

# SOURCE-DOMAIN ADAPTIVE FILTERING FOR MIMO SYSTEMS WITH APPLICATION TO ACOUSTIC ECHO CANCELLATION

Karim Helwani, Herbert Buchner, and Sascha Spors

Deutsche Telekom Laboratories, Berlin University of Technology  
 Ernst-Reuter-Platz 7, 10587 Berlin, Germany  
 {karim.helwani, herbert.buchner, sascha.spors}@telekom.de

## ABSTRACT

Combining the channels of a multiple input/multiple output (MIMO) system into suitably chosen modes by a domain transformation offers great improvements of adaptive filtering algorithms. In this paper we present an algorithm for adaptive MIMO filtering, called source-domain adaptive filtering (SDAF), with application to multichannel acoustic echo cancellation operating in an optimally adjusted transform domain without requiring a-priori knowledge about the system. Experimental results show a significant performance improvement compared to fixed transformation bases.

**Index Terms**— multichannel acoustic echo cancellation, transform-domain, basis estimation, adaptive filtering.

## 1. INTRODUCTION

Full-duplex communication in a hands-free communication scenario with multichannel setup ( $P$  loudspeakers and  $Q$  microphones) for unrestricted audio content requires acoustic echo cancellation (AEC), where  $P \cdot Q$  echo paths (near-end  $\mathbf{H}$  in Fig. 1) have to be identified by an adaptive MIMO FIR filter of length  $L$ . The fact that

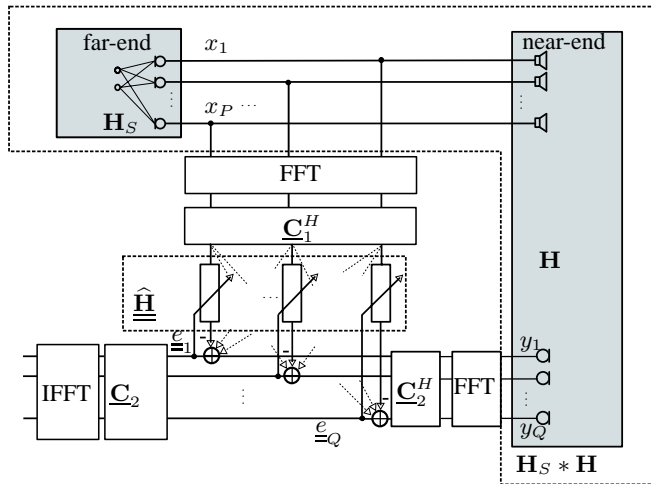


Fig. 1. Illustration of echo cancellation in transformed domain

the signals of the multichannel system are not only intrachannel correlated but typically also highly interchannel correlated results in an ill-conditioned correlation matrix in the underlying normal equation of the MIMO adaptive filter. The approach in this paper to cope with these correlation problems is based on two-stage transformations as

illustrated in Fig. 1 that allow us to perform the adaptive filtering in an adjusted spatio-temporal transform-domain.

It is known, that in the time-domain the recursive least-squares (RLS) algorithm is the optimum choice in terms of convergence speed for ill-conditioned optimization problems in adaptive filtering based on the least-squares criterion [1]. The update equation of MIMO RLS reads

$$\hat{\mathbf{H}}(n) = \hat{\mathbf{H}}(n-1) + \mathbf{R}_{\mathbf{xx}}^{-1}(n)\mathbf{x}(n)\mathbf{e}^T(n) \quad (1)$$

with

$$\mathbf{e}(n) = \mathbf{y}(n) - \hat{\mathbf{H}}^T(n-1)\mathbf{x}(n), \quad (2)$$

$$\mathbf{x}(n) = [\mathbf{x}_1^T(n), \mathbf{x}_2^T(n), \dots, \mathbf{x}_P^T(n)]^T,$$

$$\mathbf{x}_p(n) = [x_p(n), x_p(n-1), \dots, x_p(n-L+1)]^T,$$

$$\mathbf{y}(n) = [y_1(n), y_2(n), \dots, y_Q(n)]^T.$$

$\hat{\mathbf{H}}(n)$  denotes the  $PL \times Q$  MIMO coefficient matrix composed by  $P \cdot Q$  subfilters,  $\hat{\mathbf{h}}_{pq} = [\hat{h}_{pq,0}, \hat{h}_{pq,1}, \dots, \hat{h}_{pq,L-1}]^T$  and  $n$  the time instant.

