Rate-constrained binaural MWF-based noise reduction algorithms

Simon Doclo, Toby Christian Lawin-Ore, Thomas Rohdenburg

University of Oldenburg - SIGPROC, 26111 Oldenburg
E-Mail: simon.doclo@uni-oldenburg.de
Web: www.sigproc.uni-oldenburg.de

Fraunhofer IDMT - HSA, 26129 Oldenburg
E-Mail: rog@idmt.fraunhofer.de
Web: www.idmt.fraunhofer.de

Abstract

In a binaural hearing aid system, output signals for the left and the right ear are generated by exchanging information between the two hearing aids. A significant noise reduction can be achieved using the binaural multi-channel Wiener filter (MWF), which requires all microphone signals from both hearing aids. However, the limited bandwidth of the binaural link typically does not allow to transmit all microphone signals between the two hearing aids. Recently, an iterative distributed MWF-based scheme has been presented, where each hearing aid only transmits a filtered combination of its microphone signals. In this paper, the performance gain of the binaural MWF and the iterative distributed MWF scheme are analysed as a function of the available bandwidth of the binaural link.

1 Introduction

Noise reduction algorithms in hearing aids are crucial for hearing impaired persons to improve speech intelligibility in noisy environments. In a binaural hearing aid system, output signals for both ears are generated, either by operating both hearing aids independently (a bilateral system) or by exchanging information between the hearing aids (a binaural system), e.g. using a wireless link [1]-[5].

In [1, 2] an overview and a theoretical performance analysis has been presented for several MWF-based noise reduction algorithms, more in particular the binaural MWF and its extensions using partial noise estimation and incorporating interaural transfer functions. In these algorithms all microphone signals need to be transmitted between the hearing aids, requiring a large bandwidth. In [3] an iterative distributed version of the binaural MWF has been presented, where each hearing aid only transmits a filtered combination of its microphone signals and which converges to the binaural MWF in the case of a single desired source. Recently, this distributed MWF scheme has also been extended towards multiple nodes [6].

For a binaural link with limited capacity, a theoretically optimal (in an information-theoretic sense) transmission scheme has been presented in [4], however requiring knowledge about the joint statistics of the signals at both hearing aids, which is typically not available in practice. In [5] the relation between performance gain and link capacity has been studied for several suboptimal (but practical) schemes, such as transmitting one microphone signal or transmitting an estimate of the desired signal obtained at the transmitting device. In this paper, a similar relation between performance gain and link capacity will be studied for the iterative distributed MWF scheme proposed in [3].

2 Configuration and notation

Consider the binaural hearing aid configuration depicted in Figure 1, where both hearing aids have a microphone array consisting of M microphones. The mth microphone signal in the left hearing aid $Y_{0,m}(\omega)$ can be written as

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

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 is $Y_{1,m}(\omega) = X_{1,m}(\omega) + V_{1,m}(\omega)$. For conciseness we will omit the frequency-domain variable ω in the remainder of the paper. We define the M-dimensional stacked vectors $\mathbf{Y}_0$ and $\mathbf{Y}_1$ and the $2M$-dimensional signal vector $\mathbf{Y}$ as

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

The signal vector can hence be written as $\mathbf{Y} = \mathbf{X} + \mathbf{V}$. In the case of a single speech source, the speech signal vector can be written as $\mathbf{X} = \mathbf{A}S$, where the steering vector $\mathbf{A}$ containing the acoustic transfer functions between the speech source and the microphones is defined similarly as $\mathbf{Y}$ and S denotes the speech signal.

The signals transmitted from the left (right) hearing aid to the right (left) hearing aid are represented respectively by the N-dimensional vectors $\mathbf{Y}_{10}$ and $\mathbf{Y}_{01}$, typically with $N \leq M$, which are assumed to be linear combinations of the contralateral microphone signals, i.e.

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

with $\mathbf{F}_{10}$ and $\mathbf{F}_{01}$ $M \times N$-dimensional complex matrices.

