IMPROVED CHANGE PREDICTION FOR COMBINED BEAMFORMING AND ECHO CANCELLATION WITH APPLICATION TO A GENERALIZED SIDELOBE CANCELER

Stefan Kühl,[1] *Alexander Bohlender,*[2] *Matthias Schrammen,*[1] *Peter Jax*[1]

[1] RWTH Aachen University, Institute of Communication Systems, Aachen, Germany
{kuehl, schrammen, jax}@iks.rwth-aachen.de
[2] Ghent University - imec, IDLab, Dept. of Electronics and Information Systems, Ghent, Belgium
alexander.bohlender@ugent.be

ABSTRACT

Adaptive beamforming and echo cancellation are often necessary in hands-free situations in order to enhance the communication quality. Unfortunately, the combination of both algorithms leads to problems. Performing echo cancellation before the beamformer (AEC-first) leads to a high complexity. In the other case (BF-first) the echo reduction is drastically decreased due to the changes of the beamformer, which have to be tracked by the echo canceler. Recently, the authors presented the directed change prediction algorithm with directed recovery, which predicts the effective impulse response after the next beamformer change and therefore allows to maintain the low complexity of the BF-first structure and to guarantee a robust echo cancellation. However, the algorithm assumes an only slowly changing acoustical environment which can be problematic in typical time-variant scenarios. In this paper an improved change prediction is presented, which uses adaptive shadow filters to reduce the convergence time of the change prediction. For this enhanced algorithm, it is shown how it can be applied to more advanced beamformer structures like the generalized sidelobe canceler and how the information provided by the improved change prediction can also be used to enhance the performance of the overall interference cancellation.

Index Terms— Change Prediction, ChaP, Beamforming, Acoustic Echo Cancellation

1. INTRODUCTION

In hands-free speech communication scenarios an *acoustic echo canceler* (AEC) is essential to remove undesired echo components. Nowadays, many communication systems like mobile phones or teleconferencing systems often have multiple microphones. A separate AEC for each single microphone is principally required, if the AEC is the first element in the signal processing chain (AEC-first). Often adaptive beamforming techniques like a *generalized sidelobe canceler* (GSC) [1–3] are used to enhance the desired speech component and suppress undesired interference components. After the *beamformer* (BF) only one single-channel enhanced signal is present. Therefore, it seems favorable to shift the AEC behind the BF to reduce the computational complexity and potentially benefit from an already enhanced microphone signal, resulting in a so-called BF-first system. Unfortunately, if the beamformer is adaptive and changes its steering direction over time, the AEC has to track these changes in addition to changes in the acoustical environment. Without any interaction between the BF and the AEC, this impairs the adaptation performance of the AEC.

In the past, several approaches to counteract this problem have been proposed. Two approaches are to jointly adapt the parameters in the AEC and BF [4, 5] or to integrate the AEC in the BF, e.g., in the *reference* or *interference canceler path* of a GSC [6, 7]. In an AEC-first system the echo paths for the microphones can be very similar, which can also be exploited to reduce the complexity [8, 9]. When computational complexity is not the main constraint, approaches exist which try to further enhance the AEC performance [10, 11].

Recently, in [12] the *change prediction* (ChaP) method has been proposed, which utilizes the knowledge of the switching time of the beamformer and AEC estimates for different steering directions to predict a new a priori estimate for the AEC after the next BF change. Furthermore, it was shown how reliability information from an AEC, e.g., its state of convergence, can be taken into account leading to the *directed change prediction with directed recovery* algorithm (dChaP/DR).

In this paper, we show how dChaP/DR can be extended to guarantee a faster convergence over time. In addition different approaches for the integration of dChaP/DR into the frequently used GSC structure are shown. Therefore, after the signal model, in Sec. 2 the basic GSC principle is introduced and it is shown at which position the AEC can be integrated in the GSC. The integration of dChaP/DR into the *reference path* of the GSC structure, including a brief repetition of the dChaP/DR concept, is introduced in Sec. 3. Afterwards, in Sec. 4 a new shadow filtering technique is introduced and analyzed in terms of trade-off between complexity and performance. Additionally, the knowledge about the estimated echo paths can be exploited for an additional AEC in the *interference canceler* (IC) path of the GSC as well as described in Sec. 5. Afterwards the presented algorithms are evaluated in Sec. 6. Finally, the paper is concluded in Sec. 7.

