MULTICHANNEL WIENER FILTERING VIA MULTICHANNEL DECORRELATION

Philipp Thüne and Gerald Enzner

Institute of Communication Acoustics (IKA), Ruhr-Universität Bochum, D-44780 Bochum, Germany
Email: {philipp.thuene | gerald.enzner}@rub.de

ABSTRACT

Extracting a target source signal from multiple noisy observations is an essential task in many applications of signal processing such as digital communications or speech and audio processing. The multichannel Wiener filter is able to solve this task in a minimum-mean-square-error (MMSE) optimal way by applying a spatial filter succeeded by a spectral postfilter. Its direct implementation, however, is difficult due to requiring the statistics of the unobservable source and noise signals. In this paper, we apply the signal-separation-based technique of multichannel decorrelation and reveal its relation to the Wiener post-filtering component. On this basis, we present a numerically robust and efficient adaptive algorithm to find an estimate of the MMSE-optimal postfilter based on the statistics of the observable signals alone. Experimental evaluation demonstrates the validity of the proposed approach and confirms the convergence of the adaptive algorithm to the MMSE-optimal postfilter solution.

Index Terms— Multichannel Wiener filter, multichannel decorrelation, adaptive filters, post-filtering

1. INTRODUCTION

Multichannel Wiener filtering (MWF) has attracted the signal processing community for its ability to extract a target source signal with minimum mean-square error (MMSE) from multiple noisy observations, thereby utilizing the diversity information related to the multiple sensors [1, 2, 3]. Some applications, such as speech and audio signal processing, or digital communications, found means to partly detect the required *a priori* information of the noisy processes from the available signals. The field of blind channel identification [4], for instance, emerged in digital communications and further evolved in the speech community to detect the source-to-receiver transfer functions [5, 6, 7, 8, 9] needed for the implementation of the spatial minimum variance distortionless response (MVDR) beamformer part of the MWF [10]. The speech enhancement field additionally provides options to detect the signal-to-noise ratio (SNR) of the MVDR output, effectively via estimation of the covariance of the unobservable noise signals, which turns out to be crucial for the performance of MWF [11]. The frequency-dependent SNR is then used for spectral-enhancement filtering, the second part of the MWF, subsequent to the MVDR part [12, 10, 13, 14]. In speech enhancement, the noise covariance estimation is mainly performed by capitalizing on the fact that the speech signal is assumed to be less stationary than the environment noise [15, 16, 17] or, alternatively, by beamforming approaches [18]. Other approaches make use of the cross-correlation or coherence to adapt an appropriate postfilter [19, 20, 21].

In our contribution, we revert to the case of stationary Wiener filter theory, where the SNR cannot be detected from the MVDR output alone. In order to overcome this fundamental limitation, we suggest to look at MMSE-optimal individual reconstructions of the target signal at the individual sensors via the MWF, similar to the approach in hearing aid applications where linear modifications associated with binaural cues are to be preserved [22, 23, 24]. In particular, we look at the individual output errors with respect to the original individual sensor signals. These output errors can be considered estimates of the observation noise and were thus decorrelated if the original observation noises were mutually independent. Hence, we propose to enforce such decorrelation of the observable output error processes to retrieve an optimal decorrelation postfilter that is related to the MMSE-optimal spectral postfilter component of the MWF. This technique is referred to as multichannel decorrelation (MCD) and the paper describes the exact relationship of MCD and MWF, which turns out to be a simple memoryless mapping function. On this basis, we derive a numerically robust and efficient algorithm to adapt the MCD solution and exploit the found relationship to obtain an estimated postfilter that converges to the ideal MWF postfilter. Simulations illustrate how both the available number of observations in time and the number of channels contribute to this convergence.

The remainder of our paper is organized as follows. Sec. 2 will review the multichannel Wiener filter and its decomposition into a spatial filter and a spectral postfilter. In Sec. 3 we introduce the idea of multichannel decorrelation and highlight its relation to multichannel Wiener post-filtering before presenting the proposed algorithm for postfilter adaptation in Sec. 4. Sec. 5 provides experimental validation of the proposed technique and finally Sec. 6 concludes this contribution by distinguishing it from prior work.

Throughout this paper we will use uppercase letters to denote frequency domain quantities, bold-face letters for vectors, and underlined letters for matrices. $\underline{\mathbf{I}}$ and $\underline{\mathbf{0}}$ are identity and all-zero matrices of the size defined by a subscript and $\mathcal{E}\{\cdot\}$ is statistical expectation, while $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$, and $(\cdot)^{-1}$ denote complex conjugate, matrix transpose, Hermitian transpose, and inverse, respectively.