The output signals Z_0 and Z_1 for the left and the right hearing aid are obtained by filtering and summing the ipsilateral microphone signals and the transmitted signal(s) from the contralateral ear. Hence, the output signals can always be written as a linear combination of all microphone signals, i.e. $Z_0 = \mathbf{W}_0^H \mathbf{Y}$ and $Z_1 = \mathbf{W}_1^H \mathbf{Y}$, with

$$\mathbf{W}_0 = \begin{bmatrix} \mathbf{W}_{00} \\ \mathbf{W}_{01} \end{bmatrix} = \begin{bmatrix} \mathbf{W}_{00} \\ \mathbf{F}_{01}\mathbf{G}_{01} \end{bmatrix}, \mathbf{W}_1 = \begin{bmatrix} \mathbf{W}_{10} \\ \mathbf{W}_{11} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_{10}\mathbf{G}_{10} \\ \mathbf{W}_{11} \end{bmatrix}.$$

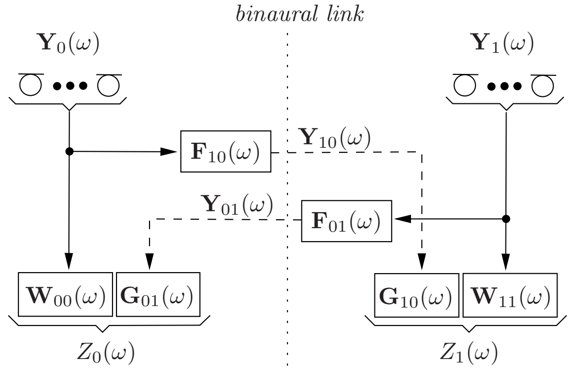

Figure 1: General binaural processing scheme

ISBN 978-3-8007-3300-2

3 Binaural MWF

The binaural MWF (B-MWF) requires all microphone signals to be transmitted, i.e. $\mathbf{F}_{10} = \mathbf{F}_{01} = \mathbf{I}_M$. The binaural MWF produces an MMSE (minimum-mean-square-error) estimate of the speech component in both hearing aids [1, 2]. In order to provide an additional trade-off between speech distortion and noise reduction, the speech distortion weighted multi-channel Wiener filter (SDW-MWF) minimises the weighted sum of the residual noise energy and the speech distortion energy, resulting in

$$\mathbf{W}_0^m = (\mathbf{\Phi}_x + \mu \mathbf{\Phi}_v)^{-1} \mathbf{\Phi}_x \mathbf{e}_0, \quad \mathbf{W}_1^m = (\mathbf{\Phi}_x + \mu \mathbf{\Phi}_v)^{-1} \mathbf{\Phi}_x \mathbf{e}_1 ,$$

where μ is a trade-off parameter, $\mathbf{\Phi}_x$ and $\mathbf{\Phi}_v$ are the speech and the noise correlation matrix, i.e. $\mathbf{\Phi}_x = \mathscr{E}\{\mathbf{X}\mathbf{X}^H\}$ and $\mathbf{\Phi}_v = \mathscr{E}\{\mathbf{V}\mathbf{V}^H\}$, and $\mathbf{e}_0$ and $\mathbf{e}_1$ are vectors of which only one element is equal to 1 and the other elements are equal to 0, with $\mathbf{e}_0(1) = 1$ and $\mathbf{e}_1(M+1) = 1$. In the case of a single speech source, it can be shown that [1, 2]

$$\mathbf{W}_1^m = \alpha \mathbf{W}_0^m , \qquad (4)$$

where $\alpha = A_{1,0}^* / A_{0,0}^*$ is the complex conjugate of the interaural transfer function of the speech component.

4 Transmission of single signal

The binaural MWF exploits all microphone signals, requiring M contralateral signals to be transmitted. However, due to power limitations, the capacity of the link typically does not allow to transmit all microphone signals. In [3, 5] different MWF-based schemes have been proposed where only one contralateral signal is transmitted, e.g. the front microphone signal, the monaural MWF output signal, or an iterative distributed scheme that converges to the optimal B-MWF solution in the case of a single speech source.

4.1 Front microphone signal (MWF-front)

In this scheme, only the front contralateral microphone signal is transmitted, i.e. $\mathbf{F}_{10} = \mathbf{F}_{01} = [\ 1 \quad 0 \quad \dots \quad 0\ ]^T$.

4.2 Contralateral MWF (MWF-contra)

In this scheme, the transmitted signal is the output of a monaural MWF, estimating the contralateral speech component using the M contralateral microphone signals. However, it has been shown in [3] that in general this solution is suboptimal, since the optimal solution is only obtained in the case of uncorrelated noise components.

