

Research Article

Rate-Constrained Beamforming in Binaural Hearing Aids

Sriram Srinivasan and Albertus C. den Brinker

Digital Signal Processing Group, Philips Research, High Tech Campus 36, 5656 AE Eindhoven, The Netherlands

Correspondence should be addressed to Sriram Srinivasan, sriram.srinivasan@philips.com

Received 1 December 2008; Accepted 6 July 2009

Recommended by Henning Puder

Recently, hearing aid systems where the left and right ear devices collaborate with one another have received much attention. Apart from supporting natural binaural hearing, such systems hold great potential for improving the intelligibility of speech in the presence of noise through beamforming algorithms. Binaural beamforming for hearing aids requires an exchange of microphone signals between the two devices over a wireless link. This paper studies two problems: which signal to transmit from one ear to the other, and at what bit-rate. The first problem is relevant as modern hearing aids usually contain multiple microphones, and the optimal choice for the signal to be transmitted is not obvious. The second problem is relevant as the capacity of the wireless link is limited by stringent power consumption constraints imposed by the limited battery life of hearing aids.

Copyright © 2009 S. Srinivasan and A. C. den Brinker. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

Modern hearing aids are capable of performing a variety of advanced digital processing tasks such as amplification and dynamic compression, sound environment classification, feedback cancellation, beamforming, and single-channel noise reduction. Improving the intelligibility and quality of speech in noise through beamforming is arguably the most sought after feature among hearing aid users [1]. Modern hearing aids typically have multiple microphones and have been proven to provide reasonable improvements in intelligibility and listening comfort.

As human hearing is binaural by nature, it is intuitive to expect an improved experience by using one hearing aid for each ear, and the number of such fittings has increased significantly [2]. These fittings however have mostly been bilateral, that is, two independently operating hearing aids on the left and right ears. To experience the benefits of binaural hearing, the two hearing aids need to collaborate with one another to ensure that binaural cues are presented in a consistent manner to the user. Furthermore, the larger spacing between microphones in binaural systems compared to monaural ones provides more flexibility for tasks such as beamforming. Such binaural systems introduce new challenges, for example, preserving binaural localization cues such as interaural time and level differences, and the

exchange of signals between the hearing aids to enable binaural processing. The former has been addressed in [3, 4]. The latter is the subject of this paper.

Binaural beamforming requires an exchange of signals between the left and right hearing aids. A wired link between the two devices is cumbersome and unacceptable from an aesthetic point of view, thus necessitating a wireless link. Wireless transmission of data is power intensive, and to preserve battery life, it becomes important to limit the number of bits exchanged over the link. A reduction in the bit-rate affects the performance of the beamformer. This paper investigates the relation between the transmission bit-rate and beamformer performance.

In the absence of bit-rate constraints, a simple practical scheme is to transmit all observed microphone signals from one ear to the other, where they are fed into a beamformer together with the locally observed signals to obtain an estimate of the desired signal. In the presence of a limited capacity link, however, an intelligent decision on what signal to transmit is necessary to effectively utilize the available bandwidth. Reduced bandwidth binaural beamforming algorithms have been discussed in [5], but the relation between bit-rate and performance was not studied.

An elegant theoretically optimal (in an information-theoretic sense) transmission scheme is presented in [6],

where the hearing aid problem is viewed as a remote Wyner-Ziv problem, and the transmitting device encodes its signals such that the receiving device obtains the MMSE estimate of the desired signal, with the locally observed signals at the left device as side information. However, it requires knowledge of the (joint) statistics of the signals observed at both the receiving and transmitting ears, which are not available in a hearing aid setup. A suboptimal but practical choice presented in [7] and shown in Figure 1 is to first obtain an estimate of the desired signal at the transmitting (right ear in this example) device, and then transmit this estimate. This choice is asymptotically (at infinite bit-rate) optimal if the desired signal is observed in the presence of spatially uncorrelated noise, but not when a localized interferer is present.

From an information point of view, transmitting only an estimate of the desired signal does not convey all the information that could be used in reconstructing the desired signal at the receiving (left ear) device. Specifically, lack of information about the interferer in the transmitted signal results in an inability to exploit the larger left-right microphone spacing (provided by the binaural setup) to cancel the interferer. This paper proposes and investigates two practical alternatives to circumvent this problem. The first approach is to obtain an estimate of the interference-plus-noise at the right hearing aid using the right ear microphone signals, and transmit this estimate to the left device. This scheme is similar to the one in Figure 1, except that the signal being estimated at the right ear is the undesired signal. Intuitively, this would enable better performance in the presence of a localized interferer as both the locally available microphone signals and the received signal can be used in the interference cancellation process, and this is indeed observed for certain situations in the simulations described later in the paper.

Following the information point of view one step further leads to the second scheme proposed in this paper, which is to just transmit one or more of the right ear microphone signals at rate R , as shown in Figure 2. The unprocessed signal conveys more information about both the desired and the undesired signal, although potentially requiring a higher bit-rate. What remains to be seen is the trade-off between rate and performance.

This paper provides a framework to quantify the performance of the two above-mentioned beamforming schemes in terms of the rate R , the location of the desired source and interferer, and the signal-to-interference ratio (SIR). The performance is then compared to both the optimal scheme discussed in [6] and the suboptimal scheme of Figure 1 at different bit-rates.

For the two-microphone system in Figure 2, given a bit-rate R , another possibility is to transmit each of the two right ear microphone signals at a rate $R/2$. This is however not considered in this paper for the following reason. In terms of interference cancellation, even at infinite rate, transmitting both signals results only in a marginal improvement in SIR compared to transmitting one signal. The reason is that the left ear already has an endfire array.

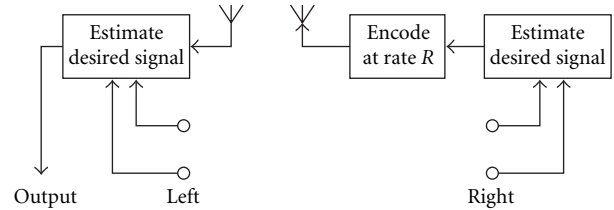


FIGURE 1: The scheme of [7]. The desired signal is first estimated from the right ear microphone signals, and then transmitted at a rate R . At the left ear, the desired signal is estimated from the received signal and the local microphone signals.

