Research Article

Combination of Adaptive Feedback Cancellation and Binaural Adaptive Filtering in Hearing Aids

Anthony Lombard, Klaus Reindl, and Walter Kellermann

Multimedia Communications and Signal Processing, University of Erlangen-Nuremberg, Cauerstr. 7, 91058 Erlangen, Germany

Correspondence should be addressed to Anthony Lombard, lombard@lnt.de

We study a system combining adaptive feedback cancellation and adaptive filtering connecting inputs from both ears for signal enhancement in hearing aids. For the first time, such a binaural system is analyzed in terms of system stability, convergence of the algorithms, and possible interaction effects. As major outcomes of this study, a new stability condition adapted to the considered binaural scenario is presented, some already existing and commonly used feedback cancellation performance measures for the unilateral case are adapted to the binaural case, and possible interaction effects between the algorithms are identified. For illustration purposes, a blind source separation algorithm has been chosen as an example for adaptive binaural spatial filtering. Experimental results for binaural hearing aids confirm the theoretical findings and the validity of the new measures.

1. Introduction

Traditionally, signal enhancement techniques for hearing aids (HAs) were mainly developed independently for each ear [1–4]. However, since the human auditory system is a binaural system combining the signals received from both ears for audio perception, providing merely bilateral systems (that operate independently for each ear) to the hearing-aid user may distort crucial binaural information needed to localize sound sources correctly and to improve speech perception in noise. Foreseeing the availability of wireless technologies for connecting the two ears, several binaural processing strategies have therefore been presented in the last decade [5–10]. In [5], a binaural adaptive noise reduction algorithm exploiting one microphone signal from each ear has been proposed. Interaural time difference cues of speech signals were preserved by processing only the high-frequency components while leaving the low frequencies unchanged. Binaural spectral subtraction is proposed in [6]. It utilizes cross-correlation analysis of the two microphone signals for a more reliable estimation of the common noise power spectrum, without requiring stationarity for the interfering noise as the single-microphone versions do. Binaural multi-channel Wiener filtering approaches preserving binaural cues were also proposed, for example, in [7–9], and signal enhancement techniques based on blind source separation (BSS) were presented in [10].

Research on feedback suppression and control system theory in general has also given rise to numerous hearing-aid specific publications in recent years. The behavior of unilateral closed-loop systems and the ability of adaptive feedback cancellation algorithms to compensate for the feedback has been extensively studied in the literature (see, e.g., [11–15]). But despite the progress in binaural signal enhancement, binaural systems have not been considered in this context. In this paper, we therefore present a theoretical analysis of a binaural system combining adaptive feedback cancellation (AFC) and binaural adaptive filtering (BAF) techniques for signal enhancement in hearing aids.

The paper is organized as follows. An efficient binaural configuration combining AFC and BAF is described in Section 2. Generic vector/matrix notations are introduced for each part of the processing chain. Interaction effects concerning the AFC are then presented in Section 3. It includes a derivation of the ideal binaural AFC solution, a convergence analysis of the AFC filters based on the binaural Wiener solution, and a stability analysis of the binaural system. Interaction effects concerning the BAF are discussed

in Section 4. Here, to illustrate our argumentation, a BSS scheme has been chosen as an example for adaptive binaural filtering. Experimental conditions and results are finally presented in Sections 5 and 6 before providing concluding remarks in Section 7.

2. Signal Model

AFC and BAF techniques can be combined in two different ways. The feedback cancellation can be performed directly on the microphone inputs, or it can be applied at a later stage, to the BAF outputs. The second variant requires in general fewer filters but it has also several drawbacks. Actually, when the AFC comes after the BAF in the processing chain, the feedback cancellation task is complicated by the necessity to follow the continuously time-varying BAF filters. It may also significantly increase the necessary length of the AFC filters. Moreover, the BAF cannot benefit from the feedback cancellation effectuated by the AFC in this case. Especially at high HA amplification levels, the presence of strong feedback components in the sensor inputs may, therefore, seriously disturb the functioning of the BAF. These are structurally the same effects as those encountered when combining adaptive beamforming with acoustic echo cancellation (AEC) [16].

In this paper, we will therefore concentrate on the "AFC-first" alternative, where AFC is followed by the BAF. Figure 1 depicts the signal model adopted in this study. Each component of the signal model will be described separately in the following and generic vector/matrix notations will be introduced to carry out a general analysis of the overall system in Sections 3 and 4.

2.1. Notations. In this paper, lower-case boldface characters represent (row) vectors capturing signals or the filters of single-input-multiple-output (SIMO) systems. Accordingly, multiple-input-single-output (MISO) systems are described by transposed vectors. Matrices denoting multiple-input-multiple-output (MIMO) systems are represented by upper-case boldface characters. The transposition of a vector or a matrix will be denoted by the superscript $\{\cdot\}^T$.

2.2. The Microphone Signals. We consider here multi-sensor hearing aid devices with P microphones at each ear (see Figure 1), where P typically ranges between one and three. Because of the reverberation in the acoustical environment, Q point source signals s_q ($q = 1, \ldots, Q$) are filtered by a MIMO mixing system (one $Q \times P$ MIMO system for each ear in the figure) modeled by finite impulse response (FIR) filters. This can be expressed in the z-domain as:

$$x_{I_p}^s(z) = \sum_{q=1}^{Q} s_q(z) h_{qI_p}(z) \quad I \in \{L, R\}, \tag{1}$$

where $x_{I_p}^s(z)$ is the z-domain representation of the received source signal mixture at the pth sensor of the left ($I = L$) and right ($I = R$) hearing aid, respectively. $h_{qL_p}(z)$ and $h_{qR_p}(z)$ denote the transfer functions (polynomes of order up to several thousands typically) between the qth source

and the pth sensor at the left and right ears, respectively. One of the point sources may be seen as the target source to be extracted, the remaining $Q - 1$ being considered as interfering point sources. For the sake of simplicity, the z-transform dependency (z) will be omitted in the rest of this paper, as long as the notation is not ambiguous.

The acoustic feedback originating from the loudspeakers (LS) u_L and u_R at the left and right ears, respectively, is modeled by four $1 \times P$ SIMO systems of FIR filters. f_{LL_p} and f_{RL_p} represent the (z-domain) transfer functions (polynomes of order up to several hundreds typically) from the loudspeakers to the pth sensor on the left side, and f_{LR_p} and f_{RR_p} represent the transfer functions from the loudspeakers to the pth sensor on the right side. The feedback components captured by the pth microphone of each ear can therefore be expressed in the z-domain as

$$x_{I_p}^u = u_L f_{LI_p} + u_R f_{RI_p} \quad I \in \{L, R\}. \tag{2}$$

Note that as long as the energy of the two LS signals are comparable, the "cross" feedback signals (traveling from one ear to the other) are negligible compared to the "direct" feedback signals (occuring on each side independently). With the feedback paths (FBP) used in this study (see the description of the evaluation data in Section 5.3), an energy difference ranging from 15 to 30 dB has been observed between the "direct" and "cross" FBP impulse responses. When the HA gains are set at similar levels in both ears, the "cross" FBPs can then be neglected. But the impact of the "cross" feedback signals becomes more significant when a large difference exists between the two HA gains. Here, therefore, we explicitly account for the two types of feedback by modelling both the "direct" paths (with transfer functions f_{LL_p} and f_{RR_p}, $p = 1, \ldots, P$) and the "cross" paths (with transfer functions f_{RL_p} and f_{LR_p}, $p = 1, \ldots, P$) by FIR filters.

Diffuse noise signals n_{L_p} and n_{R_p}, $p = 1, \ldots, P$ constitute the last microphone signal components on the left and right ears, respectively. The z-domain representation of the pth sensor signal at each ear is finally given by:

$$x_{I_p} = x_{I_p}^s + x_{I_p}^n + x_{I_p}^u \quad I \in \{L, R\}. \tag{3}$$

This can be reformulated in a compact matrix form jointly capturing the P microphone signals of each HA:

$$\mathbf{x} = \mathbf{x}^s + \mathbf{x}^n + \mathbf{x}^u = \mathbf{sH} + \mathbf{x}^n + \mathbf{uF}, \tag{4}$$

where we have used the z-domain signal vectors

$$\mathbf{s} = [s_1, \ldots, s_Q], \tag{5}$$

$$\mathbf{x}_L^s = \left[x_{L_1}^s, \ldots, x_{L_P}^s\right], \tag{6}$$

$$\mathbf{x}_R^s = \left[x_{R_1}^s, \ldots, x_{R_P}^s\right], \tag{7}$$

$$\mathbf{x}^s = \left[\mathbf{x}_L^s \quad \mathbf{x}_R^s\right], \tag{8}$$

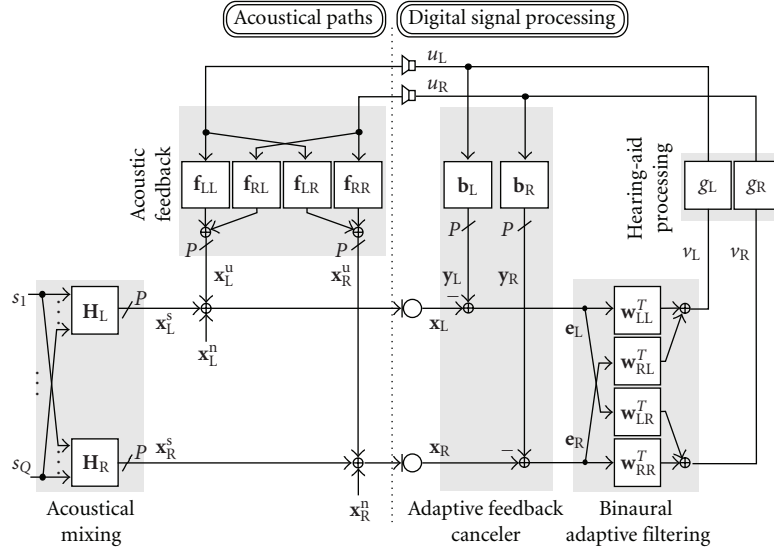$$\mathbf{u} = \left[u_L \quad u_R\right], \tag{9}$$

FIGURE 1: Signal model of the AFC-BAF combination.

as well as the *z*-domain matrices

$$\mathbf{H}_L = \begin{bmatrix} h_{1L_1} & \cdots & h_{1L_P} \\ \vdots & \ddots & \vdots \\ h_{QL_1} & \cdots & h_{QL_P} \end{bmatrix}, \tag{10}$$

$$\mathbf{H}_R = \begin{bmatrix} h_{1R_1} & \cdots & h_{1R_P} \\ \vdots & \ddots & \vdots \\ h_{QR_1} & \cdots & h_{QR_P} \end{bmatrix}, \tag{11}$$

$$\mathbf{H} = [\mathbf{H}_L \quad \mathbf{H}_R], \tag{12}$$

$$\mathbf{f}_{LL} = [f_{LL_1}, \ldots, f_{LL_P}], \tag{13}$$

$$\mathbf{f}_{RL} = [f_{RL_1}, \ldots, f_{RL_P}], \tag{14}$$

$$\mathbf{F}_L = \begin{bmatrix} \mathbf{f}_{LL}^T & \mathbf{f}_{RL}^T \end{bmatrix}^T, \tag{15}$$

$$\mathbf{f}_{LR} = [f_{LR_1}, \ldots, f_{LR_P}], \tag{16}$$

$$\mathbf{f}_{RR} = [f_{RR_1}, \ldots, f_{RR_P}], \tag{17}$$

$$\mathbf{F}_R = \begin{bmatrix} \mathbf{f}_{LR}^T & \mathbf{f}_{RR}^T \end{bmatrix}^T, \tag{18}$$