2. MULTICHANNEL WIENER FILTERING

We consider a multichannel system consisting of P channels in the frequency domain for a single frequency bin only, where we omit the frequency index for clarity of presentation. In our system, a single source signal S is transmitted via channels H_i, $i = 1 \dots P$, to receivers that pick up signals Y_i as a sum of the transmitted signals $D_i = H_i\, S$ and observation noise N_i, cf. Fig. 1. We may compactly write this in vector notation as

$$\mathbf{Y} = \mathbf{D} + \mathbf{N}\,, \tag{1}$$

where $\mathbf{D} = \mathbf{H}\, S$ and the signal and channel vectors are defined as $\mathbf{Y} = [Y_1\, Y_2\, \cdots\, Y_P]^T$, $\mathbf{N} = [N_1\, N_2\, \cdots\, N_P]^T$, and $\mathbf{H} = [H_1\, H_2\, \cdots\, H_P]^T$, respectively.

The multichannel Wiener filter (MWF) is the MMSE-optimal estimator for the source signal in this multichannel scenario [10, 12],

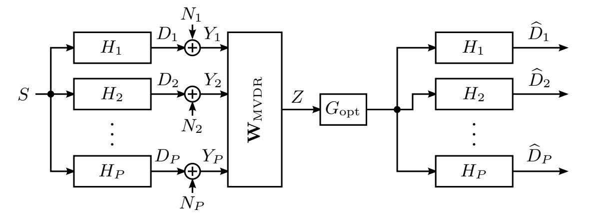


Fig. 1. Signal model and multichannel Wiener filter decomposition.

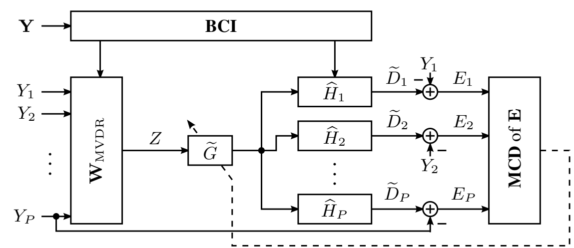


Fig. 2. Block diagram of multichannel decorrelation.

and can provide an estimate $\widehat{D}_i$ of D_i that minimizes $J_{\text{MSE}} = \sum_{i=1}^{P} \mathcal{E}\left\{(D_i - \widehat{D}_i)(D_i - \widehat{D}_i)^*\right\}$. This estimate is obtained by applying filters $\mathbf{W}_{i,\text{opt}}$ to the received signal vector $\mathbf{Y}$, i.e., $\widehat{D}_i = \mathbf{W}_{i,\text{opt}}^T \mathbf{Y}$. By minimization of J_{MSE}, this filter can be found to be a regular MWF [10], multiplied by the respective channel transfer function H_i,

$$\begin{aligned} \mathbf{W}_{i,\text{opt}} &= H_i \left(\mathbf{H}^*\mathbf{H}^T \Phi_S + \underline{\mathbf{\Phi}}_N\right)^{-1} \mathbf{H}^* \Phi_S \\ &= H_i \, \mathbf{W}_{\text{MVDR}} \, G_{\text{opt}} \,, \end{aligned} \tag{2}$$

that can be decomposed into an MVDR beamformer $\mathbf{W}_{\text{MVDR}}$ for spatial filtering and a single-channel Wiener postfilter G_{opt} for spectral enhancement [10, 12], i.e.,

$$\mathbf{W}_{\text{MVDR}} = \frac{\underline{\mathbf{\Phi}}_N^{-1}\mathbf{H}^*}{\mathbf{H}^T \underline{\mathbf{\Phi}}_N^{-1} \mathbf{H}^*} \,, \tag{3}$$

$$G_{\text{opt}} = \frac{\Phi_S}{\Phi_S + \left(\mathbf{H}^T \underline{\mathbf{\Phi}}_N^{-1} \mathbf{H}^*\right)^{-1}} \,. \tag{4}$$

The spatial and spectral filter components of the MWF require the covariances $\Phi_S = \mathcal{E}\{SS^*\}$ and $\underline{\mathbf{\Phi}}_N = \mathcal{E}\left\{\mathbf{N}\mathbf{N}^H\right\}$ of the source and observation noise signals, respectively. Fig. 1 summarizes the structural decomposition of the MWF, where the MVDR output signal is denoted $Z = \mathbf{W}_{\text{MVDR}}^T \mathbf{Y}$.

