Katholieke Universiteit Leuven

Departement Elektrotechniek ESAT-SISTA/TR 2003-115

Spatially pre-processed speech distortion weighted multi-channel Wiener filtering for noise reduction in hearing aids [1]

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

Published in Proc. of the 2003 International Workshop on Acoustic Echo and Noise Control (IWAENC 2003), Kyoto, Japan, September 8-11 2003, pp. 147-150

[1]This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/sista/spriet/reports/03-115.pdf*

[2]K.U.Leuven, Dept. of Electrical Engineering (ESAT), SISTA, Kasteelpark Arenberg 10, 3001 Leuven-Heverlee, Belgium, Tel. 32/16/32 18 99, Fax 32/16/32 19 70, WWW: *http://www.esat.kuleuven.ac.be/sista*. E-mail: *ann.spriet@esat.kuleuven.ac.be*. K.U.Leuven, Lab. Exp. ORL/ ENT-Dept., Kapucijnenvoer 33, 3000 Leuven, Belgium, Tel. 32/16/33 24 15, Fax 32/16/33 23 35, WWW: *http://www.kuleuven.ac.be/exporl/Lab/Default.htm*. Ann Spriet is a Research Assistant supported by the Fonds voor Wetenschappelijk Onderzoek (FWO) - Vlaanderen. This research work was carried out at the ESAT laboratory and Lab. Exp. ORL of the Katholieke Universiteit Leuven, in the the frame of IUAP P5/22 ('Dynamical Systems and Control: Computation, Identification and Modelling'), the Concerted Research Action GOA-MEFISTO-666 (Mathematical Engineering for Information and Communication Systems Technology)of the Flemish Government, Research Project FWO nr.G.0233.01 ('Signal processing and automatic patient fitting for advanced auditory prostheses'), IWT project 020540 ('Innovative Speech Processing Algorithms for Improved Performance of Cochlear Implants') and was partially sponsored by Cochlear. The scientific responsibility is assumed by its authors.

[3]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, WWW: *http://www.esat.kuleuven.ac.be/sista*. E-mail: *marc.moonen@esat.kuleuven.ac.be*. Marc Moonen is a professor at the Katholieke Universiteit Leuven.

[4]K.U.Leuven, Lab. Exp. ORL, Dept. Neurowetenschappen, Kapucijnenvoer 33, 3000 Leuven, Belgium, Tel. 32/16/33 23 42, Fax 32/16/33 23 35, WWW: *http://www.kuleuven.ac.be/exporl/Lab/Default.htm* E-mail: *jan.wouters@uz.kuleuven.ac.be*. Jan Wouters is a professor at the Katholieke Universiteit Leuven.

SPATIALLY PRE-PROCESSED SPEECH DISTORTION WEIGHTED MULTI-CHANNEL WIENER FILTERING FOR NOISE REDUCTION IN HEARING AIDS

Ann Spriet[1,2*], *Marc Moonen*[1], *Jan Wouters*[2]

[1]K.U. Leuven, ESAT/SCD-SISTA
Kasteelpark Arenberg 10, 3001 Leuven, Belgium
{spriet,moonen}@esat.kuleuven.ac.be

[2]K.U. Leuven - Lab. Exp. ORL
Kapucijnenvoer 33, 3000 Leuven, Belgium
jan.wouters@uz.kuleuven.ac.be

ABSTRACT

In this paper we establish a generalized noise reduction scheme, called the Spatially Pre-processed Speech Distortion Weighted Multi-channel Wiener filter (SP-SDW-MWF), that encompasses the Generalized Sidelobe Canceller (GSC) and a recently developed Multi-channel Wiener Filtering (MWF) technique as extreme cases and allows for in-between solutions. Compared to the widely studied GSC with Quadratic Inequality Constraint (QIC-GSC), the SP-SDW-MWF achieves a better noise reduction performance, for a given maximum speech distortion level.

1. INTRODUCTION

Noise reduction algorithms are crucial to improve the intelligibility for hearing impaired people in background noise. Multi-microphone systems exploit spatial in addition to temporal and spectral information of the desired and noise signal and are thus preferred to single microphone procedures.

In [1, 2, 3], an MWF technique for noise reduction has been proposed that provides a Minimum Mean Square Error (MMSE) estimate of the desired signal portion in one of the microphone signals. In contrast to the GSC [4], it does not rely on a priori assumptions about the signal model so that it is less sensitive to signal model errors such as microphone mismatch [5].