2. GSC IN BF-FIRST STRUCTURE

In Fig. 1 the block diagram of a BF-first system, consisting of a GSC followed by an AEC is shown. The microphone signals are denoted by

$$y_n(i) = d_n(i) + s_n(i) + v_n(i), \quad n = 1, \ldots, N_{\mathrm{mic}} \tag{1}$$

with the microphone related echo signal $d_n(i)$, desired speech signal $s_n(i)$, interference signal $v_n(i)$ and discrete sample index i. N_{mic} denotes the number of microphones. The acoustic transmission paths from the loudspeaker signal $x(i)$, from the desired signal $s(i)$, and from a directive interference signal $v(i)$ to the microphones are modeled by device related room impulse responses with coefficient vectors $\mathbf{h}_n(i)$, $\mathbf{g}_n(i)$, and $\mathbf{f}_n(i)$ of length N_h, N_g, and N_f, respectively. For the GSC, the microphone signals $y_n(i)$ are first delayed according to the current steering direction of the BF in order to get aligned intermediate signals $\tilde{y}_n(i)$. The delay is realized by fractional delay filters [13] with coefficient vectors $\mathbf{w}_{\tau,n}(i)$ of length

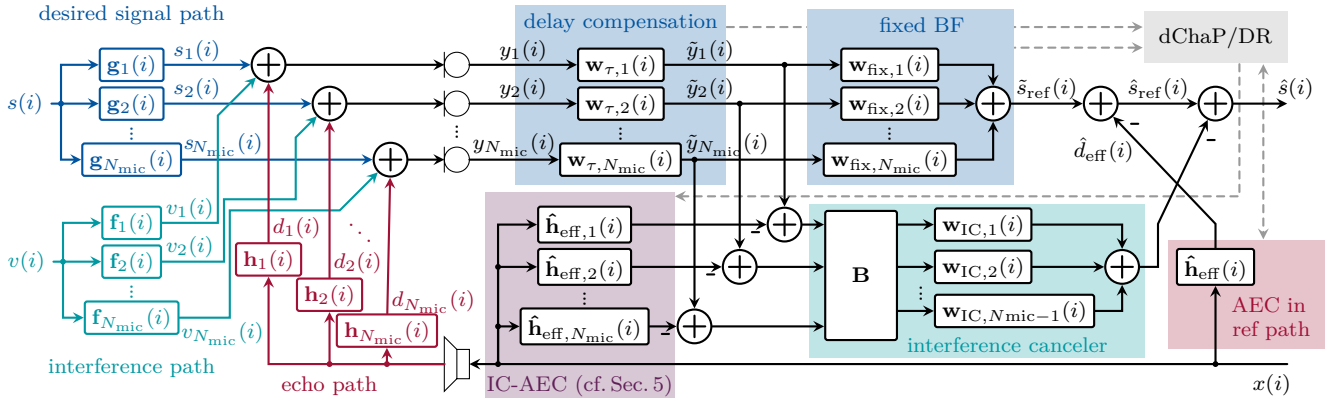

Figure 1: Block diagram of combined GSC and AEC system in BF-first configuration. For IC-AEC see Sec. 5.

N_τ. Afterwards, the signals $\tilde{y}_n(i)$ are filtered in the reference path of the GSC with fixed beamformer filters with coefficient vectors $\mathbf{w}_{\text{fix},n}(i)$ of length N_{fix} and summed up to the signal $\tilde{s}_{\text{ref}}(i)$. The combination of $\mathbf{w}_{\tau,n}(i)$ and $\mathbf{w}_{\text{fix},n}(i)$ will be termed BF weights in the following. In the parallel IC path of the GSC, the $\tilde{y}_n(i)$ are processed by a blocking matrix $\mathbf{B}$ which cancels the desired signal. As an example, when the $\mathbf{g}_n(i)$ consist of a delay only, the blocking matrix

$$\mathbf{B} = \begin{pmatrix} 1 & -1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & -1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 1 & -1 \end{pmatrix} \tag{2}$$

is optimal. There exist many approaches how to choose or adapt $\mathbf{B}$ to yield a better signal cancellation in a reverberant environment [14, 15]. Nevertheless, since the focus of this paper lies not on an optimal BF, $\mathbf{B}$ defined in (2) will be used throughout this paper.

At the output of $\mathbf{B}$, only $N_{\text{mic}} - 1$ signals are left containing only the remaining interference signals including echo components. These are filtered by IC filters with coefficient vectors $\mathbf{w}_{\text{IC},m}(i)$ of length N_{IC} with $m = 1, \ldots, N_{\text{mic}} - 1$, summed up and subtracted from the reference path output $\hat{s}_{\text{ref}}(i)$.

