MULTI-CHANNEL NOISE REDUCTION IN HEARING AIDS WITH WIRELESS ACCESS TO AN EXTERNAL REFERENCE SIGNAL

*Annelies Geusens**, *Alexander Bertrand**,†, *Bram Cornelis** *and Marc Moonen**,†

∗ KU Leuven, Dept. of Electrical Engineering-ESAT, SCD-SISTA \ † IBBT Future Health Department
Kasteelpark Arenberg 10, B-3001 Leuven, Belgium
E-mail: annelies.geusens@esat.kuleuven.be, alexander.bertrand@esat.kuleuven.be,
bram.cornelis@gmail.com, marc.moonen@esat.kuleuven.be

ABSTRACT

The standard scenario for multi-channel noise reduction in a hearing aid (HA) is a scenario with one desired speech signal in background noise. In this case, the on/off phases of the desired speech signal are detected and exploited to estimate the covariance matrices required in the noise reduction algorithms, namely the covariance matrix of the background noise and the covariance matrix of the desired speech signal. These matrices can then be used to construct a so-called Speech Distortion Weighted Multi-channel Wiener Filter (SDW-MWF). In this paper, we consider a more general scenario, where next to the desired speech signal there is also a second desired signal played by an audio device, such as, e.g., a TV set. We assume that this second source signal is transmitted to the HA over a wireless link. While the desired speech signal is an on/off signal, the second desired signal may be continuously active. This in particular disallows adopting the usual covariance matrix estimation procedures. It will be demonstrated how the external reference signal, together with the on/off phases of the desired speech signal, can be exploited to estimate the required covariance matrices to compute the SDW-MWF. This is done by decomposing the general SDW-MWF into two subproblems; a single-channel least squares (LS) filtering and a rank-1 SDW-MWF (R1-MWF). We will provide simulations that compare two different implementations of this decomposition.

Index Terms— multi-channel noise reduction, speech enhancement, wireless hearing aids

1. INTRODUCTION

State-of-the-art HA's are capable to receive and/or transmit audio signals over a wireless channel [1]. This allows the HA to communicate with external audio devices (such as a TV set) and receive the devices clean playback signal. We will refer to the latter as the 'external reference signal'. Different from standard HA scenarios with a single desired speech signal, we assume here that this playback signal is also a desired signal, in addition to a desired speech signal (e.g., from a nearby speaker). While the desired speech signal is an on/off signal, the second desired signal may be continuously active. The purpose of this paper is to investigate several approaches to reconstruct or synthesize the desired signals as they impinge on the HA of the listener, while reducing undesired background noise. A possible scenario is depicted in Fig. 1, showing a HA user listening to a desired speaker and a TV at the same time, while there is additional background noise from other directions.

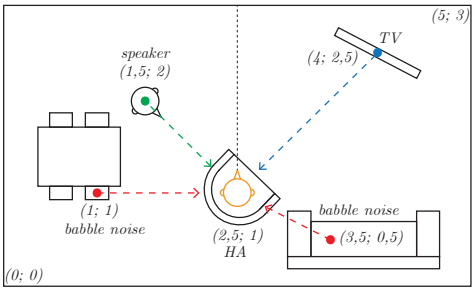


Fig. 1: Overview of a possible scenario. The numbers in brackets denote the coordinates of the elements (in m).

A commonly used algorithm for noise reduction in microphone arrays is the Speech Distortion Weighted Multi-channel Wiener Filter (SDW-MWF) [2]. An obvious approach could be to estimate both desired sources by using an SDW-MWF where the input channels consist of both the external reference signal and the local microphone signals. However, the SDW-MWF cannot be computed in practice, since one of the desired signals is assumed to be non-speech. Therefore, the input channels may not contain sufficient silent signal segments to be able to estimate the background noise statistics as required in the SDW-MWF. To resolve this, two alternative approaches are suggested that split the problem into a linear adaptive filtering problem and a rank-1 SDW-MWF (R1-MWF) problem [3]. It is shown that these alternatives are theoretically equivalent to the general SDW-MWF as mentioned earlier. However, we will explain that both alternatives have their own practical advantages and disadvantages. We use a simulated HA scenario to accurately demonstrate the performance of both approaches. We also provide audio files (online) with practical recordings for the same HA scenario.

