R_m}[k]]\}. \tag{12}$$

Unfortunately, in real life, these statistical quantities are not immediately available. Therefore, we cannot calculate the left and right Wiener filters directly. Instead, we make a least-squares approximation of the filters. This data-based approach requires a few extra definitions. Using (8), we write the input matrix $\mathbf{Y}$, which is of size K by $2MN$, as follows:

$$\mathbf{Y}[k] = \begin{bmatrix} \mathbf{y}^T[k] \\ \mathbf{y}^T[k-1] \\ \vdots \\ \mathbf{y}^T[k-K+1] \end{bmatrix}. \tag{13}$$

Analogously, the speech input matrix $\mathbf{X}[k]$ and the noise input matrix $\mathbf{V}[k]$, can be defined, where $\mathbf{Y}[k] = \mathbf{X}[k] + \mathbf{V}[k]$. Finally, we write the desired signals $\mathbf{d}_L[k]$ and $\mathbf{d}_R[k]$, which are the unknown noise input vectors, as follows:

$$\mathbf{d}_L[k] = \mathbf{v}_{L_m}[k] = \begin{bmatrix} v_{L_m}[k] \\ v_{L_m}[k-1] \\ \vdots \\ v_{L_m}[k-K+1] \end{bmatrix} \tag{14}$$

$$\mathbf{d}_R[k] = \mathbf{v}_{R_m}[k] = \begin{bmatrix} v_{R_m}[k] \\ v_{R_m}[k-1] \\ \vdots \\ v_{R_m}[k-K+1] \end{bmatrix}. \tag{15}$$

We define the desired matrix $\mathbf{D}[k]$ as $[\mathbf{d}_L[k]\mathbf{d}_R[k]]$. We can now estimate $E\{\mathbf{y}[k]\mathbf{y}^T[k]\}$ by the matrix $\mathbf{Y}^T[k]\mathbf{Y}[k]$ (up to a scaling). In order to estimate $E\{\mathbf{v}[k][v_{L_m}[k] v_{R_m}[k]]\}$ by $\mathbf{V}^T[k]\mathbf{D}[k]$ (up to the same scaling), we must use the second assumption we made in our system model, since the input noise matrix $\mathbf{V}[k]$, and therefore the desired matrix $\mathbf{D}[k]$ are not known explicitly. The assumption is that the noise signal is short-term stationary. This means that $E\{\mathbf{v}[k][v_{L_m}[k] v_{R_m}[k]]\}$ is the same whether it is calculated during noise only periods k^n or at all time instants k. The second assumption is expressed as

$$E\{\mathbf{v}[k][v_{L_m}[k] v_{R_m}[k]]\} = E\{\mathbf{v}[k^n][v_{L_m}[k^n] v_{R_m}[k^n]]\}. \tag{16}$$

Invoking the second assumption, $E\{\mathbf{v}[k][v_{L_m}[k] v_{R_m}[k]]\}$ can be estimated by $\mathbf{V}^T[k^n]\mathbf{D}[k^n]$ at time instants where only noise is present. Therefore, we can write the least-squares approximation of the Wiener filters as

$$[\mathbf{w}_{LS_{\text{Left}}}\mathbf{w}_{LS_{\text{Right}}}] = (\mathbf{Y}^T[k]\mathbf{Y}[k])^{-1}\mathbf{V}^T[k^n]\mathbf{D}[k^n]. \tag{17}$$

This least-squares approximation of the Wiener filters is what we use in practice.

Since the speech signals are estimated by subtracting the noise estimates from the original signal, the ITD cues of the speech are well preserved. On the other hand, this algorithm is not designed to preserve the ITD cues of the noise.