4.3 Distributed MWF scheme (DB-MWF)

The iterative distributed MWF scheme is depicted in Figure 2. Basically, in each iteration the filter $\mathbf{F}_{10}$ is equal to $\mathbf{W}_{00}$ from the previous iteration, and the filter $\mathbf{F}_{01}$ is equal to $\mathbf{W}_{11}$ from the previous iteration. If we denote the filters and the signals in the ith iteration with superscript i, then the iterative procedure runs as follows:

1. Transmit $Y_{01}^i = \mathbf{W}_{11}^{i,H} \mathbf{Y}_1$ to the left hearing aid.
2. Using $\mathbf{Y}_0$ and Y_{01}^i as input signals, calculate $\mathbf{W}_{00}^i$ and G_{01}^i as the MWF solution estimating the speech component in the left front microphone.
3. Transmit $Y_{10}^i = \mathbf{W}_{00}^{i,H} \mathbf{Y}_0$ to the right hearing aid.

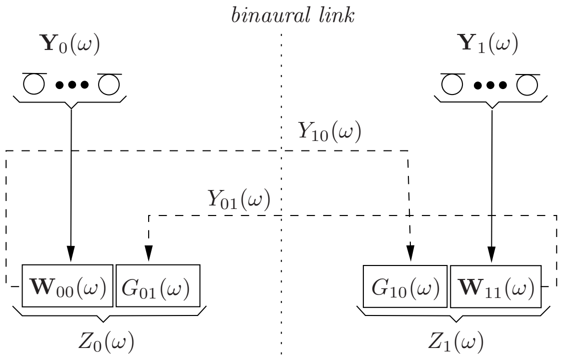

Figure 2: Distributed binaural MWF scheme (DB-MWF)

4. Using $\mathbf{Y}_1$ and Y_{10}^i as input signals, calculate $\mathbf{W}_{11}^{i+1}$ and G_{10}^{i+1} as the MWF solution estimating the speech component in the right front microphone.

It has been shown in [3] that at convergence, i.e. for $i \to \infty$,

$$\begin{bmatrix} \mathbf{W}_{10}^\infty \\ \mathbf{W}_{11}^\infty \end{bmatrix} = \begin{bmatrix} G_{10}^\infty \mathbf{W}_{00}^\infty \\ 1/G_{01}^\infty \mathbf{W}_{01}^\infty \end{bmatrix} , \qquad (5)$$

and that in the case of a single speech source, the MWF cost functions are decreasing in each iteration, i.e.

$$J_0(\mathbf{W}_0^{i+1}) \leq J_0(\mathbf{W}_0^i), \quad J_1(\mathbf{W}_1^{i+1}) \leq J_1(\mathbf{W}_1^i) . \qquad (6)$$

such that $G_{10}^\infty = \alpha$ and $G_{01}^\infty = 1/\alpha$, and the distributed binaural scheme converges to the optimal B-MWF solution.

5 Capacity of binaural link

Whereas in [3] a binaural link with infinite capacity has been assumed, in this paper we will analyse the influence of the available capacity of the link on the performance of the MWF-based algorithms, especially the B-MWF and the distributed MWF scheme. This analysis is similar to [5], where the influence on the performance of the MWF-front and MWF-contra schemes has been analysed.

From now on, we will only consider the right hearing aid as the transmitting device and analyse the performance at the left hearing aid[1]. When the transmitted signal Y_{01} from the right hearing aid is compressed at rate R (bits per sample), the following rate-distortion relation holds [7]:

$$R(\lambda) = \frac{1}{4\pi} \int_{-\infty}^{\infty} \max\left(0, \log_2 \frac{\Phi_Y^{01}(\omega)}{\lambda}\right) d\omega , \qquad (7)$$

$$D(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \min\left(\lambda, \Phi_Y^{01}(\omega)\right) d\omega , \qquad (8)$$

with the parameter λ linking the rate R and the distortion D, and Φ_Y^{01} the power spectral density (PSD) of the signal Y_{01}. The compressed signal $\tilde{Y}_{01}$ is transmitted to the left hearing aid and can be represented using the forward channel representation [7], depicted in Figure 3, as

$$\tilde{Y}_{01} = B Y_{01} + W , \qquad (9)$$

where B is a bandpass filter with frequency response

$$B = \max\left(0, \frac{\Phi_Y^{01} - \lambda}{\Phi_Y^{01}}\right) ,$$

[1] Since the processing is symmetric, obviously the same analysis holds when considering the left hearing aid as the transmitting device.