In this paper we are particularly interested in the so-called postfilter component G_{opt} of the MWF, assuming that a reasonable MVDR-estimate has already been found. Since the covariances required in (4) are those of the unobservable processes S and $\mathbf{N}$, it is subject to research how to estimate these statistics from the observable signals $\mathbf{Y}$. A good overview is provided, e.g., in [12, 13]. In contrast to estimating Φ_S and $\underline{\mathbf{\Phi}}_N$ and constructing the postfilter separately, we present in the following section a technique to adapt a postfilter in a more contained fashion based on the observable signals only.

3. MULTICHANNEL DECORRELATION

In this section we assume that the observation noise is uncorrelated and, for ease of presentation, also has equal variance Φ_N on all channels, i.e., $\underline{\mathbf{\Phi}}_N = \Phi_N \underline{\mathbf{I}}_{P\times P}$. This assumption to some extent simplifies the MWF decomposition such that (3) and (4) become

$$\mathbf{W}_{\text{MVDR}} = \frac{\mathbf{H}^*}{\|\mathbf{H}\|^2} \,, \tag{5}$$

$$G_{\text{opt}} = \frac{\Phi_S}{\Phi_S + \frac{\Phi_N}{\|\mathbf{H}\|^2}} \,, \tag{6}$$

where $\|\mathbf{H}\|^2 = \mathbf{H}^H\mathbf{H}$. The assumption of a diagonal noise covariance then allows to apply techniques from source separation to yield a multichannel decorrelation procedure that eventually relates to the estimation of the Wiener postfilter G_{opt} for signal enhancement.

3.1. Concept of Multichannel Decorrelation

In our multichannel decorrelation (MCD) approach illustrated in Fig. 2, we assume that a blind channel identification (BCI) technique has been applied to yield channel estimates $\widehat{\mathbf{H}} = \left[\widehat{H}_1 \; \widehat{H}_2 \; \cdots \; \widehat{H}_P\right]^T$. For cross-relation-based BCI algorithms [4, 25], it has been shown that the error due to the blindness of the approach can be expressed as a single channel convolutive error in time domain [26]. In our frequency domain model we may therefore describe the estimated channels as $\widehat{\mathbf{H}} = F\,\mathbf{H}$, where F denotes an arbitrary complex-valued error that is common to all channels in a certain frequency band. The blindly estimated $\widehat{\mathbf{H}}$ are then used to implement an MVDR beamformer according to (5) and for reconstructing the individual $\widetilde{D}_i$.

As depicted in Fig. 2, we aim to find the single channel filter $\widetilde{G}$ that spectrally modifies the MVDR output Z such that the error

$$\mathbf{E} = \mathbf{Y} - \widetilde{\mathbf{D}} \approx \mathbf{N} \tag{7}$$

between the observed signals $\mathbf{Y}$ and estimated signals $\widetilde{\mathbf{D}} = \widehat{\mathbf{H}}\,\widetilde{G}\,Z = \left[\widetilde{D}_1 \; \widetilde{D}_2 \; \cdots \; \widetilde{D}_P\right]^T$, is decorrelated in the sense that the error covariance $\underline{\mathbf{\Phi}}_E = \mathcal{E}\left\{\mathbf{E}\mathbf{E}^H\right\}$ is rendered diagonal.

3.2. Cost Function and Optimal MCD Postfilter

For calculating a postfilter $\widetilde{G}$ that minimizes the cross-covariance of the error signals in (7) we define the cost function

$$J_{\text{MCD}}(\widetilde{G}) = \log\det\operatorname{diag}\underline{\mathbf{\Phi}}_E - \log\det\underline{\mathbf{\Phi}}_E \tag{8}$$

as a variant of the cost function proposed in [27] for speech signal separation, which has later been generalized in an information-theoretic sense [28].

To find the optimum postfilter $\widetilde{G}$ for decorrelation, we set the derivative [29] of the cost function in (8)

$$\frac{d\,J_{\text{MCD}}(\widetilde{G})}{d\,\widetilde{G}} = \operatorname{Tr}\left[(\operatorname{diag}\underline{\mathbf{\Phi}}_E)^{-1}\frac{d\operatorname{diag}\underline{\mathbf{\Phi}}_E}{d\,\widetilde{G}}\right] - \operatorname{Tr}\left[\underline{\mathbf{\Phi}}_E^{-1}\frac{d\,\underline{\mathbf{\Phi}}_E}{d\,\widetilde{G}}\right] \tag{9}$$

to zero and insert the expanded error covariance based on (7), i.e.,

$$\underline{\mathbf{\Phi}}_E = \widetilde{G}^2\Phi_Z\widehat{\mathbf{H}}\widehat{\mathbf{H}}^H - 2\,\widetilde{G}\underline{\mathbf{\Phi}}_Y\frac{\widehat{\mathbf{H}}\widehat{\mathbf{H}}^H}{\|\widehat{\mathbf{H}}\|^2} + \underline{\mathbf{\Phi}}_Y \,, \tag{10}$$

