A General Derivation of Wave-Domain Adaptive Filtering and Application to Acoustic Echo Cancellation

Herbert Buchner and Sascha Spors

Deutsche Telekom Laboratories, Berlin University of Technology

Ernst-Reuter-Platz 7, 10587 Berlin, Germany

{Herbert.Buchner, Sascha.Spors}@telekom.de

Abstract

Wave-Domain Adaptive Filtering (WDAF) was introduced as an efficient spatio-temporal generalization of the popular Frequency-Domain Adaptive Filtering (FDAF). Through the incorporation of the mathematical foundations on wavefields, WDAF is suitable even for massive MIMO systems with highly cross-correlated broadband input signals. In this paper, we present a new rigorous derivation of WDAF leading to a whole class of powerful MIMO adaptation algorithms within a compact matrix framework. We show efficient approximations which provide both new efficient WDAF realizations and interesting links to well-known algorithms (including the various FDAF algorithms). Due to the rigorous approach, we obtain important practical design rules.

1 Introduction

Multichannel techniques for signal reproduction and acquisition offer spatial selectivity and diversity as additional degrees of freedom over single-channel systems. An important example of broadband signals in this context are speech and audio signals at the acoustic human-machine interface. Multichannel sound reproduction enhances sound realism in virtual reality and multimedia communication systems, such as teleconferencing, and aims to create a three-dimensional illusion of sound sources positioned in a virtual acoustical environment. However, advanced loudspeaker-based approaches, like the 3/2-Surround format still rely on a restrained listening area ('sweet spot'). A volume solution for a large listening space is offered by the Wave Field Synthesis (WFS) method [1] which is based on a mathematical formulation of Huygens Principle, *the Kirchhoff-Helmholtz integrals*, known from wave physics. Thus, in WFS, arrays of a large number P of individually driven loudspeakers generate a prespecified sound field. In this case, P may lie between 20 and several hundred. On the recording side of advanced acoustic human-machine interfaces, the
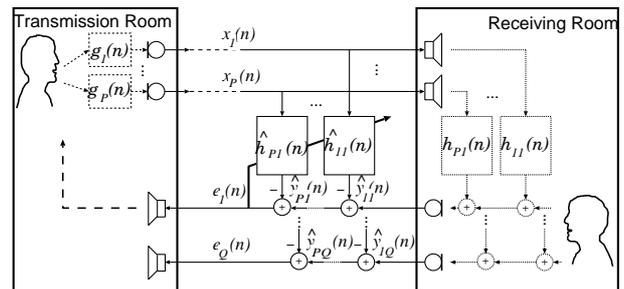

Figure 1: Multichannel acoustic echo cancellation as a prominent example for MIMO system identification.

use of microphone arrays [2], where the number Q of microphones may reach up to 1000 [3], is an effective approach to separate desired and undesired sources in the listening environment, and to cope with reverberation in the recorded signal.

A major challenge to fully exploit the potential of array processing in practical applications lies in the development of adaptive MIMO (multiple-input and multiple-output) systems that are suitable for the large number of channels in this environment. The point-to-point optimization in adaptive MIMO systems often suffers from convergence problems and high computational complexity so that some applications are beyond reach with current techniques.

An important example that we consider here is acoustic echo cancellation (AEC) [4], which is a key component in modern systems for hands-free full-duplex communication. Figure 1 shows the AEC setup for the MIMO case with P loudspeakers and Q microphones in the receiving room. In this case, $P \cdot Q$ acoustic echo paths with impulse responses $h_{pq}(n)$, $p = 1, \ldots, P$, $q = 1, \ldots, Q$ have to be continuously identified by an adaptive MIMO FIR filter with coefficients $\hat{h}_{pq,\ell}$, $\ell = 0, \ldots, L - 1$, where L denotes the filter length. The resulting echo replica are then subtracted from the received microphone signals $y_q(n)$ to

obtain the outgoing signals

$$e_q(n) = y_q(n) - \sum_{p=1}^{P} \sum_{\ell=0}^{L-1} x_p(n-\ell)\hat{h}_{pq,\ell} \qquad (1)$$

at time instant n. Unfortunately, implementing such a system for large loudspeaker arrays as mentioned above, seems to be out of reach with most of the current multichannel adaptation algorithms [5, 6]. Similar challenging adaptive MIMO filtering problems arise in other building blocks of the acoustic interface, e.g., for acoustic room compensation (ARC) on the reproduction side, where a system of suitable prefilters takes into account the actual room acoustics prior to sound reproduction by WFS, and also for adaptive interference cancellation on the recording side [2].