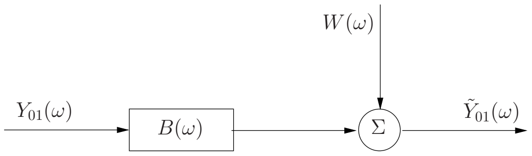

Figure 3: Forward channel representation

and W is additive Gaussian noise with PSD

$$\Phi_W = \max\left(0, \lambda \frac{\Phi_Y^{01} - \lambda}{\Phi_Y^{01}}\right) .$$

It should be noted that using such a representation in the analysis provides an upper bound on the achievable performance at rate R.

6 Experimental results

In this section the performance of the B-MWF and the DB-MWF scheme are compared for different noise scenarios as a function of the capacity of the binaural link.

6.1 Setup and performance measures

Simulations have been performed using a binaural hearing aid setup, where the number of microphones on each hearing aid is $M = 2$ and the distance between the microphones on each hearing aid is 1 cm. We consider a scenario with a single speech source S, a single interference I and spatially uncorrelated noise on each microphone, such that the microphone signal vector $\mathbf{Y}$ can be written as

$$\mathbf{Y} = \mathbf{A}_s S + \mathbf{A}_i I + \mathbf{U} , \tag{10}$$

where $\mathbf{A}_s$ and $\mathbf{A}_i$ represent the acoustic transfer functions for the speech source and the interference, respectively, and $\mathbf{U}$ represents spatially uncorrelated noise. Since the speech source, interference and noise are assumed to be uncorrelated, the correlation matrix $\mathbf{\Phi}_y$ is equal to

$$\mathbf{\Phi}_y = \Phi_s \mathbf{A}_s \mathbf{A}_s^H + \Phi_i \mathbf{A}_i \mathbf{A}_i^H + \Phi_u \mathbf{I}_{2M} , \tag{11}$$

with Φ_s, Φ_i and Φ_u the PSDs of the speech source, interference and noise, respectively. All involved PSDs are assumed to be flat in the band [-Ω,Ω], where $\Omega = 2\pi F$ and $F = 8$ kHz. The signal-to-interference ratio (SIR) and the signal-to-noise ratio (SNR) are defined as

$$\mathrm{SIR} = 10\log_{10}\frac{\Phi_s}{\Phi_i}, \quad \mathrm{SNR} = 10\log_{10}\frac{\Phi_s}{\Phi_u} . \tag{12}$$

The acoustic transfer functions $\mathbf{A}_s$ and $\mathbf{A}_i$ are modelled using the spherical head shadow model in [8] with a radius of 8.75 cm, without taking into account reverberation. The speech source is located at 0° in front of the hearing aid user, whereas the interference is located at $\theta = 30°$ to the left of the hearing aid user. Note that the PSD Φ_Y^{01} is non-flat due to the non-flat acoustic transfer functions.

As in [4, 5], the performance gain is defined as the ratio between the MSE at rate 0 and the MSE at rate R, i.e.

$$G(R) = 10\log_{10}\frac{\xi(0)}{\xi(R)} , \tag{13}$$

which represents the gain in dB due to the availability of the wireless link. As already mentioned, we will only consider the performance gain at the left hearing aid.

For all MWF-based algorithms the trade-off parameter $\mu = 1$ has been used.