In this paper, we establish a generalized scheme, called SP-SDW-MWF that encompasses the GSC and MWF as extreme cases and allows for in-between solutions such as the Speech Distortion Regularized GSC (SDR-GSC). The SDR-GSC adds robustness to the GSC by taking speech distortion explicitly into account in the design criterion of the adaptive stage. Compared to the widely studied QIC-GSC [6, 7], the SDR-GSC achieves better noise reduction for small model errors, while guaranteeing robustness against large model errors. In addition, the extra filtering of the speech reference in the SP-SDW-MWF further improves the performance. We show that, in the absence of model errors and for infinite filters, the SP-SDW-MWF corresponds to an SDR-GSC cascaded with an SDW Single-channel Wiener Filter (SDW-SWF). In contrast to the SDR-GSC and the QIC-GSC, its performance does not degrade due to microphone mismatch. The theoretical results are illustrated through experiments with a Behind-The-Ear (BTE) hearing aid.

*Ann Spriet is a Research Assistant with F.W.O.-Vlaanderen. This research was carried out at the ESAT laboratory and Lab. Exp. ORL of K.U. Leuven, in the frame of IUAP P5/22 (2002-2007), the Concerted Research Action GOA-MEFISTO-666 of the Flemish Government, FWO Project nr. G.0233.01, IWT project 020540 and was partially sponsored by Cochlear. The scientific responsibility is assumed by its authors.

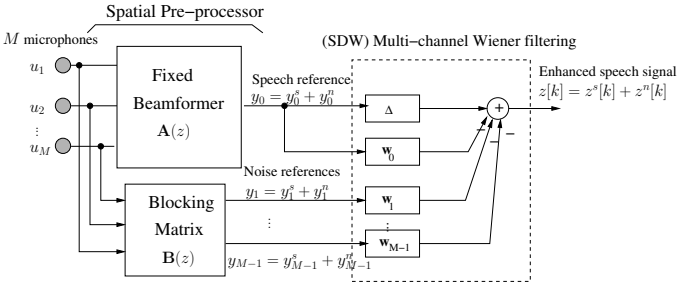


Fig. 1. Spatially pre-processed SDW MWF.

2. SPATIALLY PRE-PROCESSED SDW MWF

2.1. Concept

The SP-SDW-MWF, described in Figure 1, consists of a fixed, spatial pre-processor, i.e., a fixed beamformer $\mathbf{A}(z)$ and a blocking matrix $\mathbf{B}(z)$, and an adaptive SDW-MWF [1, 2, 8].

Given M microphone signals[1]

$$u_i[k] = u_i^s[k] + u_i^n[k],\ i = 1,\ ...,\ M \tag{1}$$

the spatial pre-processor creates a speech reference

$$y_0[k] = y_0^s[k] + y_0^n[k] \tag{2}$$

by steering a beam towards the front and $M-1$ noise references

$$y_i[k] = y_i^s[k] + y_i^n[k],\ i = 1,\ ...,\ M-1 \tag{3}$$

by steering zeroes towards the front. During periods of speech, the references $y_i[k]$ consist of speech + noise, i.e., $y_i[k] = y_i^s[k] + y_i^n[k]$, $i = 0,\ ...,\ M-1$. During periods of noise, only the noise component $y_i^n[k]$ is observed. The fixed beamformer $\mathbf{A}(z)$ should be designed so that the distortion in the speech reference $y_0^s[k]$ is minimal for all possible errors in the assumed signal model.

The adaptive SDW-MWF [1, 2, 8] $\mathbf{w}_k \in \mathbb{R}^{ML\times 1}$

$$\boxed{\mathbf{w}_k = \left(\frac{1}{\mu}\mathcal{E}\{\mathbf{y}_k^s\mathbf{y}_k^{s,T}\} + \mathcal{E}\{\mathbf{y}_k^n\mathbf{y}_k^{n,T}\}\right)^{-1}\mathcal{E}\{\mathbf{y}_k^n y_0^n[k-\Delta]\},} \tag{4}$$

with

$$\mathbf{w}_k^T = \left[\ \mathbf{w}_0^T[k] \quad \mathbf{w}_1^T[k] \quad ... \quad \mathbf{w}_{M-1}^T[k]\ \right], \tag{5}$$

$$\mathbf{w}_i[k] = \left[\ w[0] \quad w[1] \quad ... \quad w[L-1]\ \right]^T, \tag{6}$$

$$\mathbf{y}_k^T = \left[\ \mathbf{y}_0^T[k] \quad \mathbf{y}_1^T[k] \quad ... \quad \mathbf{y}_{M-1}^T[k]\ \right], \tag{7}$$

$$\mathbf{y}_i[k] = \left[\ y_i[k] \quad y_i[k-1] \quad ... \quad y_i[k-L+1]\ \right]^T, \tag{8}$$