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_L & \mathbf{F}_R \end{bmatrix} = \begin{bmatrix} \mathbf{f}_{LL} & \mathbf{f}_{LR} \\ \mathbf{f}_{RL} & \mathbf{f}_{RR} \end{bmatrix}. \tag{19}$$

Furthermore, $\mathbf{x}^n$ and $\mathbf{x}^u$ capturing the noise and feedback components present in the microphone signals are defined in a similar way to $\mathbf{x}^s$. The sensor signal decomposition (4) can be further refined by distinguishing between target and interfering sources:

$$\mathbf{x}^s = \mathbf{x}^{s^{tar}} + \mathbf{x}^{s^{int}} = s^{tar}\mathbf{h}^{tar} + \mathbf{s}^{int}\mathbf{H}^{int}. \tag{20}$$

s^{tar} refers to the target source and $\mathbf{s}^{int}$ is a subset of $\mathbf{s}$ capturing the $Q - 1$ remaining interfering sources. $\mathbf{h}^{tar}$ is a row of $\mathbf{H}$ which captures the transfer functions from the target source to the sensors and $\mathbf{H}^{int}$ is a matrix containing the remaining $Q - 1$ rows of $\mathbf{H}$. Like the other vectors and matrices defined above, these four entities can be further decomposed into their left and right subsets, labeled with the indices L and R, respectively.

2.3. The AFC Processing. As can be seen from Figure 1, we apply here AFC to remove the feedback components present in the sensor signals, before passing them to the BAF. Feedback cancellation is achieved by trying to produce replicas of these undesired components, using a set of adaptive filters. The solution adopted here consists of two $1 \times P$ SIMO systems of adaptive FIR filters, with transfer functions b_{L_p} and b_{R_p} between the left (resp. right) loudspeaker and the pth sensor on the left (resp. right) side. The output

$$y_{I_p} = u_I b_{I_p} \quad I \in \{L, R\} \tag{21}$$

of the pth filter on the left (resp. right) side is then subtracted from the pth sensor signal on the left (resp. right) side, producing a residual signal

$$e_{I_p} = x_{I_p} - y_{I_p} \quad I \in \{L, R\}, \tag{22}$$

which is, ideally, free of any feedback components. (21) and (22) can be reformulated in matrix form as follows:

$$\mathbf{e} = \mathbf{x} - \mathbf{y} = \mathbf{x} - \mathbf{u}\mathbf{B}, \tag{23}$$

with the block-diagonal constraint

$$\mathbf{B} \stackrel{!}{=} \mathbf{B}^c = \begin{bmatrix} \mathbf{b}_L & 0 \\ 0 & \mathbf{b}_R \end{bmatrix} \tag{24}$$

put on the AFC system. The vectors $\mathbf{e}$ and $\mathbf{y}$, capturing the z-domain representations of the residual and AFC output signals, respectively, are defined in analogous way to $\mathbf{x}^s$ in (8). As can be seen from (21) and (22), we perform here bilateral feedback cancellation (as opposed to binaural operations) since AFC is performed for each ear separately. This is reflected in (24), where we force the off-diagonal terms to be zero instead of reproducing the acoustic feedback system $\mathbf{F}$ with its set of four SIMO systems. The reason for this will become clear in Section 3.1. Guidelines regarding an arbitrary (i.e., unconstrained) AFC system $\mathbf{B}$ (defined similarly to $\mathbf{F}$ in this case) will also be provided at some points in the paper. The superscript $\{\cdot\}^c$ is used to distinguish constrained systems $\mathbf{B}^c$ defined by (24) from arbitrary (unconstrained) systems $\mathbf{B}$ (with possibly non-zero off-diagonal terms).

2.4. The BAF Processing. The BAF filters perform spatial filtering to enhance the signal coming from one of the Q external point sources. This is performed here binaurally, that is, by combining signals from both ears (see Figure 1). The binaural filtering operations can be described by a set of four $P \times 1$ MISO systems of adaptive FIR filters. This can be expressed in the z-domain as follows:

$$v_{\mathrm{I}} = \sum_{p=1}^{P} e_{\mathrm{L}_p} w_{\mathrm{L}_p \mathrm{I}} + e_{\mathrm{R}_p} w_{\mathrm{R}_p \mathrm{I}} \quad \mathrm{I} \in \{\mathrm{L}, \mathrm{R}\}, \qquad (25)$$

where $w_{\mathrm{L}_p\mathrm{I}}$ and $w_{\mathrm{R}_p\mathrm{I}}$, $p = 1,\ldots,P$, $\mathrm{I} \in \{\mathrm{L}, \mathrm{R}\}$ are the transfer functions applied on the pth sensor of the left and right hearing aids, respectively. To reformulate (25) in matrix form, we define the vector

$$\mathbf{v} = \begin{bmatrix} v_{\mathrm{L}} & v_{\mathrm{R}} \end{bmatrix}, \qquad (26)$$

which jointly captures the z-domain representations of the two BAF outputs, and the vector and matrices

$$\mathbf{w}_{\mathrm{LL}} = \begin{bmatrix} w_{\mathrm{L}_1\mathrm{L}}, \ldots, w_{\mathrm{L}_P\mathrm{L}} \end{bmatrix}, \qquad (27)$$

$$\mathbf{w}_{\mathrm{RL}} = \begin{bmatrix} w_{\mathrm{R}_1\mathrm{L}}, \ldots, w_{\mathrm{R}_P\mathrm{L}} \end{bmatrix}, \qquad (28)$$

$$\mathbf{w}_{\mathrm{L}} = \begin{bmatrix} \mathbf{w}_{\mathrm{LL}} & \mathbf{w}_{\mathrm{RL}} \end{bmatrix}, \qquad (29)$$

$$\mathbf{w}_{\mathrm{LR}} = \begin{bmatrix} w_{\mathrm{L}_1\mathrm{R}}, \ldots, w_{\mathrm{L}_P\mathrm{R}} \end{bmatrix}, \qquad (30)$$

$$\mathbf{w}_{\mathrm{RR}} = \begin{bmatrix} w_{\mathrm{R}_1\mathrm{R}}, \ldots, w_{\mathrm{R}_P\mathrm{R}} \end{bmatrix}, \qquad (31)$$

$$\mathbf{w}_{\mathrm{R}} = \begin{bmatrix} \mathbf{w}_{\mathrm{LR}} & \mathbf{w}_{\mathrm{RR}} \end{bmatrix}, \qquad (32)$$

$$\mathbf{W} = \begin{bmatrix} \mathbf{w}_{\mathrm{L}}^T & \mathbf{w}_{\mathrm{R}}^T \end{bmatrix} = \begin{bmatrix} \mathbf{w}_{\mathrm{LL}}^T & \mathbf{w}_{\mathrm{LR}}^T \\ \mathbf{w}_{\mathrm{RL}}^T & \mathbf{w}_{\mathrm{RR}}^T \end{bmatrix}, \qquad (33)$$

related to the transfer functions of the MIMO BAF system. We can finally express (25) as:

$$\mathbf{v} = \mathbf{e}\,\mathbf{W}. \qquad (34)$$

2.5. The Forward Paths. Conventional HA processing (mainly a gain correction) is performed on the output of the AFC-BAF combination, before being played back by the loudspeakers:

$$u_{\mathrm{I}} = v_{\mathrm{I}}\,g_{\mathrm{I}} \quad \mathrm{I} \in \{\mathrm{L}, \mathrm{R}\}, \qquad (35)$$

where g_{L} and g_{R} model the HA processing in the z-domain, at the left and right ears, respectively. In the literature, this part of the processing chain is often referred to as the forward path (in opposition to the acoustic feedback path). To facilitate the analysis, we will assume that the HA processing is linear and time-invariant (at least between two adaptation steps) in this study. (35) can be conveniently written in matrix form as:

$$\mathbf{u} = \mathbf{v}\,\mathrm{Diag}\,\{\mathbf{g}\}, \qquad (36)$$

with

$$\mathbf{g} = \begin{bmatrix} g_{\mathrm{L}} & g_{\mathrm{R}} \end{bmatrix}. \qquad (37)$$

The Diag$\{\cdot\}$ operator applied to a vector builds a diagonal matrix with the vector entries placed on the main diagonal.

Note that for simplicity, we assumed that the number of sensors P used on each device for digital signal processing was equal. The above notations as well as the following analysis are however readily applicable to asymmetrical configurations also, simply by resizing the above-defined vectors and matrices, or by setting the corresponding microphone signals and all the associated transfer functions to zero. In particular, the unilateral case can be seen as a special case of the binaural structure discussed in this paper, with one or more microphones used on one side, but none on the other side.

3. Interaction Effects on the Feedback Cancellation

The structure depicted in Figure 1 for binaural HAs mainly deviates from the well-known unilateral case by the presence of binaural spatial filtering. The binaural structure is characterized by a significantly more complex closed-loop system, possibly with multiple microphone inputs, but most importantly with two connected LS outputs, which considerably complicates the analysis of the system. However, we will see in the following how, under certain conditions, we can exploit the compact matrix notations introduced in the previous section, to describe the behavior of the closed-loop system. We will draw some interesting conclusions on the present binaural system, emphasizing its deviation from the standard unilateral case in terms of ideal cancellation solution, convergence of the AFC filters and system stability.

3.1. The Ideal Binaural AFC Solution. In the unilateral and single-channel case, the adaptation of the (single) AFC filter tries to adjust the compensation signal (the filter output) to the (single-channel) acoustic feedback signal. Under ideal conditions, this approach guarantees perfect removal of the undesired feedback components and simultaneously prevents the occurrence of howling caused by system instabilities
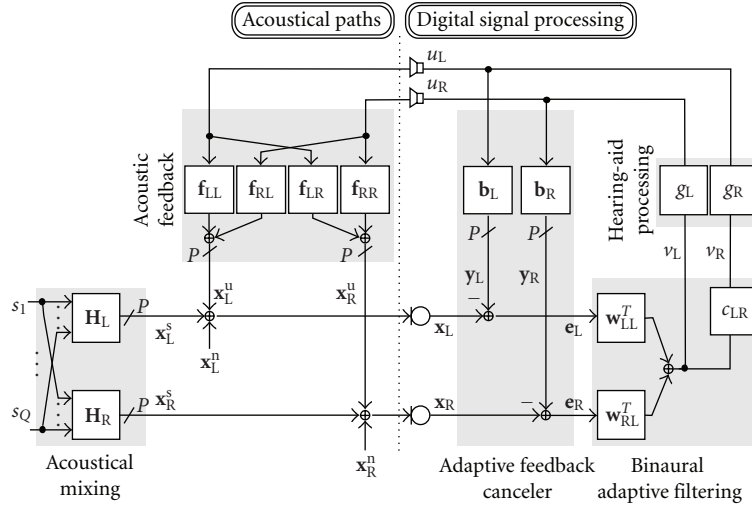
FIGURE 2: Equivalent signal model of the AFC-BAF combination under the assumption (40).