where the covariance of the received signals,

$$\begin{aligned} \underline{\mathbf{\Phi}}_Y &= \mathcal{E}\left\{\mathbf{Y}\mathbf{Y}^H\right\} \\ &= \Phi_S\mathbf{H}\mathbf{H}^H + \Phi_N\underline{\mathbf{I}}_{P\times P} \end{aligned} \tag{11}$$

is the sum of a rank-one matrix and the fully diagonal noise covariance, and the output covariance of the MVDR beamformer is

$$\begin{aligned}\Phi_Z &= \mathcal{E}\{ZZ^*\}\\ &= \frac{\widehat{\mathbf{H}}^H}{\|\widehat{\mathbf{H}}\|^2}\left(\Phi_S\mathbf{H}\mathbf{H}^H + \Phi_N\underline{\mathbf{I}}_{P\times P}\right)\frac{\widehat{\mathbf{H}}}{\|\widehat{\mathbf{H}}\|^2}\\ &= \frac{1}{|F|^2}\left(\Phi_S + \frac{\Phi_N}{\|\mathbf{H}\|^2}\right). \end{aligned} \tag{12}$$

After rearranging the final result, we find a total of three equilibrium points of the cost function in (8). The first solution, $\widetilde{G}_1^\star = 1$, turns out to be a fixed local maximum of the cost function, whereas the second and third solutions

$$\widetilde{G}_{2/3}^\star = 1 \pm \sqrt{1 - \frac{\Phi_S}{\Phi_S + \frac{\Phi_N}{\|\mathbf{H}\|^2}}} \tag{13}$$

are local minima. In fact, $\widetilde{G}_{2/3}^\star$ are global minima, which can be understood by inserting both (11) and (12) into (10) and eventually substituting (13), i.e.,

$$\begin{aligned}\underline{\mathbf{\Phi}}_E &= \left(\widetilde{G}_{2/3}^\star\right)^2\Phi_Z\widehat{\mathbf{H}}\widehat{\mathbf{H}}^H - 2\,\widetilde{G}_{2/3}^\star\underline{\mathbf{\Phi}}_Y\frac{\widehat{\mathbf{H}}\widehat{\mathbf{H}}^H}{\|\widehat{\mathbf{H}}\|^2} + \underline{\mathbf{\Phi}}_Y\\ &= \left(\widetilde{G}_{2/3}^\star\right)^2\left(\Phi_S + \frac{\Phi_N}{\|\mathbf{H}\|^2}\right)\mathbf{H}\mathbf{H}^H\\ &\quad - 2\,\widetilde{G}_{2/3}^\star\left(\Phi_S + \frac{\Phi_N}{\|\mathbf{H}\|^2}\right)\mathbf{H}\mathbf{H}^H + \Phi_S\mathbf{H}\mathbf{H}^H + \Phi_N\underline{\mathbf{I}}_{P\times P}\\ &= \Phi_N\underline{\mathbf{I}}_{P\times P}. \end{aligned} \tag{14}$$

Solutions $\widetilde{G}_{2/3}^\star$ thus fully diagonalize the error covariance and further enforce $J_{\text{MCD}}(\widetilde{G}_{2/3}^\star) = 0$. The cost function in (8) is a nonnegative function by virtue of Hadamard's inequality [30] and we therefore conclude that $\widetilde{G}_{2/3}^\star$ are global minima as anticipated. Notice also that (14), and hence the optimality of $\widetilde{G}_{2/3}^\star$ in terms of ideal decorrelation, holds regardless of the arbitrary error factor F introduced by the BCI because it cancels in all terms. The MCD solution is therefore robust to this kind of unavoidable error factor.

3.3. Relation to Multichannel Wiener Filter

While the MWF has been designed to yield MMSE-optimal estimates $\widehat{D}_i$ of the transmitted signal components D_i, the previously derived MCD approach effectively constructs $\widetilde{D}_i$ such that the corresponding error signals E_i are mutually uncorrelated. Since G_{opt} and $\widetilde{G}_{2/3}^\star$ minimize their respective cost functions J_{MSE} and J_{MCD} it is clear that neither would the MWF ideally decorrelate the output errors nor would the MCD solution be MMSE-optimal. In fact, by comparing the filter gain G_{opt} to $\widetilde{G}_3^\star$ [1] for various signal-to-noise ratios (SNR) at the MVDR output, cf. Fig. 3, we observe that $\widetilde{G}_3^\star$ is overly aggressive w.r.t. the optimal MMSE signal enhancement.