$\mathbf{x}(n)$  is the length- $PL$  input signal vector (loudspeaker signals in the near-end), and  $\mathbf{e}(n)$  denotes a length- $Q$  vector of the error signals. The  $PL \times PL$  correlation matrix  $\mathbf{R}_{\mathbf{xx}}$  contains all inter- and intrachannel correlations and is usually estimated by the recursive formula with the memory factor  $\lambda$ .

$$\mathbf{R}_{\mathbf{xx}}(n) = \lambda\mathbf{R}_{\mathbf{xx}}(n-1) + (1-\lambda)\mathbf{x}(n)\mathbf{x}^H(n). \quad (3)$$

The major problems of RLS algorithms are the very high complexity, caused by inversion of large covariance matrices and potential numerical instabilities.

Performing the adaptive filtering in convenient domains could overcome the mentioned problems. A very efficient practical way to solve the problems of the multichannel RLS-type algorithms is to formulate it in the frequency domain. The resulting algorithm, Multichannel Frequency-Domain Adaptive Filtering (MC-FDAF), presented in [2] has the most important features that in addition to the efficient use of the FFT, all the sub-matrices in the input correlation matrix are approximately diagonalized by the DFT. Thus, on the one hand it is possible to efficiently take into account the intrachannel correlations by considering only the main diagonal in the Fourier domain. On the other hand the underlying acoustic channels, modeled as linear time-invariant systems, are diagonalized in the spectral domain and the filtering can be performed as simple multiplication. Advanced acoustical virtual reality systems being developed rely on massive multichannel systems, e.g. the wave field synthesis method [3], where closely spaced arrays of a large number  $P$  of individually

driven loudspeakers generate a prespecified sound field.  $P$  may lie between 20 and several hundred. An analogous approach is possible for wave field analysis (WFA) using microphone arrays with a large number  $Q$  of microphones.

Building a full-duplex system with massive multichannel setup for unrestricted audio content might be considered as the supreme discipline of multichannel acoustic echo cancellation (MCAEC) research since in this case even the  $P \times P$  frequency bin-wise correlation matrices of the loudspeaker driving signals are generally still large and ill-conditioned after the approximate block wise diagonalization within the frequency-domain adaptive filtering (FDAF) coefficient update. It has been shown that extending the point-to-point MIMO-model by more detailed spatial consideration could highly improve the performance of adaptive filtering. Recently developed wave-domain adaptive filtering WDAF [4] exploits wave physics in an attempt to decompose the MIMO-system into independent SISO-systems using well known eigenfunctions as plane-waves and spherical harmonics, but typically, depending on the present boundary conditions the system diagonalization is only approximate. A more general but only theoretical approach is the eigenspace adaptive filtering. Here the filtering is done in the eigenspace of the system [5]. These two approaches for MIMO adaptive filtering are system-based. In general AEC is a system identification problem, and the eigenspace of the system is not known. In this paper we present a new algorithm for massive multichannel acoustic echo cancellation, that does not require a-priori knowledge about the system and offers efficiency and source fidelity by a detailed mode selective optimal filtering.

## 2. ADAPTIVE FILTERING IN TRANSFORM-DOMAIN

The different strategies for adaptive filtering in transform domains share the following stages of processing: 1. temporal decoupling (typically using the DFT)[2], 2. spatial transformation into a domain, in which the MIMO system can be considered as decoupled. Most recent algorithms use a block-based formulation, which is derived by combining  $L$  consecutive samples into blocks, formulating the error signal in terms of blocks and minimizing the error. Block-based adaptation is typically computationally less complex and therefore favorable for practical application [6, 7]. This study aims at illustrating a scheme for adaptive filtering in an optimally adjusted transform-domain, that we call source-domain adaptive filtering. For clarity of the presentation we assume in the following a temporal transformation based on DFT for large block lengths ( $L \rightarrow \infty$ ) such that each frequency bin can be treated independently.