2. PROBLEM STATEMENT

The scenario that is investigated in this paper consists of two desired sources u_{ad} (an audio device) and u_{sp} (a speaker) that propagate through an acoustic path and impinge on the local microphones on a HA. The signal u_{ad} is available to the HA through a wireless chan-

Acknowledgements: The work of A. Bertrand was supported by a Postdoctoral Fellowship of the Research Foundation - Flanders (FWO). This work was carried out at the ESAT Laboratory of KU Leuven, in the frame of KU Leuven Research Council CoE EF/05/006 'Optimization in Engineering' (OPTEC) and PFV/10/002 (OPTEC), Concerted Research Action GOA-MaNet, 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), Research Project IBBT, and Research Project FWO nr. G.0763.12 (Wireless acoustic sensor networks for extended auditory communication). The scientific responsibility is assumed by its authors.

nel, and we assume that its sampling rate is perfectly synchronized with the sampling rate of the microphones. This external reference signal can be treated as a extra (virtual) microphone. The transfer functions (TF) from the sources to the microphones yield the following two steering vectors (in the frequency domain).

$$\mathbf{a}_{ad}(\omega) = \begin{bmatrix} a_{ad,1}(\omega) & \dots & a_{ad,M}(\omega) & 1 \end{bmatrix}^T = \begin{bmatrix} \tilde{\mathbf{a}}_{ad}(\omega) & 1 \end{bmatrix} \tag{1}$$

$$\mathbf{a}_{sp}(\omega) = \begin{bmatrix} a_{sp,1}(\omega) & \dots & a_{sp,M}(\omega) & 0 \end{bmatrix}^T = \begin{bmatrix} \tilde{\mathbf{a}}_{sp}(\omega) & 0 \end{bmatrix} \tag{2}$$

with M the number of local microphones, $a_{ad,i}$ the TF from the audio device to microphone i and $a_{sp,i}$ the TF from the speaker to microphone i and where ω denotes the frequency-domain variable, which will be omitted in the sequel for the sake of conciseness.

The microphone signals are stacked in the vector $\mathbf{y}$ and can be written as

$$\mathbf{y} = \begin{bmatrix} y_1 & \dots & y_M & u_{ad} \end{bmatrix}^T \tag{3}$$

$$= \mathbf{s} + \mathbf{n} = \mathbf{a}_{ad}u_{ad} + \mathbf{a}_{sp}u_{sp} + \begin{bmatrix} \tilde{\mathbf{n}} & 0 \end{bmatrix}^T \tag{4}$$

with y_i the i-th microphone signal, $\mathbf{s}$ a vector containing the desired components and $\mathbf{n}$ a vector containing the noise components. Because the external reference signal does not contain noise, the last element of $\mathbf{n}$ is zero.

The goal of this paper is to reconstruct or synthesize the filtered versions of the two source signals as they impinge on a reference microphones at the HA. Hence, the desired output is $s_i = a_{ad,i}u_{ad} + a_{sp,i}u_{sp}$ where i refers to the reference microphone. Without loss of generality we will use the first microphone as reference. The two desired signals will be called the 'audio signal' and the 'speech signal' in the sequel of the paper.

3. MWF-BASED NOISE REDUCTION

Since there are two desired signals that both impinge on a microphone array (the local microphones), a straightforward approach could be to perform a multi-channel filtering which extracts these two signals. A suitable filter is the *Multi-channel Wiener Filter* (MWF), which minimizes the difference between the filter output and the desired output (which is the desired signal in the reference microphone):

$$\mathbf{w}_{MWF} = \arg\min_{\mathbf{w}} E\{|s_1 - \mathbf{w}^H\mathbf{y}|^2\} \tag{5}$$

where $E\{\cdot\}$ denotes the expected value operator, and superscript H denotes the conjugate transpose. The closed-form solution of this optimization problem is

$$\mathbf{w}_{MWF} = \mathbf{R}_{yy}^{-1}\mathbf{R}_{ss}\mathbf{e}_1 = (\mathbf{R}_{ss} + \mathbf{R}_{nn})^{-1}\mathbf{R}_{ss}\mathbf{e}_1 \tag{6}$$