IV. Controlled Binaural Multichannel Wiener Filtering

The controlled binaural multichannel Wiener filtering algorithm attempts to estimate the speech component and a desired amount of residual noise of the mth microphone pair. In order to accomplish this, the Wiener filters are designed to estimate only a portion λ of the noise components of the mth microphone pair. Therefore, a portion of the noise signal will arrive undistorted at the output of the algorithm. As less emphasis is placed on noise reduction, more of the noise signal arrives at the output of the algorithm unprocessed. Therefore, more noise ITD cues will arrive undistorted to the user. The amount of noise passed unprocessed to the output of the algorithm must be limited in order to maintain good noise reduction. The estimates of the noise signals at the mth microphone pair are $\tilde{v}^*_{\mathrm{L}_m}[k]$ and $\tilde{v}^*_{\mathrm{R}_m}[k]$, which estimate $\lambda v_{\mathrm{L}_m}[k]$ and $\lambda v_{\mathrm{R}_m}[k]$, respectively. The speech and residual noise estimate and the error estimate of the left microphone are

$$\begin{aligned} z_{\mathrm{L}_m} &= (x_{\mathrm{L}_m}[k] + v_{\mathrm{L}_m}[k]) - \tilde{v}^*_{\mathrm{L}_m}[k] \\ &= x_{\mathrm{L}_m}[k] + (1-\lambda)v_{\mathrm{L}_m}[k] + e^*_{\mathrm{L}_m} \quad (18) \\ e^*_{\mathrm{L}_m} &= \lambda v_{\mathrm{L}_M}[k] - \tilde{v}^*_{\mathrm{L}_M}[k]. \quad (19) \end{aligned}$$

The speech and residual noise estimate and the error estimate of the right microphone can be defined similarly. Fig. 1 depicts the approach of this algorithm.

Using the new noise estimates $\tilde{v}^*_{\mathrm{L}_m}$ and $\tilde{v}^*_{\mathrm{R}_m}$, the new cost function is defined as

$$E\{|\mathbf{y}^T[k][\mathbf{w}_{\mathrm{Left}}[k]\mathbf{w}_{\mathrm{Right}}[k]] - \lambda[v_{\mathrm{L}_m}[k]v_{\mathrm{R}_m}[k]]|^2\}. \quad (20)$$

Similarly, the Wiener filters that minimize the above cost function are defined as

$$[\mathbf{w}_{\mathrm{WF}_{\mathrm{Left}}}[k]\mathbf{w}_{\mathrm{WF}_{\mathrm{Right}}}[k]] = E\{\mathbf{y}[k]\mathbf{y}^T[k]\}^{-1}E\{\mathbf{y}[k](\lambda\,[v_{\mathrm{L}_m}[k]v_{\mathrm{R}_m}[k]])\}. \quad (21)$$

Again, we assume that the speech signal is statistically independent of the noise signal, and that the noise is short-term stationary. Since λ is a scalar, the least-squares estimate of the Wiener filters can be written as

$$\left[\mathbf{w}_{\mathrm{LS}_{\mathrm{Left}}}\mathbf{w}_{\mathrm{LS}_{\mathrm{Right}}}\right] = \lambda(\mathbf{Y}^T[k]\mathbf{Y}[k])^{-1}\mathbf{V}^T[k^n]\mathbf{D}[k^n]. \quad (22)$$

This is a scaled version of (7).

Clearly, λ controls the emphasis placed on noise reduction. If $\lambda = 1$, then the algorithm is the same as the algorithm described in Section III, and the maximum amount of noise reduction is performed. On the other hand, when $\lambda = 0$ no noise reduction is performed; the output signals are exactly the same as the input signals. Therefore, a value for $\lambda \in [0, 1]$ must be chosen that suits the current acoustical situation and user.

V. Performance

A. Experimental Setup

The recordings were made in an anechoic room. Two GN ReSound Canta behind the ear (BTE) hearing aids were placed on a CORTEX MK2 artificial head. Each hearing aid had two omnidirectional microphones. The speech and noise sources were placed 1 m from the center of the artificial head. The sound pressure level (SPL) measured at the center of the dummy head was 70-dB SPL. Speech and noise sources were recorded separately. All recordings were performed at a sampling frequency of 16 kHz. HINT sentences and HINT noise were used for the speech and noise signals [14].