### 2.1. Domain Transformation

By an optimal domain transformation the signals are combined into modes that can be considered to be processed in separated manner by the MIMO system, such that the system is decomposed into single input/single output (SISO) systems. The transformation is given by

$$\underline{\underline{\mathbf{y}}}^{(\nu)} := \underline{\underline{\mathbf{C}}}_2^{(\nu)H} \underline{\mathbf{y}}^{(\nu)}, \quad \underline{\underline{\mathbf{x}}}^{(\nu)} := \underline{\underline{\mathbf{C}}}_1^{(\nu)H} \underline{\mathbf{x}}^{(\nu)}, \quad (4)$$

<sup>1</sup> where  $\nu$  indicates a temporal frequency bin wise quantity. In the following we omit the DFT bin index ( $\nu$ ) for brevity.

<sup>1</sup>We use in this paper underlined symbols for frequency-domain quantities and double-underlined ones for spatially transformed quantities.

### 2.2. Estimation of MIMO Filter Coefficients in Spatio-Temporal Domain

Assuming least-squares as optimization criterion for echo cancellation, the Wiener-Hopf equation gives the formula for optimal filtering, this reads in the transform-domain<sup>2</sup>,

$$\underline{\underline{\hat{\mathbf{H}}}}(n) = \underline{\underline{\mathbf{T}}}_{\mathbf{xx}}^{-1}(n) \underline{\underline{\mathbf{T}}}_{\mathbf{xy}}(n), \quad (5)$$

where

$$\underline{\underline{\mathbf{T}}}_{\mathbf{xx}} = \hat{\mathcal{E}}\{\underline{\underline{\mathbf{x}}} \underline{\underline{\mathbf{x}}}^H\} = \underline{\underline{\mathbf{C}}}_1^H \underline{\underline{\mathbf{S}}}_{\mathbf{xx}} \underline{\underline{\mathbf{C}}}_1, \quad (6)$$

$$\underline{\underline{\mathbf{T}}}_{\mathbf{xy}} = \hat{\mathcal{E}}\{\underline{\underline{\mathbf{x}}} \underline{\underline{\mathbf{y}}}^H\} = \underline{\underline{\mathbf{C}}}_1^H \underline{\underline{\mathbf{S}}}_{\mathbf{xy}} \underline{\underline{\mathbf{C}}}_2, \quad (7)$$

$$\underline{\underline{\mathbf{S}}}_{\mathbf{xx}} = \hat{\mathcal{E}}\{\underline{\underline{\mathbf{x}}} \underline{\underline{\mathbf{x}}}^H\}, \quad \underline{\underline{\mathbf{S}}}_{\mathbf{xy}} = \hat{\mathcal{E}}\{\underline{\underline{\mathbf{x}}} \underline{\underline{\mathbf{y}}}^H\}. \quad (8)$$

$\hat{\mathcal{E}}\{\cdot\}$  denotes a suitable approximation of the expectation operator. Accordingly, an optimal echo cancellation by identifying minimal number of echo paths (only the diagonal elements of  $\underline{\underline{\hat{\mathbf{H}}}}$ ) can be obtained, if a basis  $\underline{\underline{\mathbf{C}}}_1$ ,  $\underline{\underline{\mathbf{C}}}_2$  can be found, that decorrelates the loudspeaker signals as well as the microphone signals in the near-end room for multiple time instances. In this case  $\underline{\underline{\mathbf{T}}}_{\mathbf{xy}}$  and  $\underline{\underline{\mathbf{T}}}_{\mathbf{xx}}$  are diagonalized. This is equivalent to the problem of separating these signals.  $\underline{\underline{\mathbf{C}}}_1$  can be considered as separation filter with respect to  $\underline{\underline{\mathbf{H}}}_S$  (the far-end room) and  $\underline{\underline{\mathbf{C}}}_2$  can be considered as source separation filter along the system cascade  $\underline{\underline{\mathbf{H}}}_S \cdot \underline{\underline{\mathbf{H}}}$ , see Fig. 1.