with $\mathbf{R}_{yy} = E\{\mathbf{y}\mathbf{y}^H\}$, $\mathbf{R}_{ss} = E\{\mathbf{s}\mathbf{s}^H\}$ and $\mathbf{R}_{nn} = E\{\mathbf{n}\mathbf{n}^H\}$ the correlation matrices of the microphone signals $\mathbf{y}$, the desired signal components $\mathbf{s}$ and the noise components $\mathbf{n}$, respectively and $\mathbf{e}_i = [0 \dots 0\ 1\ 0 \dots 0]^T$ where the 1 is the i-th entry.

The SDW-MWF is an extension of the MWF which adds a parameter μ to the solution:

$$\mathbf{w}_{SDW-MWF} = (\mathbf{R}_{ss} + \mu\mathbf{R}_{nn})^{-1}\mathbf{R}_{ss}\mathbf{e}_1 \ . \tag{7}$$

This parameter allows to make a trade-off between noise-reduction (high value for μ) and low signal distortion (low value for μ).

In practice, the correlation matrices are not known and have to be estimated. For $\mathbf{R}_{yy}$, this can be straightforwardly done based on temporal averaging. The noise correlation matrix $\mathbf{R}_{nn}$ is usually estimated during noise-only periods, detected by a *voice activity detector* (VAD). The matrix $\mathbf{R}_{ss}$ is then obtained by subtracting the microphone and noise correlation matrix: $\mathbf{R}_{ss} = \mathbf{R}_{yy} - \mathbf{R}_{nn}$ (this only holds if the noise and the desired signal are uncorrelated).

The above method introduces a problem in the scenario envisaged here. If the audio signal u_{ad} is non-speech and continuously active (which is a realistic assumption for a television or a radio), there are no *noise-only* periods and so $\mathbf{R}_{nn}$ and hence $\mathbf{R}_{ss}$ cannot be estimated. Therefore two alternative schemes are introduced in the next section, where both desired signals are estimated separately instead of a joint estimation.

4. DECOMPOSITION INTO AN LS FILTER AND AN R1-MWF

In this section we prove that the SDW-MWF can be split in two parts that estimate both desired signals separately. Before starting the derivations, some preliminary expressions are deduced. Then the SDW-MWF filter is rewritten in two different ways.

4.1. Preliminary expressions

The correlation matrix of the desired components can be written as

$$\mathbf{R}_{ss} = P_{ad}\mathbf{a}_{ad}\mathbf{a}_{ad}^H + P_{sp}\mathbf{a}_{sp}\mathbf{a}_{sp}^H \tag{8}$$

with $P_{ad} = E\{|u_{ad}|^2\}$ and $P_{sp} = E\{|u_{sp}|^2\}$. The SDW-MWF filter $\mathbf{w}$ can then be rewritten as

$$\mathbf{w} = (P_{sp}\mathbf{a}_{sp}\mathbf{a}_{sp}^H + \mathbf{A})^{-1}(P_{ad}\mathbf{a}_{ad}\mathbf{a}_{ad}^H + P_{sp}\mathbf{a}_{sp}\mathbf{a}_{sp}^H)\mathbf{e}_1 \tag{9}$$

where the matrix $\mathbf{A}$ is defined as

$$\mathbf{A} = P_{ad}\mathbf{a}_{ad}\mathbf{a}_{ad}^H + \mu\mathbf{R}_{nn} \ . \tag{10}$$

The matrix $\mathbf{A}$ can be rewritten as

$$\mathbf{A} = \left(P_{ad} \begin{bmatrix} \tilde{\mathbf{a}}_{ad} \\ 1 \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{a}}_{ad}^H & 1 \end{bmatrix} + \left[\begin{array}{c|c} \mu\tilde{\mathbf{R}}_{nn} & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline \begin{matrix} 0 & \dots & 0 \end{matrix} & 0 \end{array}\right] \right) \tag{11}$$

$$= \left[\begin{array}{c|c} \mu\tilde{\mathbf{R}}_{nn} + \tilde{\mathbf{a}}_{ad}\tilde{\mathbf{a}}_{ad}^H P_{ad} & \tilde{\mathbf{a}}_{ad}P_{ad} \\ \hline \tilde{\mathbf{a}}_{ad}^H P_{ad} & P_{ad} \end{array}\right] \tag{12}$$

where $\tilde{\mathbf{R}}_{nn} = E\{\tilde{\mathbf{n}}\tilde{\mathbf{n}}^H\}$.

