Katholieke Universiteit Leuven

Departement Elektrotechniek ESAT-SISTA/TR 2006-154

Combined feedback and noise suppression in hearing aids [1]

Ann Spriet[2], Geert Rombouts[3] ,Marc Moonen[4] ,Jan Wouters[5]

Accepted for publication in IEEE Trans. on Audio, Speech and Language Processing

[1]This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/sista/spriet/reports/06-154.pdf*

[2]K.U.Leuven, Dept. of Electrical Engineering (ESAT), SISTA, Kasteelpark Arenberg 10, 3001 Leuven-Heverlee, Belgium, Tel. 32/16/32 17 95, Fax 32/16/32 19 70. K.U.Leuven, ExpORL, Dept. Neurosciences, O.& N2, Herestraat 49 bus 721, B-3000 Leuven, Belgium, Ann Spriet is a postdoctoral researcher supported by FWO-Vlaanderen. This research work was carried out at the ESAT laboratory and ExpORL laboratory of the K. U. Leuven, in the frame of the European project HearCom ('Hearing in the communication society', WP5), the Belgian Programme on Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Office IUAP P5/22 ('Dynamical Systems and Control: Computation, Identification and Modelling'), the Concerted Research Action GOA-AMBIORICS ('Algorithms for medical and biological research, integration and software') of the Flemish Government.

[3]K.U.Leuven, Dept. of Electrical Engineering (ESAT), SISTA, Kasteelpark Arenberg 10, 3001 Leuven-Heverlee, Belgium, Tel. 32/16/32 18 56, Fax 32/16/32 19 70.

[4]K.U.Leuven, Dept. of Electrical Engineering (ESAT), SISTA, Kasteelpark Arenberg 10, 3001 Heverlee, Belgium, Tel. 32/16/32 17 09, Fax 32/16/32 19 70. Marc Moonen is a professor at the Katholieke Universiteit Leuven.

[5]K.U.Leuven, ExpORL, Dept. Neurosciences, O.& N2, Herestraat 49 bus 721, B-3000 Leuven, Belgium, Tel: +32 16 330485, Fax: +32 16 330486. Jan Wouters is a professor at the Katholieke Universiteit Leuven.

Combined feedback and noise suppression in hearing aids

Ann Spriet*[1,2], Geert Rombouts[1], Marc Moonen[1], *Member, IEEE,* Jan Wouters[2]

[1]ESAT/SCD-SISTA, K.U. Leuven

Kasteelpark Arenberg 10, B-3001 Leuven-Heverlee, Belgium

Tel.: +32 16 321795, Fax: +32 16 321970

E-mail: spriet@esat.kuleuven.be; rombouts@esat.kuleuven.be; moonen@esat.kuleuven.be

[2]ExpORL / Dept. Neurosciences, K.U. Leuven, O.&N 2

Herestraat 49 bus 721, B-3000 Leuven, Belgium

E-mail: jan.wouters@med.kuleuven.be

Manuscript received September 8, 2006; revised February 20, 2007. Ann Spriet is a postdoctoral researcher supported by FWO-Vlaanderen. This research work was carried out at the ESAT laboratory and ExpORL of the K. U. Leuven, in the frame of the European project Hearcom ('Hearing in the communication society'), the Belgian Programme on Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Office IUAP P5/22 ('Dynamical Systems and Control: Computation, Identification and Modelling'), the Concerted Research Action GOA-AMBIORICS ('Algorithms for medical and biological research, integration and software') of the Flemish Government.The associate editor coordinating the review of this paper and approving it for publication was Dr. Hiroshi Sawada.

A. Spriet, G. Rombouts and M. Moonen are with the Katholieke Universiteit Leuven, Department of Electrical Engineering (ESAT-SCD), B-3001 Leuven, Belgium (e-mail: spriet,rombouts,moonen@esat.kuleuven.be).

A. Spriet and J. Wouters are with the Katholieke Universiteit Leuven, Dept. Neurosciences (ExpORL), B-3000 Leuven, Belgium (e-mail: jan.wouters@med.kuleuven.be).

Abstract

In this paper, solutions for combined feedback and noise suppression in hearing aids are developed. The techniques presented are based on the generalized sidelobe canceller (GSC) and adaptive feedback canceller (AFC), with a prediction error method (PEM) adaptation to avoid speech distortion. Two possible cascades of GSC-based noise reduction and AFC, namely an 'AFC first' and a 'GSC first', as well as a truly integrated solution that jointly suppresses feedback and noise are discussed. The integrated solution (called PEM-GFIC) achieves optimum synergies between noise and feedback suppression at the lowest computational cost. In addition, it cancels more feedback than the generalized echo and interference canceller, a joint solution for echo and noise suppression. In the cascaded solutions, the feedback and noise suppression filters are not always optimally exploited. For high input SNRs, 'AFC first' scheme generally may achieve better feedback cancellation because of its larger number of degrees of freedom. However, noise reduction by the GSC-stage seriously affects the feedback cancellation performance. At low SNRs, 'GSC first' generally achieves more feedback cancellation than PEM-GFIC at the expense of worse noise reduction. At high hearing aid gains and/or large SNRs, the noise reduction stage however negatively affects the performance of the feedback cancellation filter, resulting in a worse feedback and noise suppression compared to PEM-GFIC.

Index Terms

Feedback cancellation; noise reduction; Generalized Sidelobe Canceller; hearing aids

I. INTRODUCTION

Background noise as well as acoustic feedback (howling) represent major problems to hearing aid users, which reduce their speech understanding as well as comfort of listening significantly. For both individual problems, several solutions have already been proposed in the literature:

- To improve speech understanding in noise, especially multi-microphone noise reduction techniques such as the generalized sidelobe canceller (GSC) have been proposed and successfully applied in hearing instruments [1]-[6]. The GSC consists of a fixed beamformer, a blocking matrix and an adaptive noise canceller (ANC). The blocking matrix is designed to avoid so-called speech leakage in the noise references such that the ANC does not introduce any speech distortion.
- A promising solution for acoustic feedback suppression is the adaptive feedback canceller (AFC). Here, a model of the feedback path is identified, which is then used to estimate and remove the (unwanted) feedback signal from the microphone signal [7]-[13]. The presence of a closed signal loop however introduces signal correlation between the desired part of the microphone signal and

the loudspeaker signal. As a result, standard adaptive filtering techniques such as those used for echo cancellation will cancel the desired signal instead of the feedback signal and hence, fail to provide a reliable feedback path model. To improve the feedback path estimation, feedback cancellation algorithms based on the prediction-error method (PEM-AFC) were proposed in [13], [14], [15]. In these algorithms, the correlation between the loudspeaker signal and the desired speech signal is removed through adaptive pre-whitening of the desired signal.

For the design and evaluation of a noise reduction or a feedback cancellation system, respectively, the presence of feedback and noise has typically been ignored. In [16], it has been shown that the presence of feedback may seriously affect the performance of the classical GSC. In the presence of feedback, the noise references contain both noise and the feedback signal. Due to correlation between the feedback signal and the speech component in the speech reference, the GSC then cancels the desired speech signal instead of the feedback and the noise. To remove the correlation and hence, to avoid the resulting speech distortion, the ANC of the GSC has been extended in [16] with pre-whitening operations, as used in the PEM-AFC [13], [14]. In the sequel, this technique will be referred to as PEM-based GSC (PEM-GSC).

In this paper, solutions for the combined feedback and noise suppression problem that exploit both the microphone and the loudspeaker signals are developed and evaluated (note that PEM-GSC uses only microphone signals). We concentrate on solutions that are based on the GSC and the AFC. A straightforward approach is to perform the noise reduction and the feedback cancellation algorithm independently, either by first canceling the acoustic feedback in all microphone signals and then performing multi-microphone noise reduction, or vice-versa, by first performing multi-microphone noise reduction, followed by a single-channel feedback canceller. To avoid speech cancellation, the PEM-GSC [16] and the PEM-AFC [13], [14] should then be used. The interaction between the two algorithms, i.e., the influence of the noise reduction stage on the feedback cancellation stage and vice-versa, will be analyzed. Furthermore, it is expected that a better performance can be obtained by truly integrating noise reduction and feedback cancellation into one multi-channel (i.e., exploiting the microphone and loudspeaker signals) signal processing algorithm. This problem resembles the combination of echo cancellation and noise reduction [17]. In [17], a joint solution based on the GSC and standard adaptive echo cancellation has been proposed, called the generalized echo and interference canceller (GEIC). However, the presence of the closed signal loop and the continuous double-talk situation in the acoustic feedback problem require alternative solutions. Because of signal correlation between the feedback and the desired speech signal, the GEIC would indeed cancel the desired speech instead of the feedback and the noise. To avoid speech cancellation, the GEIC will be extended

with adaptive pre-whitening operations.

The paper is organized as follows. Section II defines the signal model and the performance measures that are used in the paper. Section III describes a joint solution for feedback cancellation and noise reduction that directly estimates the clean speech signal (i.e., without noise and acoustic feedback) based on all the microphone signals and the loudspeaker signal. The PEM-GSC of [16] is a special case of this integrated scheme where the loudspeaker signal is then not used. The proposed solution can also be viewed as an extension of the GEIC with adaptive pre-whitening operations and will be referred to as the PEM-based generalized feedback and interference canceller (PEM-GFIC). In Section IV and Section V, cascades of the PEM-GSC [16] and PEM-AFC [13], [14] are discussed: feedback cancellation followed by noise reduction (referred to as 'AFC first') and noise reduction followed by feedback cancellation (referred to as 'GSC first') are considered. The interactions between the GSC-stage and the AFC-stage are investigated and the results are illustrated based on simulations with anechoic recordings. Finally, Section VI compares the noise and feedback suppression performance of the integrated PEM-GFIC solution with the GEIC and the cascaded solutions for realistic, reverberant recordings with a two-microphone commercial hearing aid. Compared to the GEIC, the PEM-GFIC solution achieves better feedback cancellation thanks to the adaptive pre-whitening of the desired signal. Compared to the cascaded solutions, the PEM-GFIC achieves optimum synergies between noise reduction and feedback cancellation at the lowest computational cost. In the cascaded solutions, the noise reduction and feedback cancellation stage influence each other so that they are not always optimally exploited. For high input SNRs where feedback cancellation is the major goal, the 'AFC first' scheme generally may achieve better feedback cancellation because of its larger number of degrees of freedom. However, noise reduction by the GSC-stage negatively affects the performance of the feedback cancellation stage. 'GSC first' focuses more on feedback cancellation than the PEM-GFIC. At low input SNRs, this generally results in better feedback cancellation at the expense of worse noise reduction. At high gains and/or large input SNRs, the noise reduction filter negatively influences the performance of the feedback cancellation filter, resulting in a worse feedback and noise suppression compared to the PEM-GFIC.

II. Signal model and performance measures

A. Signal model and notation

Figure 1 defines the signal model and notation that is used throughout this paper. A discrete-time filter with length L_F is represented as a polynomial $F(q)$ in q, i.e.,

$$F(q) = f_0 q^0 + \ldots + f_{L_F-1} q^{-L_F+1}, \tag{1}$$

where q^{-1} is the delay operator. This notation is adopted from system identification theory [18] and allows the following notation for the filtering of $u[k]$ with $F(q)$

$$F(q)u[k]. \tag{2}$$

The feedback path between the loudspeaker and the i-th microphone is denoted by $F_i(q)$. The gain and compression module of the hearing aid is denoted by $G(q)$. We assume that $G(q)$ contains a delay d_G of at least one sample.