Inspecting (6) and (13), however, we interestingly find that both optimal postfilters are essentially functions of the source variance Φ_S and the effective noise variance $\Phi_N/\|\mathbf{H}\|^2$ at the MVDR output. Hence, it is easily possible to come up with a memoryless nonlinear mapping between $\widetilde{G}_{2/3}^\star$ and G_{opt}, i.e.,

$$G_{\text{opt}} = 1 - \left(\widetilde{G}_{2/3}^\star - 1\right)^2. \tag{15}$$

[1] We choose $\widetilde{G}_3^\star$ over $\widetilde{G}_2^\star$ here, since $0 < \widetilde{G}_3^\star < 1$ similar to $0 < G_{\text{opt}} < 1$.

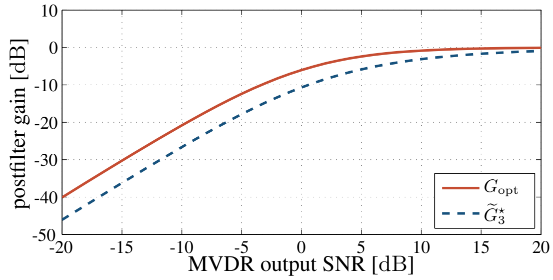


Fig. 3. MWF postfilter gain G_{opt} and MCD postfilter gain $\widetilde{G}_3^\star$ as a function of the MVDR output $\text{SNR}_{\text{MVDR}} = \Phi_S/(\Phi_N/\|\mathbf{H}\|^2)$.

It is important to notice that the decorrelating postfilter $\widetilde{G}$ is derived from a cost function based on the observable errors $\mathbf{E}$ only. We therefore propose MCD in conjunction with (15) to obtain an estimate $\widehat{G}$ of the optimal signal enhancement postfilter G_{opt} without the need to infer statistics Φ_S and Φ_N of the unobservable signals.

4. POSTFILTER ADAPTATION ALGORITHM

The MCD cost function in (8) is based on observable signals only, yet a straightforward minimization by means of gradient descent [31] turns out to be computationally inconvenient and might cause numerical problems since the gradient in (9) requires the inverse of the error covariance. We have, however, shown that the optimal MCD postfilter $\widetilde{G}_{2/3}^\star$ reduces the cost function to zero and thus makes $\underline{\mathbf{\Phi}}_E$ a fully diagonal matrix. We may therefore approach a data-driven calculation of $\widetilde{G}$ by removing the main diagonal from all involved matrices in (10) and solve

$$\begin{aligned}\underline{\mathbf{0}}_{P\times P} &= \widetilde{G}^2\Phi_Z\left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^H - \operatorname{diag}\left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^H\right)\right)\\ &\quad - 2\,\widetilde{G}\left(\underline{\mathbf{\Phi}}_Y\frac{\widehat{\mathbf{H}}\widehat{\mathbf{H}}^H}{\|\widehat{\mathbf{H}}\|^2} - \operatorname{diag}\left(\underline{\mathbf{\Phi}}_Y\frac{\widehat{\mathbf{H}}\widehat{\mathbf{H}}^H}{\|\widehat{\mathbf{H}}\|^2}\right)\right)\\ &\quad + \left(\underline{\mathbf{\Phi}}_Y - \operatorname{diag}\underline{\mathbf{\Phi}}_Y\right). \end{aligned} \tag{16}$$

Since the system of equations in (16) holds for every single element, we can likewise turn this into a vector equation

$$\mathbf{0}_{P^2} = \widetilde{G}^2\,\mathbf{a} + \widetilde{G}\,\mathbf{b} + \mathbf{c}, \tag{17}$$

in which the P^2-element vectors

$$\mathbf{a} = \Phi_Z\operatorname{vect}\left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^H - \operatorname{diag}\left(\widehat{\mathbf{H}}\widehat{\mathbf{H}}^H\right)\right), \tag{18}$$

$$\mathbf{b} = -2\operatorname{vect}\left(\underline{\mathbf{\Phi}}_Y\frac{\widehat{\mathbf{H}}\widehat{\mathbf{H}}^H}{\|\widehat{\mathbf{H}}\|^2} - \operatorname{diag}\left(\underline{\mathbf{\Phi}}_Y\frac{\widehat{\mathbf{H}}\widehat{\mathbf{H}}^H}{\|\widehat{\mathbf{H}}\|^2}\right)\right), \tag{19}$$

$$\mathbf{c} = \operatorname{vect}\left(\underline{\mathbf{\Phi}}_Y - \operatorname{diag}\underline{\mathbf{\Phi}}_Y\right) \tag{20}$$

contain the elements of the respective matrices in arbitrary order. In block-frequency domain processing, we may obtain estimates $\widehat{\underline{\mathbf{\Phi}}}_Y$ and $\widehat{\Phi}_Z$ by averaging $\mathbf{Y}\mathbf{Y}^H$ and $|Z|^2$, respectively, over a fixed number of signal frames, to approximate the vectors in (18) to (20). Eq. (17) can then be considered as a nonlinear least-squares problem and can be solved, e.g., by applying the Gauss-Newton algorithm [32]. We obtain the MCD postfilter by repeating Q times the Gauss-