[11] (the stability of the binaural closed-loop system will be discussed in Section 3.3). The adaptation of the filter coefficients towards the desired solution is usually achieved using a gradient-descent-like learning rule, in its simplest form using the least mean square (LMS) algorithm [17]. The functioning of the AFC in the binaural configuration shown in Figure 1 is similar.

The residual signal vector (23) can be decomposed into its source, noise and feedback components using (4):

$$\mathbf{e} = \mathbf{x}^s + \mathbf{x}^n + \underbrace{\mathbf{u}(\mathbf{F} - \mathbf{B})}_{\mathbf{e}^{\mathrm{FB}}}, \qquad (38)$$

where $\mathbf{B}$ denotes an arbitrary (unconstrained) AFC system matrix (Section 2.3). $\mathbf{e}^{\mathrm{FB}} = \left[\mathbf{e}_L^{\mathrm{FB}} \ \mathbf{e}_R^{\mathrm{FB}}\right] = [e_{L_1}^{\mathrm{FB}}, \ldots, e_{L_P}^{\mathrm{FB}}, e_{R_1}^{\mathrm{FB}}, \ldots, e_{R_P}^{\mathrm{FB}}]$ captures the z-domain representations of the residual feedback components to be removed by the AFC. The only way to perfectly remove the feedback components from the residual signals (i.e., $\mathbf{e}^{\mathrm{FB}} = 0$), for arbitrary output signal vectors $\mathbf{u}$, is to have

$$\mathbf{B} = \mathbf{F} \stackrel{\triangle}{=} \hat{\mathbf{B}}. \qquad (39)$$

$\hat{\mathbf{B}}$ denotes the ideal AFC solution in the *unconstrained* case. This is the binaural analogon to the ideal AFC solution in the unilateral case, where perfect cancellation is achieved by reproducing an exact replica of the acoustical FBP. In practice, this solution is however very difficult to reach adaptively because it requires the two signals u_L and u_R to be uncorrelated, which is obviously not fulfilled in our binaural HA scenario since the two HAs are connected (the correlation is actually highly desirable since the HAs should form a spatial image of the acoustic scene, which implies that the two LS signals must be correlated to reflect interaural time and level differences). This problem has been extensively described in the literature on multi-channel AEC, where it is referred to as the "non-uniqueness problem". Several attempts have been reported in the literature to partly

alleviate this issue (see, e.g., [18–20]). These techniques may be useful in the HA case also, but this is beyond the scope of the present work.

In this paper, instead of trying to solve the problem mentioned above, we explicitly account for the correlation of the two LS output signals. The relation between the HA outputs can be tracked back to the relation existing between the BAF outputs v_L and v_R (Figure 1), which are generated from the same set of sensors and aim at reproducing a binaural impression of the same acoustical scene. The relation between v_L and v_R can be described by a linear operator $c_{LR}(z)$ transforming $v_L(z)$ into $v_R(z)$ such that:

$$v_R = v_L c_{LR} \quad \forall v_L, \qquad (40)$$

which is actually perfectly true if and only if c_{LR} transforms $\mathbf{w}_L$ into $\mathbf{w}_R$:

$$\mathbf{w}_R = \mathbf{w}_L c_{LR}. \qquad (41)$$

Therefore, the assumption (40) will only be an approximation in general, except for a specific class of BAF systems satisfying (41). The BSS algorithm discussed in Section 4 belongs to this class. Figure 2 shows the equivalent signal model resulting from (40). As can be seen from the figure, c_{LR} can be equivalently considered as being part of the right forward path to further simplify the analysis. Accordingly, we then define the new vector

$$\tilde{\mathbf{g}} = \begin{bmatrix} \tilde{g}_L & \tilde{g}_R \end{bmatrix} = \begin{bmatrix} g_L & c_{LR}g_R \end{bmatrix} \qquad (42)$$

jointly capturing c_{LR} and the HA processing. Provided that g_L and g_R are linear, (41) (and hence (40)) is equivalent to assuming the existence of a linear dependency between the LS outputs, which we can express as follows:

$$\mathbf{u} = v_L \tilde{\mathbf{g}} = \frac{u_L}{\tilde{g}_L} \tilde{\mathbf{g}} = \frac{u_R}{\tilde{g}_R} \tilde{\mathbf{g}}. \qquad (43)$$

This assumption implies that only one filter (instead of two, one for each LS signal) suffices to cancel the feedback components in each sensor channel. It corresponds to the constraint (24) mentioned in Section 2.3, which forces the AFC system matrix $\mathbf{B}$ to be block-diagonal ($\mathbf{B} \overset{!}{=} \mathbf{B}^c$). The required number of AFC filters reduces accordingly from $2 \times 2P$ to $2P$.

Using the constraint (24) and the assumption (43) in (38), we can derive the *constrained* ideal AFC solution minimizing $\mathbf{e}_I^{FB}$, $I \in \{L, R\}$, considering each side separately:

$$\mathbf{e}_I^{FB} = \mathbf{u}\mathbf{F}_I - u_I \mathbf{b}_I$$

$$= \frac{u_I}{\widetilde{g}_I} \widetilde{\mathbf{g}}\mathbf{F}_I - u_I \mathbf{b}_I$$

$$= u_I \left[\underbrace{\widetilde{\mathbf{g}}\mathbf{F}_I \widetilde{g}_I^{-1}}_{\widehat{\mathbf{b}}_I} - \mathbf{b}_I \right] \quad I \in \{L, R\}. \tag{44}$$

Here, $\widehat{\mathbf{b}}_I$ denote the ideal AFC solution for the left or right HA. It can be easily verified that inserting (44) into (23) leads to the following residual signal decomposition:

$$\mathbf{e} = \mathbf{x}^s + \mathbf{x}^n + \underbrace{\mathbf{u}\left(\widehat{\mathbf{B}}^c - \mathbf{B}^c\right)}_{\mathbf{e}^{FB}}, \tag{45}$$

where

$$\widehat{\mathbf{B}}^c = \mathrm{Bdiag}\left\{\widehat{\mathbf{b}}_L, \widehat{\mathbf{b}}_R\right\} \tag{46}$$

denotes the ideal AFC solution when $\mathbf{B}$ is constrained to be block-diagonal ($\mathbf{B} \overset{!}{=} \mathbf{B}^c$) and under the assumption (43). The $\mathrm{Bdiag}\{\cdot\}$ operator is the block-wise counterpart of the $\mathrm{Diag}\{\cdot\}$ operator. Applied to a list of vectors, it builds a block-diagonal matrix with the listed vectors placed on the main diagonal of the block-matrix, respectively.

To illustrate these results, we expand the ideal AFC solution (46) using (15) and (18):

$$\widehat{\mathbf{b}}_L = (\widetilde{g}_L \mathbf{f}_{LL} + \widetilde{g}_R \mathbf{f}_{RL}) \widetilde{g}_L^{-1}$$

$$= \underbrace{\mathbf{f}_{LL}}_{\text{direct}} + \underbrace{\widetilde{g}_R/\widetilde{g}_L \mathbf{f}_{RL}}_{\text{cross}},$$

$$\widehat{\mathbf{b}}_R = (\widetilde{g}_R \mathbf{f}_{RR} + \widetilde{g}_L \mathbf{f}_{LR}) \widetilde{g}_R^{-1} \tag{47}$$

$$= \underbrace{\mathbf{f}_{RR}}_{\text{direct}} + \underbrace{\widetilde{g}_R/\widetilde{g}_L \mathbf{f}_{RL}}_{\text{cross}}.$$

For each filter, we can clearly identify two terms due to, respectively, the "direct" and "cross" FBPs (see Section 2.2). Contrary to the "direct" terms, the "cross" terms are identifiable only under the assumption (43) that the LS outputs are linearly dependent. Should this assumption not hold because of, for example, some non-linearities in the forward paths, the "cross" FBPs would not be completely identifiable. The feedback signals propagating from one ear to the other would then act as a disturbance to the AFC adaptation process. Note, however, that since the amplitude of the "cross" FBPs is negligible compared to the amplitude of the "direct" FBPs (Section 2.2), the consequences would be very limited as long as the HA gains are set to similar amplification levels, as can be seen from (47). It should also be noted that the forward path generally includes some (small) decorrelation delays D_L and D_R to help the AFC filters to converge to their desired solution (see Section 3.2). If those delays are set differently for each ear, causality of the "cross" terms in (47) will not always be guaranteed, in which case the ideal solution will not be achievable with the present scheme. This situation can be easily avoided by either setting the decorrelation delays $D_L = D_R$ equal for each ear (which appears to be the most reasonable choice to avoid artificial interaural time differences), or by delaying the LS signals (but using the non-delayed signals as AFC filter inputs). However, since it would further increase the overall delay from the microphone inputs to the LS outputs, the latter choice appears unattractive in the HA scenario.

3.2. The Binaural Wiener AFC Solution. In the configuration depicted in Figure 2, similar to the standard unilateral case (see, e.g., [12]), conventional gradient-descent-based learning rules do not lead to the ideal solution discussed in Section 3.1 but to the so-called Wiener solution [17]. Actually, instead of minimizing the feedback components $\mathbf{e}^{FB}$ in the residual signals, the AFC filters are optimized by minimizing the mean-squared error of the *overall* residual signals (38).

In the following, we conduct therefore a convergence analysis of the binaural system depicted in Figure 2, by deriving the Wiener solution of the system in the frequency domain:

$$\mathbf{b}_I^{\mathrm{Wiener}}\left(z = e^{j\omega}\right) = \mathbf{r}_{\mathbf{x}_I u_I}\left(e^{j\omega}\right) r_{u_I u_I}^{-1}\left(e^{j\omega}\right)$$

$$= \left(\mathbf{r}_{\mathbf{u} u_I}\mathbf{F}_I + \mathbf{r}_{\mathbf{x}_I^s u_I} + \mathbf{r}_{\mathbf{x}_I^n u_I}\right) r_{u_I u_I}^{-1} \tag{48}$$

$$= \underbrace{\widetilde{\mathbf{g}}\mathbf{F}_I \widetilde{g}_I^{-1}}_{\widehat{\mathbf{b}}_I(z=e^{j\omega})} + \underbrace{\mathbf{r}_{\mathbf{x}_I^s u_I} r_{u_I u_I}^{-1} + \mathbf{r}_{\mathbf{x}_I^n u_I} r_{u_I u_I}^{-1}}_{\check{\mathbf{b}}_I(z=e^{j\omega})} \quad I \in \{L, R\}, \tag{49}$$

where the frequency dependency $(e^{j\omega})$ was omitted in (48) and (49) for the sake of simplicity, like in the rest of this section. $\widehat{\mathbf{b}}_I(z = e^{j\omega})$ is recognized as the (frequency-domain) ideal AFC solution discussed in Section 3.1, and $\check{\mathbf{b}}_I(z = e^{j\omega})$ denotes a (frequency-domain) bias term. The assumption (43) has been exploited in (48) to obtain the above final result. $r_{u_I u_I}$ represents the (auto-) power spectral density of u_I, $I \in \{L, R\}$, and $\mathbf{r}_{\mathbf{x}_I u_I} = [r_{x_{I_1} u_I}, \ldots, r_{x_{I_P} u_I}]$, $I \in \{L, R\}$, is a vector capturing cross-power spectral densities. The cross-power spectral density vectors $\mathbf{r}_{\mathbf{x}_I^s u_I}$ and $\mathbf{r}_{\mathbf{x}_I^n u_I}$ are defined in a similar way.