In both sets of simulations, the location of the speech source varied from 0°, directly in front of the user, to 345° in increments of 15°. The noise source remained fixed at 90°, to the right of the user. The signals fed into the algorithms were 10 s in length. The first half of the signal consisted of noise only. A short one and a half second sentence was spoken in the second half of the signal amidst the continuous background noise.

The first set of simulations were run to compare the two versions of the binaural multichannel Wiener filtering algorithm and the binaural adaptive algorithm discussed in [8], which will be referred to as algorithm-[8] from now on. These simulations explored the influence of the parameters, λ and the cutoff frequency, on ITD cues preservation and noise reduction performance.

In order to make a fair comparison between the algorithms, only the front microphone signal from each hearing aid was used for the binaural multichannel Wiener filtering algorithm. The filter length N was fixed at 100. The same filter length was used for the controlled binaural multichannel Wiener filter, and the parameter λ was set at 0.7 and 0.6. A batch implementation of the algorithm was used. The filters were calculated using the whole signal and remained constant throughout the filtering operation. These same filter coefficients were also used to generate the filtered clean speech and noise signals necessary to calculate the performance measures.

The filter length of algorithm-[8] was 201. The filter was adapted, during periods of noise only, by a normalized LMS algorithm. Cutoff frequencies of 500, 1200, and 1500 Hz were simulated. After the filter converged, the algorithm was rerun separately on the clean speech and noise signals using the converged filter coefficients to generate the filtered clean speech and noise signals necessary to calculate the performance measures.

The second set of simulations were run to compare the controlled binaural multichannel Wiener filtering algorithm and the controlled bilateral multichannel Wiener filtering algorithm (two controlled monaural multichannel Wiener filters). These simulations also explored the influence of the parameter λ on ITD cue preservation and noise reduction performance for both algorithms.

In order to make a fair comparison between the algorithms, both microphone signals from each hearing aid were used, $M = 2$, for the binaural and bilateral multichannel Wiener filtering algorithms. For both algorithms the filter length N was fixed at 100. The parameter λ was set at 1, 0.7, and 0.6 for controlled binaural and controlled bilateral multichannel Wiener filtering algorithms. Again a batch implementation of the algorithm was used. The filters were calculated using the whole signal and remained constant throughout the filtering operation. These same filter coefficients were also used to generate the filtered clean speech and noise signals necessary to calculate the performance measures.

B. Performance Measures

To quantify the preservation of the ITD cues, ITD error, the absolute difference between the ITD of the processed signal and the ITD of the unprocessed signal, is used. Cross correlation was used to calculate the ITD between two signals. The improvement in intelligibility weighted signal-to-noise-ratio ($\mathrm{SNR}_{\mathrm{INT}}$), defined as the difference between the output $\mathrm{SNR}_{\mathrm{INT}}$ and the input $\mathrm{SNR}_{\mathrm{INT}}$, is used to evaluate the noise reduction performance of the algorithms. $\mathrm{SNR}_{\mathrm{INT}}$ is defined in [15], as follows:

$$\mathrm{SNR}_{\mathrm{INT}} = \sum_{j=1}^{J} w_j \mathrm{SNR}_j. \quad (23)$$

The weight w_j emphasizes the importance of the jth frequency band's overall contribution to intelligibility, and SNR_j is the SNR of the jth frequency band. The band definitions and the individual weights of the J frequency bands are given in [16].