Estimating a separating filter ideally exploits all information contained in the involved mixed signals. The filter estimation process should take into account all their fundamental stochastic properties, i.e. Non-Gaussianity, Nonwhiteness, and Nonstationarity.

Note that for the AEC problem the possible permutations of the separated sources after the systems  $\underline{\underline{\mathbf{C}}}_1$  and  $\underline{\underline{\mathbf{C}}}_2$  have to be aligned, which can be easily guaranteed since the separation is supervised in that sense, that the input and output signals are given.

## 3. TRANSFORM-DOMAIN ESTIMATION

The above mentioned separation problem is known in the field of blind source separation (BSS). BSS techniques aim at jointly diagonalizing the time-lagged covariance matrices of the mixed signals. As an analytical simultaneous diagonalization is given only for two matrices by the generalized eigenvalue decomposition, finding a joint diagonalization matrix must be considered as an optimization problem of a cost function on a predefined matrix norm. Most of separation algorithms in the literature take a constraint into account to restrict the set of matrices where the algorithm is searching for a solution, e.g. the unitarity of the filter matrices. An approach which solves the BSS with this constraint is [8]. Many other constraints can be found in the literature, e.g. minimum distortion principle [9] and the constraints made for the deconvolution problem.

In general the ideal transformation basis is non unitary [10] and Parseval's theorem cannot be applied unless an optimal separation solution using the unitarity constraint is chosen.

The desired real-time processing during AEC requires a method enabling us to determine the transformation basis iteratively. Hence, the criteria that the basis should fulfill can be summarized as: 1. maximal preserved variance (separation of latent variables), 2. unitarity, 3. minimal reconstruction error. These criteria lead to the

<sup>2</sup>Conventionally we use  $\mathbf{R}$  for the covariance matrix in the time domain, for its representation in the temporal frequency domain  $\underline{\underline{\mathbf{S}}}$ , and  $\underline{\underline{\mathbf{T}}}$  for the transformation of  $\underline{\underline{\mathbf{S}}}$  into the spatially transformed domain.

principal component analysis approach, but since the covariance matrices can only be estimated, the principal vectors (the source-domain basis) should be updated. Updating the basis means taking into account the new samples to find a space where the available information is embedded in an optimal way. Assuming the recursive estimation of the correlation matrices as in (3) allows us to transfer the basis update problem to an incremental rank-one modified eigenvalue problem [11].

The source-domain update can also be understood as subspace-tracking, which can be effectively achieved using the PASTd algorithms [12]. The disadvantage of this algorithm is on the one hand, that the estimated transformation matrices are only nearly unitary, it has been shown that reorthonormalizing the subspace eigenvectors results in performance degradation, and on the other hand we need in our case an immediate update of the basis when changes in the source-domain happens to overcome the non-uniqueness problem stated in [13]. This non-uniqueness problem is usually solved by manipulating the transmitted signal in the near-end room, but this manipulation should be avoided as far as possible for WFS and Ambisonics, where the loudspeakers driving functions are analytically derived. In Section 4.3 we discuss some practical aspects of the updating process.

### 3.1. Permutation Problem

As stated in Sec. 2.2 the possible permutations of the separated sources have to be aligned between the loudspeaker side and microphone side. The correctly aligned eigenvectors of both covariance matrices can be approximated by singular value decomposition of  $\underline{\mathbf{S}}_{\mathbf{x}\mathbf{y}}$ .