Common problems in such massive MIMO adaptive filtering applications are:

- high computational complexity (P and Q may be on the order of hundreds, and the filter length in each sub-channel is typically on the order of thousand)

- low adaptation performance due to the severely ill-conditioned correlation matrix in the underlying normal equation to be solved for coefficient optimization

The ill-conditioning results from the fact that the filter input signals are not only auto-correlated but are typically also highly cross-correlated [6].

For ill-conditioned optimization problems in adaptive signal processing, such as multichannel AEC, in many ways the recursive least-squares (RLS) algorithm is known to be the optimum choice in terms of convergence speed as it exhibits properties that are independent of the eigenvalue spread [7]. The update equation of the MIMO RLS algorithm reads

$$\hat{\mathbf{H}}(n) = \hat{\mathbf{H}}(n-1) + \mathbf{R}_{\mathbf{xx}}^{-1}(n)\mathbf{x}(n)\mathbf{e}^T(n), \qquad (2)$$

where $\hat{\mathbf{H}}(n)$ denotes the $PL \times Q$ MIMO coefficient matrix, $\mathbf{x}(n)$ is the length-PL input signal vector, and $\mathbf{e}(n)$ denotes a length-Q vector of the error signals according to (1). The $PL \times PL$ correlation matrix $\mathbf{R}_{\mathbf{xx}}$ takes all auto-correlations within, and - most importantly for multichannel processing - all cross-correlations between the input channels into account (see upper left corner of Fig. 2). However, the major problems of RLS algorithms are the very high computational complexity (mainly due to the large matrix inversion) and potential numerical instabilities which often limit the actual performance in practice.

An efficient and popular alternative to time-domain algorithms are *Frequency-Domain Adaptive Filtering* (FDAF) algorithms [8]. In FDAF, the adaptive filters are updated in a block-by-block fashion, using the fast Fourier transform (FFT) as a powerful vehicle. The FDAF approach has been extended to the multichannel case (MC FDAF) by a mathematically rigorous derivation based on a weighted least-squares criterion, e.g., [9]. It has been shown that there is a generic broadband frequency-domain algorithm which is equivalent to the RLS algorithm. As a result of this approach, the arithmetic complexity of multichannel algorithms can be significantly reduced compared to time-domain adaptive algorithms while the desirable RLS-like properties and the basic structure of (2) are maintained by an inherent block-diagonalization of the correlation matrix as shown in the second column of Fig. 2. This allows to perform the matrix inversion in (2) in a frequency-bin selective way using only small and better conditioned $P \times P$ matrices $\mathbf{S}_{\mathbf{xx}}^{(\nu)}$ in the bins $\nu = 0, \ldots, 2L-1$. Note that all cross-correlations are still fully taken into account by this approach.

Unfortunately, with the dramatically increased number of highly correlated loudspeaker channels in WFS-based systems, even the matrices $\mathbf{S}_{\mathbf{xx}}^{(\nu)}$ become large and ill-conditioned so that current algorithms cannot be used. Therefore, the basic idea of *Wave-Domain Adaptive Filtering* (WDAF) is to extend the conventional concept of MC FDAF by an explicit consideration of the spatial dimensions and by exploitation of wave physics as already mentioned above. This basic idea was introduced for multichannel AEC in [10] and it was shown to be effective in conjunction with actual WFS audio signals.