[1]In the sequel, the superscripts s and n are used to refer to the speech and noise contribution of a signal.

provides an estimate $\mathbf{w}_k^T\mathbf{y}_k$ of the noise contribution[2] $\mathbf{y}_0^n[k-\Delta]$ in the speech reference by minimizing the cost function $J(\mathbf{w}_k)$

$$J(\mathbf{w}_k) = \frac{1}{\mu}\underbrace{\mathcal{E}\{\left|\mathbf{w}_k^T\mathbf{y}_k^s\right|^2\}}_{\varepsilon_d^2} + \underbrace{\mathcal{E}\{\left|y_0^n[k-\Delta] - \mathbf{w}_k^T\mathbf{y}_k^n\right|^2\}}_{\varepsilon_n^2}. \quad (9)$$

The term ε_d^2 represents the speech distortion energy and ε_n^2 the residual noise energy. The term $\frac{1}{\mu}\varepsilon_d^2$ in (9) limits the possible speech distortion at the output $z[k]$ of the SP-SDW-MWF. The parameter $1/\mu \in [0,\infty)$ trades off between noise reduction and speech distortion, hence the name *speech distortion weighted* MWF: the larger $1/\mu$, the smaller the possible speech distortion. For $\mu = 0$, all emphasis is put on speech distortion and so $z[k]$ is equal to the output of the fixed beamformer $\mathbf{A}(z)$, delayed by Δ samples.

2.2. Implementation of SP-SDW-MWF

In practice, $\mathcal{E}\{\mathbf{y}_k^s\mathbf{y}_k^{s,T}\}$ in (4) is unknown. Assuming that speech and noise are uncorrelated, $\mathcal{E}\{\mathbf{y}_k^s\mathbf{y}_k^{s,T}\}$ can be estimated as

$$\mathcal{E}\{\mathbf{y}_k^s\mathbf{y}_k^{s,T}\} = \mathcal{E}\{\mathbf{y}_k\mathbf{y}_k^T\} - \mathcal{E}\{\mathbf{y}_k^n\mathbf{y}_k^{n,T}\}, \quad (10)$$

where $\mathcal{E}\{\mathbf{y}_k\mathbf{y}_k^T\}$ is estimated during speech + noise and $\mathcal{E}\{\mathbf{y}_k^n\mathbf{y}_k^{n,T}\}$ during periods of noise only. The second order statistics of the noise signal are assumed to be quite stationary so that they can be estimated during periods of noise only. Like for the GSC, a robust speech detection is thus needed.

In [1, 2] implementations of the (SDW-)MWF have been proposed based on a GSVD or QR decomposition. A subband implementation [3] results in improved intelligibility at a significantly lower cost. The same[3] techniques can be applied to implement the SP-SDW-MWF.

3. DIFFERENT PARAMETER SETTINGS

Depending on the setting of $\frac{1}{\mu}$ and the presence or absence of the filter $\mathbf{w}_0$ on the speech reference, the SP-SDW-MWF corresponds to the GSC, an (SDW-)MWF or an in-between solution, called the SDR-GSC. We distinguish between two cases, i.e., the case where no filter $\mathbf{w}_0$ is applied to the speech reference (filter length $L_0 = 0$) and the case where an additional filter $\mathbf{w}_0$ is used ($L_0 \neq 0$).