The microphone signals are $x_i[k]$, $i = 1, \ldots, M$, the loudspeaker signal is $u[k]$. Each microphone signal $x_i[k]$ can be decomposed as:

$$x_i[k] = x_i^s[k] + x_i^n[k] + \underbrace{F_i(q)u[k]}_{x_i^u[k]}, \tag{3}$$

where $x_i^s[k]$ and $x_i^n[k]$ are the direct speech and the noise component in the i-th microphone signal (i.e., not originating from the loudspeaker) and $x_i^u[k] = F_i(q)u[k]$ is the feedback component in the i-th microphone signal. Let $\mathbf{x}[k] \in \mathbb{R}^{M\times 1}$ be defined as the stacked vector

$$\mathbf{x}[k] = \begin{bmatrix} x_1[k] & x_2[k] & \ldots & x_M[k] \end{bmatrix}^T. \tag{4}$$

Then, (3) can be written as

$$\mathbf{x}[k] = \mathbf{x}^s[k] + \mathbf{x}^n[k] + \underbrace{\mathbf{F}(q)u[k]}_{\mathbf{x}^u[k]} \tag{5}$$

with $\mathbf{x}^s[k]$ and $\mathbf{x}^n[k]$ defined similarly as in (4) and with $\mathbf{F}(q)$ defined as

$$\mathbf{F}(q) = \begin{bmatrix} F_1(q) & F_2(q) & \ldots & F_M(q) \end{bmatrix}^T. \tag{6}$$

Let the combined feedback-and-noise suppression filter be denoted by the $(M+1)\times 1$ vector $\mathbf{W}(q)$ that operates on the microphone signals $\mathbf{x}[k]$ and the loudspeaker signal $u[k]$. Then, the output $e[k]$

$$e[k] = \mathbf{W}^T(q) \begin{bmatrix} \mathbf{x}[k] \\ u[k] \end{bmatrix} \tag{7}$$

of the combined algorithm can be decomposed as

$$e[k] = \underbrace{\mathbf{W}^T(q) \begin{bmatrix} \mathbf{x}^s[k] + \mathbf{x}^n[k] \\ 0 \end{bmatrix}}_{e_{NR}[k]} + \underbrace{\mathbf{W}^T(q) \begin{bmatrix} \mathbf{x}^u[k] \\ u[k] \end{bmatrix}}_{e_{FC}[k]}. \tag{8}$$

The first term $e_{NR}[k]$ is the algorithm's output in the absence of internal and acoustic feedback (i.e., after setting $G(q) = 0$ or $u[k] = 0$). Ideally, $\mathbf{W}(q)$ reduces the noise $\mathbf{x}^n[k]$ as much as possible while

preserving the direct speech. The second term $e_{FC}[k]$ is the residual feedback measured at the output of the algorithm: the smaller $e_{FC}[k]$, the more the feedback is cancelled by $\mathbf{W}(q)$.

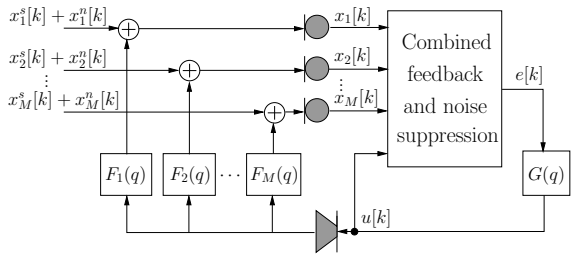


Fig. 1. Signal model and notation.

B. Performance measures

Intelligibility-weighted signal-to-noise ratio (SNR): To assess the noise reduction performance of the algorithms, the intelligibility-weighted signal-to-noise ratio (SNR) [19] is used, defined as

$$\text{SNR}_{\text{intellig}} = \sum_i I_i \text{SNR}_i. \tag{9}$$

The band importance function I_i expresses the importance of the ith one third octave band with center frequency $f_{c,i}$ for intelligibility and SNR_i is the SNR (in dB) in the i-th one third octave band. The center frequencies $f_{c,i}$ and the corresponding values I_i are defined in [20].

To merely measure the noise reduction by the algorithms, the intelligibility-weighted SNR of the output component $e_{NR}[k]$ in (8) is computed: The filters $\mathbf{W}(q)$ are first adapted; After freezing $\mathbf{W}(q)$, the closed-signal loop $G(q)$ is set to zero and the $\text{SNR}_{\text{intellig}}$ is measured. In the presence of $G(q)$, the speech and the noise are reproduced in the loudspeaker signal $u[k]$, resulting in additional speech and noise components $e_{FC}[k]$ in the output $e[k]$. The noise component in $e_{FC}[k]$ may be suppressed by the noise reduction algorithm or by the feedback canceller. It is therefore not included in the computation of the $\text{SNR}_{\text{intellig}}$ when evaluating the noise reduction performance.

Intelligibility-weighted signal-to-feedback ratio (SFR): To assess the feedback cancellation performance, the intelligibility-weighted signal-to-feedback ratio (SFR) of the speech component is used, defined as

$$\text{SFR}_{\text{intellig}} = \sum_i I_i \text{SFR}_i, \tag{10}$$

where SFR_i is the signal-to-feedback ratio of speech in the i-th one third octave band. The filters $\mathbf{W}(q)$ are first adapted; $\text{SFR}_{\text{intellig}}$ of the speech component is computed after freezing $\mathbf{W}(q)$. The $\text{SFR}_{\text{intellig}}$ at

the output of the algorithms will be compared to the SFR that would be obtained when no feedback and noise suppression is performed. The feedback noise signal is partially suppressed by the noise reduction algorithm and partially suppressed by the feedback canceller and is therefore not included in the signal-to-feedback ratio $\text{SFR}_{\text{intellig}}$.

III. INTEGRATED SOLUTION FOR FEEDBACK CANCELLATION AND NOISE REDUCTION

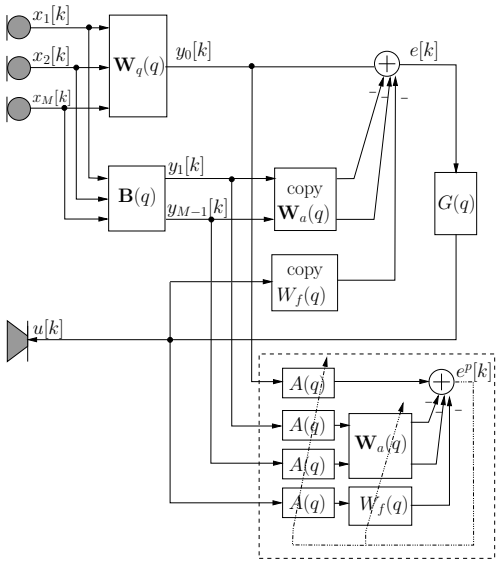


Fig. 2. PEM-based generalized feedback and interference canceller (PEM-GFIC).

A promising approach for combined feedback and noise suppression is to integrate noise reduction and feedback cancellation into one multi-channel signal processing algorithm, where acoustic feedback and noise are jointly suppressed. In this way, negative interactions between the noise reduction and feedback cancellation algorithm are avoided and optimum synergies between noise reduction and feedback cancellation are achieved at a minimum computational cost. In [17], an integrated solution for echo cancellation and noise reduction has been proposed, referred to as the GEIC. In this section, the GEIC will be extended with pre-whitening operations such that it can be used for combined feedback and noise suppression. This extension will be referred to as PEM-based Generalized Feedback and Interference Canceller (PEM-GFIC). The PEM-GSC proposed in [16] is a special case of this integrated solution where the loudspeaker signal is then not used. The PEM-GSC was proposed for applications with long acoustic feedback paths such as public address systems to reduce computational cost. Instead of modelling the feedback path (as is the case in the PEM-AFC), the PEM-GSC steers a zero to the direction of the

loudspeaker. As a result, feedback cancellation can be achieved with relatively short filters [16]. For short feedback paths as in hearing aids (i.e., 50-100 taps), additional performance may be gained at a limited extra cost by exploiting the loudspeaker signal.

A. Concept

Figure 2 depicts the integrated solution for suppressing feedback and noise, called PEM-GFIC. The scheme resembles to the GEIC proposed in [17] for combined echo cancellation and noise reduction. However, because of the closed signal loop $G(q)$, a different optimization criterion is required for the adaptive filters. The PEM-GFIC integrates adaptive feedback cancellation into the GSC. It consists of a fixed spatial pre-processor (i.e., a fixed beamformer $\mathbf{W}_q(q)$ and a blocking matrix $\mathbf{B}(q)$) and an adaptive stage. As in the GSC, the fixed beamformer $\mathbf{W}_q(q)$ creates a speech reference

$$y_0[k] = \mathbf{W}_q^T(q)\mathbf{x}[k]. \tag{11}$$

A delay d_1 is (standardly) applied to the speech reference to allow for non-causal filter taps in the adaptive stage $\mathbf{W}_a(q)$. To simplify the notation, this delay is included in $\mathbf{W}_q(q)$. The blocking matrix $\mathbf{B}(q)$ creates $M-1$ noise references

$$\mathbf{y}[k] = \left[\begin{array}{cccc} y_1[k] & y_2[k] & \dots & y_{M-1}[k] \end{array} \right]^T = \mathbf{B}^T(q)\mathbf{x}[k]. \tag{12}$$

The adaptive stage consists of an ANC $\mathbf{W}_a(q)$, which operates on the noise references $\mathbf{y}[k]$, and an adaptive feedback canceller $W_f(q)$, which operates on the loudspeaker signal $u[k]$. The loudspeaker signal $u[k]$ can be viewed as an additional noise reference to the GSC.

The output $e[k]$ of the PEM-GFIC equals

$$e[k] \;=\; y_0[k] - \left[\begin{array}{cc} \mathbf{W}_a^T(q) & W_f(q) \end{array} \right] \left[\begin{array}{c} \mathbf{y}[k] \\ u[k] \end{array} \right]. \tag{13}$$

Using (5), $e[k]$ can be decomposed into a noise reduction term $e_{NR}[k]$ and a feedback cancellation term $e_{FC}[k]$ (cf. (8)) as:

$$\begin{aligned} e[k] \;=\; & \underbrace{\left(y_0^s[k] + y_0^n[k] - \left[\begin{array}{cc} \mathbf{W}_a^T(q) & W_f(q) \end{array} \right] \left[\begin{array}{c} \mathbf{y}^s[k] + \mathbf{y}^n[k] \\ 0 \end{array} \right] \right)}_{e_{NR}[k]} \\ & + \underbrace{\left(y_0^u[k] - \left[\begin{array}{cc} \mathbf{W}_a^T(q) & W_f(q) \end{array} \right] \left[\begin{array}{c} \mathbf{y}^u[k] \\ u[k] \end{array} \right] \right)}_{e_{FC}[k]} \end{aligned} \tag{14}$$

where

$$y_0^s[k] + y_0^n[k] = \mathbf{W}_q^T(q)\left(\mathbf{x}^s[k] + \mathbf{x}^n[k]\right), \quad (15)$$

$$\mathbf{y}^s[k] + \mathbf{y}^n[k] = \mathbf{B}^T(q)\left(\mathbf{x}^s[k] + \mathbf{x}^n[k]\right), \quad (16)$$

$$y_0^u[k] = \mathbf{W}_q^T(q)\mathbf{F}(q)u[k], \quad (17)$$

$$\mathbf{y}^u[k] = \mathbf{B}^T(q)\mathbf{F}(q)u[k]. \quad (18)$$

Suppose that the adaptive stage $\begin{bmatrix} \mathbf{W}_a^T(q) & W_f(q) \end{bmatrix}$ is designed to minimize $\varepsilon\{e^2[k]\}$, as is the case in the GEIC. Then, $\begin{bmatrix} \mathbf{W}_a^T(q) & W_f(q) \end{bmatrix}$ removes the part of the speech reference signal $y_0[k]$ that is correlated with the noise references $\mathbf{y}[k]$ and $u[k]$. Ideally, we would like the filter $\mathbf{W}_a(q)$ to remove the noise $y_0^n[k]$ and the filters $\mathbf{W}_a(q)$ and $W_f(q)$ to remove the acoustic feedback $y_0^u[k]$. To achieve this goal, the loudspeaker signal $u[k]$ should be uncorrelated with the signal $e_{NR}[k]$ in (14). Then, minimization of $\varepsilon\{e^2[k]\}$ corresponds to minimization of

$$\varepsilon\{e_{NR}^2[k]\} + \varepsilon\{e_{FC}^2[k]\}, \quad (19)$$

resulting in the desired solution (assuming infinite filter lengths)

$$\mathbf{W}_a(q) = \text{argmin } \varepsilon\{e_{NR}^2[k]\} \text{ (i.e., as if there were no feedback)}, \quad (20)$$

$$W_f(q) = \text{argmin } \varepsilon\{e_{FC}^2[k]\} = \mathbf{W}_q^T(q)\mathbf{F}(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\mathbf{F}(q) \text{ (so that } e_{FC}[k] = 0). \quad (21)$$

In an echo cancellation set-up, the far-end signal $u[k]$ and the near-end signals $\mathbf{x}^s[k]$, $\mathbf{x}^n[k]$ (and hence $e_{NR}[k]$) are indeed uncorrelated, so that the desired solution (20)-(21) is achieved. In a feedback cancellation set-up, however, $u[k]$ and $e_{NR}[k]$ are typically correlated because of the presence of the closed signal loop $G(q)$. Using (5), the loudspeaker signal $u[k]$ can indeed be expressed as a function of the error signal $e_{NR}[k]$:

$$u[k] = \frac{G(q)}{1 - G(q)\left(\mathbf{W}_q^T(q)\mathbf{F}(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\mathbf{F}(q) - W_f(q)\right)} e_{NR}[k]. \quad (22)$$

Assuming that $G(q)$ contains a delay of (at least) one sample, $u[k]$ is uncorrelated with $e_{NR}[k]$ if $e_{NR}[k]$ is a white noise sequence. In practice, $e_{NR}[k]$ (i.e., the desired speech and residual noise) is spectrally colored and hence, $u[k]$ and $e_{NR}[k]$ are correlated. As a result, the adaptive filters $W_f(q)$ and $\mathbf{W}_a(q)$ will aim to cancel the desired signal $y_0^s[k]$ and the noise $y_0^n[k]$ instead of the feedback signal $y_0^u[k]$. To avoid this, a pre-whitening operation has to be included that decorrelates $u[k]$ and $e_{NR}[k]$ (cf. [13]).