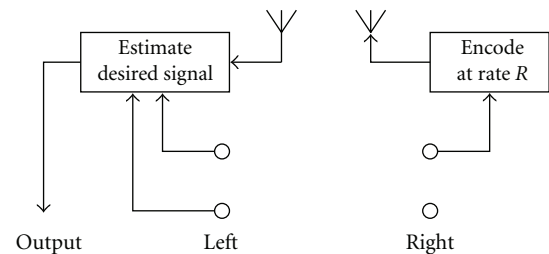


FIGURE 2: One right ear microphone signal is transmitted at a rate R . At the left ear, the desired signal is estimated from the received signal and the local microphone signals.

A large gain results from the new broadside array that is created by transmitting the signal observed at one right ear microphone. Additionally transmitting the signal from the second right ear microphone, which is located close to the first microphone, provides only a marginal gain. Thus, in the remainder of this paper, we only consider transmitting one microphone signal from the right ear.

Another aspect of the discussion on the signal to be transmitted is related to the frequency-dependence of the performance of the beamformer, especially in systems where the individual hearing aids have multiple closely spaced microphones. These small microphone arrays on the individual devices are capable of interference cancellation at high frequencies but not at low frequencies where they have a large beamwidth. As shown in [8], the benefit provided by a binaural beamformer in terms of interference cancellation is thus limited to the low-frequency part of the signal, where the monaural array performs poorly. Moreover, due to the larger size of the binaural array, spatial aliasing affects performance in high frequencies. Motivated by these reasons, this paper also investigates the effect of transmitting only the low frequencies on the required bit-rate and the resulting performance of the beamformer.

The remainder of this paper is organized as follows. Section 2 introduces the signal model and the relevant notation. The two rate-constrained transmission schemes introduced in this paper are presented in Section 3. The performance of the proposed and reference systems is compared for different scenarios in Section 4. Concluding remarks are presented in Section 5.

2. Signal Model

Consider a desired source $S(n)$ and an interfering source $I(n)$, in the presence of noise. The left ear signal model can be written as

$$X_l^k(n) = h_l^k(n) \star S(n) + g_l^k(n) \star I(n) + U_l^k(n), \quad k = 1 \cdots K, \quad (1)$$

where $h_l^k(n)$ and $g_l^k(n)$ are the transfer functions between the k th microphone on the left hearing aid and the desired and interfering sources, respectively, $U_l^k(n)$ is uncorrelated zero-mean white Gaussian noise at the k th microphone on the left hearing aid, K is the number of microphones on the left hearing aid, n is the time index, and the operator \star denotes convolution. The different sources are assumed to be zero-mean independent jointly Gaussian stationary random processes with power spectral densities (PSDs) $\Phi_S(\omega)$, $\Phi_I(\omega)$, and $\Phi_{U_l^k}(\omega)$, respectively. The above signal model allows the consideration of different scenarios, for example, desired signal in the presence of uncorrelated noise, or desired signal in the presence of a localized interferer, and so forth. Let

$$S_l^k(n) = h_l^k(n) \star S(n) \quad (2)$$

denote the desired signal at the k th microphone on the left device and let

$$W_l^k(n) = g_l^k(n) \star I(n) + U_l^k(n) \quad (3)$$

denote the undesired signal. A similar right ear model follows:

$$X_r^k(n) = \underbrace{h_r^k(n) \star S(n)}_{S_r^k(n)} + \underbrace{g_r^k(n) \star I(n) + U_r^k(n)}_{W_r^k(n)}, \quad k = 1 \cdots K, \quad (4)$$

where the relevant terms are defined analogously to the left ear. The following assumptions are made for simplicity:

$$\Phi_{U_l^k}(\omega) = \Phi_{U_r^k}(\omega) = \Phi_U(\omega), \quad k = 1 \cdots K. \quad (5)$$

Let $\mathbf{X}_l(n) = [X_l^1(n), \dots, X_l^K(n)]^T$, and $\mathbf{X}_r(n) = [X_r^1(n), \dots, X_r^K(n)]^T$. The vectors $\mathbf{U}_l(n)$ and $\mathbf{U}_r(n)$ are defined analogously. For any $X(n)$ and $Y(n)$, define $\Phi_{XY}(\omega)$ to be their cross PSD. As $U_l^k(n)$ and $U_r^k(n)$ correspond to spatially uncorrelated noise, the following holds:

$$\begin{aligned} \Phi_{U_l^j U_l^k}(\omega) &= \Phi_{U_l^j U_r^k}(\omega) = 0, \quad j, k = 1 \cdots K, \quad j \neq k, \\ \Phi_{U_l^j U_r^k}(\omega) &= \Phi_{U_l^j U_l^k}(\omega) = 0, \quad j, k = 1 \cdots K. \end{aligned} \quad (6)$$

The PSD of the microphone signal $X_l^k(n)$, is given by

$$\Phi_{X_l^k}(\omega) = |H_l^k(\omega)|^2 \Phi_S(\omega) + |G_l^k(\omega)|^2 \Phi_I(\omega) + \Phi_U(\omega), \quad (7)$$

where $H_l^k(\omega)$ is the frequency domain transfer function corresponding to $h_l^k(n)$, and $G_l^k(\omega)$ corresponds to $g_l^k(n)$. An analogous expression follows for $\Phi_{X_r^k}(\omega)$. The (j, k) th entry of the matrix $\Phi_{\mathbf{X}_l}$, which is the PSD matrix corresponding to the vector \mathbf{X}_l is given by

$$\begin{aligned} \Phi_{\mathbf{X}_l}^{jk} &= H_l^j(\omega) H_l^{k\dagger}(\omega) \Phi_S(\omega) + G_l^j(\omega) G_l^{k\dagger}(\omega) \Phi_I(\omega) \\ &\quad + \delta(j - k) \Phi_U(\omega), \end{aligned} \quad (8)$$

where the superscript \dagger denotes complex conjugate transpose.

3. What to Transmit

The problem is treated from the perspective of estimating the desired signal at the left hearing aid. Assume that the right hearing aid transmits some function of its observed microphone signals to the left hearing aid. The left device uses its locally observed microphone signals together with the signal received from the right device to obtain an estimate \hat{S}_l of the desired signal $S_l = S_l^1(n)$ (the choice $k = 1$ is arbitrary) at the left device. (The processing is symmetric and the right device similarly obtains an estimate of $S_r^1(n)$.)