3.1. SP-SDW-MWF without $\mathbf{w}_0$ (SDR-GSC)

First, consider *the case without* $\mathbf{w}_0$, i.e., $L_0 = 0$. The solution for $\bar{\mathbf{w}}_k^T = \begin{bmatrix} \mathbf{w}_1^T & \cdots & \mathbf{w}_{M-1}^T \end{bmatrix}$ in (4) then reduces to

$$\arg\min_{\bar{\mathbf{w}}_k} \frac{1}{\mu}\underbrace{\mathcal{E}\{\left|\bar{\mathbf{w}}_k^T\bar{\mathbf{y}}_k^s\right|^2\}}_{\varepsilon_d^2} + \underbrace{\mathcal{E}\{\left|y_0^n[k-\Delta] - \bar{\mathbf{w}}_k^T\bar{\mathbf{y}}_k^n\right|^2\}}_{\varepsilon_n^2}. \quad (11)$$

where $\bar{\mathbf{y}}_k^{s,n} = \begin{bmatrix} \mathbf{y}_1^{s,n}[k] & \cdots & \mathbf{y}_{M-1}^{s,n}[k] \end{bmatrix}^T$.

Compared to the ANC criterion of the GSC, i.e., minimization of the output noise power ε_n^2, a regularization term

$$\frac{1}{\mu}\mathcal{E}\{\left|\bar{\mathbf{w}}_k^T\bar{\mathbf{y}}_k^s\right|^2\} \quad (12)$$

has been added that limits the speech distortion due to model errors, hence the name *Speech Distortion Regularized GSC*. For $\mu = \infty$, distortion is ignored completely, which corresponds to the GSC-solution. Hence, the SDR-GSC encompasses the GSC as a special case.

The regularization term (12) with $\frac{1}{\mu} \neq 0$ adds robustness to the GSC, while not affecting the noise reduction performance in the absence of speech leakage:

- In the *absence of speech leakage, i.e.,* $\bar{\mathbf{y}}_k^s = \mathbf{0}$, the regularization term equals $\mathbf{0}$ for all $\bar{\mathbf{w}}$. Hence the residual noise energy ε_n^2 is effectively minimized or, in other words, the GSC solution is obtained.
- In the *presence of speech leakage*, i.e., $\bar{\mathbf{y}}_k^s \neq \mathbf{0}$, speech distortion is explicitly taken into account in the optimization criterion (11) for $\bar{\mathbf{w}}$, limiting speech distortion while reducing noise. The larger the amount of speech leakage, the more attention is paid to speech distortion.

To limit speech distortion alternatively, a QIC, i.e., $\bar{\mathbf{w}}^T\bar{\mathbf{w}} \leq \beta^2$, is often imposed on the filter $\bar{\mathbf{w}}$ [6, 7]. In contrast to the SDR-GSC, the QIC acts irrespective of the amount of speech leakage $\bar{\mathbf{y}}_k^s$ that is present. The constraint value β^2 has to be chosen based on the largest model errors that may occur. As a consequence, noise reduction performance is compromised even when no or very small model errors are present. Hence, the QIC is more conservative than the SDR-GSC (see also Section 4).

3.2. SP-SDW-MWF with filter $\mathbf{w}_0$

Since the SDW-MWF (4) takes speech distortion explicitly into account, a filter $\mathbf{w}_0$ on the speech reference $y_0[k]$ can be added (see Figure 1). The SDW-MWF then equals (4) where $\mathbf{w}_k^T = [\mathbf{w}_0^T\ \bar{\mathbf{w}}_k^T]$. Again, μ trades off speech distortion and noise reduction. For $\mu = \infty$, speech distortion is completely ignored and a zero output signal $z[k]$ is obtained. For $\mu = 1$, we obtain an MWF.

In addition, we can make the following statements:

- In the *absence of speech leakage* and for infinitely long filters $\mathbf{w}_i$, $i = 0, ..., M-1$, the SP-SDW-MWF with $\mathbf{w}_0$ corresponds to a *cascade of an SDR-GSC and an SDW Single-channel WF (SDW-SWF) postfilter* [9].
- In the *presence of speech leakage*, the SP-SDW-MWF with $\mathbf{w}_0$ tries to preserve its performance: the SP-SDW-MWF then contains extra filtering operations that compensate for the performance degradation of the SDR-GSC with SDW-SWF due to speech leakage (see Figure 2 and the proof below). In [8], e.g., we show that *for infinite filter lengths, the SP-SDW-MWF with* $\mathbf{w}_0$ *is not affected by microphone mismatch as long as the desired speech component at the output of* $\mathbf{A}(z)$ *remains unaltered.*

Proof: In case of infinite filter lengths, the SP-SDW-MWF can be represented in the frequency domain[4]. For simplicity, but without loss of generality, we assume $\Delta = 0$.