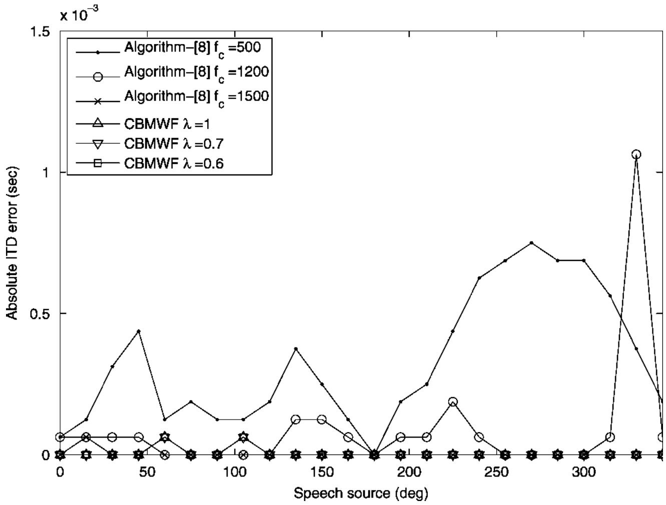

Fig. 2. ITD error speech component for speech sources between 0 and 345° with the noise source fixed at 90°.

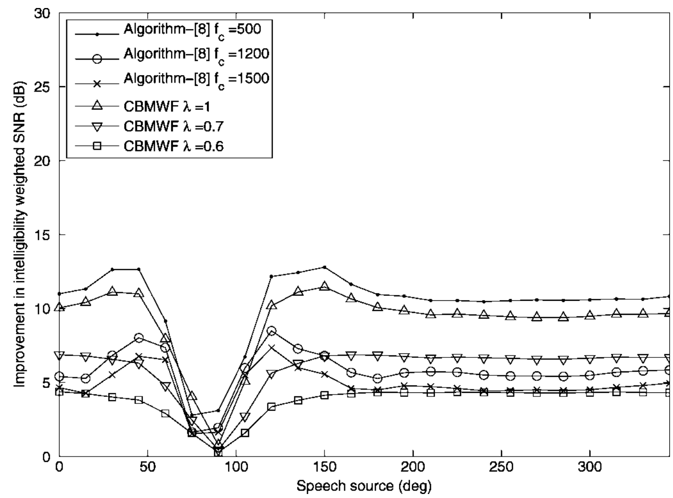

Fig. 4. Improvement in intelligibility weighted SNR for the left front microphone for speech sources between 0 and 345° with the noise source fixed at 90°.

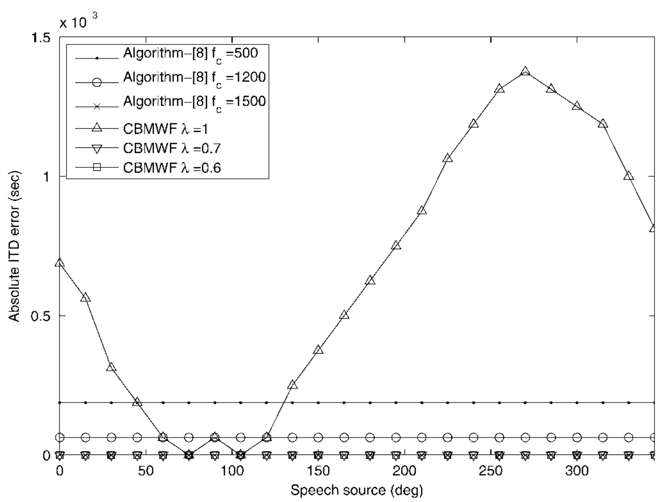

Fig. 3. ITD error noise component for speech sources between 0 and 345° with the noise source fixed at 90°.

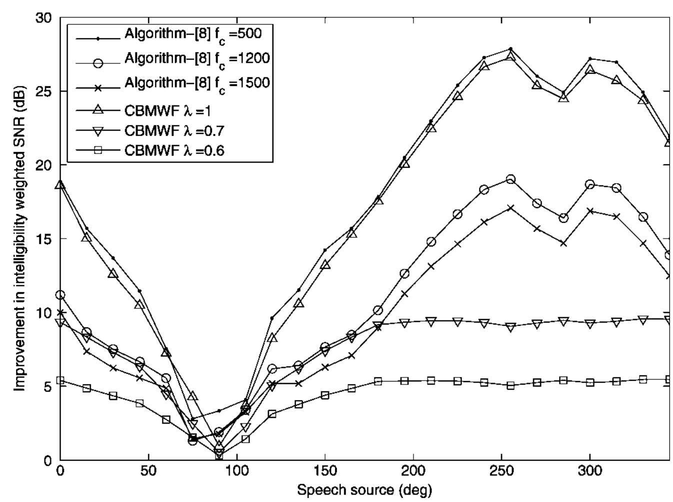

Fig. 5. Improvement in intelligibility weighted SNR for the right front microphone for speech sources between 0 and 345° with the noise source fixed at 90°.