Let us assume the following decomposition is given by a singular value decomposition for a given temporal frequency bin, then

$$\begin{aligned}\underline{\mathbf{S}}_{\mathbf{x}\mathbf{y}} &= \underline{\mathbf{C}}_1 \underline{\mathbf{T}}_{\mathbf{x}\mathbf{y}} \underline{\mathbf{C}}_2^H, \\ \underline{\mathbf{S}}_{\mathbf{x}\mathbf{y}} \underline{\mathbf{S}}_{\mathbf{x}\mathbf{y}}^H &= \underline{\mathbf{C}}_1 \underline{\mathbf{T}}_{\mathbf{x}\mathbf{y}} \underline{\mathbf{C}}_2^H \underline{\mathbf{C}}_2 \underline{\mathbf{T}}_{\mathbf{x}\mathbf{y}} \underline{\mathbf{C}}_1^H \\ &= \underline{\mathbf{C}}_1 \underline{\mathbf{T}}_{\mathbf{x}\mathbf{y}} \underline{\mathbf{T}}_{\mathbf{x}\mathbf{y}} \underline{\mathbf{C}}_1^H, \\ \underline{\mathbf{S}}_{\mathbf{x}\mathbf{y}} \underline{\mathbf{S}}_{\mathbf{x}\mathbf{y}}^H &= \hat{\mathcal{E}}\{\underline{\mathbf{x}}\underline{\mathbf{y}}^H\} \hat{\mathcal{E}}\{\underline{\mathbf{y}}\underline{\mathbf{x}}^H\} = \sigma_{\underline{\mathbf{y}}}^2 \underline{\mathbf{S}}_{\mathbf{x}\mathbf{x}}.\end{aligned}\quad (9)$$

where  $\sigma_{\underline{\mathbf{y}}}^2 := \hat{\mathcal{E}}\{\underline{\mathbf{y}}^H \underline{\mathbf{y}}\}$ . Hence, the left singular vectors of  $\underline{\mathbf{T}}_{\mathbf{x}\mathbf{y}}$  can be seen as eigenvectors of the weighted correlation matrix  $\underline{\mathbf{S}}_{\mathbf{x}\mathbf{x}}$ . In a similar way it can be shown that  $\underline{\mathbf{C}}_2^H$  contains the eigenvectors of  $\sigma_{\underline{\mathbf{x}}}^2 \underline{\mathbf{S}}_{\mathbf{y}\mathbf{y}}$ .

## 4. EFFICIENT IMPLEMENTATION

### 4.1. Algorithm

Based on [6] together with the introduction of the transition matrices  $\underline{\mathbf{G}}_{\underline{\mathbf{C}}_1}$ ,  $\underline{\mathbf{G}}_{\underline{\mathbf{C}}_2}$  for time varying  $\underline{\mathbf{C}}_1$ ,  $\underline{\mathbf{C}}_2$  analogously to [7], we can derive an exact block based MIMO algorithm. For the simplified presentation with independent frequency bin, we obtain from (1) and (3) the following set of equations.

$$\underline{\mathbf{T}}_{\mathbf{x}\mathbf{x}}(n) = \lambda \underline{\mathbf{T}}_{\mathbf{x}\mathbf{x}}(n-1) + (1-\lambda) \underline{\mathbf{x}}(n) \underline{\mathbf{x}}^H(n), \quad (11)$$

$$\underline{\mathbf{k}}(n) = (1-\lambda) \underline{\mathbf{T}}_{\mathbf{x}\mathbf{x}}^{-1}(n) \underline{\mathbf{x}}(n), \quad (12)$$

$$\underline{\mathbf{G}}_{\underline{\mathbf{C}}_1} = \underline{\mathbf{C}}_1^H(n) \underline{\mathbf{C}}_1(n-1), \quad (13)$$

$$\underline{\mathbf{G}}_{\underline{\mathbf{C}}_2} = \underline{\mathbf{C}}_2^H(n-1) \underline{\mathbf{C}}_2(n), \quad (14)$$

$$\underline{\mathbf{e}}(n) = \underline{\mathbf{y}}(n) - \left( \underline{\mathbf{G}}_{\underline{\mathbf{C}}_1} \hat{\underline{\mathbf{H}}}(n-1) \underline{\mathbf{G}}_{\underline{\mathbf{C}}_2} \right)^T \underline{\mathbf{x}}(n), \quad (15)$$

$$\hat{\underline{\mathbf{H}}}(n) = \underline{\mathbf{G}}_{\underline{\mathbf{C}}_1} \hat{\underline{\mathbf{H}}}(n-1) \underline{\mathbf{G}}_{\underline{\mathbf{C}}_2} + \underline{\mathbf{k}}(n) \underline{\mathbf{e}}^T(n). \quad (16)$$

Eqs. (13), (14) show the computational benefit of the unitarity of the transformation matrices, since otherwise  $(\cdot)^H$  had to be replaced by matrix inversions.