After the fixed BF of the GSC, an AEC is employed in order to cancel the echo. Therefore, the coefficient vector $\hat{\mathbf{h}}_{\text{eff}}(i)$ of an adaptive filter is estimated from the far-end signal $x(i)$ and $\hat{s}_{\text{ref}}(i)$. With this, an estimate of the echo signal after the reference path $\hat{d}_{\text{eff}}(i)$ is generated, which is subtracted from the fixed BF output signal $\tilde{s}_{\text{ref}}(i)$. $\mathbf{h}_{\text{eff}}(i)$ describes the effective impulse response coefficient vector of length $N_{\text{eff}} = N_h + N_\tau + N_{\text{fix}} - 2$ resulting from the combination of the acoustic transfer functions $\mathbf{h}_n(i)$ and the beamformer weights in the reference path, thus resulting in

$$\mathbf{h}_{\text{eff}}(i) = \sum_{n=1}^{N_{\text{mic}}} \mathbf{w}_{\text{fix},n}(i) * \mathbf{w}_{\tau,n}(i) * \mathbf{h}_n(i), \tag{3}$$

with $*$ denoting the convolution operation. The AEC is based on $\hat{s}_{\text{ref}}(i)$ and not on the total output signal $\hat{s}(i)$ so that the possibly highly time variant $\mathbf{w}_{\text{IC},n}(i)$ do not have an impact on the AEC performance. In this way, as long as the steering direction is not changed, the AEC only has to cope with fixed BF weights. $\hat{s}(i)$ is used to adapt the weights $\mathbf{w}_{\text{IC},m}(i)$ of the IC. In this paper, for the adaptation of $\mathbf{w}_{\text{IC},m}(i)$ a frequency-domain adaptive filter (FDAF) algorithm with fixed step-size is used [10]. For the AEC a state-space Kalman filter in the discrete Fourier transform (DFT) domain is used [16, 17]. The FDAF, the Kalman filter as well as the dChaP/DR algorithm work in the DFT domain. Therefore, from now on the DFT of all introduced vectors will be considered. Assuming frame-wise constant filters the DFT of (3) can be written as

$$\begin{aligned} H_{\text{eff}}(\mu,\lambda) &= \sum_{n=1}^{N_{\text{mic}}} W_{\text{fix},n}(\mu,\lambda) \cdot W_{\tau,n}(\mu,\lambda) \cdot H_n(\mu,\lambda) \\ &= \mathbf{W}^{\text{T}}(\mu,\lambda) \cdot \mathbf{H}(\mu,\lambda) \end{aligned} \tag{4}$$

with frequency bin index μ and frame index λ. The matrix $\mathbf{W} = [W_{\text{fix},1} \cdot W_{\tau,1}, W_{\text{fix},2} \cdot W_{\tau,2}, \ldots, W_{\text{fix},N_{\text{mic}}} \cdot W_{\tau,N_{\text{mic}}}]^{\text{T}}$ combines the channel-wise BF filters and likewise $\mathbf{H} = [H_1, H_2, \ldots, H_{N_{\text{mic}}}]^{\text{T}}$ the channel-wise acoustic transmission paths. The description of the state-space Kalman filter for the AEC can be found in [16, 17]. The AEC delivers besides the actual estimate of the effective transmission path $\hat{H}_{\text{eff}}(\mu,\lambda)$ also an estimation error covariance matrix containing reliability information $P_{\mu\mu}(\lambda)$ for every frequency bin of $\hat{H}_{\text{eff}}(\mu,\lambda)$. This can be used in the dChaP/DR algorithm to improve the performance as described in [12].

3. CHANGE PREDICTION

So far the AEC has to readapt whenever $\mathbf{W}$ changes, i.e., the steering direction of the beamformer changes. One approach to counteract this problem is the dChaP/DR method, which is recapped from [12] in the following. The main idea of ChaP is to collect estimates $\hat{H}_{\text{eff}}(\mu,\lambda)$ at several time instances right before the BF steering direction changes and use them to predict the new effective impulse response after a beamformer change. Thereby, we implicitly assume that $H_n(\lambda)$ is slowly changing for the last N_Δ BF changes. If λ_l, $l = 1, \ldots, N_\Delta$ denotes the frame indices right before the last N_Δ BF changes, it then holds that $\mathbf{H}(\mu,\lambda_1) = \mathbf{H}(\mu,\lambda_2) = \cdots = \mathbf{H}(\mu,\lambda_{N_\Delta})$. Combining (4) for all λ_l yields the linear equation system

$$\begin{bmatrix} H_{\text{eff}}(\mu,\lambda_1) \\ H_{\text{eff}}(\mu,\lambda_2) \\ \vdots \\ H_{\text{eff}}(\mu,\lambda_{N_\Delta}) \end{bmatrix} = \begin{bmatrix} \mathbf{W}^{\text{T}}(\mu,\lambda_1) \\ \mathbf{W}^{\text{T}}(\mu,\lambda_2) \\ \vdots \\ \mathbf{W}^{\text{T}}(\mu,\lambda_{N_\Delta}) \end{bmatrix} \cdot \mathbf{H}(\mu,\lambda_{N_\Delta}) \tag{5}$$