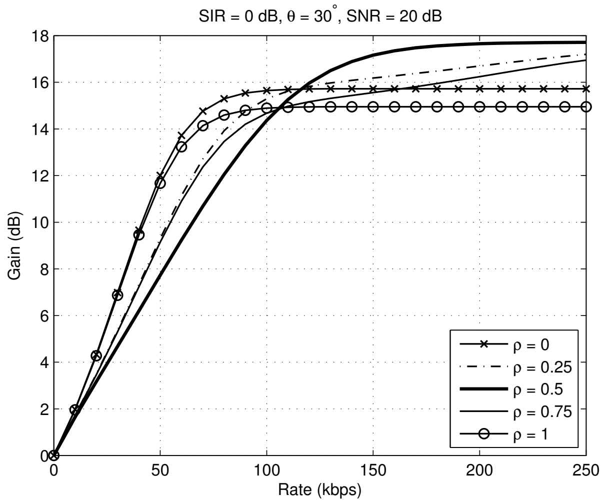

Figure 4: Performance gain as a function of link capacity for B-MWF for different compression rates of both microphone signals.

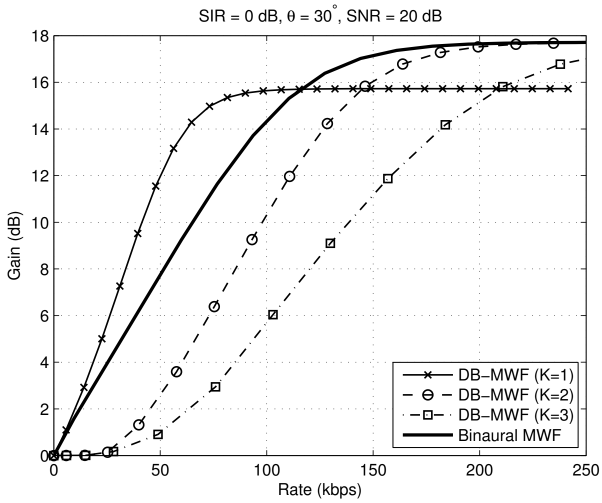

Figure 5: Performance gain as a function of link capacity for B-MWF and DB-MWF for different number of iterations (SIR= 0 dB, $\theta = 30°$, SNR= 20 dB, rate R/K).

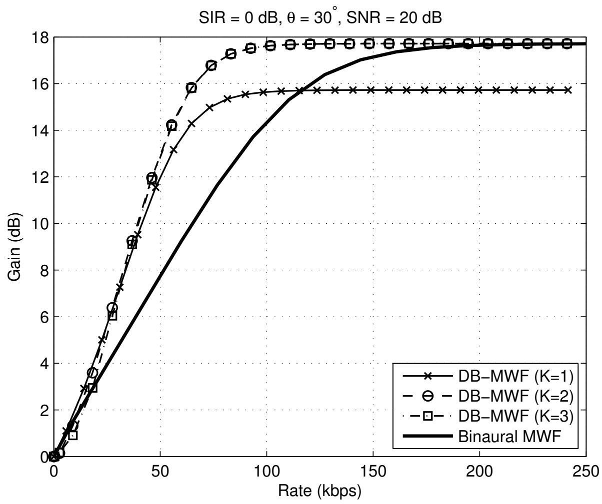

Figure 6: Performance gain as a function of link capacity for B-MWF and DB-MWF for different number of iterations (SIR= 0 dB, $\theta = 30°$, SNR= 20 dB, rate R).

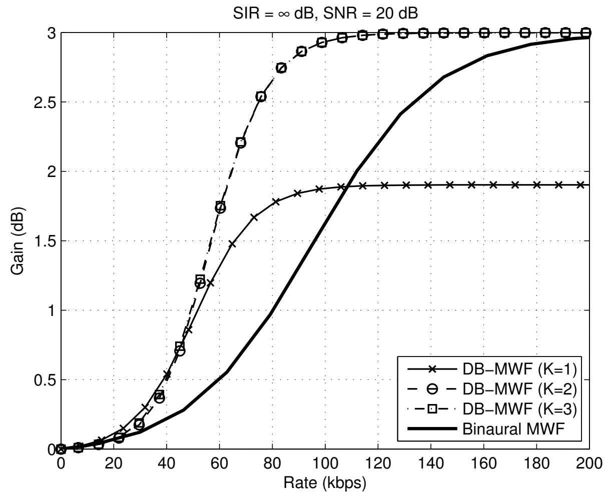

Figure 7: Performance gain as a function of link capacity for B-MWF and DB-MWF for different number of iterations (uncorrelated noise, SNR= 20 dB, rate R).