Denote the signal transmitted by the right device as $X_t(n)$, and its PSD by $\Phi_t(\omega)$. The signal $X_t(n)$ is transmitted at a rate R to the left ear. Under the assumptions in the signal model presented in Section 2, the following parametric rate-distortion relation holds [9]:

$$\begin{aligned} R(\lambda) &= \frac{1}{4\pi} \int_{-\infty}^{\infty} \max\left(0, \log_2 \frac{\Phi_t(\omega)}{\lambda}\right) d\omega, \\ D(\lambda) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \min(\lambda, \Phi_t(\omega)) d\omega, \end{aligned} \quad (9)$$

where the rate is expressed in bits per sample. The distortion here is the mean-squared error (MSE) between $X_t(n)$ and its reconstruction. Equation (9) provides the relation between the number of bits R used to represent the signal and the resulting distortion D in the reconstructed signal. As the relation between R and D cannot be obtained in closed form, it is expressed in terms of a parameter λ . Inserting a particular value of λ in (9) results in certain rate R and a corresponding distortion D . An R - D curve is obtained as λ traverses the interval $[0, \text{ess sup } \Phi_t(\omega)]$, where ess sup is the essential supremum.

Note that $X_t(n)$ is quantized without regard to the final processing steps at the left ear, and without considering the presence of the left microphone signals. Incorporating such knowledge can lead to more efficient quantization schemes, for example, by allocating bits to only those frequency components of $\Phi_t(\omega)$ that contribute to the estimation of $S_l(n)$. Such schemes however as mentioned earlier are not amenable to practical implementations as the required statistics are not available under the nonstationary conditions encountered in hearing aid applications.

Let the right device compress $X_t(n)$ at a rate R bits per sample, which corresponds to a certain λ and a distortion D . The signal received at the left ear is depicted in

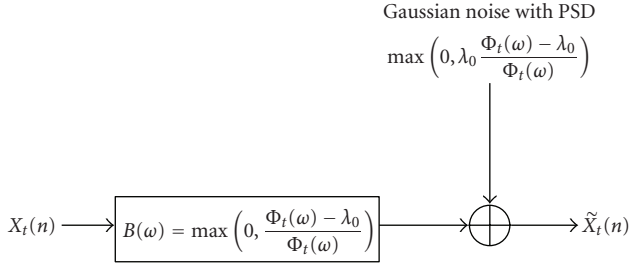


FIGURE 3: The forward channel representation.

Figure 3 using the *forward channel* representation [9]. The signal $\tilde{X}_t(n)$ is obtained by first bandlimiting $X_t(n)$ with a filter with frequency response

$$B(\omega) = \max\left(0, \frac{\Phi_t(\omega) - \lambda}{\Phi_t(\omega)}\right), \quad (10)$$

and then adding Gaussian noise with PSD given by

$$\Phi_Z(\omega) = \max\left(0, \lambda \frac{\Phi_t(\omega) - \lambda}{\Phi_t(\omega)}\right). \quad (11)$$

Note that using such a representation for $\tilde{X}_t(n)$ in the analysis provides an upper bound on the achievable performance at rate R . Define

$$\mathbf{X}(n) = [\mathbf{X}_l^T(n), \tilde{X}_t^T(n)]^T. \quad (12)$$

The MMSE estimate of the desired signal S_l is given by

$$\hat{S}_l = E\{S_l | \mathbf{X}\}, \quad (13)$$

and the corresponding MSE by

$$\xi(R) = \frac{1}{2\pi} \int_0^{2\pi} [\Phi_{S_l}(\omega) - \Phi_{S_l\mathbf{X}}(\omega)\Phi_{\mathbf{X}}^{-1}(\omega)\Phi_{\mathbf{X}S_l}(\omega)] d\omega, \quad (14)$$

where $\Phi_{S_l}(\omega) = |H_l^1(\omega)|^2\Phi_s$, $\Phi_{S_l\mathbf{X}}(\omega) = [\Phi_{S_lX_l}(\omega), \Phi_{S_l\tilde{X}_t}(\omega)]$, and $\Phi_{\mathbf{X}}$ is the PSD matrix corresponding to the vector \mathbf{X} . The MSE $\xi(R)$ can be rewritten in an intuitively appealing form in terms of the MSE resulting when estimation is performed using only X_l and a reduction term due to the availability of the innovation process $X_i = X_t - E\{X_t | \mathbf{X}_l\}$. The following theorem follows by applying results from linear estimation theory [10, Chapter 4].

Theorem 1. Let $X_i = X_t - E\{X_t | \mathbf{X}_l\}$. X_i represents the innovation or the “new” information at the left ear provided by the wireless link. Then, the MSE ξ can be written as

$$\xi(R) = \xi_l - \frac{1}{2\pi} \int_0^{2\pi} [\Phi_{S_l}(\omega) - \xi_{lr}(\omega)] d\omega, \quad (15)$$

where

$$\xi_l = \frac{1}{2\pi} \int_0^{2\pi} [\Phi_{S_l}(\omega) - \Phi_{S_lX_l}(\omega)\Phi_{X_l}^{-1}(\omega)\Phi_{X_lS_l}(\omega)] d\omega \quad (16)$$

is the error in estimating S_l from X_l alone, and

$$\xi_{lr}(\omega) = \Phi_{S_l}(\omega) - \Phi_{S_lX_i}(\omega)\Phi_{X_i}^{-1}(\omega)\Phi_{X_iS_l}(\omega) \quad (17)$$

is the error in estimating S_l from X_i .