With the Woodbury identity [4] the inverse of $\mathbf{A}$ can be computed from (12) as

$$\mathbf{A}^{-1} = \left[\begin{array}{c|c} \frac{1}{\mu}\tilde{\mathbf{R}}_{nn}^{-1} & -\frac{1}{\mu}\tilde{\mathbf{R}}_{nn}^{-1}\tilde{\mathbf{a}}_{ad} \\ \hline -\frac{1}{\mu}\tilde{\mathbf{a}}_{ad}^H\tilde{\mathbf{R}}_{nn}^{-1} & \frac{1}{P_{ad}} + \frac{1}{\mu}\tilde{\mathbf{a}}_{ad}^H\tilde{\mathbf{R}}_{nn}^{-1}\tilde{\mathbf{a}}_{ad} \end{array}\right] \tag{13}$$

and

$$(P_{sp}\mathbf{a}_{sp}\mathbf{a}_{sp}^H + \mathbf{A})^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{a}_{sp}\mathbf{a}_{sp}^H\mathbf{A}^{-1}}{\frac{1}{P_{sp}} + \mathbf{a}_{sp}^H\mathbf{A}^{-1}\mathbf{a}_{sp}} \, . \tag{14}$$

Before continuing, we first note that

$$\mathbf{A}^{-1}\mathbf{a}_{ad} = \begin{bmatrix} 0 & \dots & 0 \end{bmatrix} \left| \begin{matrix} \frac{1}{P_{ad}} \end{matrix} \right. \Big]^T \tag{15}$$

and

$$\mathbf{A}^{-1}\mathbf{a}_{sp} = \begin{bmatrix} \frac{1}{\mu}\tilde{\mathbf{R}}_{nn}^{-1}\tilde{\mathbf{a}}_{sp} \\ \hline -\frac{1}{\mu}\tilde{\mathbf{a}}_{ad}^{H}\tilde{\mathbf{R}}_{nn}^{-1}\tilde{\mathbf{a}}_{sp} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_M \\ \hline -\tilde{\mathbf{a}}_{ad}^{H} \end{bmatrix} \frac{1}{\mu}\tilde{\mathbf{R}}_{nn}^{-1}\tilde{\mathbf{a}}_{sp} \,. \tag{16}$$

with $\mathbf{I}_M$ the identity matrix of dimension M. Accordingly

$$\mathbf{a}_{sp}^{H}\mathbf{A}^{-1}\mathbf{a}_{ad} = 0 \,, \tag{17}$$

$$\mathbf{a}_{sp}^{H}\mathbf{A}^{-1}\mathbf{a}_{sp} = \frac{1}{\mu}\tilde{\mathbf{a}}_{sp}^{H}\tilde{\mathbf{R}}_{nn}^{-1}\tilde{\mathbf{a}}_{sp} \,. \tag{18}$$

By combining (14), (15) and (17), we obtain

$$(P_{sp}\mathbf{a}_{sp}\mathbf{a}_{sp}^{H} + \mathbf{A})^{-1}\mathbf{a}_{ad} = [\ 0 \quad \dots \quad 0 \ \big|\ \tfrac{1}{P_{ad}}\ ]^{T} \,. \tag{19}$$

Furthermore, by combining (14), (18) and (16), we obtain

$$(P_{sp}\mathbf{a}_{sp}\mathbf{a}_{sp}^{H} + \mathbf{A})^{-1}\mathbf{a}_{sp} = \frac{\mathbf{A}^{-1}\mathbf{a}_{sp}}{1 + \mathbf{a}_{sp}^{H}\mathbf{A}^{-1}\mathbf{a}_{sp}P_{sp}} \tag{20}$$

$$= \begin{bmatrix} \mathbf{I}_M \\ \hline -\tilde{\mathbf{a}}_{ad}^{H} \end{bmatrix} \frac{\tilde{\mathbf{R}}_{nn}^{-1}\tilde{\mathbf{a}}_{sp}}{\mu + \tilde{\mathbf{a}}_{sp}^{H}\tilde{\mathbf{R}}_{nn}^{-1}\tilde{\mathbf{a}}_{sp}P_{sp}} \,. \tag{21}$$