or written compactly

$$\mathbf{H}_{\text{eff}}(\mu,\boldsymbol{\lambda}_{N_\Delta}) = \boldsymbol{\mathcal{W}}(\mu,\boldsymbol{\lambda}_{N_\Delta}) \cdot \mathbf{H}(\mu,\lambda_{N_\Delta}) \tag{6}$$

with $\boldsymbol{\lambda}_{N_\Delta}$ denoting the vector containing all λ_l. Replacing in (6) the true $H_{\text{eff}}(\mu,\lambda)$ by the estimates of the AEC $\hat{H}_{\text{eff}}(\mu,\lambda)$ yields

$$\hat{\mathbf{H}}_{\text{eff}}(\mu,\boldsymbol{\lambda}_{N_\Delta}) = \boldsymbol{\mathcal{W}}(\mu,\boldsymbol{\lambda}_{N_\Delta}) \cdot \hat{\mathbf{H}}(\mu,\lambda_{N_\Delta}), \tag{7}$$

which can be solved by

$$\hat{\mathbf{H}}(\mu, \lambda_{N_\Delta}) = \boldsymbol{\mathcal{W}}^{\dagger}(\mu, \boldsymbol{\lambda}_{N_\Delta}) \cdot \hat{\mathbf{H}}_{\text{eff}}(\mu, \boldsymbol{\lambda}_{N_\Delta}) \tag{8}$$

with $(\cdot)^{\dagger}$ denoting a *regularized Moore-Penrose pseudoinverse*. For regularization the *effective Rank* $\mathcal{R}_{\text{eff}}$ of $\boldsymbol{\mathcal{W}}$ is calculated as defined in [18] and the pseudoinversion is calculated by a truncated singular value decomposition using only the largest $\lfloor \mathcal{R}_{\text{eff}}\,(\boldsymbol{\mathcal{W}}(\mu, \boldsymbol{\lambda}_{N_\Delta}))\rceil$ singular values with $\lfloor\cdot\rceil$ denoting the nearest integer. Due to this regularization the pseudoinversion is more robust against a poorly conditioned least squares problem.

Inserting (8) in (4) yields the a-priori estimate

$$\hat{H}^{-}_{\text{eff}}(\mu,\lambda_{N_\Delta}{+}1){=}\mathbf{W}^{\mathrm{T}}(\mu,\lambda_{N_\Delta}{+}1){\cdot}\boldsymbol{\mathcal{W}}^{\dagger}(\mu,\boldsymbol{\lambda}_{N_\Delta}){\cdot}\hat{\mathbf{H}}_{\text{eff}}(\mu,\boldsymbol{\lambda}_{N_\Delta}) \tag{9}$$

which is used for the AEC right after the BF change. This method is called the *change prediction* algorithm (ChaP).

As the Kalman filter based AEC tracks the reliability of its estimate, the reliability of the estimate from (9) can be calculated and additionally handed over to the AEC for a *directed recovery* (DR). Therefore, whenever the steering direction of the BF is changed it is calculated how similar the new steering direction is to already observed ones by

$$\Delta\mathcal{R}_{\text{eff}}(\mu,\boldsymbol{\lambda}_{N_\Delta}){=}\mathcal{R}_{\text{eff}}(\boldsymbol{\mathcal{W}}(\mu,\boldsymbol{\lambda}_{N_\Delta+1})){-}\mathcal{R}_{\text{eff}}(\boldsymbol{\mathcal{W}}(\mu,\boldsymbol{\lambda}_{N_\Delta})) \tag{10}$$

with $\boldsymbol{\mathcal{W}}(\mu, \boldsymbol{\lambda}_{N_\Delta+1})$ denoting the matrix containing the already observed steering directions and the new one. A completely unrelated new observation results in $\Delta\mathcal{R}_{\text{eff}} = 1$. With the *effective rank difference* $\text{ERD} = \max\{1 - \Delta\mathcal{R}_{\text{eff}}, 0\}$, the entries of the estimation error covariance matrix in the Kalman filter of the AEC can be empirically set to

$$P^{-}_{\mu\mu}(\lambda_{N_\Delta}{+}1){=}1{-}\Big(1{-}\min_{l}\{P_{\mu\mu}(\lambda_l)\}\Big){\cdot}\text{ERD}(\mu,\lambda_{N_\Delta}) \tag{11}$$

after each BF change.