The Wiener solution (49) shows that the optimal solution is biased due to the correlation of the different source contributions $\mathbf{x}^s$ and $\mathbf{x}^n$ with the reference inputs $\mathbf{u}_I$, $I \in \{L, R\}$ (i.e., the LS outputs), of the AFC filters. The bias term $\check{\mathbf{b}}_I$ in (49) can be further decomposed like in (20),

distinguishing between desired (target source) and undesired (interfering point sources and diffuse noise) sound sources:

$$\check{\mathbf{b}}_{\mathrm{I}}^{\mathrm{Wiener}}\left(e^{j\omega}\right) = \underbrace{\mathbf{r}_{\mathbf{x}_{\mathrm{I}}^{\mathrm{star}} u_{\mathrm{I}}} r_{u_{\mathrm{I}} u_{\mathrm{I}}}^{-1}}_{\text{due to target source}}$$

$$+ \underbrace{\mathbf{r}_{\mathbf{x}_{\mathrm{I}}^{\mathrm{sint}} u_{\mathrm{I}}} r_{u_{\mathrm{I}} u_{\mathrm{I}}}^{-1} + \mathbf{r}_{\mathbf{x}_{\mathrm{I}}^{\mathrm{n}} u_{\mathrm{I}}} r_{u_{\mathrm{I}} u_{\mathrm{I}}}^{-1}}_{\text{due to undesired sources}} \quad \mathrm{I} \in \{\mathrm{L}, \mathrm{R}\}.$$

$$(50)$$

By nature, the spatially uncorrelated diffuse noise components $\mathbf{x}^{\mathrm{n}}$ will be only weakly correlated with the LS outputs. The third bias term will have therefore only a limited impact on the convergence of the AFC filters. The diffuse noise sources will mainly act as a disturbance. Depending on the signal enhancement technique used, they might even be partly removed. But above all, the (multi-channel) BAF performs spatial filtering, which mainly affects the interfering point sources. Ideally, the interfering sources may even vanish from the LS outputs, in which case the second bias term would simply disappear. In practice, the interference sources will never be completely removed. Hence the amount of bias introduced by the interfering sources will largely depend on the interference rejection performance of the BAF. However, like in the unilateral hearing aids, the main source of estimation errors comes from the target source. Actually, since the BAF aims at producing outputs which are as close as possible to the original target source signal, the first bias term due to the (spectrally colored) target source will be much more problematic.

One simple way to reduce the correlation between the target source and the LS outputs is to insert some delays D_{L} and D_{R} in the forward paths [12]. The benefit of this method is however very limited in the HA scenario where only tiny processing delays (5 to 10 ms for moderate hearing losses) are allowed to avoid noticeable effects due to unprocessed signals leaking into the ear canal and interfering with the processed signals. Other more complicated approaches applying a prewhitening of the AFC inputs have been proposed for the unilateral case [21, 22], which could also help in the binaural case. We may also recall a well-known result from the feedback cancellation literature: the bias of the AFC solution decreases when the HA gain increases, that is, when the signal-to-feedback ratio (SFR) at the AFC inputs (the microphones) decreases. This statement also applies to the binaural case. This can be easily seen from (50) where the auto-power spectral density $r_{u_{\mathrm{I}} u_{\mathrm{I}}}^{-1}$ decreases quadratically whereas the cross-power spectral densities increase only linearly with increasing LS signal levels.

Note that the above derivation of the Wiener solution has been performed under the assumption (43) that the LS outputs are linearly dependent. When this assumption does not hold, an additional term appears in the Wiener solution. We may illustrate this exemplarily for the left side, starting from (48):

$$\mathbf{b}_{\mathrm{L}}^{\mathrm{Wiener}}\left(e^{j\omega}\right) = \underbrace{\mathbf{f}_{\mathrm{LL}} + r_{u_{\mathrm{R}} u_{\mathrm{L}}} r_{u_{\mathrm{L}} u_{\mathrm{L}}}^{-1} \mathbf{f}_{\mathrm{RL}}}_{\text{desired solution}} + \underbrace{\mathbf{r}_{\mathbf{x}_{\mathrm{L}}^{\mathrm{s}} u_{\mathrm{L}}} r_{u_{\mathrm{L}} u_{\mathrm{L}}}^{-1} + \mathbf{r}_{\mathbf{x}_{\mathrm{L}}^{\mathrm{n}} u_{\mathrm{L}}} r_{u_{\mathrm{L}} u_{\mathrm{L}}}^{-1}}_{\text{bias}}.$$

$$(51)$$

The bias term is identical to the one already obtained in (50), while the desired term is now split into two parts. The first one is related to the "direct" FBPs. The second term involves the "cross" FBPs and shows that gradient-based optimization algorithms will try to exploit the correlation of the LS outputs (when existing) to remove the feedback signal components traveling from one ear to the other. In the extreme case that the two LS signals are totally decorrelated (i.e., $r_{u_{\mathrm{R}} u_{\mathrm{L}}} = 0$), this term disappears and the "cross" feedback signals cannot be compensated. Note, however, that it would only have a very limited impact as long as the HA gains are set to similar amplification levels, as we saw in Section 3.1.

3.3. The Binaural Stability Condition. In this section, we formulate the stability condition of the binaural closed-loop system, starting from the general case before applying the block-diagonal constraint (24). We first need to express the responses u_{L} and u_{R} of the binaural system (Figure 1) on the left and right side, respectively, to an external excitation $\mathbf{x}^{\mathrm{s}} + \mathbf{x}^{\mathrm{n}}$. This can be done in the z-domain as follows:

$$u_{\mathrm{L}} = [\mathbf{x}^{\mathrm{s}} + \mathbf{x}^{\mathrm{n}} + \mathbf{u}(\mathbf{F} - \mathbf{B})]\mathbf{w}_{\mathrm{L}}^{T} g_{\mathrm{L}}$$

$$= \underbrace{(\mathbf{x}^{\mathrm{s}} + \mathbf{x}^{\mathrm{n}})\mathbf{w}_{\mathrm{L}}^{T} g_{\mathrm{L}}}_{\hat{u}_{\mathrm{L}}} + u_{\mathrm{L}}\underbrace{(\mathbf{F}_{\mathrm{L:}} - \mathbf{B}_{\mathrm{L:}})\mathbf{w}_{\mathrm{L}}^{T} g_{\mathrm{L}}}_{k_{\mathrm{LL}}}$$

$$+ u_{\mathrm{R}}\underbrace{(\mathbf{F}_{\mathrm{R:}} - \mathbf{B}_{\mathrm{R:}})\mathbf{w}_{\mathrm{L}}^{T} g_{\mathrm{L}}}_{k_{\mathrm{RL}}}$$

$$= \frac{\hat{u}_{\mathrm{L}} + u_{\mathrm{R}} k_{\mathrm{RL}}}{1 - k_{\mathrm{LL}}}, \quad (52)$$

$$u_{\mathrm{R}} = [\mathbf{x}^{\mathrm{s}} + \mathbf{x}^{\mathrm{n}} + \mathbf{u}(\mathbf{F} - \mathbf{B})]\mathbf{w}_{\mathrm{R}}^{T} g_{\mathrm{R}}$$

$$= \underbrace{(\mathbf{x}^{\mathrm{s}} + \mathbf{x}^{\mathrm{n}})\mathbf{w}_{\mathrm{R}}^{T} g_{\mathrm{R}}}_{\hat{u}_{\mathrm{R}}} + u_{\mathrm{L}}\underbrace{(\mathbf{F}_{\mathrm{L:}} - \mathbf{B}_{\mathrm{L:}})\mathbf{w}_{\mathrm{R}}^{T} g_{\mathrm{R}}}_{k_{\mathrm{LR}}}$$

$$+ u_{\mathrm{R}}\underbrace{(\mathbf{F}_{\mathrm{R:}} - \mathbf{B}_{\mathrm{R:}})\mathbf{w}_{\mathrm{R}}^{T} g_{\mathrm{R}}}_{k_{\mathrm{RR}}}$$

$$= \frac{\hat{u}_{\mathrm{R}} + u_{\mathrm{L}} k_{\mathrm{LR}}}{1 - k_{\mathrm{RR}}}, \quad (53)$$

where $\mathbf{F}_{\mathrm{L:}}$ and $\mathbf{B}_{\mathrm{L:}}$ denote the first row of $\mathbf{F}$ and $\mathbf{B}$, respectively, that is, the transfer functions applied to the left LS signal. $\mathbf{F}_{\mathrm{R:}}$ and $\mathbf{B}_{\mathrm{R:}}$ denote the second row of $\mathbf{F}$ and $\mathbf{B}$, respectively, that is, the transfer functions applied to the right LS signal. $\hat{u}_{\mathrm{L}}$ and $\hat{u}_{\mathrm{R}}$ represent the z-domain representations of the ideal system responses, once the feedback signals have been completely removed:

$$\hat{\mathbf{u}} = \begin{bmatrix} \hat{u}_{\mathrm{L}} & \hat{u}_{\mathrm{R}} \end{bmatrix} = (\mathbf{x}^{\mathrm{s}} + \mathbf{x}^{\mathrm{n}})\mathbf{W}\,\mathrm{Diag}\,\{\mathbf{g}\}. \quad (54)$$

k_{LL}, k_{RL}, k_{LR}, and k_{RR} can be interpreted as the open-loop transfer functions (OLTFs) of the system. They can be seen as the entries of the OLTF matrix $\mathbf{K}$ defined as follows:

$$\mathbf{K} = \begin{bmatrix} k_{\mathrm{LL}} & k_{\mathrm{LR}} \\ k_{\mathrm{RL}} & k_{\mathrm{RR}} \end{bmatrix} = (\mathbf{F} - \mathbf{B})\mathbf{W}\,\mathrm{Diag}\,\{\mathbf{g}\}. \quad (55)$$

Combining (52) and (53) finally yields the relations:

$$u_{\mathrm{L}} = \frac{(1 - k_{\mathrm{RR}})\hat{u}_{\mathrm{L}} + k_{\mathrm{RL}}\hat{u}_{\mathrm{R}}}{1 - k},$$

$$u_{\mathrm{R}} = \frac{(1 - k_{\mathrm{LL}})\hat{u}_{\mathrm{R}} + k_{\mathrm{LR}}\hat{u}_{\mathrm{L}}}{1 - k}, \tag{56}$$

with

$$
\begin{aligned}
k &= k_{\mathrm{LL}} + k_{\mathrm{RR}} + k_{\mathrm{LR}}k_{\mathrm{RL}} - k_{\mathrm{LL}}k_{\mathrm{RR}} \\
&= \operatorname{tr}\{\mathbf{K}\} - \det\{\mathbf{K}\},
\end{aligned} \tag{57}
$$

where the operators $\operatorname{tr}\{\cdot\}$ and $\det\{\cdot\}$ denote the trace and determinant of a matrix, respectively.