In the present paper we develop this concept further and give a more rigorous derivation of WDAF using a compact matrix framework similar as for the multichannel FDAF in [9]. From a physical point of view, the nice properties of FDAF result from the orthogonality property of the DFT basis functions, i.e., the complex exponentials which also separate the temporal dimension of the wave equation. WDAF is based on a suitable spatio-temporal transform domain based on orthogonal basis functions that allow not only a decomposition among the temporal frequencies as in MC FDAF, but also a (perfect or approximate) spatial decomposition as illustrated by the third column of Fig. 2. These basis functions must also fulfill the wave equation. Moreover, since the wave field is sampled only on a contour enclosing the listening area, the transformations have to take into account the so-called Kirchhoff-Helmholtz integrals so that these

transformations depend on the array geometries. In this way, various possible transformations which fulfill the above conditions can be derived, e.g., based on wave field decompositions in circular harmonics, spherical harmonics etc. As an example, circular arrays are known to perform well in wave field analysis [11] and lead to an efficient WDAF solution in cylindrical coordinates, e.g., [10, 12]. In this paper,
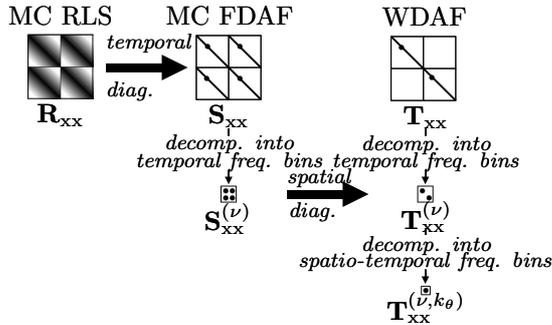

Figure 2: Illustration of the WDAF concept and its relation to conventional algorithms.

we focus on the general WDAF adaptation procedure for a given set of spatial transformations. Based on a generic formulation of suitable transformations, we derive in this paper a whole class of efficient adaptive MIMO filtering algorithms in the wave domain based on an exact broadband recursive least-squares criterion. Then, from the corresponding normal equation, we rigorously derive a generic WDAF algorithm that we formulate in different ways. As we will see, selective approximations within the generic algorithm provide both a basis for the development of new efficient algorithms and interesting links to some well-known algorithms, including all the major single-channel and multichannel FDAF variants [9], i.e., with both unconstrained (e.g., [14]) and constrained (e.g., [13]) coefficient updates, and also with partitioned coefficient updates, also known as the multidelay filter (MDF) [15, 16, 17, 18]. Due to the rigorous approach shown in this paper, we obtain various important design rules for practical system implementation and for performance optimization.

The organization of this paper is as follows. In Sect. 2 we show how to develop an exact temporal DFT-based transformation of the system in Fig. 1. These developments and associated matrix definitions serve both as the basis for the brief review of the generic MC FDAF in Sect. 3, and as the first transformation stage for the WDAF. In Sect. 4 we extend the results of Sect. 2 by the spatial transformation stage to obtain a compact matrix formulation. Finally, the derivation and discussion of the generic WDAF adap-

tation algorithm is presented in Sect. 5.

2 Block Formulation and Temporal Transformation

In order to subsequently present the generic FDAF and WDAF algorithms and to introduce some definitions, we show in this section an exact broadband formulation of the MIMO error equation (1) using the temporal DFT. Thereby we follow the notation in [9].

By partitioning the impulse response $\hat{h}_{pq}$ of length L into K segments of integer length $N = L/K$ as in [15], (1) can be written as

$$
\begin{aligned}
e_q(n) &= y_q(n) \\
&\quad - \sum_{p=1}^{P} \sum_{k=0}^{K-1} \sum_{\ell=0}^{N-1} x_p(n - Nk - \ell)\hat{h}_{pq,Nk+\ell} \\
&= y_q(n) - \sum_{p=1}^{P} \sum_{k=0}^{K-1} \mathbf{x}_{p,k}^T(n)\hat{\mathbf{h}}_{pq,k}, \quad (3)
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbf{x}_{p,k}(n) &= [x_p(n-Nk), x_p(n-Nk-1), \dots \\
&\qquad \dots, x_p(n-Nk-N+1)]^T, \\
\hat{\mathbf{h}}_{pq,k} &= [\hat{h}_{pq,Nk}, \hat{h}_{p,Nk+1}, \dots, \hat{h}_{pq,Nk+N-1}]^T.
\end{aligned}
$$

Superscript ^{T} denotes transposition of a vector or a matrix.