Furthermore, the reliability from the AEC can also be used in (9) for a *directed* ChaP. Therefore, for the last N_Δ observations the corresponding reliability information is stored in a diagonal matrix $\boldsymbol{\Psi}(\lambda_{N_\Delta})$ with the diagonal elements

$$\Psi_l(\mu) = 1 - \sqrt{P_{\mu\mu}(\lambda_l)}. \tag{12}$$

$\boldsymbol{\Psi}(\lambda_{N_\Delta})$ is then used in (9) as a weighting matrix yielding

$$\begin{aligned}\hat{H}^{-}_{\text{eff}}(\mu,\lambda_{N_\Delta}{+}1){=}\mathbf{W}^{\mathrm{T}}(\mu,\lambda_{N_\Delta}{+}1){\cdot}(\boldsymbol{\Psi}(\lambda_{N_\Delta})\boldsymbol{\mathcal{W}}(\mu,\boldsymbol{\lambda}_{N_\Delta}))^{\dagger}\\ {\cdot}\boldsymbol{\Psi}(\lambda_{N_\Delta})\hat{\mathbf{H}}_{\text{eff}}(\mu,\boldsymbol{\lambda}_{N_\Delta}).\end{aligned} \tag{13}$$

Using both (11) and (13) for the communication between AEC and BF is denoted as *directed change prediction with directed recovery* (dChaP/DR) algorithm.

4. SHADOW FILTER

The matrix $\boldsymbol{\mathcal{W}}(\mu, \boldsymbol{\lambda}_{N_\Delta})$ contains the weights for the μ-th frequency bin for the N_{mic} BF filters for the last N_Δ steering directions. When $N_\Delta < N_{\text{mic}}$, the matrix $\boldsymbol{\mathcal{W}}(\mu, \boldsymbol{\lambda}_{N_\Delta})$ has not full column rank and therefore it is not guaranteed that the pseudoinversion in (13) will result in a meaningful estimate. Therefore, in order to guarantee a meaningful estimate by (13), $\mathcal{R}_{\text{eff}}\,(\boldsymbol{\mathcal{W}}(\mu, \boldsymbol{\lambda}_{N_\Delta}))$ should approximately be equal to N_{mic}, i.e., at least N_{mic} different steering directions have to be observed by the dChaP/DR algorithm. As in modern devices the microphone arrays can consist of many microphones, it can take some time to observe N_{mic} different steering directions. If during this time the acoustical transmission paths $\mathbf{h}_n(i)$ change a lot, the assumption of approximately equal $H_n(\lambda)$ is violated and in the worst case the dChaP/DR algorithm will not deliver meaningful estimates at all.

One approach to circumvent this problem is to spend some computational complexity on the usage of N_{sh} shadow BF and AEC systems with the purpose to collect additional estimates $\hat{H}_{\text{eff}}(\mu, \lambda)$ for different steering directions at the same time. The steering directions of the shadow systems can be changed systematically after some switching time, which has to be long enough to allow the AEC to converge. In this way, the necessary time for collecting enough observations can be drastically reduced.

The complexity of the dChaP/DR algorithm has already been presented in [12]. The increase in complexity introduced by the shadow filters is the complexity of an additional BF and AEC system for each shadow filter. This is because the dChaP/DR algorithm is still only performed whenever the steering direction of the main BF changes.

5. AEC IN INTERFERENCE CANCELING PATH

The echo signal will not be canceled by the blocking matrix since it is regarded as interference to the desired signal. Therefore, the IC filters $\mathbf{w}_{\text{IC},m}$ will use some of its degrees of freedom to cancel the echo component in the IC path. Thus, removing the echo before the blocking matrix would allow the IC to better attenuate the other interference components.

So far the dChaP/DR algorithm only has an impact on the AEC performance in the reference path and not directly on the IC adaptation. As in (8), estimates of $H_n(\mu, \lambda_{N_\Delta})$ are available as intermediate result in the dChaP/DR algorithm, these can be combined with the delay filter weights $W_{\tau,n}(\mu, \lambda_{N_\Delta})$ to get in the IC path effective weights

$$\hat{H}_{\text{eff},n}(\mu, \lambda_{N_\Delta}) = \hat{H}_n(\mu, \lambda_{N_\Delta}) \cdot W_{\tau,n}(\mu, \lambda_{N_\Delta}). \tag{14}$$

These are used to subtract the predicted echo signal in the IC path as shown in Fig. 1. In the following this will be termed AEC *in interference canceler path* (IC-AEC). This enables the IC adaptation to use more of its degrees of freedom for the cancellation of the other interference components. Principally, (14) could be calculated in every frame to obtain the most recent estimate. However, this would increase the computational complexity dramatically as for the dChaP/DR a matrix inversion would have to be performed in every frame. Nevertheless, recalculating (14) only whenever the BF steering direction changes already improves the performance as shown in Sec. 6.2. The additional increase in complexity is mainly the calculation of the filtering with the $\hat{H}_{\text{eff},n}(\mu, \lambda_{N_\Delta})$ as (8) has to be calculated anyway for the dChaP/DR algorithm.