### 4.2. Adaptation Control

To achieve a trade-off between complexity and adaptation performance it is necessary to ensure the quality of the chosen adaptation domain.

#### 4.2.1. Rank Estimation

To ensure an optimal embedding of the loudspeakers and microphones signals in the chosen transform-domain a measure of reconstruction error should be defined. This is a consequence of the fact that covariance matrices  $\underline{\mathbf{S}}_{\mathbf{x}\mathbf{x}}$  and  $\underline{\mathbf{S}}_{\mathbf{y}\mathbf{y}}$  are in general rank deficient and  $\underline{\mathbf{C}}_1$ ,  $\underline{\mathbf{C}}_2$  have the dimensions  $(P \times r_1, r_1 \leq P)$ , resp.  $(Q \times r_2, r_2 \leq Q)$ , therefore the reconstruction error should be computed whenever new data is available, e.g.,

$$J_1 := \left\| \underline{\mathbf{x}} - \underline{\mathbf{C}}_1 \underline{\mathbf{C}}_1^H \underline{\mathbf{x}} \right\|_2^2. \quad (17)$$

If the error is greater than the noise level then basis should be updated and the rank newly defined.

#### 4.2.2. Compactness Measure

A measure of decorrelation between the modes should be used to ensure the optimality of the adaptation, that assume diagonal covariance matrices  $\underline{\mathbf{T}}_{\mathbf{x}\mathbf{x}}$ .

An adequate and often used measure is the Frobenius norm of the off-diagonal matrix:

$$J_2 := \left\| \text{off}(\underline{\mathbf{x}}\underline{\mathbf{x}}^H) \right\|_F^2. \quad (18)$$

### 4.3. Iterative Computation of the SVD

In terms of an iterative computation of the SVD based on the QR-algorithm we propose a new method for rank-one incremental SVD-update. The literature gives many algorithms for the updating process. The theory behind the updating process is presented in [11], where the author suggested the Newton's algorithm to determine the roots of the characteristic polynomial. In [14] a sophisticated interpolation approach was introduced. These two methods could converge slowly when the new roots are close to each other. In [15] the update strategy is optimized for appending more than one Vector to a matrix.

The basic idea of the algorithm presented here is to precondition the update problem to solve it efficiently with the QR-Algorithm. From (7) and (3) we obtain

$$\begin{aligned}\underline{\mathbf{T}}'_{\mathbf{x}\mathbf{y}}(n) &:= \underline{\mathbf{C}}_1^H(n-1) (\underline{\mathbf{S}}_{\mathbf{x}\mathbf{y}}(n)) \underline{\mathbf{C}}_2(n-1) \\ &= \underline{\mathbf{C}}_1^H(n-1) \left( \lambda \underline{\mathbf{S}}_{\mathbf{x}\mathbf{y}}(n-1) \right. \\ &\quad \left. + (1-\lambda) \underline{\mathbf{x}}(n) \underline{\mathbf{y}}(n)^H \right) \underline{\mathbf{C}}_2(n-1) \\ &= \lambda \underline{\mathbf{T}}_{\mathbf{x}\mathbf{y}}(n-1) + (1-\lambda) \underline{\mathbf{x}}'(n) \underline{\mathbf{y}}'(n)^H.\end{aligned}\quad (19)$$