Similar to the unilateral case [11], (56) indicate that the binaural closed-loop system is stable as long as the magnitude of $k(z = e^{j\omega})$ does not exceed one for any angular frequency ω:

$$\left| k\left(z = e^{j\omega}\right) \right| < 1, \quad \forall \omega. \tag{58}$$

Here, the phase condition has been ignored, as usual in the literature on AFC [14]. Note that the function k in (57) and hence the stability of the binaural system, depend on the current state of the BAF filters.

The above derivations are valid in the general case. No particular assumption has been made and the AFC system has not been constrained to be block-diagonal. In the following, we will consider the class of algorithms satisfying the assumption (41), implying that the two BAF outputs are linearly dependent. In this case, the ideal system output vector (54) becomes

$$\hat{\mathbf{u}} = (\mathbf{x}^{\mathrm{s}} + \mathbf{x}^{\mathrm{n}})\mathbf{w}_{\mathrm{L}}^{T}\widetilde{\mathbf{g}}. \tag{59}$$

Furthermore, it can easily be verified that the following relations are satisfied in this case:

$$k_{\mathrm{RL}}\hat{u}_{\mathrm{R}} = k_{\mathrm{RR}}\hat{u}_{\mathrm{L}}, \tag{60}$$

$$k_{\mathrm{LR}}\hat{u}_{\mathrm{L}} = k_{\mathrm{LL}}\hat{u}_{\mathrm{R}}, \tag{61}$$

$$\det\{\mathbf{K}\} = 0. \tag{62}$$

The closed-loop response (56) of the binaural system simplifies, therefore, in this case to

$$\mathbf{u} = \frac{1}{1 - k}\,\hat{\mathbf{u}}, \tag{63}$$

where k, defined in (57), reduces to

$$k = \operatorname{tr}\{\mathbf{K}\}. \tag{64}$$

Finally, when applying additionally the block-diagonal constraint (24) on the AFC system, (64) further simplifies to

$$k = \widetilde{\mathbf{g}}\left(\hat{\mathbf{B}}^{\mathrm{c}} - \mathbf{B}^{\mathrm{c}}\right)\mathbf{w}_{\mathrm{L}}^{T}. \tag{65}$$

The stability condition (58) formulated on k for the general case still applies here.

The above results show that in the unconstrained (constrained, resp.) case, when the AFC filters reach their ideal solution $\mathbf{B} = \mathbf{F}$ ($\mathbf{B}^{\mathrm{c}} = \hat{\mathbf{B}}^{\mathrm{c}}$, resp.), the function k in (57) ((65), resp.) is equal to zero. Hence the stability condition (58) is always fulfilled, regardless of the HA amplification levels used, and the LS outputs become ideal, with $\mathbf{u} = \hat{\mathbf{u}}$ as expected.

4. Interaction Effects on the Binaural Adaptive Filtering

The presence of feedback in the microphone signals is usually not taken into account when developing signal enhancement techniques for hearing aids. In this section, we consider the configuration depicted in Figure 1 and focus exemplarily on BSS techniques as possible candidates to implement the BAF, thereby analyzing the impact of feedback on BSS and discussing possible interaction effects with an AFC algorithm.

4.1. Overview on Blind Source Separation. The aim of blind source separation is to recover the original source signals from an observed set of signal mixtures. The term "blind" implies that the mixing process and the original source signals are unknown. In acoustical scenarios, like in the hearing-aid application, the source signals are mixed in a convolutive manner. The (convolutive) acoustical mixing system can be modeled as a MIMO system $\mathbf{H}$ of FIR filters (see Section 2.2). The case where the number Q of (simultaneously active) sources is equal to the number $2 \times P$ of microphones (assuming P channels for each ear (see Section 2.2)) is referred to as the *determined* case. The case where $Q < 2 \times P$ is called *overdetermined*, while $Q > 2 \times P$ is denoted as *underdetermined*.

The *underdetermined* BSS problem can be handled based on time-frequency masking techniques, which rely on the sparseness of the sound sources (see, e.g., [23, 24]). In this paper, we assume that the number of sources does not exceed the number of microphones. Separation can then be performed using independent component analysis (ICA) methods, merely under the assumption of statistical independence of the original source signals [25]. ICA achieves separation by applying a demixing MIMO system $\mathbf{A}$ of FIR filters on the microphone signals, hence providing an estimate of each source at the outputs of the demixing system. This is achieved by adapting the weights of the demixing filters to force the output signals to become statistically independent. Because of the adaptation criterion exploiting the independence of the sources, a distinction between desired and undesired sources is unnecessary. Adaptation of the BSS filters is therefore possible even when all sources are simultaneously active, in contrast to more conventional techniques based on Wiener filtering [8] or adaptive beamforming [26].

One way to solve the BSS problem is to transform the mixtures to the frequency domain using the discrete Fourier transform (DFT) and apply ICA techniques in each DFT-bin

independently (see e.g., [27, 28]). This approach is referred to as the *narrowband* approach, in contrast with *broadband* approaches which process all frequency bins simultaneously. Narrowband approaches are conceptually simpler but they suffer from a permutation and scaling ambiguity in each frequency bin, which must be tackled by additional heuristic mechanisms. Note however that to solve the permutation problem, information on the sensor positions is usually required and free-field sound wave propagation is assumed (see, e.g., [29, 30]). Unfortunately, in the binaural HA application, the distance between the microphones on each side of the head will generally not be known exactly and head shadowing effects will cause a disturbance of the wavefront. In this paper, we consider a *broadband* ICA approach [31, 32] based on the TRINICON framework [33]. Separation is performed exploiting second-order statistics, under the assumption that the (mutually independent) source signals are non-white and non-stationary (like speech). Since this broadband approach does not rely on accurate knowledge of the sensor placement, it is robust against unknown microphone array deformations or disturbance of the wavefront. It has already been used for binaural HAs in [10, 34].

Since BSS allows the reconstruction of the original source signals up to an unknown permutation, we cannot know a-priori which output contains the target source. Here, it is assumed that the target source is located approximately in front of the HA user, which is a standard assumption in state-of-the-art HAs. Based on the approach presented in [35], the output containing the most frontal source is then selected after estimating the time-difference-of-arrival (TDOA) of each separated source. This is done by exploiting the ability of the broadband BSS algorithm [31, 32] to perform blind system identification of the acoustical mixing system. Figure 3 illustrates the resulting AFC-BSS combination. Note that the BSS algorithm can be embedded into the general binaural configuration depicted in Figure 1, with the BAF filters $\mathbf{w}_L$ and $\mathbf{w}_R$ set identically to the BSS filters producing the selected (monaural) BSS output:

$$\mathbf{w}_L = \mathbf{w}_R = \begin{bmatrix} \mathbf{a}_{LL} & \mathbf{a}_{RL} \end{bmatrix} \quad \text{if the left output is selected,} \quad (66)$$

$$\mathbf{w}_L = \mathbf{w}_R = \begin{bmatrix} \mathbf{a}_{LR} & \mathbf{a}_{RR} \end{bmatrix} \quad \text{if the right output is selected.} \quad (67)$$

The BSS algorithm satisfies, therefore, the assumption (41) and the AFC-BSS combination can be equivalently described by Figure 2, with $c_{LR} = 1$. In the following, $v \,\hat{=}\, v_L = v_R$ refers to the selected BSS output presented (after amplification in the forward paths) to the HA user at both ears, and $\mathbf{w} \,\hat{=}\, \mathbf{w}_L = \mathbf{w}_R$ denotes the transfer functions of the selected BSS filters (common to both LS outputs). Note finally that post-processing filters may be used to recover spatial cues [10]. They can be modelled as being part of the forward paths $\widetilde{g}_L$ and $\widetilde{g}_R$.

4.2. Discussion. In the HA scenario, since the LS output signals feed back into the microphones, the closed-loop system formed by the HAs participates in the source mixing process, together with the acoustical mixing system. Therefore, the

BSS inputs result from a mixture of the external sources and the feedback signals coming from the loudspeakers. But because of the closed-loop system bringing the HA inputs to the two LS outputs, the feedback signals are correlated with the original external source signals. To understand the impact of feedback on the separation performance of a BSS algorithm, we describe below the overall mixing process.

The closed-loop transfer function from the external sources (the point sources and the diffuse noise sources) to the BSS inputs (i.e, the residual signals after AFC) can be expressed in the z-domain by inserting (59) and (63) into (45):

$$\mathbf{e} = (\mathbf{x}^s + \mathbf{x}^n) + \frac{1}{1-k}(\mathbf{x}^s + \mathbf{x}^n)\mathbf{w}^T\widetilde{\mathbf{g}}\left(\hat{\mathbf{B}}^c - \mathbf{B}^c\right)$$

$$= \underbrace{\mathbf{s}\left[\mathbf{H} + \frac{1}{1-k}\mathbf{H}\mathbf{w}^T\widetilde{\mathbf{g}}(\hat{\mathbf{B}}^c - \mathbf{B}^c)\right]}_{\mathbf{e}^s}$$

$$\quad + \underbrace{\mathbf{x}^n\left[\mathbf{I} + \frac{1}{1-k}\mathbf{w}^T\widetilde{\mathbf{g}}(\hat{\mathbf{B}}^c - \mathbf{B}^c)\right]}_{\mathbf{e}^n}, \quad (68)$$

where $\mathbf{B}^c$ and $\hat{\mathbf{B}}^c$ refer to the AFC system and its ideal solution (46), respectively, under the block-diagonal constraint (24). k characterizes the stability of the binaural closed-loop system and is defined by (65). From (68), we can identify two independent components $\mathbf{e}^s$ and $\mathbf{e}^n$ present in the BSS inputs and originating from the external point sources and from the diffuse noise, respectively. As mentioned in Section 4.1, the BSS algorithm allows to separate point sources, additional diffuse noise having only a limited impact on the separation performance [32]. We therefore concentrate on the first term in (68):

$$\mathbf{e}^s = \mathbf{s}\mathbf{H} + \mathbf{s}\underbrace{\frac{1}{1-k}\mathbf{H}\mathbf{w}^T\widetilde{\mathbf{g}}(\hat{\mathbf{B}}^c - \mathbf{B}^c)}_{\breve{\mathbf{H}}}, \quad (69)$$

which produces an additional mixing system $\breve{\mathbf{H}}$ introduced by the acoustical feedback (and the required AFC filters). Ideally, the BSS filters should converge to a solution which minimizes the contribution $v^{s^{int}}$ of the interfering point sources $\mathbf{s}^{int}$ at the BSS output v, that is,

$$v^{s^{int}} = \underbrace{\mathbf{s}^{int}\mathbf{H}^{int}\mathbf{w}^T}_{\text{acoustical mixing}} + \underbrace{\mathbf{s}^{int}\breve{\mathbf{H}}^{int}\mathbf{w}^T}_{\text{feedback loop}} \stackrel{!}{=} 0. \quad (70)$$

$\mathbf{H}^{int}$ refers to the acoustical mixing of the interfering sources $\mathbf{s}^{int}$, as defined in Section 2.2. $\breve{\mathbf{H}}^{int}$ can be defined in a similar way and describes the mixing of the interfering sources introduced by the feedback loop.