C. Results

For the first set of simulations Figs. 2 and 3 show the absolute difference between the input ITD and the output ITD of the speech component and the noise component. The noise reduction performance of the algorithms can be seen in Figs. 4 and 5.

Looking closely at Fig. 2, we see that there is no speech ITD error (except for a few slight errors that probably arise from the ITD calculation) for the binaural multichannel Wiener filtering algorithm (Controlled binaural multichannel Wiener filtering with $\lambda = 1$). In other words, the speech ITD cues are consistently preserved. Naturally, the speech ITD cues are preserved for the controlled binaural multichannel Wiener filtering algorithm, since the processing of the speech component remains the same regardless of the value of λ.

Despite preserving the speech ITD cues, the processing carried out by the binaural multichannel Wiener filtering algorithm does affect the ITD cues of the noise component. The controlled binaural multichannel Wiener filtering algorithm is designed to combat that. From Fig. 3, it is clear that as λ decreases from 1 to 0.7 and again to 0.6, the error of the noise ITD also decreases. Unfortunately, this comes at a price. Looking at Figs. 4 and 5, we see that the noise reduction performance of the algorithm degrades as λ decreases.

On the other hand, algorithm-[8] must sacrifice noise reduction performance in order to preserve the speech ITD cues. With a cutoff frequency equal to 500 Hz, there are still large ITD errors for the speech component. In order to preserve the speech ITD cues the cutoff frequency must be increased to 1200 Hz and even on to 1500 Hz. Such a high cutoff frequency causes poor noise reduction performance. From Fig. 3, it can be seen that algorithm-[8] also distorts noise ITD cues. Naturally, increasing the cutoff frequency also leads to a decrease in ITD error of the noise component.

If the amount of improvement in $\mathrm{SNR_{INT}}$ needed so the hearing aid user's understanding in noise approaches that of an unaided normal hearing person is smaller than the amount of noise reduction performed by the controlled binaural multichannel Wiener filtering algorithm when $\lambda = 1$, then λ can be decreased. If λ can be sufficiently decreased, noise ITD cues will be preserved. However, if the user and acoustical situation call for a large SNR improvement noise ITD cues may be distorted, but speech ITD cues will always be preserved. On the other hand, for algorithm-[8], the user and the acoustical situation

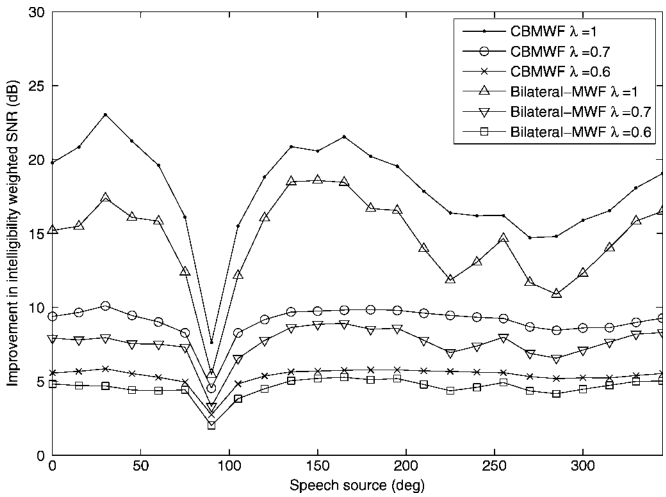

Fig. 6. Improvement in intelligibility weighted SNR for the left front microphone for speech sources between 0 and 345° with the noise source fixed at 90°.