4.2. First alternative: R1-MWF + single channel LS filter

If we further investigate equation (9) we get

$$\mathbf{w} = (P_{sp}\mathbf{a}_{sp}\mathbf{a}_{sp}^{H} + \mathbf{A})^{-1}P_{ad}\mathbf{a}_{ad}a_{ad,1}^{*} \tag{22}$$

$$+ (P_{sp}\mathbf{a}_{sp}\mathbf{a}_{sp}^{H} + \mathbf{A})^{-1}P_{sp}\mathbf{a}_{sp}a_{sp,1}^{*} \tag{23}$$

with x^* the complex conjugate of x. With (19) and (20) this becomes

$$\mathbf{w} = [\ 0 \quad \dots \quad 0 \ \big|\ a_{ad,1}^{*}\ ] + \frac{\mathbf{A}^{-1}\mathbf{a}_{sp}P_{sp}a_{sp,1}^{*}}{\mu + \mathbf{a}_{sp}^{H}\mathbf{A}^{-1}\mathbf{a}_{sp}P_{sp}} \,. \tag{24}$$

This expression consists of two terms that estimate the two desired signals as they impinge on the reference microphone. The first term uses the TF $a_{ad,1}$ from the external reference signal to the reference microphone, which can be estimated using a 1-tap adaptive LS filter.

The second term in (24) estimates the speech source using a $(M+1)$-channel noise reduction filter. We recognize the formula for R1-MWF [3]. This is a special case of SDW-MWF with only one desired localized source, yielding a rank-1 speech correlation matrix. In this expression the matrix $\mathbf{A}$ serves as the noise correlation matrix. However, from formula (10) we know that this matrix is the correlation matrix of the audio signal and the original noise together. This means the audio signal is also treated as noise for the estimation of the speech source in this scheme. A block diagram of formula (24) is presented in Fig. 2, which we refer to as scheme A.

The above result is not unexpected. It is already mentioned that the approach to estimate both signals at the same time with SDW-MWF is not possible due to the absence of noise-only periods. A logical solution to counter this would be to estimate both signals separately, which is exactly what is done in this scheme. For the audio signal we only need to estimate the TF $a_{ad,1}$, because the clean signal is already available through the wireless link. For the speech source, the clean signal is not available. Therefore a suitable solution would be to estimate it using all microphone signals.

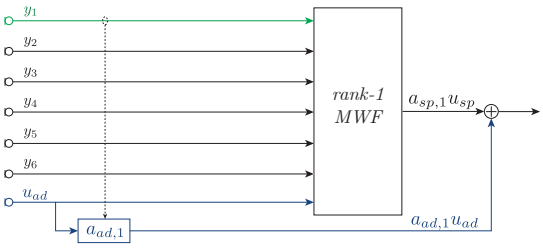


Fig. 2: Scheme A

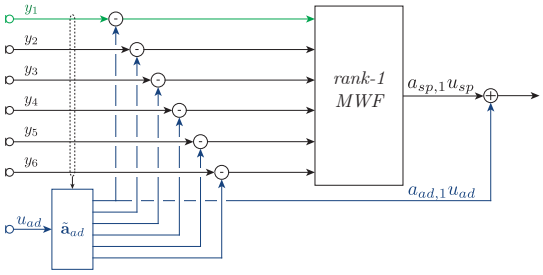


Fig. 3: Scheme B

The external reference signal does not contain any information about the speech source. However, it is useful to include it in the filtering, because it contains information about a noise source (recall that the audio signal is treated as noise for the speech source estimation). The filter can use this information to successfully suppress this noise source.

The absence of noise-only periods in the original approach does not hold anymore in this scheme since the signal u_{ad} is assumed to be noise in the R1-MWF. This approach also has a second advantage. The multi-channel filter which was a general SDW-MWF in the original scheme, is now transformed to a rank-1 SDW-MWF, which is numerically more robust when computed with the R1-MWF formula [5].