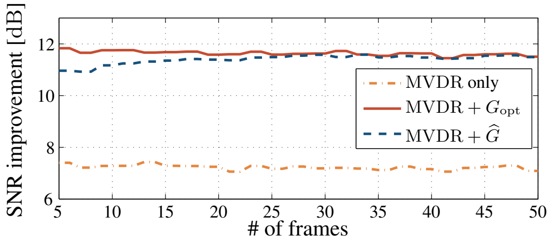


Fig. 4. Comparison of SNR improvement for different numbers of signal frames used for $\widehat{\underline{\mathbf{\Phi}}}_Y$ calculation.

Newton iteration with iteration index q

$$\mathbf{r}^{(q)} = \mathbf{0} - \left((\widetilde{G}^{(q)})^2 \, \mathbf{a} + \widetilde{G}^{(q)} \, \mathbf{b} + \mathbf{c} \right) , \tag{21}$$

$$\mathbf{j}^{(q)} = \frac{d\,\mathbf{r}^{(q)}}{d\,\widetilde{G}^{(q)}} = -2\,\widetilde{G}^{(q)}\,\mathbf{a} - \mathbf{b} \,, \tag{22}$$

$$\widetilde{G}^{(q+1)} = \widetilde{G}^{(q)} - \frac{(\mathbf{j}^{(q)})^H \mathbf{r}^{(q)}}{\|\mathbf{j}^{(q)}\|^2} \tag{23}$$

and initializing $0 < G^{(1)} < 1$. Notice that, since the residual $\mathbf{r}^{(q)}$ in (21) and the Jacobian $\mathbf{j}^{(q)}$ in (22) are vectors, the update in (23) can be implemented efficiently. Finally, an estimate $\widehat{G}$ of the signal enhancement postfilter is generated by substituting the final Gauss-Newton solution $\widetilde{G}^{(Q)}$ into (15), i.e.,

$$\widehat{G} = 1 - \left(\widetilde{G}^{(Q)} - 1 \right)^2 . \tag{24}$$

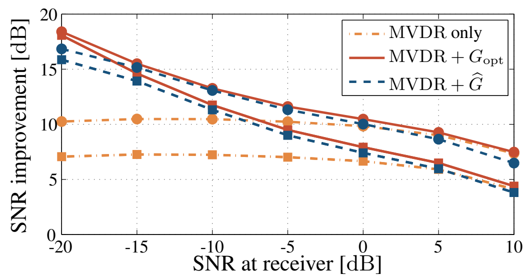
5. EXPERIMENTAL RESULTS

For experimental validation of the proposed approach we generate multiple received signals in time domain using transmission channels with 256 exponentially decaying random numbers to model, e.g., room impulse responses, and white noise as a single source signal. Uncorrelated observation noise is added to obtain a defined signal-to-noise ratio (SNR) at the receivers. The BCI channel estimates are simulated by applying a common convolutive error to the channels and these estimates are used for calculation of the MVDR.

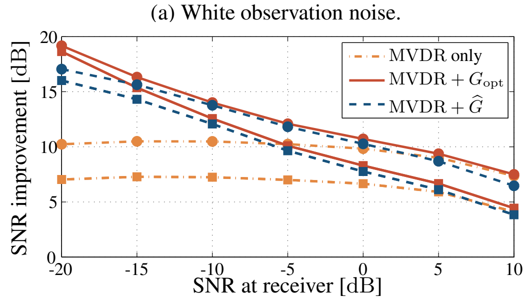
All processing is performed in overlap-save-based block-frequency domain using a frame length of 512 and a frame advance of 256 [10]. For postfilter adaptation, the number of Gauss-Newton iterations is set to $Q = 5$ and $\widehat{\Phi}_Z$ is calculated over 10 signal frames. In all experiments, we conduct 50 Monte-Carlo runs and present the averaged results illustrating the performance of the adaptive postfilter $\widehat{G}$ in (24) and the theoretically optimal postfilter $G_{\rm opt}$ in (4) as a reference, which is calculated using the true source and noise covariance information.