In the absence of feedback (and of AFC filters), the second term in (70) disappears and BSS can extract the target source by unraveling the acoustical mixing system $\mathbf{H}$, which is the desired solution. Note that this solution also allows to estimate the position of each source, which is necessary to select the output of interest, as discussed in Section 4.1. However, when strong feedback signal components are
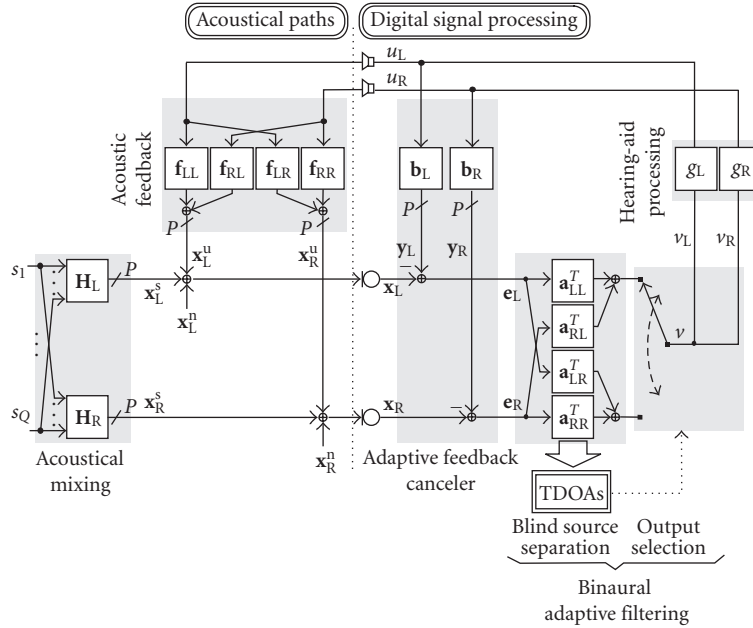
FIGURE 3: Signal model of the AFC-BSS combination.

present at the BSS inputs, the BSS solution becomes biased since the algorithm will try to unravel the feedback loop $\check{\mathbf{H}}$ instead of targetting the acoustical mixing system $\mathbf{H}$ only. The importance of the bias depends on the magnitude response of the filters captured by $\check{\mathbf{H}}$ in (70), relative to the magnitude response of the filters captured by $\mathbf{H}$. Contrary to the AFC bias encountered in Section 3.2, the BSS bias therefore decreases with increasing SFR.

The above discussion concerning BSS algorithms can be generalized to any signal enhancement techniques involving adaptive filters. The presence of feedback at the algorithm's inputs will always cause some adaptation problems. Fortunately, placing an AFC in front of the BAF like in Figure 1 can help increasing the SFR at the BAF inputs. In particular, when the AFC filters reach their ideal solution (i.e., $\mathbf{B}^c = \hat{\mathbf{B}}^c$), then $\check{\mathbf{H}}$ becomes zero and the bias term due to the feedback loop in (70) disappears, regardless of the amount of sound amplification applied in the forward paths.

5. Evaluation Setup

To validate the theoretical analysis conducted in Sections 3 and 4, the binaural configuration depicted in Figure 3 was experimentally evaluated for the combination of a feedback canceler and the blind source separation algorithm introduced in Section 4.1.

5.1. Algorithms. The BSS processing was performed using a two-channel version of the algorithm introduced in Section 4.1, picking up the front microphone at each ear (i.e., $P = 1$). Four adaptive BSS filters needed to be computed at each adaptation step. The output containing the target source (the most frontal one) was selected based on BSS-internal source localization (see Section 4.1, and [35]). To obtain

meaningful results which are, as far as possible, independent of the AFC implementation used, the AFC filter update was performed based on the frequency-domain adaptive filtering (FDAF) algorithm [36]. The FDAF algorithm allows for an individual step-size control for each DFT bin and a bin-wise optimum control mechanism of the step-size parameter, derived from [13, 37]. In practice, this optimum step-size control mechanism is inappropriate since it requires the knowledge of signals which are not available under real conditions, but it allows us to minimize the impact of a particular AFC implementation by providing useful information on the achievable AFC performance. Since we used two microphones, the (block-diagonal constrained) AFC consisted of two adaptive filters (see Figure 3).

Finally, to avoid other sources of interaction effects and concentrate on the AFC-BSS combination, we considered a simple linear time-invariant frequency-independent hearing-aid processing in the forward paths (i.e., $g_L(z) = g_L$ and $g_R(z) = g_R$). Furthermore, in all the results presented in Section 4, the same HA gains $g_L = g_R \stackrel{!}{=} g$ and decorrelation delays (see Section 3.2) $D_L = D_R = D$ were applied at both ears. The selected BSS output was therefore amplified by a factor g, delayed by D and played back at the two LS outputs.

5.2. Performance Measures. We saw in the previous sections that our binaural configuration significantly differs from what can usually be found in the literature on unilateral HAs. To be able to objectively evaluate the algorithms' performance in this context, especially concerning the AFC, we need to adapt some of the already existing and commonly used performance measures to the new binaural configuration. This issue is discussed in the following, based on the outcomes of the theoretical analysis presented in Sections 3 and 4.
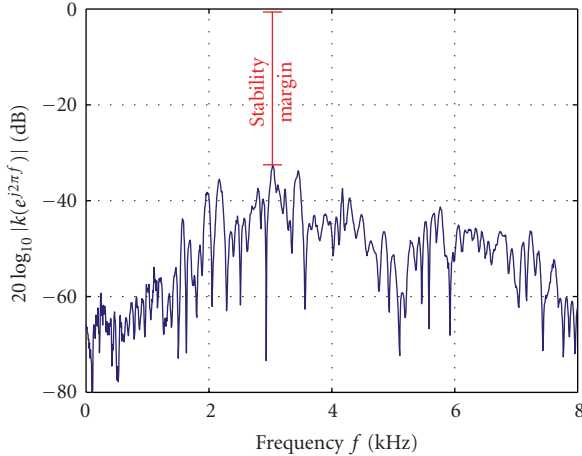
FIGURE 4: Illustration of the stability margin.

5.2.1. Feedback Cancellation Performance Measures. In the conventional unilateral case, the feedback cancellation performance is usually measured in terms of misalignment between the (single) FBP estimate and the true (single) FBP (which is the ideal solution in the unilateral case), as well as in terms of Added Stable Gain (ASG) reflecting the benefit of AFC for the user [14].

In the binaural configuration considered in this study, the misalignment should measure the mismatch between each AFC filter and its corresponding ideal solution. This can be computed in the frequency domain as follows:

$$\triangle b_{I_p} = 10 \log_{10} \frac{\sum_\omega \left| b_{I_p}(e^{j\omega}) - \hat{b}_{I_p}(e^{j\omega}) \right|^2}{\sum_\omega \left| \hat{b}_{I_p}(e^{j\omega}) \right|^2} \quad I \in \{L, R\}. \tag{71}$$

The ideal binaural AFC solution has been defined in (39) for the general case, and in (44) under the block-diagonal constraint (24) and assumption (43). In the results presented in Section 4, the misalignment has been averaged over all AFC filters (two filters in our case).

In general, it is not possible to calculate an ASG in the binaural case since the function $k(e^{j\omega})$ characterizing the stability of the system depends on both gains g_L and g_R (Section 3.3). It is however possible to calculate an added stability margin (ASM) measuring the additional gain margin (the distance of $20\log_{10}|k(e^{j\omega})|$ to 0dB, see Figure 4) obtained by the AFC

$$\text{ASM} = 20 \log_{10}\left(\min_\omega \frac{1}{|k(e^{j\omega})|} \right) - 20 \log_{10}\left(\min_\omega \frac{1}{|k(e^{j\omega})|_{\mathbf{B}=0}} \right), \tag{72}$$

where $k(z)$ has been defined in (57) and $|k(e^{j\omega})|_{\mathbf{B}=0}$ is the initial magnitude of $k(z = e^{j\omega})$, without AFC. Since the assumption (41) is valid in our case (with $c_{LR} = 1$) and since we force our AFC system to be block-diagonal, we can alternatively use the simplified expression of k given by (65). Note

that the initial stability margin $20\log_{10}(\min_\omega 1/|k(e^{j\omega})|_{\mathbf{B}=0})$ as well as the margin with AFC $20 \log_{10}(\min_\omega 1/|k(e^{j\omega})|)$, and hence the ASM, depend not only on the acoustical (binaural) FBPs, but also on the current state of the BAF filters. Also, when $g_L = g_R \stackrel{!}{=} g$, k becomes directly proportional to g and the ASM can be interpreted as an ASG.

Additionally, the SFR measured at the BSS and AFC inputs should be taken into account when assessing the AFC-BSS combination since it directly influences the performance of the algorithms. The SFR is defined in the following as the signal power ratio between the components coming from the external sources (without distinction between desired and interfering sources), and the components coming from loudspeakers (i.e., the feedback signals).

5.2.2. Signal Enhancement Performance Measures. The separation performance of the BSS algorithm is evaluated in terms of signal-to-interference ratio (SIR), that is, the signal power ratio between the components coming from the target source and the components coming from the interfering source(s). Although the feedback components $\mathbf{x}^u$ and the AFC filter outputs $\mathbf{y}$ (i.e., the compensation signals) contain some signal coming from the external sources $\mathbf{s}$ (which causes a bias of the BSS solution, as discussed in Section 4), we will ignore them in the SIR calculation since these components are undesired. An SIR gain can then be obtained as the difference between the SIR at the BSS inputs and the SIR at the BSS outputs. It reflects the ability of BSS to extract the desired components from the signal mixture $\mathbf{x}^s$, regardless of the amount of feedback (or background noise) present. Since only one BSS output is presented to the HA user (Section 4.1), we average the input SIR over all BSS input channels (here two), but we consider only the selected BSS output for the output SIR calculations.

5.3. Experimental Conditions. Since a two-channel ICA-based BSS algorithm can only separate two point sources (Section 4.1), no diffuse noise has been added to the sensor signal mixture (i.e., $\mathbf{x}^n = 0$) and only two point sources were considered (one target source and one interfering source).

Head-related impulse responses (HRIR) were measured using a pair of Siemens Life (BTE) hearing aid cases with two microphones and a single receiver (loudspeaker) inside each device (no processor). The cases were mounted on a real person and connected, via a pre-amplifier box, to a (laptop) PC equipped with a multi-channel RME Multiface sound card. Measurements were made in the following environments:

  (i) a low-reverberation chamber ($T_{60} \approx 50$ ms),

  (ii) a living-room-like environment ($T_{60} \approx 300$ ms).