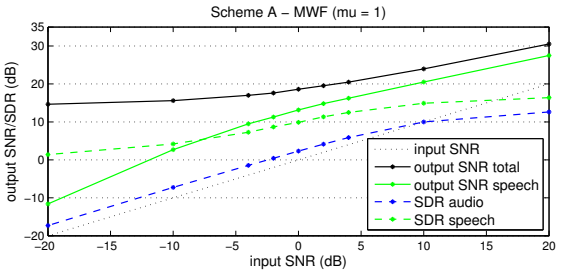
4.3. Second alternative: R1-MWF with multi-channel LS-filter

Formula (23) can also be rewritten using formula (19) and (21):

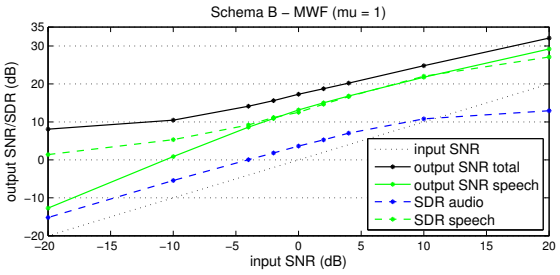
$$\mathbf{w} = [\ 0 \quad \dots \quad 0 \ \big|\ a_{ad,1}^{*}\ ]^{T} \tag{25}$$

$$+ \begin{bmatrix} \mathbf{I}_M \\ \hline -\tilde{\mathbf{a}}_{ad}^{H} \end{bmatrix} \frac{\tilde{\mathbf{R}}_{nn}^{-1}\tilde{\mathbf{a}}_{sp}P_{sp}a_{sp,1}^{*}}{\mu + \tilde{\mathbf{a}}_{sp}^{H}\tilde{\mathbf{R}}_{nn}^{-1}\tilde{\mathbf{a}}_{sp}P_{sp}} \,. \tag{26}$$

The first term is the same as in (24). The second term however is different, and it consists of two parts. First the component of the audio signal in every microphone is subtracted from the respective microphone signal by using differently filtered signals of the external reference signal , based on the TF's in $\tilde{\mathbf{a}}_{ad}$. This means that the remaining M signals will consist of the speech source and the background noise. In the second part of the term, we recognize again the formula for R1-MWF, where $\tilde{\mathbf{R}}_{nn}^{-1}$ serves as the noise correlation matrix. This time the noise does not contain the audio signal because it is already subtracted from the signals. This approach is summarized in Fig. 3, which we refer to as scheme B. In theory both schemes are equivalent, but every scheme has some practical advantages. On the one hand, scheme A is computationally cheaper than scheme B; the first step (LS filtering) has to be performed on



(a) Scheme A



(b) Scheme B

Fig. 4: Comparison of the performance of the two schemes. These simulations are performed in batch mode.

only one channel. On the other hand, the noise correlation matrix $\tilde{\mathbf{R}}_{nn}$ in scheme B is better conditioned than the matrix $\mathbf{A}$, i.e. the noise correlation matrix in scheme A. Indeed, the matrix $\mathbf{A}$ may be dominated by the audio signal, and as a result the noise correlation matrix will be closer to a rank-1 matrix, having a larger eigenvalue spread.

5. SIMULATIONS

Some test signals were generated using a RIR (room impulse response) generator [6] based on the scenario in Fig. 1 with a T60 reverberation time of 0,3 seconds. The speech signal u_{sp} was taken from the HINT database [7], and the TV-signal u_{ad} consists of a music track[1] (also containing speech from the singer). We consider a scenario with binaural HA's, each having 3 microphones. We assume that both HA's have access to each other's microphone signals (these may be exchanged through a wireless link). The 6 microphone signals where generated using head related transfer functions (HRTF's) from the MIT database [8].