Givens rotations  $\mathbf{G} := \mathbf{G}_1^H \cdots \mathbf{G}_{P-1}^H$  after [16] can be found such that  $\mathbf{G}$  is in Hessenberg form and

$$\mathbf{G} \cdot \underline{\mathbf{x}}'(n) = \left\| \underline{\mathbf{x}}'(n) \right\| \mathbf{e}_1, \quad (20)$$

where  $\mathbf{e}_1 = [1, 0, \dots, 0]^T$ . Substituting into (19) leads to

$$\underline{\mathbf{T}}'_{\mathbf{x}\mathbf{y}}(n) = \mathbf{G}^H (\lambda \mathbf{G} \underline{\mathbf{T}}_{\mathbf{x}\mathbf{y}}(n-1) + (1-\lambda) \left\| \underline{\mathbf{x}}'(n) \right\| \mathbf{e}_1 \mathbf{y}'(n)^H). \quad (21)$$

$\mathbf{G} \underline{\mathbf{T}}_{\mathbf{x}\mathbf{y}}(n-1)$  is in Hessenberg form because it is a product of a diagonal matrix with a Hessenberg's matrix and  $\mathbf{G}$  is unitary. In [16, Algorithm 5.2.3] it is shown how to compute the QR factorization of an upper Hessenberg matrix in  $O(P^2)$  flops, so we have a QR factorization of  $\underline{\mathbf{T}}'_{\mathbf{x}\mathbf{y}}(n)$  and the well known QR-algorithm can be applied to get the update matrices  $\underline{\mathbf{C}}'_1(n)$  and  $\underline{\mathbf{C}}'_2(n)$ . The basis update is then given by

$$\underline{\mathbf{C}}_1(n) = \underline{\mathbf{C}}_1(n-1) \underline{\mathbf{C}}'_1(n), \quad \underline{\mathbf{C}}_2(n) = \underline{\mathbf{C}}_2(n-1) \underline{\mathbf{C}}'_2(n). \quad (22)$$

## 5. EXPERIMENTS

To illustrate the properties of the developed SDAF, a multichannel AEC application scenario will be considered.

The implementation of the algorithm was block-based, the corresponding formulation is similar to the formulation in [6, 7]. The simulated geometrical setup consists of a near-end room with size  $6 \times 6 \times 3$  meters containing a linear loudspeaker array with  $P = 8$  loudspeakers, and in parallel in a distance of 5 meters is a linear microphone array with  $Q = 8$  microphones, acoustically modeled by the image source method with an acoustic reflection factor at the walls of  $\rho = 0.9$  and filter length  $L = 1024$  at a sampling rate of  $f_s = 8$  kHz. The spacing in both arrays is 20 cm. Noise with a level of approximately -50 dB with respect to the echo was added to the microphone signals, in order to simulate microphone and other noise sources at the near-end. The far-end is a WFS system rendering 4 randomly located virtual sources of uncorrelated white noise. The filter length of the driving functions (far-end in Fig. 1) was 128. After 3.7 seconds one source in the the far-end jumped 5 meters along the axis parallel to the loudspeakers array. The lowermost curve in Fig. 2 shows the echo return loss enhancement (ERLE) for the simulated scenario using a constant transform-domain, that was estimated by averaging the covariance matrices the loudspeakers and microphones signals over the time and diagonalizing them. The dashed curve depicts the performance of the presented SDAF. The enhancement reached by the adaptation of the estimation basis can be clearly seen. The fall of the curve is caused by moving the source, but the curve rises rapidly again to reach the room SNR because of adapting the transform-domain. Note that in these simulations we did not apply any pre-processing [13]. As a reference, the bold curve is produced by eigenspace adaptive filtering, where the basis was computed by the singular value decomposition of the estimated system by the presented SDAF. This curve is theoretical, because it requires the eigenspace of the estimated system to be known. The uppermost curve shows the theoretically ideal ERLE. This is the curve that can be reached by EAF, when the eigenspace of the system is known. The simulations have proved, that the presented SDAF is an efficient technically realizable algorithm that converges to the theoretically ideal eigenspace estimation.