$$\mathbf{W}(f) = \arg\min_{\mathbf{W}} \mathcal{E}\left\{\left|\begin{bmatrix}(1-W_0^*) & -\bar{\mathbf{W}}^H\end{bmatrix}\begin{bmatrix}Y_0^n(f)\\ \bar{\mathbf{Y}}^n(f)\end{bmatrix}\right|^2\right\} + \frac{1}{\mu}\mathcal{E}\left\{\left|\begin{bmatrix} W_0^* & \bar{\mathbf{W}}^H\end{bmatrix}\begin{bmatrix} Y_0^s(f)\\ \bar{\mathbf{Y}}^s(f)\end{bmatrix}\right|^2\right\} \quad (13)$$

Decompose $\bar{\mathbf{W}}(f)$ as

$$\bar{\mathbf{W}}(f) = (1 - W_0(f))\,\bar{\mathbf{W}}_d(f) \quad (14)$$

[2] The delay Δ is applied to the speech reference to make $\mathbf{w}$ non-causal.

[3] The GSVD-based implementation can only be used if $\mathbf{w}_0 \neq \mathbf{0}$.

[4] The frequency parameter f is often omitted in the sequel for the sake of conciseness.

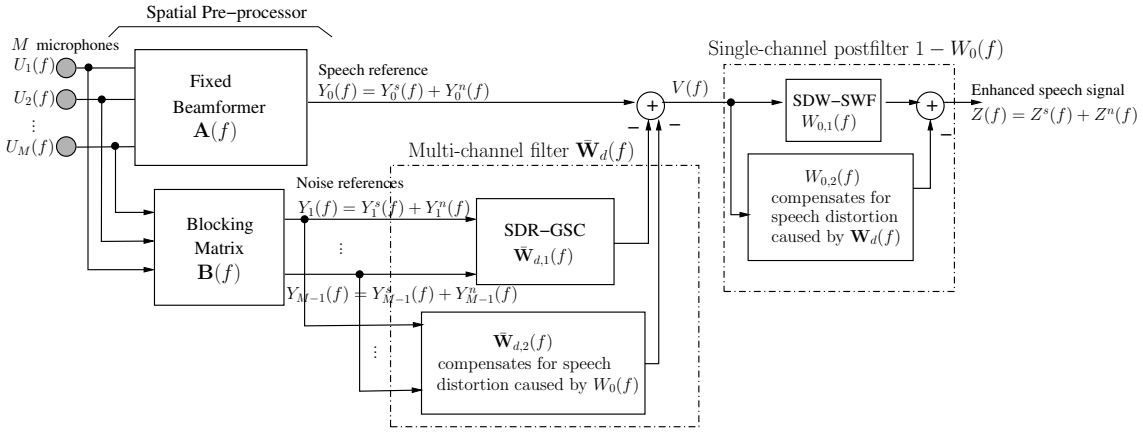


Fig. 2. Decomposition of SP-SDW-MWF with $W_0(f)$ in a multi-channel filter $\bar{\mathbf{W}}_d(f)$ and single-channel postfilter $1 - W_0(f)$.

with $W_0(f)$ the single-channel filter applied to the speech reference and $\bar{\mathbf{W}}_d(f)$ a multi-channel filter and define an intermediate output $V(f)$ (see also Figure 2) as

$$V(f) = Y_0(f) - \bar{\mathbf{W}}_d^H(f)\bar{\mathbf{Y}}(f). \quad (15)$$

Then, the cost function $J(W_0, \bar{\mathbf{W}}_d)$ of (13) can be re-written as

$$J = \mathcal{E}\left\{|(1 - W_0^*)\, V^n|^2\right\} + \frac{1}{\mu}\mathcal{E}\left\{\left|W_0^* V^s + \bar{\mathbf{W}}_d^H \bar{\mathbf{Y}}^s\right|^2\right\}. \quad (16)$$

From $\frac{\partial}{\partial W_0} J(W_0, \bar{\mathbf{W}}_d) = 0$, we find

$$W_0 = \underbrace{\left(\mathcal{E}\{V^n V^{n,*}\} + \frac{1}{\mu}\mathcal{E}\{V^s V^{s,*}\}\right)^{-1}\mathcal{E}\{V^n V^{n,*}\}}_{W_{0,1}(f)}$$
$$\underbrace{-\left(\mu\mathcal{E}\{V^n V^{n,*}\} + \mathcal{E}\{V^s V^{s,*}\}\right)^{-1}\mathcal{E}\{V^s \bar{\mathbf{Y}}^{s,H}\bar{\mathbf{W}}_d\}}_{W_{0,2}(f)}, \quad (17)$$

This *single-channel filter* $W_0(f)$ thus consists of two terms.