The MSE resulting from the two proposed schemes discussed in Section 1 can be computed by setting $X_t(n)$ appropriately, and then using (14). In the first scheme, an estimate of the undesired signal obtained using $X_r(n)$ is transmitted to aid in better interference cancellation using the larger left-right microphone spacing. The resulting MSE is given by

$$\xi_{\text{int}}(R) = \xi(R) |_{X_r=E\{W_l^1 | X_r\}}. \quad (18)$$

In the second scheme, one of the raw microphone signals at the right ear, without loss of generality $X_r^1(n)$, is transmitted. The resulting MSE is given by

$$\xi_{\text{raw}}(R) = \xi(R) |_{X_r=X_r^1}. \quad (19)$$

As an example, the relevant entities required in this case ($X_t = X_r^1$) to compute the MSE using (14) are given as follows:

$$B(\omega) = \max\left(0, \frac{\Phi_r^1(\omega) - \lambda}{\Phi_r^1(\omega)}\right),$$

$$\Phi_Z(\omega) = \max\left(0, \lambda \frac{\Phi_r^1(\omega) - \lambda}{\Phi_r^1(\omega)}\right),$$

$$\Phi_{S_l\mathbf{X}}(\omega) = [H_l^1(\omega)\mathbf{H}_l^{\dagger}(\omega)\Phi_s(\omega), B(\omega)H_l^1(\omega)H_r^{1\dagger}(\omega)\Phi_s(\omega)],$$

$$\Phi_{\mathbf{X}} = \begin{pmatrix} \Phi_{X_l} & \Phi_{X_l\tilde{X}_t} \\ \Phi_{\tilde{X}_tX_l} & \Phi_{\tilde{X}_t} \end{pmatrix}, \quad (20)$$

where the submatrix Φ_{X_l} in (20) is given by (8), the PSD $\Phi_{\tilde{X}_t}(\omega)$ of the received signal $\tilde{X}_t(n)$ is given by

$$\Phi_{\tilde{X}_t}(\omega) = |B(\omega)|^2\Phi_{X_l}(\omega) + \Phi_Z(\omega), \quad (21)$$

and the cross PSD $\Phi_{X_l\tilde{X}_t}$ is given by

$$\Phi_{X_l\tilde{X}_t} = B(\omega)(\mathbf{H}_l(\omega)H_r^{1\dagger}(\omega)\Phi_s(\omega) + \mathbf{G}_l(\omega)G_r^{1\dagger}(\omega)\Phi_i(\omega)), \quad (22)$$

with $\mathbf{H}_l(\omega) = [H_l^1(\omega), \dots, H_l^K(\omega)]^T$ and $\mathbf{G}_l(\omega) = [G_l^1(\omega), \dots, G_l^K(\omega)]^T$. The MSEs $\xi_{\text{int}}(R)$ and $\xi_{\text{raw}}(R)$ are evaluated for different bit-rates and compared to $\xi_{\text{opt}}(R)$, which is the MSE resulting from the optimal scheme of [6], and $\xi_{\text{sig}}(R)$, which is the MSE resulting from the reference scheme [7], where a local estimate of the desired signal is transmitted.

Before analyzing the performance as a function of the bit-rate, it is instructive to examine the asymptotic performance (at infinite bit-rate) of the different schemes.

An interesting contrast results by studying the case when only one microphone is present on each hearing aid, and the case with multiple microphones on each device. It is convenient to formulate the analysis in the frequency domain. In the single microphone case, the transmitted signal can be expressed as $X_t(\omega) = A(\omega)X_r^1(\omega)$, where $X_t(\omega)$ and $X_r^1(\omega)$ are obtained by applying the discrete Fourier transform (DFT) to the respective time domain entities $X_t(n)$, and $X_r^1(n)$, and $A(\omega)$ is a nonzero scalar. In the method of [7] where a local estimate of the desired signal is transmitted, $A(\omega) = \Phi_{S_r X_r}(\omega)\Phi_{X_r}^{-1}(\omega)$. Note that in the single-microphone case, $\mathbf{X}_r(\omega) = X_r^1(\omega)$. In the first proposed scheme where an estimate of the undesired signal is transmitted, $A(\omega) = \Phi_{W_r X_r}(\omega)\Phi_{X_r}^{-1}(\omega)$, and in the second proposed scheme where the first microphone signal is transmitted, $A(\omega) = 1$.

For the estimation at the left ear using both the locally observed signal and the transmitted signal, it can readily be seen that

$$E\{S_l(\omega) | X_r^1(\omega), A(\omega)X_r^1(\omega)\} = E\{S_l(\omega) | X_r^1(\omega), X_r^1(\omega)\}, \quad (23)$$

where $S_l(\omega)$ and $X_r^1(\omega)$ are obtained from their respective time domain entities $S_l(n)$ and $X_r^1(n)$ by applying the DFT. Thus, at infinite-rate in the single-microphone case, all three schemes reach the performance of the optimal scheme where both microphone signals are available at the left ear, regardless of the correlation properties of the undesired signals at the left and right ear.

The case when each hearing aid has multiple microphones, however, offers a contrasting result. In this case, the transmitted signal is given by $X_t(\omega) = \mathbf{A}(\omega)\mathbf{X}_r(\omega)$, where $\mathbf{A}(\omega)$ is a $1 \times K$ vector and assumes different values depending on the transmission scheme. Here, $X_t(\omega)$ is a down-mix of K different signals into a single signal, resulting in a potential loss of information since in a practical scheme the down-mix at the right ear is performed without knowledge of the left ear signals. In this case, even at infinite bit-rate, the three schemes may not achieve optimal performance. One exception is when the undesired signal at the different microphones is uncorrelated, and transmitting a local estimate of the desired signal provides optimal performance, asymptotically.

4. Performance Analysis

In this section, the performance of the different schemes discussed above is compared for different locations of the interferer, different SIRs, and as a function of the bit-rate. All the involved PSDs are assumed to be known to establish theoretical upper bounds on performance. First, the performance measure used to evaluate the different schemes is introduced. The experimental setup used for the performance analysis is then described. Two cases are then considered: one where the desired signal is observed in the presence of uncorrelated (e.g., sensor) noise, and the second where the desired signal is observed in the presence of a localized interferer in addition to uncorrelated noise.

4.1. Performance Measure. As in [6, 7], the performance gain is defined as the ratio between the MSE at rate 0 and the MSE at rate R :

$$G(R) = 10 \log_{10} \frac{\xi(0)}{\xi(R)}, \quad (24)$$

which represents the gain in dB due to the availability of the wireless link. The quantities $G_{\text{opt}}(R)$, $G_{\text{sig}}(R)$, $G_{\text{int}}(R)$, and $G_{\text{raw}}(R)$, corresponding to the four different transmission schemes, are computed according to (24) from their respective average MSE values $\xi_{\text{opt}}(R)$, $\xi_{\text{sig}}(R)$, $\xi_{\text{int}}(R)$, and $\xi_{\text{raw}}(R)$. $\xi(0)$ remains the same in all four cases as this corresponds to the average MSE at rate zero, which is the MSE in estimating the desired signal using only the microphone signals on the left ear.