PEM-based cost function

First, consider the simplified situation where noise is absent, i.e., $\mathbf{x}^n[k] = 0$. Assuming a perfect blocking matrix,

$$e_{NR}[k] = y_0^s[k]. \tag{23}$$

Suppose that the desired speech signal $y_0^s[k]$ can be modelled as an auto-regressive (AR) process

$$y_0^s[k] = \frac{1}{P(q)} w[k] \tag{24}$$

with $w[k]$ a white noise sequence[1] and $P(q)$ a finite-order polynomial. Assume that $P(q)$ is known.

Then, optimization of the PEM-based cost function

$$\{\mathbf{W}_a(q),\, W_f(q)\} \;\; = \;\; \text{argmin } \varepsilon\{(e^p[k])^2\} \tag{25}$$

with $e^p[k]$ the prediction-error

$$e^p[k] \;\; = \;\; A(q)\left(\mathbf{W}_q^T(q)\mathbf{x}[k] - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\mathbf{x}[k] - W_f(q)u[k]\right) \tag{26}$$

$$A(q) \;\; = \;\; P(q), \tag{27}$$

leads to the desired solution (21) for the feedback cancellation filter $W_f(q)$.

Proof: Indeed, using $\mathbf{B}^T\mathbf{x}^s[k] = 0$ and (22),

$$e^p[k] \;\; = \;\; A(q)\mathbf{W}_q^T(q)\mathbf{x}^s[k] + G(q)\left(\mathbf{W}_q^T(q)\mathbf{F}(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\mathbf{F}(q) - W_f(q)\right)A(q)u[k] \tag{28}$$

with

$$u[k] = \frac{G(q)\mathbf{W}_q^T(q)\mathbf{x}^s[k]}{1 - G(q)\left(\mathbf{W}_q^T(q)\mathbf{F}(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\mathbf{F}(q) - W_f(q)\right)}. \tag{29}$$

Since $A(q)\mathbf{W}_q^T(q)\mathbf{x}^s[k] = w[k]$ and $G(q)$ contains a delay $d_G \geq 1$, $A(q)u[k]$ is uncorrelated with $A(q)\mathbf{W}_q^T(q)\mathbf{x}^s[k]$. Hence, minimization of $\varepsilon\{(e^p[k])^2\}$ results in the desired solution (21) for $W_f(q)$. ■

In practice, noise may be present and $y_0^s[k]$ is unknown and time-varying. Hence, the AR-model $A(q)$ should be estimated online on the output $e[k]$ of the PEM-GFIC. Below, we identify the conditions under which

$$\min_{A(q),\mathbf{W}_a(q),W_f(q)} \varepsilon\{(e^p[k])^2\}, \tag{30}$$

where $\mathbf{x}^n[k] \neq 0$ and where $A(q)$ is then also an optimization variable, results in the unique desired solution for $A(q)$, $\mathbf{W}_a(q)$ and $W_f(q)$. To avoid the trivial solution $A(q) = 0$, the first tap a_0 is set to 1.

[1]For voiced speech, the excitation approximates a pulse train. The pulse train excitation can be modelled by cascading a short-term AR-model $P_{ST}(q)$ with a long-term predictor $P_{LT}(q)$, i.e., $P(q) = P_{ST}(q)P_{LT}(q)$ [14].

B. Conditions for identifiability

Using (5) and (22), the error signal $e^p[k]$ can be expressed as:

$$e^p[k] \;=\; (A(q)+A(q)G(q)Q(q))\underbrace{\left(\mathbf{W}_q^T(q)-\mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\left(\mathbf{x}^s[k]+\mathbf{x}^n[k]\right)}_{e_{NR}[k]}, \tag{31}$$

with

$$Q(q)=\frac{\mathbf{W}_q^T(q)\mathbf{F}(q)-\mathbf{W}_a^T(q)\mathbf{B}^T(q)\mathbf{F}(q)-W_f(q)}{1-G(q)\left(\mathbf{W}_q^T(q)\mathbf{F}(q)-\mathbf{W}_a^T(q)\mathbf{B}^T(q)\mathbf{F}(q)-W_f(q)\right)} \tag{32}$$

a causal filter. Since the first tap a_0 of $A(q)$ equals 1 and $G(q)$ has a delay of at least one sample, $A(q)\left(1+G(q)Q(q)\right)$ can be interpreted as a linear prediction filter of $e_{NR}[k]$. If $G(q)=q^{-d_G}\bar{G}(q)$ with $d_G \geq L_A$, where L_A is the filter length of $A(q)$, the linear prediction coefficients $A(q)+A(q)G(q)Q(q)$ uniquely determine the $A(q)$ and $Q(q)$. In the sequel, two cases are considered, namely the case of a perfect blocking matrix and the case of an imperfect blocking matrix.

a) Perfect blocking matrix:

Lemma 1: Assume that $y_0^s[k]=\mathbf{W}_q^T(q)\mathbf{x}^s[k]$ can be modelled as

$$y_0^s[k]=\frac{1}{P(q)}w[k] \tag{33}$$

with $P(q)$ a finite order polynomial with length L_P and $w[k]$ a white noise sequence and assume that perfect noise reduction can be achieved by $\mathbf{W}_a(q)$. Then, assuming a perfect blocking matrix and the infinite order case for $W_f(q)$, $\mathbf{W}_a(q)$ and $L_P \leq L_A$ with L_A the filter length of $A(q)$,

$$\min_{A(q),\mathbf{W}_a(q),W_f(q)} \varepsilon\{(e^p[k])^2\} \tag{34}$$

results in the unique desired solution

$$A(q) \;=\; P(q) \tag{35}$$

$$\mathbf{W}_a(q) \;=\; \text{argmin } \varepsilon\{e_{NR}^2[k]\} \text{ (as if there were no feedback)}, \tag{36}$$

$$W_f(q) \;=\; \mathbf{W}_q^T(q)\mathbf{F}(q)-\mathbf{W}_a^T(q)\mathbf{B}^T(q)\mathbf{F}(q) \text{ (hence, } e_{FC}[k]=0), \tag{37}$$

if the delay d_G of $G(q)$ satisfies $d_G \geq L_A$.

Proof: Assuming a *perfect blocking matrix*, i.e., $\mathbf{B}^T(q)\mathbf{x}^s[k]=0$, (31) can be re-written as:

$$\begin{aligned} e^p[k] \;=\; & \underbrace{A(q)\left(1+G(q)Q(q)\right)\mathbf{W}_q^T(q)\mathbf{x}^s[k]}_{e^{p,s}[k]} \\ & +\underbrace{A(q)\left(1+G(q)Q(q)\right)\left(\mathbf{W}_q^T(q)-\mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\mathbf{x}^n[k]}_{e^{p,n}[k]}. \end{aligned} \tag{38}$$

Since $\mathbf{x}^s[k]$ and $\mathbf{x}^n[k]$ are uncorrelated, minimizing $\varepsilon\{(e^p[k])^2\}$ is equivalent to minimizing $\varepsilon\{(e^{p,s}[k])^2\}+\varepsilon\{(e^{p,n}[k])^2\}$. Minimization of $\varepsilon\{(e^{p,s}[k])^2\}$ with $a_0 = 1$ and $d_G \geq 1$ corresponds to a linear prediction of $y_0^s[k] = \mathbf{W}_q^T\mathbf{x}^s[k]$ and thus results in

$$A(q) + A(q)G(q)Q(q) = P(q). \tag{39}$$

If $d_G \geq L_A$, the solution of (39) is unique and equals

$$A(q) = P(q), \tag{40}$$

$$Q(q) = 0. \tag{41}$$

Using (32), this corresponds to the desired solution (37) for $W_f(q)$. The filter $A(q)$ models the speech reference $y_0^s[k]$ and $W_f(q)$ cancels the acoustic feedback. Note that in the absence of noise (i.e., $\mathbf{x}^n[k] = 0$), multiple solutions for $W_f(q)$ and $\mathbf{W}_a(q)$ exist, but they will all result in a perfect feedback cancellation. Using (39),

$$e^{p,n}[k] = P(q)\left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\mathbf{x}^n[k], \tag{42}$$

$$= P(q)e_{NR}[k]. \tag{43}$$

Hence, if perfect noise reduction can be achieved by $\mathbf{W}_a(q)$, $\varepsilon\{(e^{p,n}[k])^2\}$ and $\varepsilon\{(e^{p,s}[k])^2\}$ are both minimized if and only if

$$\mathbf{W}_q^T(q)\mathbf{F}(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\mathbf{F}(q) - W_f(q) = 0 \text{ (i.e., perfect feedback cancellation)}$$

$$\left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\mathbf{x}^n[k] = 0 \text{ (i.e., perfect noise reduction).}$$

■

If the noise cannot be completely reduced by $\mathbf{W}_a(q)$,

$$\mathbf{W}_a(q) = \text{argmin } \varepsilon\{e_{NR}^2[k]\} \tag{44}$$

and $A(q)\left(1 + G(q)Q(q)\right)$ will model the speech and residual noise signal $e_{NR}[k]$. Generally, $A(q)$ will not be able to perfectly whiten both the speech and the residual noise so that there will be some residual bias in $W_f(q)$.

b) Imperfect blocking matrix: If the *blocking matrix is imperfect*, i.e., $\mathbf{B}^T(q)\mathbf{x}^s[k] \neq 0$, the desired speech signal leaks into the noise references so that $e_p[k]$ equals

$$e_p[k] = A(q)\left(1 + G(q)Q(q)\right)\left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\left(\mathbf{x}^s[k] + \mathbf{x}^n[k]\right). \tag{45}$$

The filter $\mathbf{W}_a(q)$ will not only aim to suppress the noise and feedback, but also the desired speech signal $y_0^s[k]$ so that the desired speech gets distorted: the larger the speech leakage with respect to the noise in the noise references, the larger the speech cancellation. The filter $A(q)(1+G(q)Q(q))$ will now whiten the residual speech and noise signal. If $A(q)$ perfectly whitens the residual speech and noise, the feedback path estimate $W_f(q)$ will be unbiased. To limit the speech distortion by $\mathbf{W}_a(q)$, a quadratic inequality constraint can be imposed on $\mathbf{W}_a(q)$, as in [21].