We now define the block error signal of length N. Based on (3) we write

$$
\mathbf{e}_q(m) = \mathbf{y}_q(m) - \sum_{p=1}^{P} \sum_{k=0}^{K-1} \mathbf{U}_{p,k}^T(m)\hat{\mathbf{h}}_{pq,k}, \quad (4)
$$

where m is the block time index, and

$$
\begin{aligned}
\mathbf{e}_q(m) &= [e_q(mN), \dots, e_q(mN+N-1)]^T, \\
\mathbf{y}_q(m) &= [y_q(mN), \dots, y_q(mN+N-1)]^T, \\
\mathbf{U}_{p,k}(m) &= [\mathbf{x}_{p,k}(mN), \dots, \mathbf{x}_{p,k}(mN+N-1)].
\end{aligned}
$$

To derive the frequency-domain algorithm, the block error signal (4) is transformed by a DFT matrix to its frequency-domain counterpart. The matrices $\mathbf{U}_{p,k}(m)$, $k = 0, \dots, K-1$, $p = 1, \dots, P$ are Toeplitz matrices of size $N \times N$. Since a Toeplitz matrix $\mathbf{U}_{p,k}(m)$ can be transformed, by doubling its size, to a circulant matrix of size $2N \times 2N$, and a circulant matrix can be diagonalized using the $2N \times 2N$-DFT matrix $\mathbf{F}_{2N}$ with elements $e^{-j2\pi\nu n/(2N)}$ ($\nu, n = 0, \dots, 2N-1$), we have

$$
\mathbf{U}_{p,k}^T(m) =
\underbrace{[\mathbf{0}_{N\times N}, \mathbf{I}_{N\times N}]}_{=:\mathbf{W}_{N\times 2N}^{01}} \mathbf{F}_{2N}^{-1}\underline{\mathbf{X}}_{p,k}(m)\mathbf{F}_{2N} \underbrace{[\mathbf{I}_{N\times N}, \mathbf{0}_{N\times N}]^T}_{=:\mathbf{W}_{2N\times N}^{10}}
$$

with the diagonal matrices

$$\underline{\mathbf{X}}_{p,k}(m) = \mathrm{diag}\{\mathbf{F}_{2N}[x_p(mN - Nk - N),\ldots$$
$$\ldots, x_p(mN - Nk + N - 1)]^T\}. \quad (5)$$

Underlined symbols denote frequency-domain quantities. The superscript indices '01' and '10' of the window matrices $\mathbf{W}_{N \times 2N}^{01}$ and $\mathbf{W}_{2N \times N}^{10}$ describe the relative positions of the $N \times N$ identity matrix and the $N \times N$ zero matrix within the windows. This finally leads to the following MIMO block error signal

$$\mathbf{E}(m) = \mathbf{Y}(m) - \mathbf{W}_{N \times 2N}^{01}\mathbf{F}_{2N}^{-1}\underline{\mathbf{X}}(m)\mathbf{G}_{2LP \times LP}^{10}\hat{\underline{\mathbf{H}}}, \quad (6)$$