The source signal components $\mathbf{x}^s$ were then generated by convolving speech signals with the recorded HRIRs, with the target and interfering sources placed at azimuths 0° (in front of the HA user) and 90° (facing the right ear), respectively. The target and interfering sources were approximately of equal (long-time) signal power.
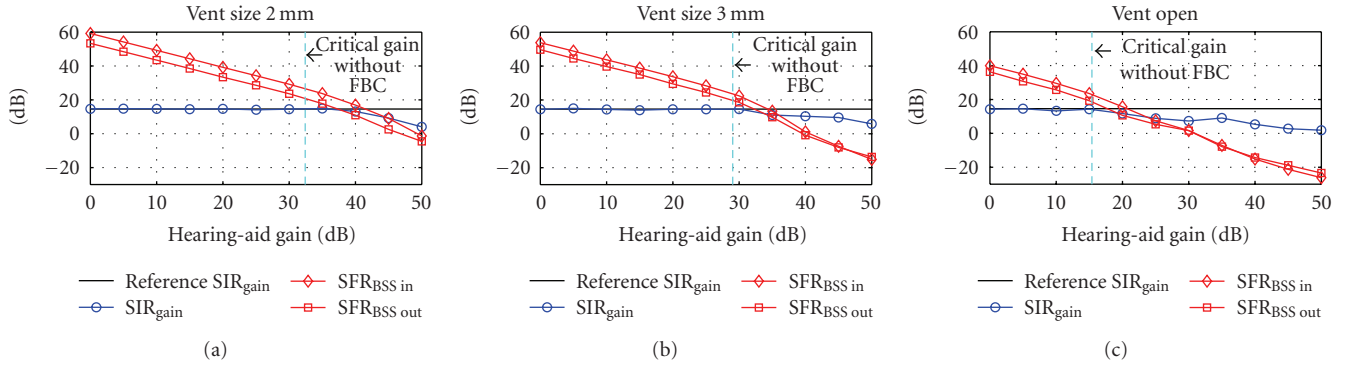
Figure 5: BSS performance for increasing HA gain, in a low-reverberation chamber.

To generate the feedback components $\mathbf{x}^u$, binaural FBPs ("direct" and "cross" FBPs, see Section 2.2) measured from Siemens BTE hearing aids were used. These recordings have been made for different vent sizes: 2 mm, 3 mm and open and in the following scenario:

(i) left HA mounted on a manikin without obstruction,

(ii) right HA mounted on a manikin with a telephone as obstruction.

The digital signal processing was performed at a sampling frequency of 16 kHz, picking up the front microphone at each ear (i.e., $P = 1$).

6. Experimental Results

In the following, experimental results involving the combination of AFC and BSS are shown and discussed. BSS filters of 1024 coefficients each were applied, the AFC filter length was set to 256 and decorrelation delays of 5 ms were included in the forward paths.

6.1. Impact of Feedback on BSS. The discussion of Section 4 indicates that a deterioration of the BSS performance is expected at low input SFR, due to a bias introduced by the feedback loop. To determine to which extent the amount of feedback deteriorates the achievable source separation, the performance of the (adaptive) BSS algorithm was experimentally evaluated for different amounts of feedback by varying the amplification level g. Preliminary tests in the absence of AFC showed that the feedback had almost no impact on the BSS performance as long as the system was stable (i.e., as long as $|k(e^{j\omega})|_{\mathbf{B}=0} < 1, \forall \omega$ because the SFR at the BSS inputs was kept high (greater than 20 dB). This basically validates the traditional way signal enhancement techniques for hearing aids have been developed, ignoring the presence of feedback.

Signal enhancement algorithms, however, can be subject to higher input SFR levels when an AFC is used to stabilize the system. To be able to further increase the gains and the amount of feedback signal in the microphone inputs while preventing system instability, the feedback components present at the BSS output v were artificially suppressed. This

is equivalent to performing AFC on the BSS output, under ideal conditions. It guarantees the stability of the system (with ASM $= +\infty$), regardless of the HA amplification level, but does not reduce the SFR at the BSS inputs. The results after convergence of the BSS algorithm are presented in Figures 5 and 6 for different rooms and vent sizes. The reference lines show the gain in SIR achieved by BSS in the absence of feedback (and hence in the absence of AFC). The critical gain depicted by vertical dashed lines in the figures, corresponds to the maximum stable gain without AFC, that is, the gain for which the initial stability margin $20 \log_{10}(\min_\omega 1/|k(e^{j\omega})|_{\mathbf{B}=0})$ becomes zero.

At low gains, the feedback has very little impact on the SIR gain because the input SFR is sufficiently high in all tested scenarios. We see also that the interference rejection causes a decrease in SFR (from the BSS inputs to the output) since parts of the external source components are attenuated. This should be beneficial to an AFC algorithm since it reduces the bias of the AFC Wiener solution due to the interfering source, as discussed in Section 3.2. However, at high gains, where the input SFR is low (less than 10 dB), the large amount of feedback causes a significant deterioration of the interference rejection performance. Moreover, it should be noted that at low gains, the input SFR decreases proportionally to the gain, as expected. We see, however, from the figures that the input SFR suddenly drops at higher gains, when the amount of feedback becomes significant (see, e.g., the transition from $g = 20$ dB to $g = 25$ dB, in Figure 6, for an open vent). Since BSS has no influence on the signal power of the external sources (the "S" component in the SFR), it means that BSS amplifies the LS signals (and hence the feedback signals at the microphones, that is, the "F" component in the SFR). This undesired effect is due to the bias introduced by the feedback loop (Section 4.2) and can be interpreted as follows: two mechanisms enter into play. The first one unravels the acoustical mixing system. It produces LS signals which are dominated by the target source (see the positive SIR gains in the figures), as desired. The second mechanism consists in amplifying the sensor signals. As long as the feedback level is small, this second mechanism is almost invisible since it would amplify signals coming from both sources. But at higher gains, where the amount of feedback in the BSS inputs become more significant, this second mechanism becomes
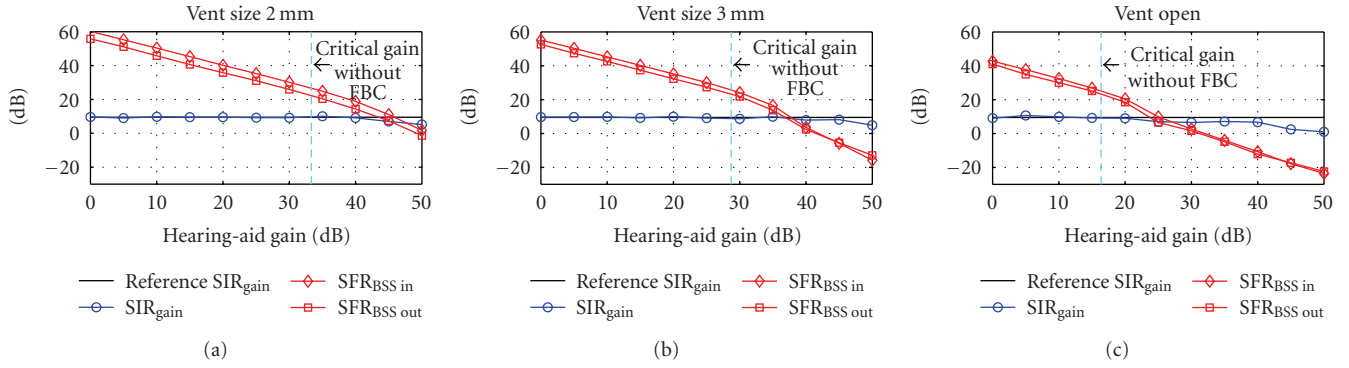
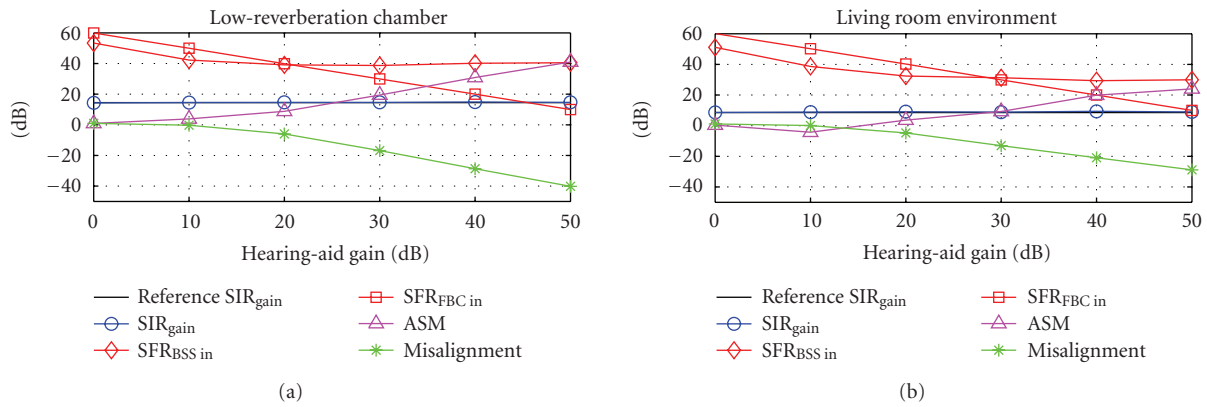FIGURE 6: BSS performance for increasing HA gain, in a living-room environment.

FIGURE 7: Performance of the AFC-BSS combination in two acoustical environments. The measured FPBs for a vent of size 2 mm were used.

more important since it acts mainly in favor of the target source. This second mechanism illustrates the impact of the feedback loop on the BSS algorithm at high feedback levels. It shows the necessity to have the AFC placed before BSS, so that BSS can benefit from a higher input SFR.

6.2. Overall Behavior of the AFC-BSS Combination. The full AFC-BSS combination has been evaluated for a vent size of 2 mm, in the low-reverberation chamber as well as in the living-room-like environment (Section 5.3). Figure 7 depicts the BSS and AFC performance obtained after convergence. Like in Figures 5 and 6, the reference lines show the gain in SIR achieved by BSS in the absence of feedback (and hence in the absence of AFC).

The results confirm the observations made in the previous section. With the AFC applied directly on the sensor signals, the BSS algorithm could indeed benefit from the ability of the AFC to keep the SFR at the BSS inputs at high levels for every considered HA gains. Therefore, BSS always provided SIR gains which were very close to the reference SIR gain obtained without feedback, even at high gains. This contrasts with the results obtained in Figures 5 and 6, where an ideal AFC was applied at the BSS output instead of being applied first.

Note that the SFR at the AFC outputs correspond here to the SFR at the BSS inputs. The gain in SFR ($SFR_{BSS_{in}} - SFR_{AFC_{in}}$, i.e., the feedback attenuation) achieved by the AFC

algorithm can be therefore directly visualized from Figure 7. As expected from the discussion presented in Section 3.1, the two AFC filters used were sufficient to efficiently compensate both the "direct" and "cross" feedback signals, and hence avoid instability of the binaural closed-loop system. Like in the unilateral case and as expected from the convergence analysis conducted in Section 3.2, the best AFC results were obtained at low input SFR levels, that is, at high gains. The AFC performance was also better in the low-reverberation chamber than in the living-room-like environment, as can be seen from the higher SFR levels at the BSS inputs, the higher ASM values and the lower misalignments. This result seems surprising at the first sight, since the FBPs were identical in both environments. It can be however easily justified by the analytical results presented in Section 3.2. We saw actually that the correlation between the external source signals and the LS signals introduce a bias of the AFC Wiener solution. The bias due to the target source is barely influenced by the BSS results since BSS left the target signal (almost) unchanged in both environments. But the BSS performance influences directly the amount of residual interfering signal present at the LS outputs, and hence the bias of the AFC Wiener solution due to the interfering source. In general, since reverberation increases the length of the acoustical mixing filters (and hence the necessary BSS filter length, typically), the more reverberant the environment, the lower the achieved separation results (for a given BSS filter length).