The IC-AEC requires a meaningful estimate of the individual transfer functions $\hat{H}_n(\mu, \lambda_{N_\Delta})$. To guarantee this, more observations than for the single-channel estimate in (13) have to be available.

For wavelengths larger than the microphone array aperture the impulse responses for each channel are very similar. Thus, the inversion in (14) is poorly conditioned for low frequencies. This results in large estimation errors for the single channels. For (13) the individual estimation errors per channel are not very severe since at these frequencies effectively an averaging over all channels is performed. Nevertheless, in (14) the estimates for the individual channels are of interest, so this problem has to be taken into account. One effective method is to average over all channels up to some frequency bin μ_c and use the same information in every channel. As at this frequencies the impulse responses are very similar anyway, this can well be justified. Thus,

$$\hat{H}^{\text{avg}}_{\text{eff},n}(\mu,\lambda_{N_\Delta}){=}\begin{cases}\frac{1}{N_{\text{mic}}}\sum\limits_{n=1}^{N_{\text{mic}}}\hat{H}_{\text{eff},n}(\mu,\lambda_{N_\Delta}) & \text{for } \mu{\leq}\mu_c\\ \hat{H}_{\text{eff},n}(\mu,\lambda_{N_\Delta}) & \text{otherwise}\end{cases} \tag{15}$$

will be used instead of (14) in the following.

6. EVALUATION

In order to demonstrate the principle mode of action of the improved dChaP/DR algorithm, a simulation of the system in Fig. 1 is carried out at a sampling frequency of $f_s = 16\,\mathrm{kHz}$ in a controlled environment. A circular microphone array consisting of 13 microphones placed on top of a rigid cylinder is considered. 12 microphones are equally distributed on a circle with radius $4.5\,\mathrm{cm}$ and one is placed at the origin of the circle. The room impulse responses are generated with a room impulse response generator using the image source method combined with real measured device-related impulse responses for the shadowing effects of the device. The dimensions of the room are $5.5\,\mathrm{m} \times 4.5\,\mathrm{m} \times 2.7\,\mathrm{m}$ and the array is slightly offset from the center. The reverberation time was set to $0.6\,\mathrm{s}$. The frame length and shift parameters of the Kalman filter based AEC are set to 1024 and 256, respectively, resulting in $\hat{N}_{\mathrm{eff}} = 769$. The forgetting factor in the state-space Markov model of the Kalman algorithm is chosen to $A = 0.998$. The delay filters $\mathbf{w}_{\tau,n}(i)$ are fractional delay filters adjusted to the current steering direction with an additional delay of 32 samples. Together with the fixed BF filters $\mathbf{w}_{\mathrm{fix},n}(i)$, which are constant delays of 32 taps, the reference path realizes a delay-and-sum BF which is causal for all steering directions. For the IC path the blocking matrix defined in (2) is used and the adaptation is performed by an unconstrained FDAF with fixed stepsize $\mu = 0.1$ and length $N_{\mathrm{IC}} = 1024$. To reduce the impact of undermodeling effects by an insufficient adaptive filter length of the AEC, the simulated acoustic transmission paths are truncated to $N_h = 704$ taps. The microphone signals $y_n(i)$ are then calculated by the convolution of the source signals with the respective impulse responses. For the sake of simplicity all considered sources are at an elevation angle of 0° at a distance of $0.6\,\mathrm{m}$ from the array. The far-end source, responsible for the echo signal, is located at a constant azimuth angle $\varphi_x = -20^\circ$. The position of the desired signal $\varphi_s(i)$ and a directed interference signal $\varphi_v(i)$ changes every second as shown in the bottom part of Fig. 2. The simulation is carried out for the time of $20\,\mathrm{s}$ with white Gaussian noise signals for all source signals $x(i)$, $s(i)$, and $v(i)$. The first channel is used as the reference channel. The echo-to-signal-plus-interference ratio $\mathrm{ESIR} = 10\lg\left(\mathrm{E}\left\{d_1^2(i)\right\}/\mathrm{E}\left\{(s_1(i)+v_1(i))^2\right\}\right)$ is set to $15\,\mathrm{dB}$ and the signal-to-interference ratio for the reference channel $\mathrm{SIR} = 10\lg\left(\mathrm{E}\left\{s_1^2(i)\right\}/\mathrm{E}\left\{v_1^2(i)\right\}\right)$ to $-10\,\mathrm{dB}$. The number of considered last beamformer changes N_Δ is chosen to 25.