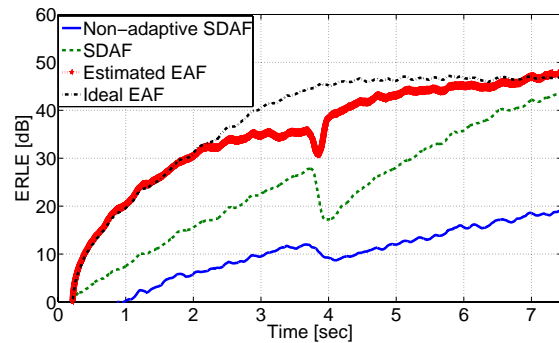


Fig. 2. Echo return loss enhancement.

## 6. CONCLUSION

In this paper we presented a new algorithm for massive multichannel acoustic echo cancellation and we outlined a scheme for an adaptive basis estimation allowing us performing optimal adaptive filtering even in absence of a-priori knowledge about the system.

## 7. REFERENCES

- [1] S. Haykin, "Adaptive filter theory," *Prentice-HALL, Inc.*, 1991.
- [2] H. Buchner, J. Benesty, and W. Kellermann, "Multichannel frequency-domain adaptive algorithms with application to acoustic echo cancellation," *Adaptive signal processing: Application to real-world problems*, Springer-Verlag, Berlin/Heidelberg, 2003.
- [3] A.J. Berkhout, D. De Vries, and P. Vogel, "Acoustic control by wave field synthesis," *The Journal of the Acoustical Society of America*, vol. 93, pp. 2764, 1993.
- [4] H. Buchner, S. Spors, and W. Kellermann, "Wave-domain adaptive filtering: Acoustic echo cancellation for full-duplex systems based on wave-field synthesis," in *Proc. IEEE ICASSP*, 2004, vol. 4.
- [5] S. Spors, H. Buchner, and R. Rabenstein, "Eigenspace adaptive filtering for efficient pre-equalization of acoustic MIMO systems," in *Proc. EUSIPCO*, 2006.
- [6] H. Buchner and S. Spors, "A general derivation of wave-domain adaptive filtering and application to acoustic echo cancellation," in *Proc. Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA*, Oct 2008.
- [7] S. Spors, H. Buchner, and K. Helwani, "Block-based multichannel transform-domain adaptive filtering," in *Proc. EUSIPCO*, 2009.
- [8] J.F. Cardoso and A. Souloumiac, "Jacobi angles for simultaneous diagonalization," *SIAM Journal on Matrix Analysis and Applications*, vol. 17, no. 1, pp. 161–164, 1996.
- [9] K. Matsuoka, "Minimal distortion principle for blind source separation," in *SICE 2002. Proceedings of the 41st SICE Annual Conference*, 2002, vol. 4.
- [10] H. Buchner, R. Aichner, and W. Kellermann, "TRINICON-based blind system identification with application to multiple-source localization and separation," in *Blind Speech Separation*. 2007, pp. 101 – 147, Springer.
- [11] G.H. Golub, "Some modified matrix eigenvalue problems," *Siam Review*, p. 318334, 1973.
- [12] B. Yang, "Projection approximation subspace tracking," *Transactions on Signal Processing, IEEE*, vol. 43, no. 1, pp. 95–107, 1995.
- [13] J. Benesty, D.R. Morgan, and M.M. Sondhi, "A better understanding and an improved solution to the specific problems of stereophonic acoustic echo cancellation," *IEEE Transactions on Speech and Audio Processing*, vol. 6, no. 2, pp. 156–165, 1998.
- [14] J.R. Bunch, Ch.P. Nielsen, and D.C. Sorensen, "Rank-one modification of the symmetric eigenproblem," *Numerische Mathematik*, vol. 31, no. 1, pp. 31–48, 1978.
- [15] M. Brand, "Incremental singular value decomposition of uncertain data with missing values," in *Proc. Computer Vision ECCV 2002*, pp. 707–720, 2002.
- [16] G.H. Golub and C.F. Van Loan, *Matrix computations*, Johns Hopkins University Press, 1996.