where

$$\begin{aligned}
\mathbf{E}(m) &= [\mathbf{e}_1(m),\ldots,\mathbf{e}_Q(m)], \\
\mathbf{Y}(m) &= [\mathbf{y}_1(m),\ldots,\mathbf{y}_Q(m)], \\
\underline{\mathbf{X}}(m) &= [\underline{\mathbf{X}}_1(m),\ldots,\underline{\mathbf{X}}_P(m)], \\
\underline{\mathbf{X}}_p(m) &= [\underline{\mathbf{X}}_{p,0}(m),\ldots,\underline{\mathbf{X}}_{p,K-1}(m)], \\
\mathbf{G}_{2LP \times LP}^{10} &= \mathrm{Bdiag}\{\mathbf{G}_{2N \times N}^{10},\ldots,\mathbf{G}_{2N \times N}^{10}\}, \quad (7) \\
\mathbf{G}_{2N \times N}^{10} &= \mathbf{F}_{2N}\mathbf{W}_{2N \times N}^{10}\mathbf{F}_N^{-1}, \\
\hat{\underline{\mathbf{H}}} &= \begin{bmatrix} \hat{\underline{\mathbf{h}}}_{11} & \cdots & \hat{\underline{\mathbf{h}}}_{1Q} \\ \vdots & \ddots & \vdots \\ \hat{\underline{\mathbf{h}}}_{P1} & \cdots & \hat{\underline{\mathbf{h}}}_{PQ} \end{bmatrix}, \\
\hat{\underline{\mathbf{h}}}_{pq} &= \left[\hat{\underline{\mathbf{h}}}_{pq,0}^T,\ldots,\hat{\underline{\mathbf{h}}}_{pq,K-1}^T\right]^T, \\
\hat{\underline{\mathbf{h}}}_{pq,k} &= \mathbf{F}_N\hat{\mathbf{h}}_{pq,k}.
\end{aligned}$$

The Bdiag operator in (7) forms a block-diagonal matrix with the listed submatrices on its diagonal.

3 Review of the Generic MC FDAF Algorithm

The generic FDAF algorithm in its MIMO formulation can be derived from (6) by minimizing the following optimization criterion [9] w.r.t. $\hat{\underline{\mathbf{H}}}$:

$$J(m) = (1 - \lambda)\sum_{i=0}^{m}\lambda^{m-i}\mathrm{tr}[\mathbf{E}^H(i)\mathbf{E}(i)], \quad (8)$$

where λ $(0 < \lambda < 1)$ is an exponential forgetting factor. The so-called normal equation is obtained by setting the gradient equal to zero,

$$\nabla J(m,\hat{\underline{\mathbf{H}}}) = 2\frac{\partial}{\partial\hat{\underline{\mathbf{H}}}^*}J(m,\hat{\underline{\mathbf{H}}}) = \mathbf{0}. \quad (9)$$

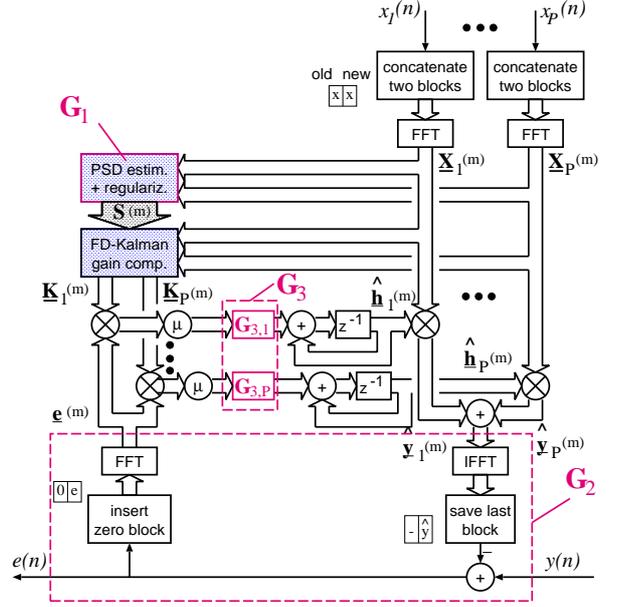The generic FDAF provides the *exact recursive solution* of this normal equation. As shown in [9], it can

Figure 3: Principle of the MC FDAF ($N = L$, $Q = 1$).