Remark 1: In the above reasoning, we assumed that speech was continuously present, i.e., $\mathbf{x}^s[k] \neq 0$. A special situation occurs if $\mathbf{x}^s[k] = 0$, i.e., during speech pauses. In this case, the filters $A(q)$, $\mathbf{W}_a(q)$ and $W_f(q)$ minimize $\varepsilon\{(e^{p,n}[k])^2\}$. If perfect noise reduction can be achieved by the filter $\mathbf{W}_a(q)$, the solution for $A(q)\,(1+G(q)Q(q))$ is not determined: an infinite number of solutions for $A(q)$ and $W_f(q)$ then exist, including solutions $A(q)$ and $W_f(q)$ that result in an unstable closed-loop system. In practice, reverberation and sensor noise will always be present so that not all the noise can be suppressed by $\mathbf{W}_a(q)$. As a result, some residual noise (and hence, feedback of the noise) will always remain. The filter $A(q)$ will then whiten the residual noise. If the whitening is perfect, the filters $\mathbf{W}_a(q)$ and $W_f(q)$ will aim to optimally suppress the feedback and noise.

C. Simulations

To illustrate the performance of the PEM-GFIC, Figure 3 depicts the noise reduction and feedback cancellation performance as a function of the SNR at the first microphone of a 2-microphone BTE. For details on the implementation, the speech and noise source positions and the acoustic feedback paths, we refer to Section VI. The speech and the noise signals $x_i^s[k]$ and $x_i^n[k]$ are anechoic recordings of stationary speech weighted noise (generated as white Gaussian noise passed through a 20th order all-pole filter) with a commercial 2-microphone BTE. To avoid effects caused by speech leakage, a perfect blocking matrix $\mathbf{B}(q)$ is used in the GSC, i.e., $\mathbf{y}^s[k] = 0$. The performance is assessed for two different gains in the forward path $G(q) = Gq^{-d_G}$: a gain below the stability margin, i.e., $G = 6$ and a gain above the stability margin, i.e., $G = 15$, corresponding to an output $\text{SFR}_{\text{intellig}}$ of 17 dB and $-\infty$ (because of instability), respectively, in the absence of the filters $W_f(q)$ and $\mathbf{W}_a(q)$.

a) Performance of PEM-GFIC: Figure 3 shows the improvement in $\text{SNR}_{\text{intellig}}$ ($\Delta\text{SNR}_{\text{intellig}}$) with respect to the speech reference and the output $\text{SFR}_{\text{intellig}}$ as a function of the input SNR. The stars and diamonds in Figure 3(b) depict the $\text{SFR}_{\text{intellig}}$ of the speech reference $y_0[k]$. For comparison, also the performance measures for the PEM-GSC (i.e., the integrated scheme without filter $W_f(q)$) are shown. The PEM-GFIC clearly trades off between noise reduction and feedback cancellation. At high gains and

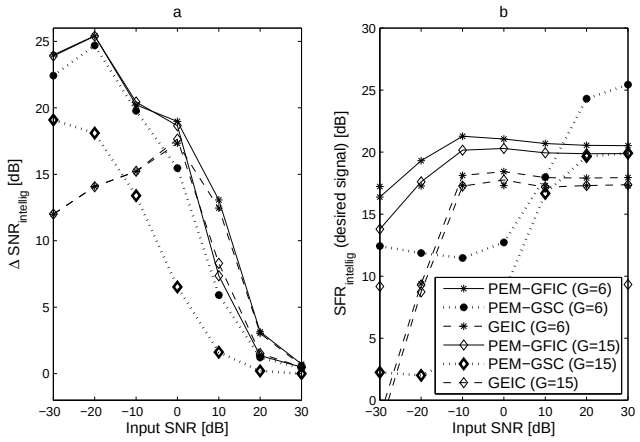


Fig. 3. Performance of the PEM-GFIC , the PEM-GSC and the GEIC: a) $\Delta\text{SNR}_{\text{intellig}}$ (dB); (b) output $\text{SFR}_{\text{intellig}}$ of speech (dB) ; The stars ($G = 6$) and diamonds ($G = 15$) denote the intelligibility-weighted speech-to-feedback ratio of the speech reference $y_0[k]$ for the PEM-GFIC.

high SNRs, more feedback suppression is achieved than at low gains and low SNRs at the expense of a reduced noise suppression. Indeed, the improvement in $\text{SFR}_{\text{intellig}}$ compared to the speech reference $y_0[k]$ increases with increasing gain and input SNR. Without feedback cancellation filter $W_f(q)$, the filter $\mathbf{W}_a(q)$ has to suppress both the noise and the feedback. At low input SNRs, $\mathbf{W}_a(q)$ concentrates on noise reduction, at the expense of a reduced feedback suppression. At high input SNRs, $\mathbf{W}_a(q)$ focuses on feedback cancellation. Because of the smaller number of degrees of freedom and the short acoustic feedback paths, the noise and feedback suppression obtained with the PEM-GSC is generally worse than with PEM-GFIC. In public-address systems where the feedback paths are much longer, the PEM-GSC may achieve similar performance than the PEM-GFIC for the same computational cost [16].

b) Comparison with GEIC: Figure 3 also depicts the $\Delta\text{SNR}_{\text{intellig}}$ and the output $\text{SFR}_{\text{intellig}}$ of the GEIC, respectively, as a function of the input SNR. Due to correlation between the loudspeaker signal $u[k]$ and $y_0^s[k]+y_0^n[k]$, the filter $W_f(q)$ and $\mathbf{W}_a(q)$ aim to remove $y_0^s[k]$ and $y_0^n[k]$ instead of the feedback signal $y_0^u[k]$, resulting in a biased feedback path estimate. This explains the worse feedback cancellation performance compared to the PEM-GFIC. The bias in $W_f(q)$ may result in a closed-loop system that is close to instability. In Figure 3, this is the case at low input SNRs, explaining the very low output $\text{SFR}_{\text{intellig}}$. At low input SNRs, the filters $W_f(q)$ and $\mathbf{W}_a(q)$ both concentrate on noise reduction. The feedback cancellation filter $W_f(q)$ in addition to $\mathbf{W}_a(q)$ reduce the noise $y_0^n[k]$, which explains why the noise reduction by $\mathbf{W}_a(q)$ decreases with decreasing input SNR.

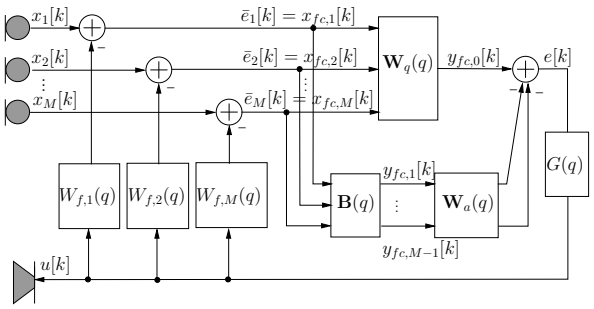


Fig. 4. Feedback cancellation followed by noise reduction ('AFC first').

IV. Feedback cancellation followed by noise reduction ('AFC first')

In this and the next section, cascades of GSC-based noise reduction and PEM-based adaptive feedback cancellation are discussed. These cascades will be compared with the integrated PEM-GFIC solution of Section III. To avoid desired signal cancellation, the PEM-GSC of [16] is used as a noise reduction stage in both cascaded schemes. Interactions between the noise reduction stage and feedback cancellation stage will be studied.

A straightforward approach is to first cancel the feedback in all microphone signals independently and then to perform noise reduction on the feedback-compensated microphone signals. This is depicted in Figure 4. In the sequel, this approach is referred to as 'AFC first'. The feedback cancellation stage consists of M adaptive feedback cancellers, resulting in a high computational cost.

A. Concept

Feedback cancellation stage

In a first stage, the feedback in each microphone signal $x_i[k]$ is removed independently through a PEM-AFC $W_{f,i}(q)$ [13]. The $W_{f,i}(q)$, $i = 1, \ldots, M$ minimize the prediction-errors

$$\bar{e}_i^p[k] \;=\; \bar{A}_i(q)\left(x_i^s[k] + x_i^n[k] + \left(F_i(q) - W_{f,i}(q)\right)u[k]\right), \tag{46}$$

where $\bar{A}_i(q)$ is a finite order polynomial with length $L_{\bar{A}_i}$, which whitens the clean (i.e., without feedback) signal component $x_i^s[k] + x_i^n[k]$ in the i-th microphone [13].

The loudspeaker signal $u[k]$ equals

$$u[k] = \frac{G(q)\left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T\right)\left(\mathbf{x}^s[k] + \mathbf{x}^n[k]\right)}{1 - G(q)\left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\left(\mathbf{F}(q) - \mathbf{W}_f(q)\right)}. \tag{47}$$

Define $\mathbf{H}^s(q)$ and $\mathbf{H}^n(q)$ as the vectors containing the transfer function ratios from the speech and the noise source to the microphones, i.e.,

$$\mathbf{H}^s(q) = \left[\begin{array}{ccc} H_1^s(q) & \ldots & H_M^s(q) \end{array} \right]^T, \tag{48}$$

$$\mathbf{H}^n(q) = \left[\begin{array}{ccc} H_1^s(q) & \ldots & H_M^s(q) \end{array} \right]^T. \tag{49}$$

Then, using

$$\mathbf{x}^s[k] = \frac{\mathbf{H}^s(q)}{H_i^s(q)} x_i^s[k], \quad \mathbf{x}^n[k] = \frac{\mathbf{H}^n(q)}{H_i^n(q)} x_i^n[k], \tag{50}$$

$u[k]$ can be expressed as a function of $x_i^s[k]$ and $x_i^n[k]$:

$$u[k] = \frac{G(q)\left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\left(\frac{\mathbf{H}^s(q)}{H_i^s(q)} x_i^s[k] + \frac{\mathbf{H}^n(q)}{H_i^n(q)} x_i^n[k]\right)}{1 - G(q)\left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\left(\mathbf{F}(q) - \mathbf{W}_f(q)\right)}. \tag{51}$$

Since $x_i^s[k]$ and $x_i^n[k]$ are uncorrelated, minimization of $\varepsilon\{(\bar{e}_i^p[k])^2\}$ is equivalent to minimization of

$$\varepsilon\{(\bar{e}_i^{p,s}[k])^2\} + \varepsilon\{(\bar{e}_i^{p,n}[k])^2\}, \tag{52}$$

where

$$\bar{e}_i^{p,s}[k] = \bar{A}_i(q)\left(x_i^s[k] + \left(F_i(q) - W_{f,i}(q)\right) u^s[k]\right), \tag{53}$$

$$\bar{e}_i^{p,n}[k] = \bar{A}_i(q)\left(x_i^n[k] + \left(F_i(q) - W_{f,i}(q)\right) u^n[k]\right), \tag{54}$$

and $u^s[k]$ and $u^n[k]$ are the speech and the noise component of the loudspeaker signal. According to [13], the following lemma applies.

Lemma 2: Assume that

$$x_i^s[k] = \frac{1}{P_i^s(q)} w^s[k] \text{ and } x_i^n[k] = \frac{1}{P_i^n(q)} w^n[k] \tag{55}$$

with $w^s[k]$, $w^n[k]$ white noise sequences and $P_i^s(q)$ and $P_i^n(q)$ finite order polynomials with length $L_{p_i^s}$ and $L_{P_i^n}$, respectively, and assume that $\bar{A}_i(q)$ has a filter length $L_{\bar{A}_i} \geq \max(L_{P_i^s}, L_{P_i^n})$. If $G(q)\frac{\mathbf{H}^s(q)}{H_i^s(q)}$ and $G(q)\frac{\mathbf{H}^n(q)}{H_i^n(q)}$ contain a delay $\geq L_{P_i^s}$ and $\geq L_{P_i^n}$, respectively, minimization of $\varepsilon\{(\bar{e}_i^{p,s}[k])^2\}$ results in the solution (cf. [13])

$$\bar{A}_i(q) = P_i^s(q); \quad W_{f,i}(q) = F_i(q), \tag{56}$$

while minimization of $\varepsilon\{(\bar{e}_i^{p,n}[k])^2\}$ results in

$$\bar{A}_i(q) = P_i^n(q); \quad W_{f,i}(q) = F_i(q). \tag{57}$$

Hence, if only speech or noise is present, $\bar{A}_i(q)$ and $W_{f,i}(q)$ converge to the desired solution. If, however, both speech and noise are present, $\bar{A}_i(q)$ should both model the speech model $P_i^s(q)$ and the noise model $P_i^n(q)$. Generally $P_i^s(q) \neq P_i^n(q)$ so that a residual bias in the feedback path estimate $W_{f,i}(q)$ will exist.