- The *first term* $W_{0,1}(f)$ estimates the noise component $V^n(f)$ in the intermediate output $V(f)$. The filter $1 - W_{0,1}$ then corresponds to an SDW-SWF that estimates the speech component $V^s(f)$ in the intermediate output $V(f)$.
- The *second term* $W_{0,2}(f)$ estimates the speech leakage filtered by $\bar{\mathbf{W}}_d(f)$, i.e., $-\bar{\mathbf{W}}_d^H\bar{\mathbf{Y}}^s$. The speech component in the intermediate output $V(f)$ equals $V^s(f) = Y_0^s - \bar{\mathbf{W}}_d^H\bar{\mathbf{Y}}^s$. The filter $W_{0,2}(f)$ thus tries to compensate for the distortion $-\bar{\mathbf{W}}_d^H\bar{\mathbf{Y}}^s$ by adding an estimate of $\bar{\mathbf{W}}_d^H\bar{\mathbf{Y}}^s$ to the output of the SDW-SWF. In the absence of speech leakage (i.e., $\bar{\mathbf{Y}}^s(f) = \mathbf{0}$), $W_{0,2}(f) = 0$.

From $\frac{\partial}{\partial \bar{\mathbf{W}}_d} J(W_0, \bar{\mathbf{W}}_d) = 0$, we find:

$$\bar{\mathbf{W}}_d = \underbrace{\left(\mathcal{E}\{\bar{\mathbf{Y}}^n\bar{\mathbf{Y}}^{n,H}\} + \frac{1}{\mu}\mathcal{E}\{\bar{\mathbf{Y}}^s\bar{\mathbf{Y}}^{s,H}\}\right)^{-1}\mathcal{E}\{\bar{\mathbf{Y}}^n Y_0^{n,*}\}}_{\bar{\mathbf{W}}_{d,1}}$$
$$\underbrace{-\left(\mu\mathcal{E}\{\bar{\mathbf{Y}}^n\bar{\mathbf{Y}}^{n,H}\} + \mathcal{E}\{\bar{\mathbf{Y}}^s\bar{\mathbf{Y}}^{s,H}\}\right)^{-1}\mathcal{E}\{\bar{\mathbf{Y}}^s Y_0^{s,*}\frac{W_0}{1 - W_0}\}}_{\bar{\mathbf{W}}_{d,2}}. \quad (18)$$

This *multi-channel filter* $\bar{\mathbf{W}}_d(f)$ consists of two terms.

- The *first term* $\bar{\mathbf{W}}_{d,1}(f)$ corresponds to the SDR GSC and estimates the noise component $Y_0^n(f)$ at the output of the fixed beamformer $\mathbf{A}(f)$.
- The *second term* $\bar{\mathbf{W}}_{d,2}(f)$ tries to compensate for the speech distortion $-W_0^*(f)Y_0^s(f)$ caused by $W_0(f)$ by adding an estimate of $\frac{W_0^*(f)}{1-W_0^*(f)}Y_0^s(f)$ to the output of the SDR-GSC. Note that this corresponds to adding an estimate of $W_0^*(f)Y_0^s(f)$ to the output $Z(f)$ of the SP-SDW-MWF. In the absence of speech leakage, $\bar{\mathbf{W}}_{d,2}(f) = \mathbf{0}$.

Figure 2 graphically illustrates the solution for $\bar{\mathbf{W}}_d(f)$ and $W_0(f)$. In the *absence of speech leakage*, the filters $W_{0,2}(f)$ and $\bar{\mathbf{W}}_{d,2}(f)$ equal 0, hence, the SP-SDW-MWF corresponds to an SDR-GSC (or GSC) cascaded with a SDW-SWF. In the *presence of speech leakage*, the SP-SDW-MWF with $\mathbf{w}_0$ tries to preserve its performance: the SP-SDW-MWF then contains extra filtering operations (i.e., $W_{0,2}(f)$ and $\bar{\mathbf{W}}_{d,2}(f)$) that compensate for the performance degradation of the SDR-GSC with SDW-SWF due to speech leakage ■

4. ILLUSTRATION THROUGH EXPERIMENTS

This section illustrates the theoretical results of Section 3 through experimental results with an BTE hearing aid.