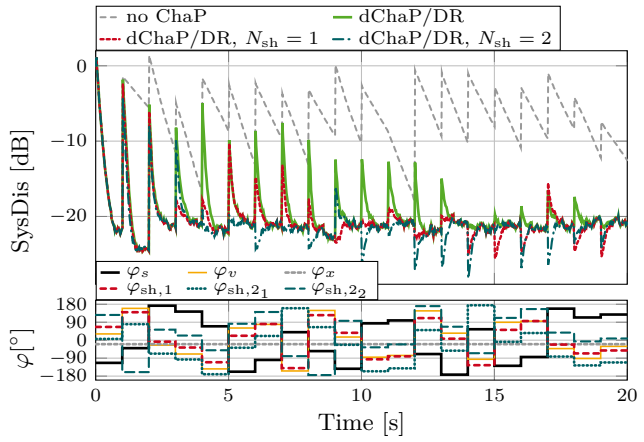

Figure 2: Top: System distance for the AEC in reference path without and with up to two shadow filters. Bottom: azimuth angle φ_{az} for source and steering directions as in Sec. 6.1.

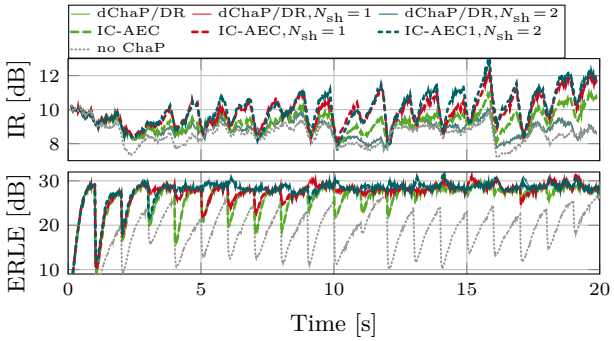

Figure 3: IR and ERLE at the output of the system without and with IC-AEC and shadow filters (cf. Sec. 6.2).

6.1. Shadow Filter

For the evaluation of the shadow filter approach, the AEC performance in the reference path is considered first. Therefore, the system distance $\mathrm{SysDis}(i) = 10\lg(\|\mathbf{h}_{\mathrm{eff}}(i) - \hat{\mathbf{h}}_{\mathrm{eff}}(i)\|^2/\|\mathbf{h}_{\mathrm{eff}}(i)\|^2)$ is used as performance metric. In the top part of Fig. 2 the system distance for the case without and with dChaP/DR and with additional shadow filters is shown. In the bottom part, the steering directions of the shadow beamformers φ_{sh} are shown. The steering direction of the main BF always matches the angles of incidence of the desired signal φ_s. It can be seen that there is a significant gain in performance when dChaP/DR is used compared to the case without. Nevertheless, without shadow filters it takes up to approximately $13\,\mathrm{s}$ until the system has observed enough different directions to yield a robust estimate for the next BF change. By employing one or more shadow filters, the time to converge can be reduced significantly, in this example to $7\,\mathrm{s}$ and $4\,\mathrm{s}$ for one and two shadow filters, respectively.

6.2. IC-AEC

For the evaluation of the IC-AEC, in Fig. 3 the interference rejection $\mathrm{IR} = 10\lg\{\mathrm{E}\left\{v_1^2(i)\right\}/\mathrm{E}\left\{v_{\mathrm{out}}^2(i)\right\}\}$ and the echo return loss enhancement $\mathrm{ERLE} = 10\lg\{\mathrm{E}\left\{d_1^2(i)\right\}/\mathrm{E}\left\{d_{\mathrm{out}}^2(i)\right\}\}$ with v_{out} and d_{out} denoting the interference and echo components at the output of the system are shown. In (15) $\mu_c = 27$ is used. From the top part of Fig. 3, it can be seen that dChaP/DR improves the interference rejection in every case. As expected, the IC-AEC leads to further improvements. Also, by the one or more shadow filters the convergence can be enhanced, so that in this example improvements up to $3\,\mathrm{dB}$ are gained. In the lower part the ERLE at the output is shown. The system distance of the AEC in the *reference path* is not changed by IC-AEC as the IC has no impact on $\hat{s}_{\mathrm{ref}}(i)$. As expected by the analysis of the system distance, the ERLE gets better for a larger number of shadow filters. Comparing the results with and without IC-AEC shows no significant difference, showing that the IC-AEC does not impair the echo reduction performance of the IC. Therefore, using IC-AEC is an attractive extension to dChaP/DR.