Remark 2: Given the small microphone interspacings in hearing aids,

$$x_i^s[k] \approx x_j^s[k] \text{ and } x_i^n[k] \approx x_j^n[k], \tag{58}$$

so that a common AR-model $\bar{A}(q)$ can be used in the feedback cancellers $W_{f,i}(q)$. In this paper, this common AR-model $\bar{A}(q)$ is estimated by minimizing the sum of the mean squared errors $\varepsilon\{\big(\bar{A}(q)x_{fc,i}[k]\big)^2\}$, i.e.,

$$\sum_{i=1}^{M} \varepsilon\{\big(\bar{A}(q)x_{fc,i}[k]\big)^2\}, \tag{59}$$

where

$$x_{fc,i}[k] = x_i[k] - W_{f,i}(q)u_i[k].$$

The common AR-model $\bar{A}(q)$ gives results similar to those obtained with an individual AR-model $\bar{A}_i(q)$ for each microphone signal. To remove possible microphone mismatch, the microphones were calibrated.

Noise reduction stage

In the second stage, the noise is removed from the feedback-compensated microphone signals $x_{fc,i}[k]$ by a GSC, which consists of a fixed beamformer $\mathbf{W}_q(q)$, a blocking matrix $\mathbf{B}(q)$ and an ANC $\mathbf{W}_a(q)$ (see Figure 4). The ANC $\mathbf{W}_a(q)$ removes the part of the speech reference $y_0[k]$ that is correlated with the noise references $\mathbf{y}[k]$. We will assume a perfect blocking matrix, i.e., $\mathbf{B}^T\mathbf{x}^s[k] = 0$. If the acoustic feedback is completely removed by the feedback cancellation stage, a standard ANC can be used in the second stage. In practice, however, there will be some residual feedback in the feedback-compensated signals $x_{fc,i}[k]$, $i = 1, \ldots, M$ and hence, in the noise references. This results in speech leakage so that desired signal distortion or cancellation may occur (cf. Section III) [16]. To avoid such signal distortion, the PEM-GSC of [16] is used instead of the standard GSC.

The optimization criterion of the filters $A(q)$ and $\mathbf{W}_a(q)$ corresponds to the optimization criterion of a PEM-GFIC with $W_f(q) = 0$ that now operates on the feedback compensated microphone signals $\mathbf{x}_{fc}[k] = \mathbf{x}[k] - \mathbf{W}_f(q)u[k]$ instead of the original microphone signals $\mathbf{x}[k]$, i.e.,

$$\{A(q),\, \mathbf{W}_a(q)\} = \arg\min \varepsilon\{(e^p[k])^2\}, \tag{60}$$

with

$$e^p[k] = A(q)\left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\underbrace{\left(\mathbf{x}^s[k] + \mathbf{x}^n[k] + \left(\mathbf{F}(q) - \mathbf{W}_f(q)\right)u[k]\right)}_{\mathbf{x}_{fc}[k]}, \tag{61}$$

and

$$u[k] = \frac{G(q)\left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\left(\mathbf{x}^s[k] + \mathbf{x}^n[k]\right)}{1 - G(q)\left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\left(\mathbf{F}(q) - \mathbf{W}_f(q)\right)}. \quad (62)$$

Lemma 3: Assume that $y_0^s[k] = \mathbf{W}_q^T(q)\mathbf{x}^s[k]$ can be modelled as

$$y_0^s[k] = \frac{1}{P(q)}w[k] \quad (63)$$

with $P(q)$ a finite order polynomial with length L_P and assume that $\mathbf{W}_a(q)$ has sufficient degrees of freedom to suppress both the noise and the residual feedback. Then, assuming a perfect blocking matrix, the infinite order case for $\mathbf{W}_a(q)$ and $L_P \leq L_A$,

$$\min_{A(q),\mathbf{W}_a(q)} \varepsilon\{(e^p[k])^2\} \quad (64)$$

results in the unique desired solution

$$A(q) = P(q), \quad (65)$$

$$\left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\mathbf{x}^n[k] = 0 \text{ (i.e.,perfect noise reduction)}, \quad (66)$$

$$\left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\left(\mathbf{F}(q) - \mathbf{W}_f(q)\right) = 0 \text{ (i.e., perfect feedback cancellation)}, \quad (67)$$

if $d_G \geq L_A$.

Proof: The proof is similar to the one of Section III with now $W_f(q) = 0$. Assuming a perfect blocking matrix, the prediction error $e^p[k]$ equals:

$$\begin{aligned} e^p[k] = & \underbrace{A(q)\left(1 + G(q)Q(q)\right)\mathbf{W}_q^T(q)\mathbf{x}^s[k]}_{e^{p,s}[k]} \\ & + \underbrace{A(q)\left(1 + G(q)Q(q)\right)\left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\mathbf{x}^n[k]}_{e^{p,n}[k]} \end{aligned} \quad (68)$$

with

$$Q(q) = \frac{\left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\left(\mathbf{F}(q) - \mathbf{W}_f(q)\right)}{1 - G(q)\left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\left(\mathbf{F}(q) - \mathbf{W}_f(q)\right)} \quad (69)$$

a causal filter. Minimization of $\varepsilon\{(e^{p,s}[k])^2\}$ with $a_0 = 1$ and $d_G \geq L_A$ results in the desired solution (65) and (67) for $A(q)$ and $W_f(q)$ (cf. Section III). The term $\varepsilon\{(e^{p,n}[k])^2\}$ is minimized if

$$\left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\mathbf{x}^n[k] = 0 \text{ (i.e., perfect noise reduction)}.$$

Hence, if the filter $\mathbf{W}_a(q)$ has sufficient degrees of freedom to completely suppress the noise and the residual feedback, the desired solution (65)-(67) is achieved. ■

If $\mathbf{W}_a(q)$ has insufficient degrees of freedom to suppress the noise and the residual feedback, $\mathbf{W}_a(q)$ trades off between noise reduction and feedback cancellation and $A(q)$ will not only aim to whiten the speech but also the residual noise.

B. Interaction between the noise reduction and feedback cancellation stage

The noise reduction stage affects the performance of the feedback cancellation stage and vice-versa[2].

The performance of the feedback cancellers $W_{f,i}(q)$, $i = 1, \ldots, M$ is determined by the energy and SNR of the feedback signal $F_i(q)u[k]$ compared to the energy and SNR of the near-end signal $x_i^s[k] + x_i^n[k]$: the larger the energy ratio, the smaller the misadjustment and the bias in case of imperfect decorrelation by $\bar{A}(q)$. The noise canceller $\mathbf{W}_a(q)$ reduces the output noise power and hence the noise energy in the loudspeaker signal $u[k]$, especially at low SNR (see (51)). As a result, the energy ratio of the feedback signal to the near-end signal decreases, degrading the performance of the feedback cancellers $W_{f,i}(q)$. The feedback to near-end signal ratio may especially be small for low input SNRs (e.g., during speech pauses) and large SNR improvements by $\mathbf{W}_a(q)$, causing large misadjustments of the feedback cancellers $W_{f,i}(q)$. On the other hand, the time-variance and a possible delay in the noise cancellation stage reduces the correlation between the loudspeaker signal $u[k]$ and the near-end signal $x_i^s[k] + x_i^n[k]$, which may then improve the performance of the feedback cancellers.

The performance of the noise cancellation stage similarly depends on the feedback cancellers $\mathbf{W}_f(q)$. Typically, the filter $\mathbf{W}_a(q)$ trades off between noise reduction and feedback cancellation. The more effective the feedback cancellers $\mathbf{W}_f(q)$, the more the filter $\mathbf{W}_a(q)$ concentrates on noise reduction.

C. Simulations

a) Interactions between the feedback cancellation stage and the noise reduction stage: To illustrate the interactions between the feedback cancellation stage and the noise reduction stage, Figure 5 depicts the SNR-improvement, $\Delta\text{SNR}_{\text{intellig}}$, with respect to the speech reference and the output $\text{SFR}_{\text{intellig}}$ of 'AFC first' as a function of the input SNR. The same set-up as in Section III has been used. For comparison, the $\text{SFR}_{\text{intellig}}$ of the speech reference $y_{0,fc}[k] = \mathbf{W}_q^T(q)\mathbf{x}_{fc}[k]$ is also depicted (dotted lines), illustrating the feedback cancellation performance achieved by the feedback cancellation stage. The noise reduction performance generally increases with decreasing input SNR and is rather independent of the gain G. Most of the feedback in the speech reference is removed by the feedback cancellation stage so that the second stage basically focuses on noise reduction.

The feedback suppression by the feedback cancellation stage increases with increasing input SNR and with increasing gain. Indeed, the improvement in $\text{SFR}_{\text{intellig}}$ compared to the uncompensated speech reference $\mathbf{W}_q^T(q)\mathbf{x}[k]$ (represented in Figure 5 by the stars and diamonds) increases with increasing gain

[2]Note that the former is not the case for combined echo cancellation and noise reduction because of the absence of $G(q)$.

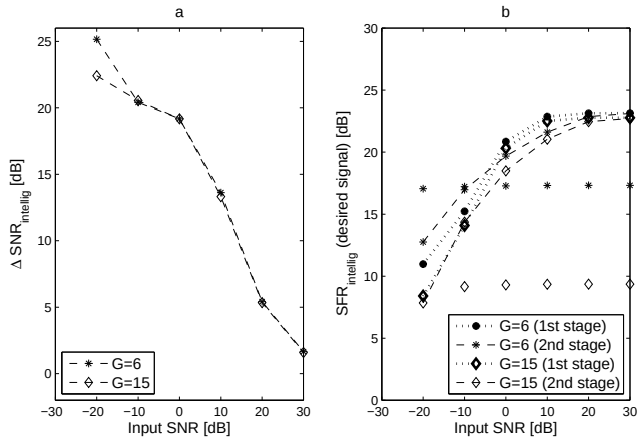


Fig. 5. Performance of 'AFC first': a) ΔSNR$_{\text{intellig}}$ (dB); (b) Output SFR$_{\text{intellig}}$ of speech (dB). The stars and diamonds denote the SFR$_{\text{intellig}}$ of $\mathbf{W}_q^T(q)\mathbf{x}[k]$ (i.e., the speech reference without feedback compensation).

and input SNR. The larger the input SNR, the larger the feedback to near-end signal ratio and hence, the higher the accuracy of the feedback cancellation stage. At low SNRs, the noise reduction by the GSC reduces the energy of the loudspeaker signal $u[k]$ and hence, reduces the feedback to near-end signal ratio at the input of the feedback cancellers $W_{f,i}(q)$. This results in a larger misadjustment of the feedback cancellers $W_{f,i}(q)$: For the input SNRs of -30 dB instability even occurs.

A smaller feedback suppression in the first stage results in more focus towards feedback suppression in the second stage. For the input SNR of -10 dB and -20 dB and $|G(q)| = 6$, the drop in feedback suppression by the first stage at low SNR is (partially) compensated by the second stage. At very high SNRs and high gains (where the first stage achieves good feedback cancellation), the second stage basically concentrates on noise reduction, so that no improvement or even, a small decrease in SFR$_{\text{intellig}}$ is achieved.

b) Comparison with PEM-GFIC: For high input SNRs, 'AFC first' obviously outperforms the PEM-GFIC in terms of feedback cancellation: at high input SNRs, feedback cancellation is the most important aim and 'AFC first' has an M times larger number of degrees of freedom. At high gains and very high SNRs, 'AFC first' also achieves more noise reduction. Because of the effective feedback cancellation by the first stage, the noise reduction stage of 'AFC first' concentrates more on noise reduction than the PEM-GFIC (especially for high gains). In contrast to 'AFC first', noise reduction by the PEM-GFIC does not degrade the feedback to near-end signal ratio at the input of the feedback canceller so that the PEM-GFIC outperforms 'AFC first' at low input SNRs.