4.1. Set-up and performance measures

The performance of the SP-SDW-MWF has been assessed for different parameter settings based on recordings in an office room with a three-microphone BTE, mounted on a dummy head. The desired signal and the noise signals are uncorrelated, stationary and speech-like. The desired signal and the total noise signal both have a level of 70 dB SPL at the center of the head. The desired source is positioned in front of the head. Five noise sources are positioned at 75°, 120°, 180°, 240° and 285°. For evaluation purposes, the speech and noise signal have been recorded separately. In the experiments, the microphones have been calibrated in an anechoic room while the BTE was mounted on the head. A delay-and-sum beamformer is used as a fixed beamformer. For small-sized arrays, this beamformer offers sufficient robustness against signal model errors as it minimizes the white noise gain[5]. The blocking matrix $\mathbf{B}$ pairwise subtracts the time aligned calibrated microphone signals.

To investigate the effect of the different parameter settings (i.e., μ, $\mathbf{w}_0$) on the performance of the SP-SDW-MWF, the filter coef-

[5]The white noise gain, defined as the ratio of the spatially white noise gain to the gain of the desired signal, is often used to quantify the sensitivity of an algorithm against errors in the assumed signal model [6].

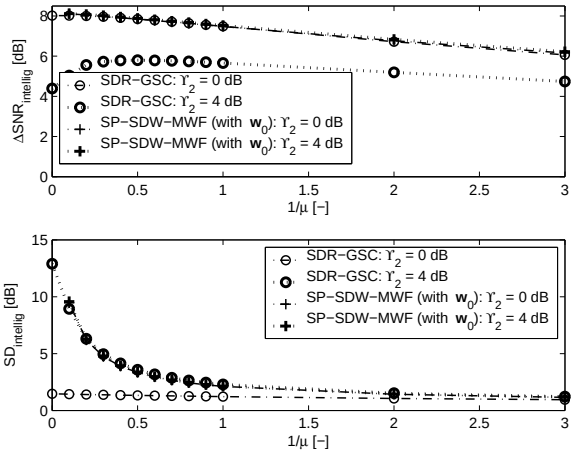


Fig. 3. Performance of SDR-GSC and SP-SDW-MWF.

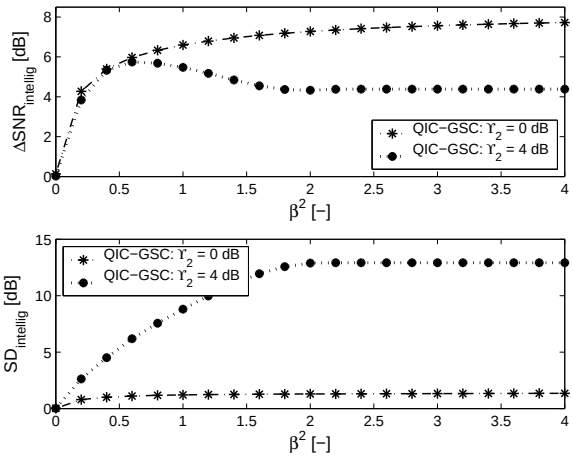


Fig. 4. Performance of QIC-GSC.

ficients are computed using (4) where $\mathcal{E}\{\mathbf{y}_k^s \mathbf{y}_k^{s,T}\}$ is estimated by means of the clean speech contributions of the microphone signals. In practice, $\mathcal{E}\{\mathbf{y}_k^s \mathbf{y}_k^{s,T}\}$ is approximated using (10). The effect of approximation (10) on the performance was found to be small for the given data set. The QIC-GSC is implemented using variable loading RLS [7]. The filter length $L = 96$.

To assess the performance, the intelligibility weighted signal-to-noise ratio improvement $\Delta\text{SNR}_{\text{intellig}}$ is used, defined as

$$\Delta\text{SNR}_{\text{intellig}} = \sum_i I_i(\text{SNR}_{i,\text{out}} - \text{SNR}_{i,\text{in}}), \qquad (19)$$

where the band importance function I_i expresses the importance of the i-th one-third octave band with center frequency f_i^c for intelligibility [10], and where $\text{SNR}_{i,\text{out}}$ and $\text{SNR}_{i,in}$ is the output and input SNR (in dB) in the i-th one-third octave band, respectively. Similarly, we define an intelligibility weighted spectral distortion measure (in dB), called $\text{SD}_{\text{intellig}}$, of the desired signal as