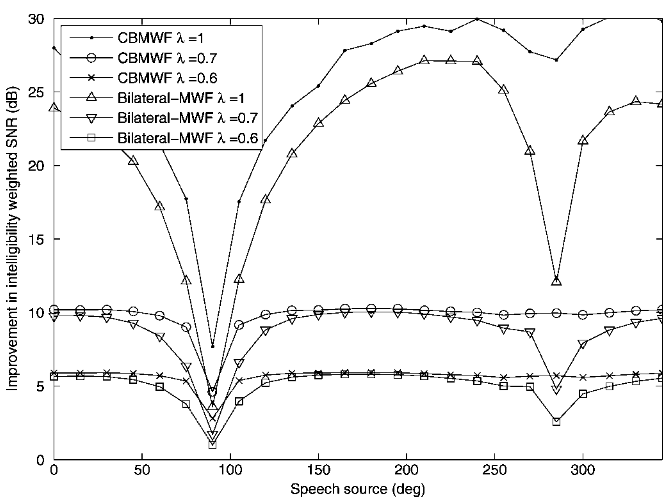

Fig. 7. Improvement in intelligibility weighted SNR for the right front microphone for speech sources between 0 and 345° with the noise source fixed at 90°.

may require an improvement in $\mathrm{SNR}_{\mathrm{INT}}$ that causes both speech and noise ITD cues to be distorted. Therefore, the binaural multichannel Wiener filtering and controlled multichannel Wiener filtering have an advantage over algorithm-[8].

For the second set of simulations, Figs. 6 and 7 show that the controlled binaural multichannel Wiener filtering algorithm outperforms the controlled bilateral multichannel Wiener filtering algorithm in terms of noise reduction. Moreover, the binaural algorithm outperforms the bilateral algorithm for all values of λ. This was expected since the binaural approach has access to all four microphones and the best ear. This particular binaural benefit is visible in Fig. 6 when the speech source is located at 270° and the noise source at 90°. In this extreme case we see a huge difference in noise reduction performance between the binaural and monaural multichannel Wiener filtering algorithms.

The ITD error of the speech component was zero for both the bilateral and binaural algorithms and all values of λ. Despite preserving the speech ITD cues, again the processing carried out by both algorithms with $\lambda = 1$ distorts the ITD cues of the noise component. These plots are not included owing to their similarity to Figs. 2 and 3. As λ was decreased from 1 to 0.7 and again to 0.6 the ITD error of the noise component decreased to zero for both the bilateral and binaural algorithms. Unfortunately, this comes at a price. Looking at Figs. 6 and 7, we see that the noise reduction performance of both algorithms degrades as λ decreases.

The results of this study are promising; however, we must note that these are initial results and there is much work to be done. First, the noise scenario was relatively simple. In addition, the signals were anechoic recordings and no head movements were present. In a real-time implementation of the controlled binaural multichannel Wiener filter, these head movements may need to be taken into account when adapting the filters. Furthermore, we assumed that we had access to a perfect VAD. In real life this is not the case. Clearly, the performance of both algorithm—[8] and controlled binaural multichannel Wiener filtering relies on VAD. A discussion of the influence of VAD errors on monaural multichannel Wiener filtering techniques can be found in [17]. Future work should look into the influence of VAD errors on the binaural extension. Moreover, we only explored the effect of the noise reduction techniques on ITD cues using an error metric. Further research must be conducted to explore the controlled binaural multichannel Wiener filtering technique perceptually, specifically its effect on speech reception thresholds and localization performance. This would also lead to further insight in the effect of the parameter λ.