V. NOISE REDUCTION FOLLOWED BY FEEDBACK CANCELLATION ('GSC FIRST')

A second approach is to first perform noise reduction, followed by a single feedback canceller $W_f(q)$ (cf. Figure 6). Compared to 'AFC first', the computational cost of the feedback cancellation stage is then reduced to a single feedback canceller. In addition, the noise reduction stage reduces the disturbance to the feedback cancellation stage and hence, the misadjustment of the feedback canceller $W_f(q)$. However, in this set-up, the feedback canceller should not only model the feedback paths $F_i(q)$, but also the adaptive noise reduction stage $\mathbf{W}_a(q)$. For the feedback canceller $W_f(q)$ to be able to converge, the noise reduction stage adaptation should then be slower than the feedback canceller adaptation.

A. Concept

Noise reduction stage

First, the noise is removed from the microphone signals by means of a GSC. Because of the presence of acoustic feedback in the microphone signals, a classical GSC cannot be used: the acoustic feedback would then result in desired signal distortion or cancellation (cf. Section III) [16]. To avoid such speech distortion, the PEM-GSC [16] is used. The optimization criterion of the PEM-GSC $\{\bar{A}(q),\, \mathbf{W}_a(q)\}$ with $\bar{a}_0 = 1$ corresponds to the optimization criterion of a PEM-GFIC with $W_f(q) = 0$:

$$\left\{\bar{A}(q),\, \mathbf{W}_a(q)\right\} = \arg\min \varepsilon\{(\bar{e}^p[k])^2\}, \tag{70}$$

with

$$\bar{e}^p[k] = \bar{A}(q)\left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\left(\mathbf{x}^s[k] + \mathbf{x}^n[k] + \mathbf{F}(q)u[k]\right), \tag{71}$$

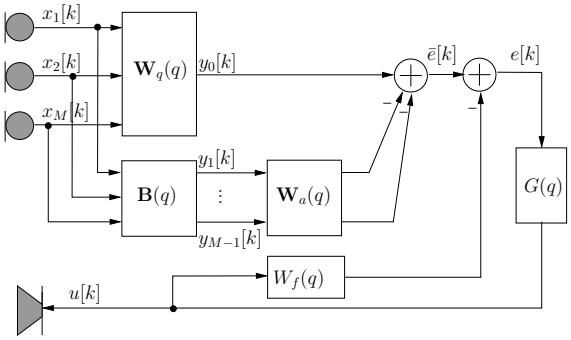


Fig. 6. Noise reduction followed by feedback cancellation ('GSC first').

and

$$u[k] = \frac{G(q)\left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\left(\mathbf{x}^s[k] + \mathbf{x}^n[k]\right)}{1 - G(q)\left[\left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\mathbf{F}(q) - W_f(q)\right]}. \tag{72}$$

The following lemma applies to the solution of (70) (cf. Section IV).

Lemma 4: Assume that $y_0^s[k] = \mathbf{W}_q^T(q)\mathbf{x}^s[k]$ can be modelled as

$$y_0^s[k] = \frac{1}{P(q)}w[k] \tag{73}$$

with $P(q)$ a finite order polynomial with length L_P and assume that $\mathbf{W}_a(q)$ has sufficient degrees of freedom to suppress both the noise and the feedback. Then, assuming a perfect blocking matrix, the infinite order case for $\mathbf{W}_a(q)$ and $L_P \leq L_{\bar{A}}$ with $L_{\bar{A}}$ the filter length of $\bar{A}(q)$ (with $a_0 = 1$),

$$\min_{\bar{A}(q),\mathbf{W}_a(q)} \varepsilon\{(\bar{e}^p[k])^2\} \tag{74}$$

results in the unique desired solution

$$\bar{A}(q) = P(q), \tag{75}$$

$$\left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\mathbf{x}^n[k] = 0 \text{ (i.e.,perfect noise reduction)}, \tag{76}$$

$$\left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\mathbf{F}(q) = 0 \text{ (i.e., perfect feedback cancellation)}, \tag{77}$$

if $d_G \geq L_{\bar{A}}$.

Proof: The proof is similar to the proof of Lemma 3 and is therefore be omitted. ■

If $\mathbf{W}_a(q)$ has insufficient degrees of freedom to suppress both the feedback and the noise, $\mathbf{W}_a(q)$ trades off between noise reduction and feedback cancellation. The larger $\bar{A}(q)y_0^n[k]$ compared to $\bar{A}(q)y_0^u[k]$, the more focus on noise reduction at the expense of less feedback cancellation and vice-versa.

Feedback cancellation stage

In a second stage, the (residual) acoustic feedback in the GSC-output

$$\begin{aligned}\bar{e}[k] &= y_0[k] - \mathbf{W}_a^T(q)\mathbf{y}[k] \\ &= \left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\left(\mathbf{x}^s[k] + \mathbf{x}^n[k] + \mathbf{F}(q)u[k]\right).\end{aligned} \tag{78}$$

is further removed by a single PEM-AFC $W_f(q)$ [15]. To cancel the residual feedback, the filter $W_f(q)$ should be given as

$$W_f(q) = \left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\mathbf{F}(q). \tag{79}$$

Hence, $W_f(q)$ should not only model the feedback paths $\mathbf{F}(q)$, but also the time-varying noise reduction filter $\mathbf{W}_a(q)$, which may complicate the feedback cancellation. The PEM-AFC $W_f(q)$ minimizes the prediction-error

$$e^p[k] = A(q)\left[\bar{e}^s[k] + \bar{e}^n[k] + \bar{e}^u[k] - W_f(q)u[k]\right], \tag{80}$$

with

$$\bar{e}^s[k] + \bar{e}^n[k] = \left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\left(\mathbf{x}^s[k] + \mathbf{x}^n[k]\right), \tag{81}$$

$$\bar{e}^u[k] = \left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\mathbf{F}(q)u[k], \tag{82}$$

$$u[k] = \frac{G(q)\left(\mathbf{W}_q^T(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\right)\left(\mathbf{x}^s[k] + \mathbf{x}^n[k]\right)}{1 - G(q)\left(\mathbf{W}_q^T(q)\mathbf{F}(q) - \mathbf{W}_a^T(q)\mathbf{B}^T(q)\mathbf{F}(q) - W_f(q)\right)}. \tag{83}$$

If $G(q)$ contains a delay $d_G \geq L_A$, $A(q)$ whitens he direct speech and noise component $\bar{e}^s[k] + \bar{e}^n[k]$ and hence, removes the correlation between $\bar{e}^s[k] + \bar{e}^n[k]$ and $u[k]$ [13]. If the whitening is perfect, the feedback canceller $W_f(q)$ will be unbiased.

B. Interaction between the noise reduction stage and the feedback cancellation stage

The performance of the noise reduction stage influences the performance of the feedback canceller $W_f(q)$ and vice-versa.

The performance of the feedback canceller $W_f(q)$ is determined by the energy ratio of the feedback component $\bar{e}^u[k]$ in the GSC output with respect to the direct speech and noise component $\bar{e}^s[k] + \bar{e}^n[k]$ and hence, depends on the GSC-stage $\mathbf{W}_a(q)$: the larger the energy ratio, the smaller misadjustment of the feedback path estimate $W_f(q)$. In contrast to 'AFC first', noise reduction by the GSC not only reduces the energy of $u[k]$ but also the energy of the disturbance $\bar{e}^s[k] + \bar{e}^n[k]$. At low SNRs, the feedback to near-end ratio at the input of the feedback canceller $W_f(q)$ will thus be larger than at the input of the feedback cancellers of 'AFC first', resulting in a smaller misadjustment.

The more the noise reduction stage focuses on feedback cancellation instead of noise reduction, the lower the energy of the feedback component $\bar{e}^u[k]$ and hence, the lower the feedback to near-end ratio at the input of $W_f(q)$. As a result, the misadjustment of the feedback canceller $W_f(q)$ increases with increasing gain and input SNR. At high gains and high SNRs where both $W_f(q)$ and $\mathbf{W}_a(q)$ focus on feedback cancellation, part of the feedback cancellation by the first stage may therefore be lost in the second stage. At low SNRs where $\mathbf{W}_a(q)$ concentrates on noise reduction, the noise reduction by the GSC may decrease or increase the feedback to near-end ratio depending on the noise and feedback

scenario and hence, negatively or positively influence the feedback canceller $W_f(q)$. Finally, because of its time-variance, the noise reduction stage may complicate the feedback cancellation by $W_f(q)$ (e.g., in time-varying scenarios).

The performance of the feedback canceller $W_f(q)$ similarly influences the noise reduction stage. The amount of feedback $y_0^u[k] = \mathbf{W}_q^T(q)\mathbf{F}(q)u[k]$ in the speech reference depends on the feedback canceller $W_f(q)$: the higher the feedback suppression by $W_f(q)$, the smaller $y_0^u[k]$ and hence, the more $\mathbf{W}_a(q)$ concentrates on noise reduction. As a result, 'GSC first' achieves more noise reduction than the PEM-GSC alone [16], especially at high gains G.

C. Simulations

a) Interactions between the noise reduction and feedback cancellation stage: Figure 7 shows the SNR-improvement, $\Delta\text{SNR}_{\text{intellig}}$, and the output $\text{SFR}_{\text{intellig}}$ of 'GSC first' as a function of the input SNR. For comparison, the $\text{SFR}_{\text{intellig}}$ at the output of the PEM-GSC stage is also depicted (dotted lines). Figure 7(a) shows that the amount of noise reduction decreases with increasing SNR and increasing gain: the larger the input SNR and the gain, the more the GSC-stage concentrates on feedback cancellation (see also the good feedback cancellation performance of the first stage at high input SNRs), at the expense of a reduced noise suppression. As a result, less noise reduction is achieved at high input SNRs and high gains than with 'AFC first' and with PEM-GFIC. For low input SNRs and low gains, the first stage concentrates on noise reduction, resulting in a $\Delta\text{SNR}_{\text{intellig}}$ similar to those obtained with 'AFC first' and PEM-GFIC.

The larger the feedback cancellation achieved by the GSC-stage, the lower the feedback to near-end signal ratio at the input of $W_f(q)$ and hence, the worse the feedback cancellation by the feedback canceller $W_f(q)$. At high input SNRs, $W_f(q)$ even degrades the feedback cancellation performance. This illustrates that the filters $\mathbf{W}_a(q)$ and $W_f(q)$ are not optimally exploited by 'GSC first'. In contrast to 'AFC first', the noise reduction by the GSC does not negatively affect the performance of the feedback canceller so that better feedback cancellation is achieved at low input SNRs. For this noise and feedback scenario, the first stage of 'GSC first' increases the amount of feedback $\bar{e}^u[k]$ in the GSC-output compared to the speech reference $y_0^u[k]$[3] and hence, improves the feedback to near-end ratio at the input of the feedback canceller $W_f(q)$.

[3]Note that the $\text{SFR}_{\text{intellig}}$ at the output of the first GSC stage is lower than the $\text{SFR}_{\text{intellig}}$ of the speech reference $y_0[k] = \mathbf{W}_q^T(q)\mathbf{x}[k]$, represented in Figure 7(b) by the stars and the diamonds.

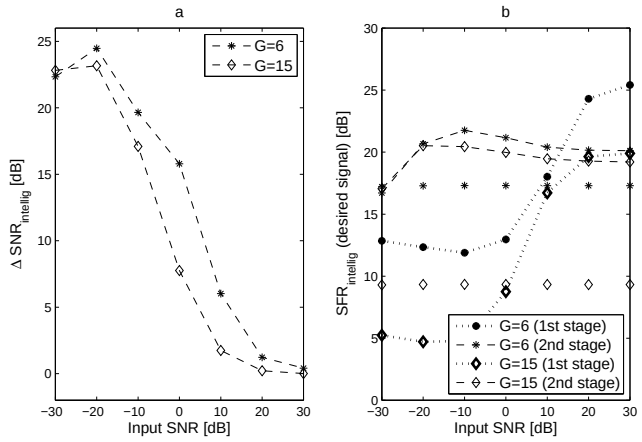


Fig. 7. Performance of 'GSC first': a) $\Delta\text{SNR}_{\text{intellig}}$ (dB); (b) Output $\text{SFR}_{\text{intellig}}$ of speech (dB). The stars and diamonds denote the intelligibility-weighted signal-to-feedback ratio of the speech reference $y_0[k]$.

b) Comparison with PEM-GFIC: At high input SNRs (i.e., above 0 dB) and high gains, PEM-GFIC outperforms the 'GSC first' approach in terms of feedback and noise suppression. For SNRs below -10 dB, 'GSC first' suppresses more feedback than PEM-GFIC at the expense of less noise reduction.