$$\text{SD}_{\text{intellig}} = \sum_i I_i \text{SD}_\text{i} \qquad (20)$$

with SD_i the average spectral distortion (dB) in i-th one-third band, calculated as

$$\text{SD}_i = \frac{1}{(2^{1/6} - 2^{-1/6})\, f_i^c} \int_{2^{-1/6} f_i^c}^{2^{1/6} f_i^c} |10 \log_{10} G^s(f)|\, df, \qquad (21)$$

with $G^s(f)$ the power transfer function of speech from the input to the output of the noise reduction algorithm. To exclude the effect of the spatial pre-processor, the performance measures are calculated with respect to the output of the fixed beamformer.

4.2. Experimental results

Figure 3 depicts $\Delta\text{SNR}_{\text{intellig}}$ and $\text{SD}_{\text{intellig}}$ of the SDR-GSC and the SP-SDW-MWF with respect to the output of the fixed beamformer as a function of the trade-off parameter $\frac{1}{\mu}$. The effect of a gain mismatch Υ_2 of 4 dB at the second microphone is depicted too. For comparison, Figure 4 plots the performance of the QIC-GSC with QIC $\bar{\mathbf{w}}^T \bar{\mathbf{w}} \leq \beta^2$, as a function of β^2.

Both, the SP-SDW-MWF and the QIC-GSC increase the robustness of the GSC (i.e., the SDR-GSC with $1/\mu = 0$): the speech distortion in the presence of model errors is reduced by increasing $1/\mu$ or decreasing β^2. For the given set-up, a value $1/\mu$ between 0.4 and 0.8 seems appropriate for guaranteeing good performance for a gain mismatch up to 4 dB.

For a given maximum allowable distortion $\text{SD}_{\text{intellig}}$, the SDR-GSC and the SP-SDW-MWF with $\mathbf{w}_0$ achieve a better noise reduction performance than the QIC-GSC. The SDR-GSC outperforms the QIC-GSC for small model errors, while guaranteeing robustness against large model errors. The performance of the SP-SDW-MWF with $\mathbf{w}_0$ is -in contrast to the SDR-GSC and the QIC-GSC- not affected by microphone mismatch.

5. REFERENCES

[1] S. Doclo and M. Moonen, "GSVD-Based Optimal Filtering for Single and Multimicrophone Speech Enhancement," *IEEE Trans. SP*, vol. 50, no. 9, pp. 2230–2244, Sept. 2002.

[2] G. Rombouts and M. Moonen, "QRD-based optimal filtering for acoustic noise reduction," in *Proc. of EUSIPCO*, 2002, vol. 3, pp. 301–304.

[3] A. Spriet, M. Moonen, and J. Wouters, "A multi-channel subband GSVD approach to speech enhancement," *ETT*, vol. 13, no. 2, pp. 149–158, 2002.

[4] L. J. Griffiths and C. W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. AP*, vol. 30, pp. 27–34, Jan. 1982.

[5] A. Spriet, M. Moonen, and J. Wouters, "Robustness analysis of GSVD based optimal filtering and GSC for hearing aid applications," in *Proc. of WASPAA*, 2001.

[6] H. Cox, R. M. Zeskind, and M. M. Owen, "Robust Adaptive Beamforming," *IEEE Trans. ASSP*, vol. 35, no. 10, pp. 1365–1376, Oct. 1987.

[7] Z. Tian, K.L. Bell, and H.L. Van Trees, "A Recursive Least Squares Implementation for LCMP Beamforming Under Quadratic Constraint," *IEEE Trans. SP*, vol. 49, no. 6, pp. 1138–1145, June 2001.

[8] A. Spriet, M. Moonen, and J. Wouters, "Spatially pre-processed speech distortion weighted multi-channel Wiener filtering for noise reduction," Tech. Rep. ESAT-SISTA/TR 03-46, ESAT/SISTA, K.U. Leuven (Belgium), 2003.

[9] C. Marro, Y. Mahieux, and K. U. Simmer, "Analysis of Noise Reduction and Dereverberation Techniques Based on Microphone Arrays with Postfiltering," *IEEE Trans. SAP*, vol. 6, no. 3, pp. 240–259, May 1998.

[10] Acoustical Society of America, "ANSI S3.5-1997 American National Standard Methods for Calculation of the Speech Intelligibility Index," June 1997.