7. CONCLUSION

For a time-variant beamformer an acoustic echo canceler in the BF-first structure has an insufficient performance. In this paper an improved change prediction algorithm was presented which predicts the effective impulse response after a beamformer change and therefore allows the operation in the BF-first structure. It was shown how the change prediction can be improved by a shadow filter technique. Furthermore the integration of the method in a GSC has been shown both for the task of echo reduction and for the improvement of the overall interference cancellation.

8. REFERENCES

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

[2] B. R. Breed and J. Strauss, "A short proof of the equivalence of LCMV and GSC beamforming," *IEEE Signal Processing Letters*, vol. 9, no. 6, pp. 168–169, June 2002.

[3] E. Vincent, T. Virtanen, and S. Gannot, *Audio Source Separation and Speech Enhancement*, 1st ed. Wiley Publishing, 2018.

[4] W. Herbordt, W. Kellermann, and S. Nakamura, "Joint optimization of LCMV beamforming and acoustic echo cancellation," in *12th European Signal Processing Conference (EUSIPCO)*, Sept. 2004, pp. 2003–2006.

[5] M. H. Maruo, J. C. M. Bermudez, and L. S. Resende, "Statistical Analysis of a Jointly Optimized Beamformer-Assisted Acoustic Echo Canceler," *IEEE Transactions on Signal Processing*, vol. 62, no. 1, pp. 252–265, Jan. 2014.

[6] W. Herbordt and W. Kellermann, "GSAEC - Acoustic echo cancellation embedded into the generalized sidelobe canceller," in *10th European Signal Processing Conference (EUSIPCO)*, Sept. 2000, pp. 1–4.

[7] W. Herbordt, *Sound Capture for Human/Machine Interfaces*, ser. Lecture Notes in Control and Information Sciences. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, vol. 315.

[8] A. Cohen, A. Barnov, S. Markovich-Golan, and P. Kroon, "Joint Beamforming and Echo Cancellation Combining QRD Based Multichannel AEC and MVDR for Reducing Noise and Non-Linear Echo," in *26th European Signal Processing Conference (EUSIPCO)*, Sept. 2018, pp. 6–10.

[9] M. L. Valero and E. A. P. Habets, "Low-Complexity Multi-Microphone Acoustic Echo Control in the Short-Time Fourier Transform Domain," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no. 3, pp. 595–609, Mar. 2019.

[10] G. Reuven, S. Gannot, and I. Cohen, "Joint noise reduction and acoustic echo cancellation using the transfer-function generalized sidelobe canceller," *Speech Communication*, vol. 49, no. 7–8, pp. 623–635, July 2007.

[11] G. Reuven, S. Gannot, and I. Cohen, "Multichannel acoustic echo cancellation and noise reduction in reverberant environments using the transfer-function GSC," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 1, Apr. 2007, pp. I–81–I–84.

[12] M. Schrammen, A. Bohlender, S. Kühl, and P. Jax, "Change Prediction for Low Complexity Combined Beamforming and Acoustic Echo Cancellation," in *27th European Signal Processing Conference (EUSIPCO)*, A Coruña, Spain, Sept. 2019, submitted.

[13] T. I. Laakso, V. Valimaki, M. Karjalainen, and U. K. Laine, "Splitting the unit delay [FIR/all pass filters design]," *IEEE Signal Processing Magazine*, vol. 13, no. 1, pp. 30–60, 1996.

[14] O. Hoshuyama, A. Sugiyama, and A. Hirano, "A robust adaptive beamformer for microphone arrays with a blocking matrix using constrained adaptive filters," *IEEE Transactions on Signal Processing*, vol. 47, no. 10, pp. 2677–2684, Oct. 1999.

[15] M. Brandstein, D. Ward, A. Lacroix, and A. Venetsanopoulos, Eds., *Microphone Arrays*, 1st ed., ser. Digital Signal Processing. Berlin, Heidelberg: Springer Berlin Heidelberg, 2001.

[16] G. Enzner and P. Vary, "Frequency-domain adaptive Kalman filter for acoustic echo control in hands-free telephones," *Signal Processing*, vol. 86, no. 6, pp. 1140–1156, June 2006.

[17] G. Enzner, "A Model-Based Optimum Filtering Approach to Acoustic Echo Control: Theory and Practice," Dissertation, IND, RWTH Aachen, Aachen, Germany, Apr. 2006.

[18] O. Roy and M. Vetterli, "The effective rank: A measure of effective dimensionality," in *15th European Signal Processing Conference (EUSIPCO)*, 2007, pp. 606–610.