4.2. Experimental Setup. In the analysis, the number of microphones on each hearing aid was set to a maximum of two, that is, $K = 2$. Simulations were performed both for $K = 1$ and $K = 2$. The spherical head shadow model described in [11] was used to obtain the transfer functions $H_l^k(\omega)$, $H_r^k(\omega)$, $G_l^k(\omega)$, and $G_r^k(\omega)$, for $k = 1, 2$. The distance between microphones on a single hearing aid was assumed to be 0.01 m. The radius of the sphere was set to 0.0875 m. The desired, interfering, and noise sources were assumed to have flat PSDs Φ_s , Φ_i , and Φ_u , respectively, in the band $[-\Omega, \Omega]$, where $\Omega = 2\pi F$, and $F = 8000$ Hz. Note that Φ_t is not flat due to the nonflat transfer functions.

4.3. Desired Source in Uncorrelated Noise. The desired source is assumed to be located at 0° in front of the hearing aid user. This is a common assumption in hearing aids [1]. The signal-to-noise ratio (SNR), computed as $10 \log_{10} \Phi_s / \Phi_u$, is assumed to be 20 dB. The SIR, computed as $10 \log_{10} \Phi_s / \Phi_i$, is assumed to be infinity, that is, $\Phi_i = 0$. Thus the only undesired signal in the system is uncorrelated noise.

Figure 4(a) plots the gain due to the availability of the wireless link for $K = 1$; that is, each hearing aid has only one microphone. $G_{\text{sig}}(R)$, $G_{\text{int}}(R)$, and $G_{\text{raw}}(R)$ are almost identical, and reach $G_{\text{opt}}(R)$ at $R = \infty$, as expected from (23). At low rates, the theoretically optimal scheme $G_{\text{opt}}(R)$ performs better than the three suboptimal schemes as it uses information about signal statistics at the remote device. The gain is 3 dB, corresponding to the familiar gain in uncorrelated noise resulting from the doubling of microphones from one to two. Clearly, in this case transmitting the raw microphone signal is a good choice as the computational load and delay in first obtaining a local estimate can be avoided.

Figure 4(b) plots the gain for $K = 2$, and as discussed at the end of Section 3, the contrast with $K = 1$ is evident. At high rate, both $G_{\text{opt}}(R)$ and $G_{\text{sig}}(R)$ approach 3 dB, again due to a doubling of microphones, now from two to four. $G_{\text{raw}}(R)$ saturates at a lower value as there are only three microphone signals available for the estimation. Finally, transmitting an estimate of the undesired signal leads to zero gain in this case as the noise is spatially uncorrelated, and thus the transmitted signal does not contribute to the estimation of

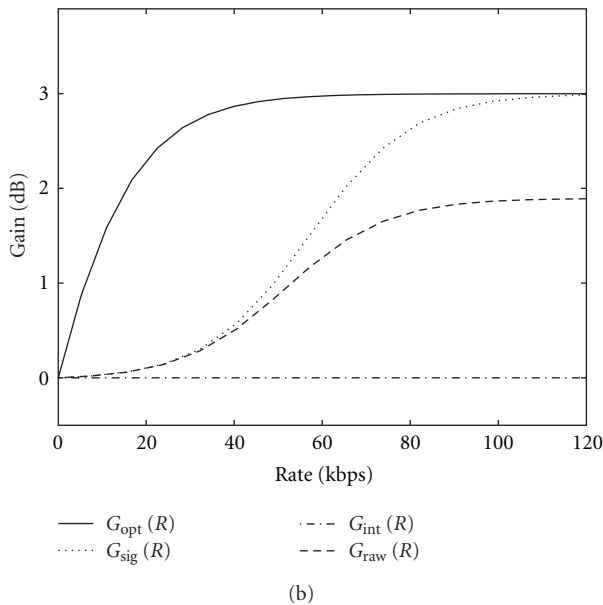
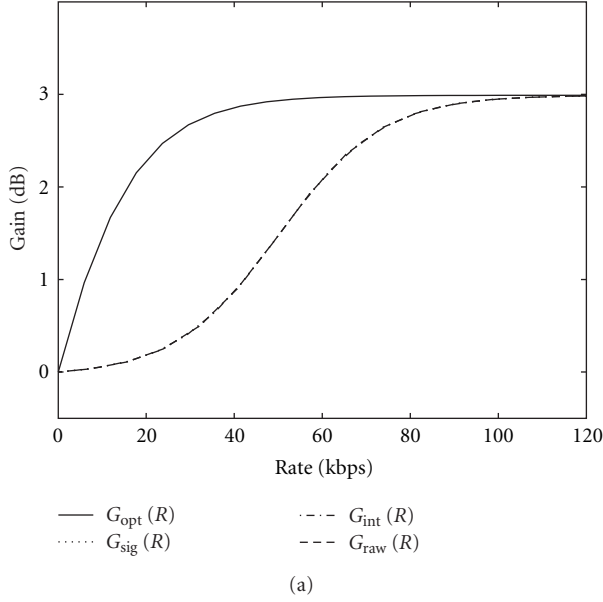


FIGURE 4: Performance gain for the three schemes when a desired signal is observed in the presence of uncorrelated noise (i.e., $SIR = \infty$). (a) $K = 1$, (b) $K = 2$.

the desired signal at the left ear. It is interesting to note that for $K = 1$, transmitting an estimate of the undesired signal led to an improvement, which can be explained by (23).

When comparing the results for $K = 1$ with $K = 2$, it needs to be noted that the figures plot the improvement compared to rate $R = 0$ in each case, and not the absolute SNR gain. This applies to the results shown in the subsequent sections as well. For uncorrelated noise, the absolute SNR gain at infinite bit-rate in the four microphone system compared to the SNR at a single microphone is 6 dB with $K = 2$ and 3 dB with $K = 1$.