Our first experiment investigates the deviation of $\widehat{G}$ from $G_{\rm opt}$ as a function of the number of signal frames used for $\widehat{\underline{\mathbf{\Phi}}}_Y$ calculation. For $P = 5$ and a receiver SNR of $-10\,$dB, Fig. 4 depicts the SNR improvement, i.e., the difference between output SNR and receiver SNR in decibel.[2] When no postfilter is applied, the MVDR alone achieves a plausible SNR improvement of roughly $(10\,\log_{10} P)\,$dB. The upper bound is given by the fully informed postfilter $G_{\rm opt}$. It is approached by the adaptive postfilter $\widehat{G}$ and already a small number of signal frames (data points) seems to be sufficient for good convergence of $\widehat{G}$ to $G_{\rm opt}$.

[2] Receiver: $\mathrm{SNR_{rec}} = \Phi_{\mathbf{D}}/\Phi_{\mathbf{N}}$; output: $\mathrm{SNR_{out}} = \Phi_{\mathbf{D}}/\Phi_{(\mathbf{D}-\widehat{\mathbf{D}})}$.



(a) White observation noise.



(b) Colored observation noise.

Fig. 5. SNR improvement vs. receiver SNR for $P = 5\,(\blacksquare)$ and $P = 10\,(\bullet)$ channels.

In a second experiment, we investigate the signal enhancement of the proposed postfilter $\widehat{G}$ over a wide range of receiver SNRs with $\widehat{\underline{\mathbf{\Phi}}}_Y$ averaged over 30 signal frames. Fig. 5(a) illustrates the SNR improvement for $P = 5$ and $P = 10$ when white observation noise is considered while the results of the same experiment but with colored low-pass observation noise are depicted in Fig. 5(b). We observe that, while the spatial MVDR filter alone is providing a fixed SNR improvement, the postfilter significantly improves the SNR for negative receiver SNRs. There is a strong agreement in the signal enhancement properties of the informed and adapted postfilters, and comparing Figs. 5(a) and (b) reveals that the adaptation works similarly well for different noise types. For positive receiver SNR, the improvement is reduced for all approaches due to the imperfection of broadband equalization via narrowband MVDR [14]. In mid-SNR range, the adaptive algorithm closely touches the informed solution.

6. CONCLUSIONS AND RELATION TO PRIOR WORK

In this contribution we have presented a novel approach to postfilter adaptation for MMSE-optimal multichannel Wiener filtering. The proposed solution is based on the relation between multichannel decorrelation and optimal filtering that is exploited in the proposed algorithm in a numerically robust and efficient way.

Previous approaches [15, 16, 17, 12, 13] for postfilter adaptation are based on the idea of noise covariance estimation by relying on the different degrees of stationarity in source signal and noise. The quality of the noise covariance estimates, however, is crucial [11] and the estimation itself is difficult since the noise is unobservable. In contrast, the proposed approach exploits a new criterion via the multichannel decorrelation, which can be performed on the received signals and the MVDR output alone, both of which are readily available to the algorithm. It is therefore not restricted to the assumptions of a non-stationary source in a stationary noise environment as is shown by simulations with fully stationary setups.

7. REFERENCES

[1] J. Chen, J. Benesty, Y. Huang, and S. Doclo, "New insights into the noise reduction Wiener filter," *IEEE Trans. Audio, Speech, Language Process.*, vol. 14, no. 4, pp. 1218–1234, July 2006.

[2] Y. Huang, J. Benesty, and J. Chen, "Analysis and comparison of multichannel noise reduction methods in a common framework," *IEEE Trans. Audio, Speech, Language Process.*, vol. 16, no. 5, pp. 957–967, July 2008.

[3] R. C. Hendriks, R. Heusdens, U. Kjems, and J. Jensen, "On optimal multichannel mean-squared error estimators for speech enhancement," *IEEE Signal Process. Lett.*, vol. 16, no. 10, pp. 885–888, Oct. 2009.

[4] K. Abed-Meraim, W. Qiu, and Y. Hua, "Blind system identification," *Proc. IEEE*, vol. 85, no. 8, pp. 1310–1322, Aug. 1997.

[5] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Process.*, vol. 43, no. 2, pp. 516–525, Feb. 1995.

[6] G. Xu, H. Liu, L. Tong, and T. Kailath, "A least-squares approach to blind channel identification," *IEEE Trans. Signal Process.*, vol. 43, no. 12, pp. 2982–2993, Dec. 1995.

[7] B. Yang, "Projection approximation subspace tracking," *IEEE Trans. Signal Process.*, vol. 43, no. 1, pp. 95–107, Jan. 1995.