VI. Simulations

The algorithms for combined feedback and noise suppression have been evaluated for acoustic feedback paths and microphone signals measured in a commercial two-microphone BTE hearing aid of GN Resound. In Section III, Section IV and Section V, anechoic microphone recordings of speech weighted noise and a perfect blocking matrix were used to illustrate the optimal performance of the algorithms and study interaction effects. In this section, the performance will be compared for more realistic, reverberant microphone recordings.

A. Set-up and implementation

The acoustic feedback paths from the loudspeaker to the BTE microphones were measured in an office room. The reverberation time of the office room equals 500 ms. During the measurements, the hearing aid was attached to the right ear of a Cortex MK II artificial head. An earmold with a vent size of 2 mm was used. The sampling frequency f_s equals 16 kHz. The forward path $G(q) = Gq^{-d_G}$ with G a frequency-independent gain and $d_G = 2$ ms.

The desired signal is a real speech signal consisting of HINT-sentences produced by a male speaker [22]. The noise signal is a speech weighted noise signal generated by passing white Gaussian noise

through a 20th order all-pole filter. The speech and the noise source were positioned at a distance of 1 m from the artificial head: the speech source was positioned at 0°, the noise source at 90°. The performance measures were averaged over 20 trials with different speech and noise signals.

In the simulations, partitioned-block frequency-domain (PBFD) implementations are used for the adaptive feedback and noise suppression filters of the different algorithms [15]. The FFT-length M was set to 64, the partition length to 32. The block length L of the PBFD implementation was set to 16, resulting in an algorithmic delay of 2 ms. The step sizes $\mu_m[k]$, $m = 0, \ldots, M-1$ of the PBFD adaptive filters are normalized according to the sum of the input and error power in each frequency-bin and each channel, i.e.,

$$\mu_m[k] = \frac{\bar{\mu}_m}{P_{u,m}[k] + P_{e,m}[k]}, \tag{84}$$

where $P_{u,m}[k]$ and $P_{e,m}[k]$ denote the input and error power in the mth frequency bin, respectively [15], [23]. This normalization reduces the step size when the desired signal is strong and hence, mitigates the effect of a strong desired signal segment on the excess mean squared error. In the simulations, the step sizes $\bar{\mu}_m$ are set to 0.01. The feedback and noise suppression filters are continuously adapted so that changes in the feedback and noise scenario are continuously tracked. To avoid speech cancellation, the quadratic inequality constraint $\|\mathbf{w}_a[k]\|_2^2 \leq 4$ is imposed on the ANC filters $\mathbf{W}_a(q)$ by means of the scaled projection algorithm [21].

Similarly as in the single-channel PEM-AFC [14], [15], the AR-model $A(q)$ (or $\bar{A}(q)$) and the noise and feedback suppression filters $\mathbf{W}_a(q)$ and $\mathbf{W}_f(q)$ are computed in two steps. First, the AR-model $A(q)$ (or $\bar{A}(q)$) is estimated on the error signal $e[k]$ (or $\bar{e}[k]$) of the noise and/or feedback suppression filters: the AR-model is estimated on subsequent frames of $N = 160$ samples (i.e., 10 ms) through the Levinson-Durbin algorithm [24]. In a second step, the AR-model is used to pre-filter the input and reference signal of the adaptive noise and feedback suppression filters, for the same time frame. The resulting residuals $e^p[k]$ (or $\bar{e}^p[k]$) are then used to update the adaptive filters $\mathbf{W}_f(q)$ and/or $\mathbf{W}_a(q)$.

In the experiments, the microphones were calibrated and time-aligned based on anechoic recordings of a speech weighted noise signal positioned at 0° through 20-taps calibration filters. A delay-and-sum beamformer is used as a fixed beamformer, i.e.,

$$\mathbf{W}_q(q) = \frac{1}{2} q^{-d_1} \begin{bmatrix} 1 & 1 \end{bmatrix}^T.$$

The blocking matrix $\mathbf{B}(q)$ pairwise subtracts the time-aligned calibrated microphone signals. The delay d_1 is set to 16 to allow for a-causal filtering by the ANC $\mathbf{W}_a(q)$.

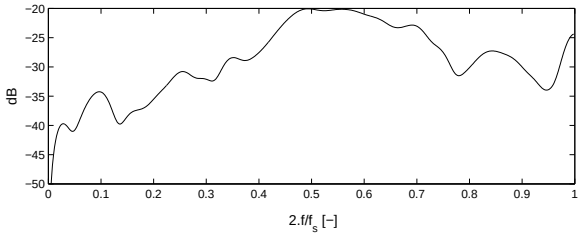


Fig. 8. Acoustic feedback path $\mathbf{W}_q^T(q)\mathbf{F}(q)$.

Figure 8 depicts the feedback path $\mathbf{W}_q^T(q)\mathbf{F}(q)$ from the loudspeaker $u[k]$ to the input of the gain module $G(q)$, showing that instability occurs for $G \geq 10$ if no noise and feedback suppression is performed, i.e., $\mathbf{W}_a(q)$ and $\mathbf{W}_f(q)$ equal zero.

B. Simulation results

Figure 9 compares the performance of the PEM-GFIC with GEIC, 'AFC first' and 'GSC first' for reverberant microphone recordings of a real speech signal and a speech weighted noise signal. In contrast to the simulations in Section III and Section V, some speech leakage is now present in the noise references because of reverberation. The forward path gain $G = 15$ (i.e., above the stability margin).

For all algorithms, the amount of noise reduction decreases with increasing input SNR, while the amount of feedback cancellation generally increases with increasing input SNR[4].

In terms of feedback cancellation performance, 'AFC first' performs worse than the PEM-GFIC and 'GSC first', except at very large input SNRs: the noise reduction by the GSC stage negatively affects the performance of the feedback cancellation stage. 'GSC first' generally achieves better feedback cancellation than PEM-GFIC, except for very large input SNR. However, in terms of noise reduction, 'GSC first' always performs worse than the PEM-GFIC and the 'AFC first' scheme. The GEIC achieves good noise reduction, but worse feedback cancellation. Because of signal correlation, the GEIC tends to remove the speech and the noise instead of the acoustic feedback.

To illustrate the effective benefit of the different algorithms for a hearing aid user, the maximum stable gain (MSG) of the algorithms was determined by increasing the gain G until audible oscillations occurred. The initial gain was set at 1 dB, i.e., far below the stability margin. After 5 s of convergence, the gain was gradually and slowly increased in steps of 0.1 dB each 1 s, while the adaptive algorithm was

[4]At high input SNRs, the feedback cancellation performance slightly decreases because it is more difficult to whiten the real speech signal than the stationary noise signal with an AR-model.

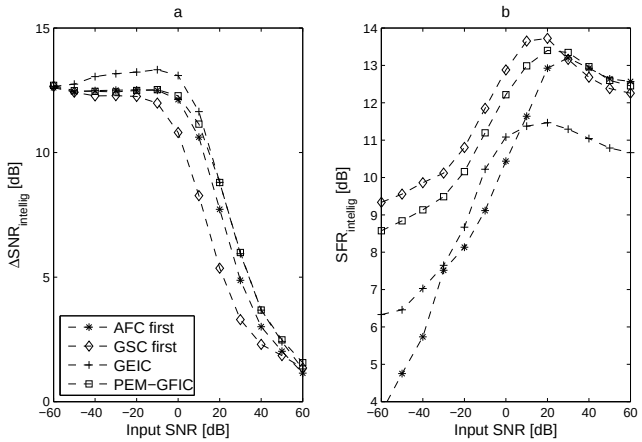


Fig. 9. $\Delta\text{SNR}_{\text{intellig}}$ (a) and output $\text{SFR}_{\text{intellig}}$ (b) of the different combined feedback and noise suppression algorithms for reverberant recordings, real speech as a desired signal and $G = 15$.

input-SNR	Algorithm	MSG			$G = 20$ dB		$G = 30$ dB	
		MSG [dB]	$\Delta\text{SNR}_{\text{intellig}}$ [dB]	$\text{SFR}_{\text{intellig}}$ [dB]	$\Delta\text{SNR}_{\text{intellig}}$ [dB]	$\text{SFR}_{\text{intellig}}$ [dB]	$\Delta\text{SNR}_{\text{intellig}}$ [dB]	$\text{SFR}_{\text{intellig}}$ [dB]
∞	'AFC first'	41	-	9.7	-	12.7	-	12.3
	'GSC first'	40	-	8.1	-	12.2	-	11.4
	GEIC	33	-	9.6	-	11.3	-	9.0
	PEM-GFIC	43	-	8.5	-	12.2	-	11.6
30 dB	'AFC first'	42	4.4	11.3	4.7	14.2	4.5	12.8
	'GSC first'	43	-0.2	9.7	4.2	13.3	1.1	12.6
	GEIC	36	4.6	9.7	5.4	11.6	5.8	9.9
	PEM-GFIC	45	1.5	9.8	5.7	13.4	5.8	12.3
0 dB	'AFC first'	39	12.7	8.5	12.8	11.4	12.6	9.0
	'GSC first'	48	-0.2	8.0	12.3	13.2	6.0	13.0
	GEIC	40	12.9	9.1	13.0	12.1	13.1	11.0
	PEM-GFIC	47	8.2	11.0	12.9	12.6	12.9	12.3
-10 dB	'AFC first'	27	13.2	8.9	13.3	9.7	13.1	8.0
	'GSC first'	48	-0.2	9.6	13.0	12.2	8.7	12.1
	GEIC	38	13.4	10.6	13.3	10.3	13.3	10.7
	PEM-GFIC	47	9.9	9.6	13.3	11.3	13.2	10.7

TABLE I

MEASUREMENT OF MAXIMUM STABLE GAIN FOR DIFFERENT INPUT SNRS.

running. Table I compares the MSG of the algorithms for different input SNRs at the first microphone. The improvement $\Delta\text{SNR}_{\text{intellig}}$ and $\text{SFR}_{\text{intellig}}$ measured at this MSG are also shown. In addition, Table I depicts the noise reduction performance $\Delta\text{SNR}_{\text{intellig}}$ and feedback cancellation performance $\text{SFR}_{\text{intellig}}$ at a gain of 20 dB and 30 dB. For all algorithms, the amount of noise reduction decreases with increasing input SNR. At very high input SNR, the 'AFC first' scheme achieves good feedback and noise suppression performance. At intermediate and low SNRs, the feedback cancellation performance is however negatively affected by the noise reduction stage: compared to the 'GSC first' scheme and the PEM-GFIC, a smaller MSG is obtained. The 'GSC first' scheme generally achieves a large MSG. However, no noise is reduced at the MSG, even not for low input SNR. The integrated PEM-GFIC obtains a comparable MSG while still providing noise reduction. At lower gains $G = 20$ dB and $G = 30$ dB, 'GSC first' provides some noise reduction, but less than the PEM-GFIC. For high input SNRs, PEM-GFIC outperforms the 'GSC first' in terms of both feedback and noise suppression performance. The noise reduction by the GEIC at $G = 20$ dB and $G = 30$ dB is comparable to PEM-GFIC, but the amount of feedback cancellation is worse. This is also manifested by the lower MSG that can be obtained with the GEIC.

The simulations above compared the steady-state performance of the different algorithms. A comparison of the convergence and tracking performance is the scope of further study. Because of the joint adaptation of the noise and feedback suppression filters, the PEM-GFIC is expected to have the fastest convergence and the best tracking performance.

VII. Conclusions

In this paper, solutions for the problem of combined feedback and noise suppression in hearing aids have been developed and evaluated. Cascades of GSC-based noise reduction and PEM-based feedback cancellation as well as a truly integrated solution for feedback cancellation and noise reduction, called PEM-GFIC, have been discussed.