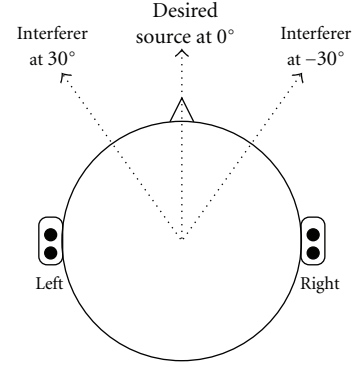


FIGURE 5: Location of desired and interfering sources. For an interferer located at 30° , the SIR at the left ear is lower than at the right ear due to head shadow.

4.4. Desired and Interfering Sources in Uncorrelated Noise. The behavior of the different schemes in the presence of a localized interferer is of interest in the hearing aid scenario. As before, a desired source is assumed to be located at 0° (front of the user), and the SNR is set to 20 dB. In addition, an interferer is assumed to be located at -30° (i.e., front, 30° to the right, see Figure 5), and the SIR is set to 0 dB. Figure 6 compares the four schemes for this case. For $K = 1$, Figure 6(a) shows that the different schemes exhibit similar performance.

For $K = 2$, Figure 6(b) provides useful insights. It is evident from the dotted curve that transmitting an estimate of the desired signal leads to poor performance. Transmitting an estimate of the interferer, interestingly, results in a higher gain as seen from the dash-dot curve and can be explained as follows. At high rates, the interferer is well preserved in the transmitted signal. Better interference suppression is now possible using the binaural array (larger spacing) than with the closely spaced monaural array, and thus the improved performance. Transmitting the unprocessed signal results in an even higher gain $G_{\text{raw}}(R)$ that approaches the gain resulting from the optimal scheme. In this case, not only is better interference rejection possible but also better estimation of the desired signal as the transmitted signal contains both the desired and undesired signals. Again, it is important to note that the figure only shows the gain due to the presence of the wireless link, and not the absolute SNR gain, which is higher for $K = 2$ than for $K = 1$ due to the higher number of microphones.

Figure 7 considers the case when the interferer is located at 30° instead of -30° , which leads to an interesting result. Again, we focus on $K = 2$. The behavior of $G_{\text{opt}}(R)$ and $G_{\text{raw}}(R)$ in Figure 7(b) is similar to Figure 6(b), but the curves $G_{\text{sig}}(R)$ and $G_{\text{int}}(R)$ appear to be almost interchanged with respect to Figure 6(b). This reversal in performance can be intuitively explained by the head shadow effect. Note that the performance gain is measured at the left ear. When the interferer is located at 30° , the SIR at the left ear is lower than the SIR at the right ear as the interferer is closer to the left ear, and shadowed by the head at the right ear; see Figure 5. Thus at the right ear, it is possible to obtain a

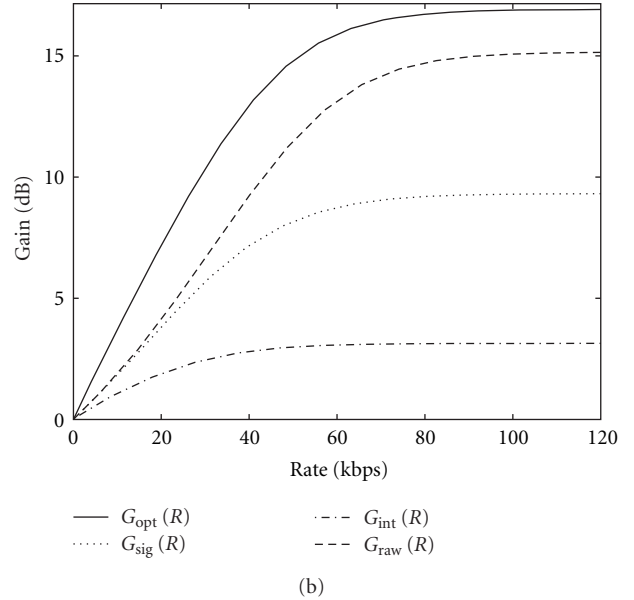
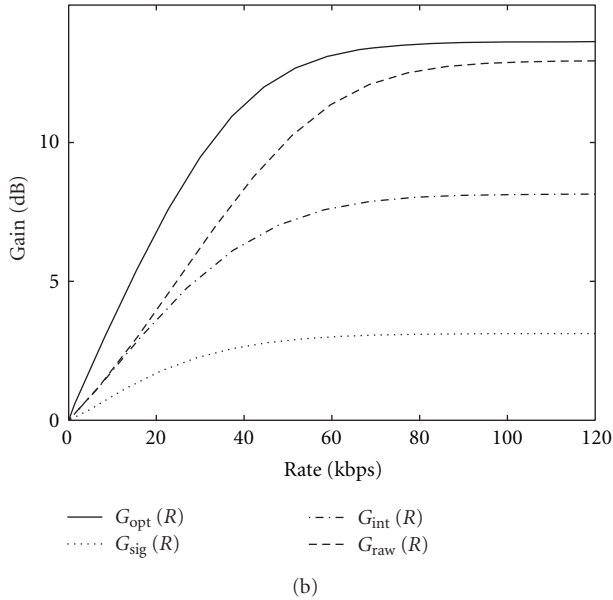
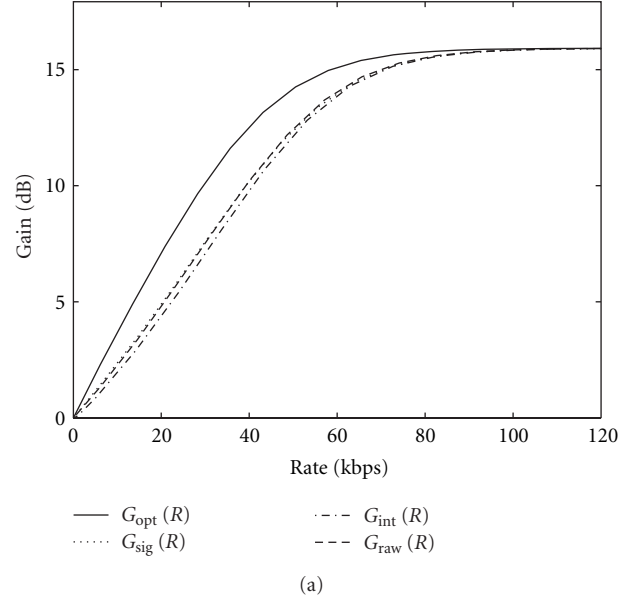
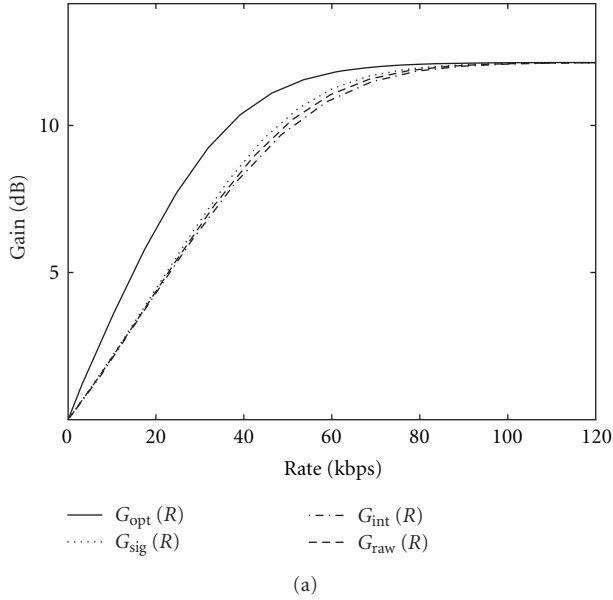