[8] Y. Huang and J. Benesty, "A class of frequency-domain adaptive approaches to blind multichannel identification," *IEEE Trans. Signal Process.*, vol. 51, no. 1, pp. 11–24, Jan. 2003.

[9] D. Schmid, G. Enzner, S. Malik, D. Kolossa, and R. Martin, "Variational Bayesian inference for multichannel dereverberation and noise reduction," *IEEE Trans. Audio, Speech, Language Process.*, vol. 22, no. 8, pp. 1320–1335, Aug. 2014.

[10] P. Vary and R. Martin, *Digital Speech Transmission: Enhancement, Coding and Error Concealment*, John Wiley & Sons, 2006.

[11] B. Cornelis, M. Moonen, and J. Wouters, "Performance analysis of multichannel Wiener filter-based noise reduction in hearing aids under second order statistics estimation errors," *IEEE Trans. Audio, Speech, Language Process.*, vol. 19, no. 5, pp. 1368–1381, July 2011.

[12] K. U. Simmer, J. Bitzer, and C. Marro, "Post-filtering techniques," in *Microphone Arrays*, M. Brandstein and D. Ward, Eds., chapter 3, pp. 39–60. Springer, 2001.

[13] S. Gannot and I. Cohen, "Adaptive beamforming and postfiltering," in *Springer Handbook of Speech Processing*, J. Benesty, M. M. Sondhi, and Y. Huang, Eds., chapter 47, pp. 945–978. Springer, 2008.

[14] E. A. P. Habets, J. Benesty, I. Cohen, S. Gannot, and J. Dmochowski, "New insights into the MVDR beamformer in room acoustics," *IEEE Trans. Audio, Speech, Language Process.*, vol. 18, no. 1, pp. 158–170, Jan. 2010.

[15] 3GPP TS 126.094 - V4.0.0, "Universal mobile telecommunications system (UMTS); mandatory speech codec speech processing functions AMR speech codec; voice activity detector (VAD)," 2001.

[16] R. Martin, "Noise power spectral density estimation based on optimal smoothing and minimum statistics," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 5, pp. 504–512, July 2001.

[17] I. Cohen, "Noise spectrum estimation in adverse environments: Improved minima controlled recursive averaging," *IEEE Trans. Speech Audio Process.*, vol. 11, no. 5, pp. 466–475, Sept. 2003.

[18] R. C. Hendriks and T. Gerkmann, "Noise correlation matrix estimation for multi-microphone speech enhancement," *IEEE Trans. Audio, Speech, Language Process.*, vol. 1, no. 1, pp. 223–233, Jan. 2012.

[19] R. Zelinski, "A microphone array with adaptive post-filtering for noise reduction in reverberant rooms," in *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, New York, NY, USA, Apr. 1988, vol. 5, pp. 2578–2581.

[20] M. Dörbecker and S. Ernst, "Combination of two-channel spectral subtraction and adaptive Wiener post-filtering for noise reduction and dereverberation," in *European Signal Processing Conf.*, Trieste, Italy, Sept. 1996, pp. 995–998.

[21] I. A. McCowan and H. Bourlard, "Microphone array post-filter based on noise field coherence," *IEEE Trans. Speech Audio Process.*, vol. 11, no. 6, pp. 709–716, Nov. 2003.

[22] T. J. Klasen, S. Doclo, T. van den Bogaert, M. Moonen, and J. Wouters, "Binaural multich-channel Wiener filtering for hearing aids: Preserving interaural time and level differences," in *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Toulouse, France, May 2006.

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

[24] D. Marquardt, V. Hohmann, and S. Doclo, "Binaural cue preservation for hearing aids using multi-channel Wiener filter with instantaneous ITF preservation," in *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Kyoto, Japan, Mar. 2012, pp. 21–24.

[25] Y. Huang, J. Benesty, and J. Chen, "Adaptive blind multichannel identification," in *Springer Handbook of Speech Processing*, J. Benesty, M. M. Sondhi, and Y. Huang, Eds., chapter 13, pp. 259–282. Springer, 2008.

[26] D. Schmid and G. Enzner, "Cross-relation-based blind SIMO identifiability in the presence of near-common zeros and noise," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 60–72, Jan. 2012.

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

[28] H. Buchner, R. Aichner, and W. Kellermann, "TRINICON-based blind system identification with application to multiple-source localization and separation," in *Blind Speech Separation*, S. Makino, T.-W. Lee, and H. Sawada, Eds., chapter 4, pp. 101–147. Springer, 2007.

[29] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, John Wiley & Sons, 2007.

[30] L. L. Scharf, *Statistical Signal Processing*, Addison-Wesley, 1991.

[31] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, 2002.

[32] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.