6.2 Binaural MWF

For SIR= 0 dB and SNR= 20 dB, Figure 4 shows the performance gain of the B-MWF, where the 2 microphone signals from the right hearing aid are transmitted to the left hearing aid and where the total link capacity R is distributed between the rates for compressing the front and the back microphone signal, i.e. the front signal is allocated a rate R_f and the back signal a rate R_b, where

$$R_f = (1-\rho)R, \quad R_b = \rho R\,, \tag{14}$$

with $0 \leq \rho \leq 1$. This means that for $\rho = 0$ only the front microphone signal is compressed at rate R, for $\rho = 1$ only the back microphone signal is compressed at rate R, and for $\rho = 0.5$ both microphone signals are compressed with the same rate $R/2$. As can be seen from Figure 4, for low rates the highest performance gain is achieved by transmitting just a single microphone signal (in this case the front signal), whereas from a certain rate on it pays off compressing and transmitting both microphone signals.

6.3 Distributed MWF scheme

Figures 5-7 compare the performance gain of the distributed MWF scheme with the binaural MWF, for a scenario with and without interference. For the first iteration ($K = 1$), the distributed MWF scheme is initialised with $\mathbf{W}_{11}^{1,H} = [\ 1 \quad 0\ ]$, i.e. the front microphone signal from the right hearing aid is transmitted to the left hearing aid.

For the scenario with interference (SIR= 0 dB, SNR= 20 dB), Figure 5 shows the performance gain of the B-MWF and the DB-MWF scheme for a different number of iterations K, where the total link capacity R is evenly distributed between the iterations, i.e. in each iteration the signal Y_{01} is compressed with rate R/K. As already noted in Figure 4, for low rates the highest performance gain is achieved by transmitting just a single microphone signal, corresponding in this figure to $K = 1$. Performing more iterations only leads to an improved performance at high rates. As shown in [3], for an infinite (i.e. sufficiently high) rate the DB-MWF scheme converges to the B-MWF solution, typically requiring only a small number of iterations (in this case $K = 2$).

For many applications it can however be assumed that the signal statistics remain stationary over a (small) number of signal frames, such that the iterations of the DB-MWF scheme can be spread over subsequent frames, instead of performing several iterations on the same frame, as in Figure 5. For the scenario with interference (SIR= 0 dB, SNR= 20 dB), Figure 6 shows the performance gain of the B-MWF and the DB-MWF scheme, where now in each iteration the signal Y_{01} is compressed with rate R. This figure shows that for the considered scenario the DB-MWF scheme converges after $K = 2$ iterations at all rates, moreover achieving the highest performance gain. Similarly, Figure 7 shows the performance gain for the scenario without interference (SIR= ∞ dB, SNR= 20 dB), where the performance gain is smaller than for the scenario with interferer but the same conclusions hold.

7 Conclusion

In this paper, the performance gain of the binaural MWF and the DB-MWF scheme have been analysed as a function of the capacity of the binaural link. For the binaural MWF, it has been shown that the optimal distribution of the rate between the microphone signals depends on the total capacity. For the DB-MWF scheme, it has been shown that when the iterations can be spread over subsequent frames, the iterative DB-MWF scheme yields the highest performance gain after only a small number of iterations.

References

[1] S. Doclo, S. Gannot, M. Moonen, and A. Spriet, *Acoustic beamforming for hearing aid applications*, chapter 9 in "Handbook on Array Processing and Sensor Networks", pp. 269–302, Wiley, 2010.

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

[3] S. Doclo, T. Van den Bogaert, M. Moonen, and J. Wouters, "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.

[4] O. Roy and M. Vetterli, "Rate-constrained collaborative noise reduction for wireless hearing aids," *IEEE Trans. Signal Processing*, vol. 57, no. 2, pp. 645–657, Feb. 2009.

[5] S. Srinivasan and A. C. den Brinker, "Rate-constrained beamforming in binaural hearing aids," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, Article ID 257197, 2009.

[6] A. Bertrand and M. Moonen, "Robust distributed noise reduction in hearing aids with external acoustic sensor nodes," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, Article ID 530435, 2009.

[7] T. Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression*, Prentice Hall, 1971.

[8] R. O. Duda and W. L. Martens, "Range dependence of the response of a spherical head model," *Journal of the Acoustical Society of America*, vol. 104, no. 5, pp. 3048–3058, 1998.