FIGURE 6: Performance gain for the different schemes when a desired signal is observed in the presence of uncorrelated noise at 20 dB SNR, and an interfering source at 0 dB SIR located at -30° . (a) $K = 1$, (b) $K = 2$.

FIGURE 7: Performance gain for the different schemes when a desired signal is observed in the presence of uncorrelated noise at 20 dB SNR and an interfering source at 0 dB SIR located at 30° . (a) $K = 1$, (b) $K = 2$.

good estimate of the desired signal but not of the interferer. So, transmitting an estimate of the desired signal leads to better performance than transmitting an estimate of the interferer. For an interferer located at -30° , the interference-to-signal ratio is higher at the right ear, and thus it is possible to obtain a better estimate of the interferer than possible at the left ear. Transmitting this estimate to the left ear provides information that can be exploited for interference cancellation.

From the above analysis, it can be concluded that a decision on which signal to transmit needs to be made depending on the SIR. At high SIRs ($SIR = \infty$ in the limit,

thus only uncorrelated noise), transmitting an estimate of the desired signal is better than transmitting the raw microphone signal. At low SIRs, the converse holds. A simple rule of thumb is to always transmit the unprocessed microphone signal as the penalty at high SIRs is negligible (see Figure 4) compared to the potential gains at low SIRs (see Figures 6 and 7). In addition, such a scheme results in a lower computational load and reduced delay.

It may be noted that this paper considers theoretical upper bounds on the performance of the different schemes. In a practical scheme where the unprocessed signal is coded and transmitted, only the PSD of the noisy signal is required.

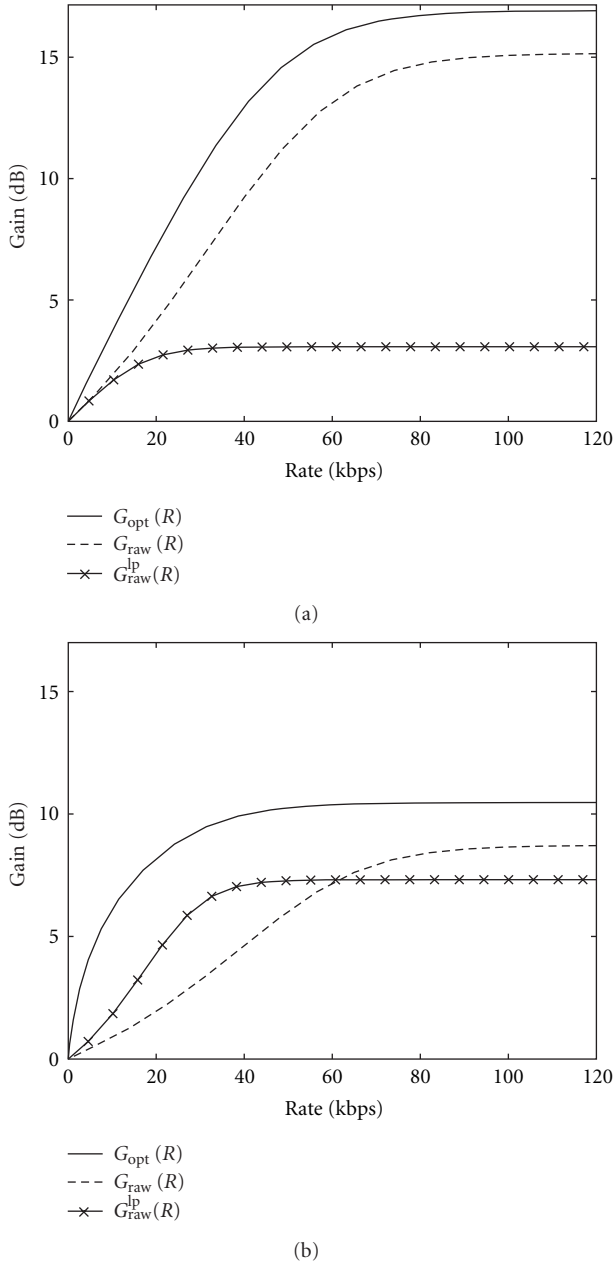


FIGURE 8: Performance gain for the different schemes when a desired signal is observed in the presence of uncorrelated noise at 20 dB SNR, $K = 2$, and only the low-frequency portion is transmitted (below 4 kHz). (a) Interfering source at 0 dB SIR located at 30° . (b) Interfering source at 0 dB SIR located at 120° .

On the other hand, the values of $G_{\text{sig}}(R)$ and $G_{\text{int}}(R)$ could be lower in practice than the presented theoretical upper bounds as they depend on knowledge of the PSD of the desired and interfering sources, respectively, which need to be estimated from the noisy signal. This makes transmitting the unprocessed signal an attractive choice.