This is confirmed here by the SIR results shown in the figures. The difference in AFC performance comes therefore from the higher amount of residual interfering signal present at the LS outputs in the living-room-like environment, which increases the bias of the AFC Wiener solution.

The AFC does not suffer from any particular negative interactions with the BSS algorithm since it comes first in the processing chain, but rather benefits from BSS, especially in the low-reverberation chamber, as we just saw. Note that the situation is very different when the AFC is applied after BSS. In this case, the AFC filters need to quickly follow the continuously time-varying BSS filters, which prevents proper convergence of the AFC filters, even with time-invariant FBPs.

7. Conclusions

An analysis of a system combining adaptive feedback cancellation and adaptive binaural filtering for signal enhancement in hearing aids was presented. To illustrate our study, a blind source separation algorithm was chosen as an example for adaptive binaural filtering. A number of interaction effects could be identified. Moreover, to correctly understand the behavior of the AFC, the system was described and analyzed in detail. A new stability condition adapted to the binaural configuration could be derived, and adequate performance measures were proposed which account for the specificities of the binaural system. Experimental evaluations confirmed and illustrated the theoretical findings.

The ideal AFC solution in the new binaural configuration could be identified but a steady-state analysis showed that the AFC suffers from a bias in its optimum (Wiener) solution. This bias, similar to the unilateral case, is due to the correlation between feedback and external source signals. It was also demonstrated theoretically as well as experimentally that a signal enhancement algorithm could help reducing this bias. The correlation between feedback and external source signals also causes a bias of the BAF solution. But contrary to the bias encountered by the AFC, the BAF bias increases with increasing HA amplification levels. Fortunately, this bias can be reduced by applying AFC on the sensor signals directly, instead of applying it on the BAF outputs.

The analysis also showed that two SIMO AFC systems of adaptive filters can effectively compensate for the four SIMO FBP systems when the outputs are sufficiently correlated (see Section 3.1). Should this condition not be fulfilled because of, for example, some non-linearities in the forward paths, the "cross" feedback signals travelling from one ear to the other would not be completely identifiable. But we saw that since the amplitude of the "cross" FBPs is usually negligible compared to the amplitude of the "direct" FBPs, the consequences would be very limited as long as the HA gains are set to similar amplification levels.

Acknowledgments

References

[1] S. Doclo and M. Moonen, "GSVD-based optimal filtering for single and multimicrophone speech enhancement," *IEEE Transactions on Signal Processing*, vol. 50, no. 9, pp. 2230–2244, 2002.

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

[3] S. Doclo, A. Spriet, M. Moonen, and J. Wouters, "Design of a robust multi-microphone noise reduction algorithm for hearing instruments," in *Proceedings of the International Symposium on Mathematical Theory of Networks and Systems (MTNS '04)*, pp. 1–9, Leuven, Belgium, July 2004.

[4] J. Vanden Berghe and J. Wouters, "An adaptive noise canceller for hearing aids using two nearby microphones," *Journal of the Acoustical Society of America*, vol. 103, no. 6, pp. 3621–3626, 1998.

[5] D. P. Welker, J. E. Greenberg, J. G. Desloge, and P. M. Zurek, "Microphone-array hearing aids with binaural output—part II: a two-microphone adaptive system," *IEEE Transactions on Speech and Audio Processing*, vol. 5, no. 6, pp. 543–551, 1997.

[6] M. Doerbecker and S. Ernst, "Combination of two-channel spectral subtraction and adaptive Wiener post-filtering for noise reduction and dereverberation," in *Proceedings of the 8th European Signal Processing Conference (EUSIPCO '96)*, pp. 995–998, Trieste, Italy, September 1996.

[7] T. J. Klasen, 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.

[8] S. Doclo, R. Dong, T. J. Klasen, J. Wouters, S. Haykin, and M. Moonen, "Extension of the multi-channel Wiener filter with ITD cues for noise reduction in binaural hearing aids," in *Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA '05)*, pp. 70–73, New Paltz, NY, USA, October 2005.

[9] T. J. Klasen, S. Doclo, T. van den Bogaert, M. Moonen, and J. Wouters, "Binaural multi-channel Wiener filtering for hearing aids: preserving interaural time and level differences," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '06)*, vol. 5, pp. 145–148, Toulouse, France, May 2006.

[10] R. Aichner, H. Buchner, M. Zourub, and W. Kellermann, "Multi-channel source separation preserving spatial information," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '07)*, vol. 1, pp. 5–8, Honolulu, Hawaii, USA, May 2007.

[11] D. P. Egolf, "Review of the acoustic feedback literature for a control systems point of view," in *The Vanderbilt Hearing-Aid Report*, pp. 94–103, York Press, London, UK, 1982.

[12] M. G. Siqueira and A. Alwan, "Steady-state analysis of continuous adaptation in acoustic feedback reduction systems for hearing-aids," *IEEE Transactions on Speech and Audio Processing*, vol. 8, no. 4, pp. 443–453, 2000.

[13] H. Puder and B. Beimel, "Controlling the adaptation of feedback cancellation filters—problem analysis and solution approaches," in *Proceedings of the 12th European Conference on*

Signal Processing (EUSIPCO '04), pp. 25–28, Vienna, Austria, September 2004.

[14] A. Spriet, *Adaptive filtering techniques for noise reduction and acoustic feedback cancellation in hearing aids*, Ph.D. thesis, Katholieke Universiteit Leuven, Leuven, Belgium, 2004.

[15] A. Spriet, G. Rombouts, M. Moonen, and J. Wouters, "Combined feedback and noise suppression in hearing aids," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 6, pp. 1777–1790, 2007.

[16] W. Kellermann, "Strategies for combining acoustic echo cancellation and adaptive beamforming microphone arrays," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '97)*, vol. 1, pp. 219–222, Munich, Germany, April 1997.

[17] S. Haykin, *Adaptive Filter Theory*, Prentice Hall Information and System Sciences Series, Prentice Hall, Upper Saddle River, NJ, USA, 4th edition, 2002.

[18] M. M. Sondhi, D. R. Morgan, and J. L. Hall, "Stereophonic acoustic echo cancellation—an overview of the fundamental problem," *IEEE Signal Processing Letters*, vol. 2, no. 8, pp. 148–151, 1995.

[19] J. Benesty, D. R. Morgan, and M. M. Sondhi, "A better understanding and an improved solution to the specific problems of stereophonic acoustic echo cancellation," *IEEE Transactions on Speech and Audio Processing*, vol. 6, no. 2, pp. 156–165, 1998.

[20] D. R. Morgan, J. L. Hall, and J. Benesty, "Investigation of several types of nonlinearities for use in stereo acoustic echo cancellation," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 6, pp. 686–696, 2001.

[21] A. Spriet, I. Proudler, M. Moonen, and J. Wouters, "An instrumental variable method for adaptive feedback cancellation in hearing aids," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '05)*, vol. 3, pp. 129–132, Philadelphia, Pa, USA, March 2005.

[22] T. van Waterschoot and M. Moonen, "Adaptive feedback cancellation for audio signals using a warped all-pole near-end signal model," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '08)*, pp. 269–272, Las Vegas, Nev, USA, April 2008.

[23] S. Araki, H. Sawada, R. Mukai, and S. Makino, "A novel blind source separation method with observation vector clustering," in *Proceedings of the International Workshop on Acoustic Echo and Noise Control (IWAENC '05)*, pp. 117–120, Eindhoven, The Netherlands, September 2005.

[24] J. Cermak, S. Araki, H. Sawada, and S. Makino, "Blind source separation based on a beamformer array and time frequency binary masking," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '07)*, vol. 1, pp. 145–148, Honolulu, Hawaii, USA, May 2007.

[25] A. Hyvaerinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley & Sons, New York, NY, USA, 2001.

[26] W. Herbordt, *Sound Capture for Human / Machine Interfaces*, Springer, Berlin, Germany, 2005.

[27] L. Parra and C. Spence, "Convolutive blind separation of non-stationary sources," *IEEE Transactions on Speech and Audio Processing*, vol. 8, no. 3, pp. 320–327, 2000.

[28] H. Sawada, R. Mukai, S. Araki, and S. Makino, "Polar coordinate based nonlinear function for frequency-domain blind source separation," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '02)*, vol. 1, pp. 1001–1004, Orlando, Fla, USA, May 2002.

[29] S. Kurita, H. Saruwatari, S. Kajita, K. Takeda, and F. Itakura, "Evaluation of blind signal separation method using directivity pattern under reverberant conditions," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '00)*, vol. 5, pp. 3140–3143, Istanbul, Turkey, June 2000.

[30] H. Sawada, R. Mukai, S. Araki, and S. Makino, "A robust and precise method for solving the permutation problem of frequency-domain blind source separation," *IEEE Transactions on Speech and Audio Processing*, vol. 12, no. 5, pp. 530–538, 2004.

[31] H. Buchner, R. Aichner, and W. Kellermann, "A generalization of blind source separation algorithms for convolutive mixtures based on second-order statistics," *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 1, pp. 120–134, 2005.

[32] R. Aichner, H. Buchner, F. Yan, and W. Kellermann, "A real-time blind source separation scheme and its application to reverberant and noisy acoustic environments," *Signal Processing*, vol. 86, no. 6, pp. 1260–1277, 2006.

[33] H. Buchner, R. Aichner, and W. Kellermann, "TRINICON: a versatile framework for multichannel blind signal processing," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '04)*, vol. 3, pp. 889–892, Montreal, Canada, May 2004.

[34] K. Eneman, H. Luts, J. Wouters, et al., "Evaluation of signal enhancement algorithms for hearing instruments," in *Proceedings of the 16th European Signal Processing Conference (EUSIPCO '08)*, pp. 1–5, Lausanne, Switzerland, August 2008.

[35] H. Buchner, R. Aichner, J. Stenglein, H. Teutsch, and W. Kellennann, "Simultaneous localization of multiple sound sources using blind adaptive MIMO filtering," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '05)*, vol. 3, pp. 97–100, Philadelphia, Pa, USA, March 2005.

[36] J. J. Shynk, "Frequency-domain and multirate adaptive filtering," *IEEE Signal Processing Magazine*, vol. 9, no. 1, pp. 14–37, 1992.

[37] A. Mader, H. Puder, and G. U. Schmidt, "Step-size control for acoustic echo cancellation filters—an overview," *Signal Processing*, vol. 80, no. 9, pp. 1697–1719, 2000.