In the cascaded solutions 'AFC first' and 'GSC first', the feedback cancellation stage influences the noise reduction stage and vice-versa. As a result, the noise reduction and feedback cancellation filters are not always optimally exploited. For very high input SNRs where feedback cancellation is the major goal, the 'AFC first' scheme generally may provide the best feedback and noise suppression performance because of its larger degrees of freedom. The computational cost is however high, since each microphone needs an adaptive feedback canceller. In addition, noise reduction by the GSC-stage negatively affects the performance of the adaptive feedback cancellers. As a result, 'AFC first' suffers from degraded feedback cancellation performance, especially if a high amount of noise suppression is achieved.

'GSC first' focuses more on feedback cancellation than the PEM-GFIC. At low input SNRs, this generally results in better feedback cancellation at the expense of worse noise reduction. At high gains and/or large input SNRs, the noise reduction filter negatively influences the performance of the feedback cancellation filter, resulting in a worse feedback and noise suppression compared to the PEM-GFIC. Compared to 'AFC first', a much lower computational cost is achieved.

The integrated solution PEM-GFIC achieves optimum synergies between noise reduction and feedback cancellation at the lowest computational cost. Compared to the 'GSC first' scheme, more noise is reduced. At high input SNRs, also a better feedback cancellation is achieved. At low input SNRs, the integrated solution outperforms 'AFC first'. Compared to the GEIC proposed for echo cancellation, the PEM-GFIC achieves better feedback cancellation performance thanks to the adaptive pre-whitening operations.

REFERENCES

[1] L. J. Griffiths and C. W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Antennas Propag.*, vol. 30, no. 1, pp. 27–34, Jan. 1982.

[2] K. M. Buckley, "Broad-band beamforming and the Generalized Sidelobe Canceller," *IEEE Trans. Acoust., Speech, and Signal Processing*, vol. 34, no. 5, pp. 1322–1323, Oct. 1986.

[3] J. E. Greenberg and P. M. Zurek, "Evaluation of an adaptive beamforming method for hearing aids," *J. Acoustic. Soc. of Am.*, vol. 91, no. 3, pp. 1662–1676, Mar. 1992.

[4] F. Luo, J. Yang, C. Pavlovic, and A. Nehorai, "Adaptive null-forming scheme in digital hearing aids," *IEEE Trans. Signal Processing*, vol. 50, no. 7, pp. 1583–1590, July 2002.

[5] A. Spriet, L. Van Deun, K. Eftaxiadis, J. Laneau, M. Moonen, B. van Dijk, A. van Wieringen, and J. Wouters, "Speech understanding in background noise with the two-microphone adaptive beamformer beam^{TM} in the nucleus freedom^{TM} cochlear implant system," *Ear and Hearing*, 2006, To appear.

[6] J. Vanden Berghe and J. Wouters, "An adaptive noise canceller for hearing aids using two nearby microphones," *J. Acoustic. Soc. Am.*, vol. 103, pp. 3621–3626, June 1998.

[7] H.-F. Chi, S. X. Goa, S. D. Soli, and A. Alwan, "Band-limited feedback cancellation with a modified filtered-X LMS algorithm for hearing aids," *Speech Communication*, vol. 39, no. 1-2, pp. 147–161, Jan. 2003.

[8] J. E. Greenberg, P. M. Zurek, and M. Brantley, "Evaluation of feedback-reduction algorithms for hearing aids," *J. Acoustic. Soc. Am.*, vol. 108, no. 5, pp. 2366–2376, Nov. 2000.

[9] J. Hellgren, "Analysis of feedback cancellation in hearing aids with filtered-X LMS and the direct method of closed loop identification," *IEEE Trans. Speech, Audio Processing*, vol. 10, no. 2, pp. 119–131, Feb. 2002.

[10] J. M. Kates, *Adaptive feedback cancellation in hearing aids*, chapter 2 in "Adaptive signal processing: Applications to real-world problems" (Benesty, J. and Huang, Y.), pp. 23–55, Springer-Verlag, 2003.

[11] J. A. Maxwell and P. M. Zurek, "Reducing acoustic feedback in hearing aids," *IEEE Trans. Speech, Audio Processing*, vol. 3, no. 4, pp. 304–313, July 1995.

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

[13] A. Spriet, I. Proudler, M. Moonen, and J. Wouters, "Adaptive feedback cancellation in hearing aids with linear prediction of the desired signal," *IEEE Transactions on Signal Processing*, vol. 53, no. 10 (part 1), pp. 3749–3763, Oct. 2005.

[14] G. Rombouts, T. van Waterschoot, K. Struyve, and M. Moonen, "Acoustic feedback suppression for long acoustic paths using a nonstationary source model," *IEEE Trans. Signal Processing*, vol. 54, no. 9, pp. 3426–3434, Sep. 2006.

[15] A. Spriet, G. Rombouts, M. Moonen, and J. Wouters, "Adaptive feedback cancellation in hearing aids," *Journal of the Franklin Institute*, vol. 343, no. 6, pp. 545–573, Sept. 2006.

[16] G. Rombouts, A. Spriet, and M. Moonen, "Generalized sidelobe canceller based acoustic feedback cancellation," in *Proc. European Signal Processing Conf. (EUSIPCO)*, Firenze, Italy, Sept. 2006.

[17] W. Herbordt, W. Kellermann, and S. Nakamura, "Joint optimization of LCMV beamforming and acoustic echo cancellation," in *Proc. European Signal Processing Conf. (EUSIPCO)*, Wien, Austria, Sept. 2004, pp. 2003–2006.

[18] L. Ljung and T. Söderström, *Theory and Practice of Recursive Identification.*, MIT Press, Cambridge, MA, 1983.

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

[20] Acoustical Society of America, "ANSI S3.5-1997 American National Standard Methods for calculation of the speech intelligibility index," June 1997.

[21] H. Cox, R. M. Zeskind, and M. M. Owen, "Robust Adaptive Beamforming," *IEEE Trans. Acoust., Speech, and Signal Processing*, vol. 35, no. 10, pp. 1365–1376, 1987.

[22] M. Nilsson, S. D. Soli, and A. Sullivan, "Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise," *J. Acoustic. Soc. Am.*, vol. 95, no. 2, pp. 1085–1099, Feb. 1994.

[23] J. E. Greenberg, "Modified LMS algorithms for speech processing with an adaptive noise canceller," *IEEE Trans. Speech, Audio Processing*, vol. 6, no. 4, pp. 338–350, July 1998.

[24] J. R. Deller, J. G. Proakis, and J. H. L. Hansen, *Discrete-Time Processing of Speech Signals*, Macmillan Publishing Company, Englewood Cliffs, New Jersey, 1993.



Ann Spriet was born in Izegem, Belgium, in 1976. She received the Electrical Engineering degree and the Ph.D. degree in applied sciences from the Katholieke Universiteit Leuven, Leuven, Belgium, in 1999 and 2004, respectively. Currently, she is a Postdoctoral Fellow of the Fund for Scientific Research-Flanders, affiliated with the Electrical Engineering Department (ESAT)and the Laboratory Experimental ORL of the Katholieke Universiteit Leuven. Her research interests are in digital signal processing for speech enhancement in hearing aids and cochlear implants. She received the 2001 Lapperre Award, a Best Student Paper Award at the International Workshop on Acousctic Echo and Noise Control in 2003 and a Best Student Paper Award at the International Conference on Acoustics, Speech, and Signal Processing in 2005.



Geert Rombouts was born in Turnhout, Belgium in 1973. He received a master degree in Electrical Engineering in 1997 from the Katholieke Universiteit Leuven, Belgium and a PhD degree in Engineering (Toegepaste Wetenschappen) from the same university in 2003. Currently, he is a post-doctoral researcher at the ESAT-SCD laboratories of the Katholieke Universiteit Leuven.



Marc Moonen (M'94) received the Electrical Engineering degree and the Ph.D. degree in applied sciences from the Katholieke Universiteit Leuven, Leuven, Belgium, in 1986 and 1990, respectively. Since 2004 he has been a Full Professor with the Electrical Engineering Department, Katholieke Universiteit Leuven, where he is currently heading a research team of 16 Ph.D. candidates and postdocs, working in the area of numerical algorithms and signal processing for digital communications, wireless communications, DSL, and audio signal processing. Dr. Moonen received the 1994 KU Leuven Research Council Award, the 1997 Alcatel Bell (Belgium) Award (with P. Vandaele), the 2004 Alcatel Bell (Belgium) Award (with R. Centrillon) and was a 1997 "Laureate of the Belgium Royal Academy of Science." He received a journal best paper award from the IEEE Transactions on Signal Processing (with G. Leus) and from Elsevier Signal Processing (with S. Doclo). He was Chairman of the IEEE Benelux Signal Processing Chapter (1998-2002), and is currently a EURASIP AdCom Member (European Association for Signal, Speech, and Signal Processing, 2000) and a member of the IEEE Signal Processing Society Technical Committe on Signal Processing for Communications. He served as Editor-in Chief for the *EURASIP Journal on Applied Signal Processing* (2003-2005), and was a member of the editorial board of *IEEE Transactions on Circuits and Systems II (2002-2003)* and the *IEEE Signal Processing Magazine* (2003-2005). He is currently a member of the editorial board of *Integration, the VLSI Journal, the EURASIP Journal on Advances in Signal Processing, the EURASIP Journal on Wireless Communications and Networking,* and *Signal Processing.*



Jan Wouters was born in Leuven, Belgium, in 1960. He received the physics degree and the Ph.D. degree in sciences / physics from the Katholieke Universiteit Leuven, Leuven, Belgium, in 1982 and 1989, respectively. From 1989 till 1992 he has been a Research Fellow with the Belgian NFWO (National Fund for Scientific Research) at the Institute of Nuclear Physics (UCL Louvain-la-Neuve and KU Leuven) and at NASA Goddard Space Flight Center (USA). From 1993 he is a Professor at the Neurosciences Department of the Katholieke Universiteit Leuven. His research activities are about audiology and the auditory system, signal processing for cochlear implants and hearing aids. Dr. Wouters received in 1989 an Award of the Flemish Ministry, in 1992 a Fullbright Award and a NATO Research Fellowship, the 1996 Flemish VVL Speech therapy - Audiology Award. He is member of the International Collegium for Rehabilitative Audiology, a Board Member of the NAG (Dutch Acoustical Society) and responsible for the Lab. Exp. ORL, author of 110 articles in international peer-review journals, reviewer for several international journals.

LIST OF FIGURES

1 Signal model and notation. 6

2 PEM-based generalized feedback and interference canceller (PEM-GFIC). 7

3 Performance of the PEM-GFIC , the PEM-GSC and the GEIC: a) $\Delta\text{SNR}_{\text{intellig}}$ (dB); (b) output $\text{SFR}_{\text{intellig}}$ of speech (dB) ; The stars ($G = 6$) and diamonds ($G = 15$) denote the intelligibility-weighted speech-to-feedback ratio of the speech reference $y_0[k]$ for the PEM-GFIC. 14

4 Feedback cancellation followed by noise reduction ('AFC first'). 15

5 Performance of 'AFC first': a) $\Delta\text{SNR}_{\text{intellig}}$ (dB); (b) Output $\text{SFR}_{\text{intellig}}$ of speech (dB). The stars and diamonds denote the $\text{SFR}_{\text{intellig}}$ of $\mathbf{W}_q^T(q)\mathbf{x}[k]$ (i.e., the speech reference without feedback compensation). 20

6 Noise reduction followed by feedback cancellation ('GSC first'). 21

7 Performance of 'GSC first': a) $\Delta\text{SNR}_{\text{intellig}}$ (dB); (b) Output $\text{SFR}_{\text{intellig}}$ of speech (dB). The stars and diamonds denote the intelligibility-weighted signal-to-feedback ratio of the speech reference $y_0[k]$. 25

8 Acoustic feedback path $\mathbf{W}_q^T(q)\mathbf{F}(q)$. 27

9 $\Delta\text{SNR}_{\text{intellig}}$ (a) and output $\text{SFR}_{\text{intellig}}$ (b) of the different combined feedback and noise suppression algorithms for reverberant recordings, real speech as a desired signal and $G = 15$. 28

LIST OF TABLES

I Measurement of maximum stable gain for different input SNRs. 28