4.5. Transmitting Only Low Frequencies. It is well known that a closely spaced microphone array offers good performance

at high frequencies, and an array with a larger microphone spacing performs well at low frequencies. This observation can be exploited in binaural beamforming [8]. Figure 8 depicts the performance when only the low-frequency content of one microphone signal (up to 4 kHz) is transmitted from the right ear. As $G_{\text{raw}}(R)$ provided the best performance in the analysis so far, only this scheme is considered in this experiment. Each hearing aid is assumed to have two microphones. At the left ear, the low-frequency portion of the desired signal is estimated using the two locally available microphone signals and the transmitted signal. The high-frequency portion is estimated using only the local signals. The gain achieved in this setup at a rate R is denoted $G_{\text{raw}}^{\text{lp}}(R)$.

When an interferer is in the front half plane at, for example, 30° as in Figure 8(a), transmitting the low-frequency part alone results in poor performance. This is because the small microphone array at the left ear cannot distinguish between desired and interfering sources that are located close together. In this case, the binaural array is useful even at high frequencies. When the interferer is located in the rear half plane at, for example, 120° as in Figure 8(b), transmitting just the low-frequency part results in good performance. $G_{\text{raw}}^{\text{lp}}(R)$ reaches its limit at a lower bit-rate than $G_{\text{raw}}(R)$ as the high-frequency content need not be transmitted. At infinite bit-rate, $G_{\text{raw}}^{\text{lp}}(R)$ is lower than $G_{\text{raw}}(R)$, but such a scheme allows a trade-off between bit-rate and performance.

5. Conclusions

In a wireless binaural hearing aid system where each hearing aid has multiple microphones, the choice of the signal that one hearing aid needs to transmit to the other is not obvious. Optimally, the right hearing aid needs to transmit the part of the desired signal that can be predicted by the right ear signals but not the left ear signals, and vice versa for the left device [6]. At the receiving end, an estimate of the desired signal is obtained using the signal received over the wireless link and the locally observed signals. However, such an optimal scheme requires that the right device is aware of the joint statistics of the signals at both the left and right devices, which is impractical in the nonstationary conditions encountered in hearing aid applications. Suboptimal practical schemes are thus required.

Transmitting an estimate of the desired signal obtained at one hearing aid to the other is asymptotically optimal when the only undesired signal in the system is spatially uncorrelated noise [7]. In the presence of a localized interfering sound source the undesired signal is correlated across the different microphones and it has been seen that such a scheme is no longer optimal. Two alternative schemes have been proposed and investigated in this paper. The first is transmitting an estimate of the undesired signal, which performs better than transmitting an estimate of the desired signal depending on the location of the interfering sound source. The second is to simply transmit one of the unprocessed microphone signals from one device to the other. In the presence of a localized interferer or equivalently at low SIRs, the second scheme provides a significant gain

compared to the other suboptimal schemes. In the presence of uncorrelated noise or equivalently at high SIRs, however, there is approximately a 1 dB loss in performance compared to the method of [7]. While it is possible to change the transmission scheme depending on the SIR, a simple rule of thumb is to always transmit the unprocessed microphone signal as the penalty at high SIRs is negligible (see Figure 4) compared to the potential gains at low SIRs (see Figures 6 and 7). Furthermore, not having to obtain an estimate before transmission results in a lower computational load, and reduced delay, both of which are critical in hearing aid applications. It is to be noted that the results discussed in this paper apply when only a single interferer is present. Performance in the presence of multiple interferers is a topic for further study.

As a microphone array with a large interelement spacing as in a binaural hearing aid systems performs well only in the low frequencies, the effect of transmitting only the low-frequency content from the right hearing aid was also investigated. For interferers located in the rear half plane, a lower bit-rate is sufficient to maintain a similar level of performance as when the whole frequency range is transmitted. As the entire frequency range is not transmitted, the asymptotic performance is lower than the full-band transmission. Such a scheme, however, provides a trade-off between the required bit-rate and achievable beamforming gain.

References

- [1] V. Hamacher, J. Chalupper, J. Eggers, et al., "Signal processing in high-end hearing aids: state of the art, challenges, and future trends," *EURASIP Journal on Applied Signal Processing*, vol. 2005, no. 18, pp. 2915–2929, 2005.
- [2] S. Kochkin, "MarkeTrak VII: customer satisfaction with hearing instruments in the digital age," *The Hearing Journal*, vol. 58, no. 9, pp. 30–43, 2005.
- [3] T. J. Klases, T. van den Bogaert, M. Moonen, and J. Wouters, "Binaural noise reduction algorithms for hearing aids that preserve interaural time delay cues," *IEEE Transactions on Signal Processing*, vol. 55, no. 4, pp. 1579–1585, 2007.
- [4] S. Doclo, T. J. Klases, T. van den Bogaert, J. Wouters, and M. Moonen, "Theoretical analysis of binaural cue preservation using multi-channel Wiener filtering and interaural transfer functions," in *Proceedings of International Workshop on Acoustic Echo and Noise Control (IWAENC '06)*, Paris, France, September 2006.
- [5] S. Doclo, T. van den Bogaert, J. Wouters, and M. Moonen, "Comparison of reduced-bandwidth MWF-based noise reduction algorithms for binaural hearing aids," in *Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA '07)*, pp. 223–226, October 2007.
- [6] O. Roy and M. Vetterli, "Rate-constrained beamforming for collaborating hearing aids," in *Proceedings of IEEE International Symposium on Information Theory (ISIT '06)*, pp. 2809–2813, 2006.
- [7] O. Roy and M. Vetterli, "Collaborating hearing aids," in *Proceedings of MSRI Workshop on Mathematics of Relaying and Cooperation in Communication Networks*, April 2006.
- [8] S. Srinivasan, "Low-bandwidth binaural beamforming," *Electronics Letters*, vol. 44, no. 22, pp. 1292–1294, 2008.
- [9] T. Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression*, Information and System Sciences Series, T. Kailath, Ed, Prentice Hall, Upper Saddle River, NJ, USA, 1971.
- [10] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*, Prentice Hall, Upper Saddle River, NJ, USA, 2000.
- [11] R. O. Duda and W. L. Martens, "Range dependence of the response of a spherical head model," *The Journal of the Acoustical Society of America*, vol. 104, no. 5, pp. 3048–3058, 1998.