The two schemes were tested in batch mode using a perfect VAD to isolate VAD-errors in the assessment[2]. Figure 4.a and 4.b show the performance as a function of the input SNR for scheme A and scheme B, respectively. The *signal-to-noise ratio* (SNR) and *signal-to-distortion ratio* (SDR) are calculated as follows:

$x_1 =$ audio signal at output of the R1-MWF

$x_2 =$ speech source at output of the R1-MWF: $\mathbf{w}^H \mathbf{a}_{sp} u_{sp}$

$x_n =$ noise component at output of the R1-MWF: $\mathbf{w}^H \mathbf{n}$

$f_{est} =$ output of LS-filter

$$\text{total output SNR} = 10 \log_{10} \left(\frac{E\{||x_1 + x_2 + f_{est}||^2\}}{E\{||x_n||^2\}} \right)$$

$$\text{speech output SNR} = 10 \log_{10} \left(\frac{E\{||x_2||^2\}}{E\{||x_n||^2\}} \right)$$

$$\text{audio SDR} = 10 \log_{10} \left(\frac{E\{||a_1 u_{ad}||^2\}}{E\{||a_1 u_{ad} - f_{est} - x_1||^2\}} \right)$$

$$\text{speech SDR} = 10 \log_{10} \left(\frac{E\{||b_1 u_{sp}||^2\}}{E\{||b_1 u_{sp} - x_2||^2\}} \right)$$

At low input SNR (-20 dB) both schemes have more or less the same SNR and SDR performance for the speech signal. However, when increasing the input SNR to more realistic levels, the SDR curve and SNR curve (of the speech) are a bit higher in scheme B compared to scheme A. The higher the input SNR, the clearer this is visible. This is mainly because of the better conditioning of the noise correlation matrix in scheme B, and this effect increases together with input SNR since the audio device then dominates more over the background noise. The estimated performance for the audio device also differs significantly between both schemes. For low input SNRs the SDR curve of the TV-signal differs about 5 dB. This is due to the higher audio-residue in scheme A at the output of the R1-MWF ($x1$) and which contributes to the distortion. For low input SNRs the total SNR is higher in scheme A. This is also due to the audio-residue at the output of the R1-MWF (because it is in the numerator of the SNR). For higher input SNRs the total output SNR of scheme B is better, because of the better speaker SNR.

[1] *Them there eyes* of *Billie Holiday*

[2] Audio examples of an adaptive implementation of our approach (including a practical energy-based VAD), applied to real-life (i.e., unsimulated) recordings are available on line at ftp://ftp.esat.kuleuven.be/pub/SISTA/abertran/papers_website/IWAENC2012.html

6. CONCLUSION

In this paper we have investigated a particular scenario for a HA noise reduction with two desired sources (an audio device and speaker). Simple extraction of both sources with SDW-MWF was not possible when the audio signal is continuously active. It was shown in this paper that the ordinary SDW-MWF can be decomposed into an LS filter and an R1-MWF. Two (theoretically the same) approaches were proposed and compared. Both schemes have their own advantages and disadvantages with respect to each other; there is a trade-off between a better conditioned problem and less computational complexity. We have compared both approaches in a simulated scenario.

7. REFERENCES

[1] Harvey Dillon, *Hearing aids*, Boomerang Press, Thieme, Sydney, 2001.

[2] M. Moonen and S. Doclo, "On the output SNR of the speech-distortion weighted multichannel Wiener filter," *IEEE Signal Processing Letters*, vol. 12, no. 12, Dec. 2005.

[3] M. Souden, J. Benesty, and S. Affes, "On optimal frequency-domain multichannel linear filtering for noise reduction," *IEEE Transactions on audio, speech, and language processing*, vol. 18, no. 2, Feb. 2010.

[4] G.H. Golub and C.F. Van Loan, *Matrix Computations*, Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, 1996.

[5] B. Cornelis, M. Moonen, and J. Wouters, "Performance analysis of multichannel Wiener filter-based noise reduction in hearing aids under second order statistics estimation errors," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 5, pp. 1368 –1381, july 2011.

[6] J.B. Allen and D.A. Berkley, "Image method for efficiently simulating small-room acoustics," *The Journal of the Acoustical Society of America*, vol. 65, no. 4, Feb. 1979.

[7] M. Nilsson, S. D. Soli, and J. A. Sullivan, "Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise," *The Journal of the Acoustical Society of America*, vol. 95, no. 2, 1994.

[8] Bill Gardner and Keith Martin, "HRTF measurements of a kemar dummy-head microphone," 1994.