VI. Conclusion

This correspondence presented a binaural extension of the monaural multichannel Wiener filter discussed in [9]–[11]. Simulations showed the advantage of a binaural algorithm to a bilateral one in terms of noise reduction performance. Binaural multichannel Wiener filtering algorithms preserve the speech ITD cues without sacrificing noise reduction performance. Conversely, algorithm-[8] sacrifices noise reduction performance in order to preserve the speech ITD cues. In order to preserve noise ITD cues some of the noise signal is passed to the output of the algorithm unprocessed. Correspondingly, some noise reduction performance is sacrificed. In the controlled binaural multichannel Wiener filtering algorithm, the parameter λ controls the amount of noise reduction performed by the algorithm; accordingly, the parameter λ also controls the distortion of the noise ITD cues. Similarly, as the cutoff frequency of algorithm-[8] increases, more speech and noise ITD cues arrive undistorted to the user. Therefore, noise reduction performance decreases as the cutoff frequency increases. Binaural multichannel Wiener filtering has a clear advantage over algorithm-[8] since it always preserves speech ITD cues.

References

[1] T. Van den Bogaert, T. Klasen, L. Van Deun, J. Wouters, and M. Moonen, "Horizontal localization with bilateral hearing aids: Without is better than with," *J. Acoust. Soc. Amer.*, vol. 119, no. 1, pp. 515–526, Jan. 2006.

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

[3] N. Erber, "Auditory-visual perception of speech," *J. Speech Hearing Dis.*, vol. 40, pp. 481–492, 1975.

[4] M. L. Hawley, R. Y. Litovsky, and J. F. Culling, "The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer," *J. Acoust. Soc. Amer.*, vol. 115, no. 2, pp. 833–843, Feb. 2004.

[5] A. Markides, *Binaural Hearing Aids*. London, U.K.: Academic, 1977.

[6] F. Wightman and D. Kistler, "The dominant role of low-frequency interaural time differences in sound localization," *J. Acoust. Soc. Amer.*, vol. 91, no. 3, pp. 1648–1661, Mar. 1992.

[7] W. Hartmann, "How we localize sound," *Phys. Today*, pp. 24–29, Nov. 1999.

[8] D. Welker, J. Greenburg, J. Desloge, and P. Zurek, "Microphone-array hearing aids with binaural output—Part II: A two-microphone adaptive system," *IEEE Trans. Speech Audio Process.*, vol. 5, no. 6, pp. 543–551, Nov. 1997.

[9] A. Spriet, M. Moonen, and J. Wouters, "Spatially pre-processed speech distortion weighted multi-channel Wiener filtering for noise reduction," *Signal Process.*, vol. 84, no. 12, pp. 2367–2387, Dec. 2004.

[10] S. Doclo and M. Moonen, "GSVD-based optimal filtering for single and multi-microphone speech enhancement," *IEEE Trans. Signal Process.*, vol. 50, no. 9, pp. 2230–2244, Sep. 2002.

[11] G. Rombouts and M. Moonen, "QRD-based unconstrained optimal filtering for acoustic noise reduction," *Signal Process.*, vol. 83, no. 9, Sep. 2003.

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

[13] J. Peissig and B. Kollmeier, "Directivity of binaural noise reduction in spatial multiple noise-source arrangements for normal and impaired listeners," *J. Acoust. Soc. Amer.*, vol. 101, no. 3, pp. 1660–1670, 1997.

[14] M. Nilsson, S. Soli, and J. Sullivan, "Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise," *J. Acoust. Soc. Amer.*, vol. 95, pp. 1085–1096, 1994.

[15] J. Greenberg, P. Peterson, and P. M. Zurek, "Intelligibility-weighted measures of speech-to-interference ratio and speech system performance," *J. Acoust. Soc. Amer.*, vol. 94, no. 5, pp. 3009–3010, Nov. 1993.

[16] *American National Standard Methods for Calculation of the Speech Intelligibility Index*, ANSI S3.5-1997, 1997, Acoustical Society of America.

[17] A. Spriet, M. Moonen, and J. Wouters, "The impact of speech detection errors on the noise reduction performance of multi-channel wiener filtering and generalized sidelobe cancellation," *Signal Process.*, vol. 85, no. 6, pp. 1073–1088, Jun. 2005.