be summarized as

$$\begin{aligned}
\mathbf{S}(m) &= \lambda\mathbf{S}(m-1) \\
&\quad + (1 - \lambda)\underline{\mathbf{X}}^H(m)\mathbf{G}_1\underline{\mathbf{X}}(m), \quad (10) \\
\underline{\mathbf{K}}(m) &= (1 - \lambda)\mathbf{S}^{-1}(m)\underline{\mathbf{X}}^H(m), \quad (11) \\
\underline{\mathbf{E}}'(m) &= \underline{\mathbf{Y}}'(m) - \mathbf{G}_2\underline{\mathbf{X}}(m)\hat{\underline{\mathbf{H}}}'(m-1), \quad (12) \\
\hat{\underline{\mathbf{H}}}'(m) &= \hat{\underline{\mathbf{H}}}'(m-1) + \mathbf{G}_3\underline{\mathbf{K}}(m)\underline{\mathbf{E}}'(m) \quad (13)
\end{aligned}$$

with the constraint matrices

$$\begin{aligned}
\mathbf{G}_1 = \mathbf{G}_2 &= \mathbf{G}_{2N \times 2N}^{01} \\
&= \mathbf{F}_{2N}\mathrm{Bdiag}\{\mathbf{0}_{N \times N}, \mathbf{I}_{N \times N}\}\mathbf{F}_{2N}^{-1}, \quad (14) \\
\mathbf{G}_3 &= \mathbf{G}_{2LP \times 2LP}^{10} \\
&= \mathrm{Bdiag}\{\mathbf{G}_{2N \times 2N}^{10},\ldots,\mathbf{G}_{2N \times 2N}^{10}\}, (15) \\
\mathbf{G}_{2N \times 2N}^{10} &= \mathbf{F}_{2N}\mathrm{Bdiag}\{\mathbf{I}_{N \times N}, \mathbf{0}_{N \times N}\}\mathbf{F}_{2N}^{-1}, \quad (16)
\end{aligned}$$

and the zero-padded DFT-domain quantities

$$\begin{aligned}
\underline{\mathbf{E}}'(m) &= \mathbf{F}_{2N}\left[\mathbf{0}_{Q \times N}, \mathbf{E}^T(m)\right]^T, \quad (17) \\
\underline{\mathbf{Y}}'(m) &= \mathbf{F}_{2N}\left[\mathbf{0}_{Q \times N}, \mathbf{Y}^T(m)\right]^T, \quad (18) \\
\hat{\underline{\mathbf{H}}}'(m) &= \mathbf{G}_{2LP \times LP}^{10}\hat{\underline{\mathbf{H}}}(m). \quad (19)
\end{aligned}$$

Due to the formal similarity of Eqs. (10)-(13) to the RLS algorithm in the time domain (2), we call the matrix $\underline{\mathbf{K}}(m)$ the frequency-domain Kalman gain [7].

3.1 Approximation and Special Cases

The algorithm (10)-(13) is strictly equivalent to the RLS algorithm in the time domain for a block length $N = 1$. Unfortunately, the matrix $\mathbf{S}(m)$ in (10) is not sparse, so the above generic algorithm still has a high computational complexity due to the matrix inversion in (11). In [9] it has been shown that $\mathbf{G}_1 = \mathbf{I}/2$ provides a very good approximation for sufficiently large N, and the resulting algorithms still converge to the Wiener solution.

This approximation leads to a blockwise diagonal structure of matrix $\mathbf{S}(m)$ in (10) with diagonal sub-matrices so that the matrix inversion can be performed in a bin-wise fashion. For $K > 1$ partitions the resulting algorithm is called the extended multidelay filter (EMDF) [16]. The classical multidelay filter (MDF), originally introduced in [15] for the single-channel case, is obtained by further approximating $\mathbf{S}(m)$ by dropping the off-diagonal components, i.e., the inter-partition correlations. Moreover, the so-called generalized multidelay filter (GMDF) follows by an additional overlap of the input data blocks so that the adaptation can be performed more frequently [17, 18].

Approximating $\mathbf{G}_3$ in (13) as the unit matrix leads to the so-called *unconstrained* FDAF, introduced for the single-channel case in [14], in contrast to the constrained single-channel version in [13].

Some efficient multichannel implementations were presented in [9]. Note that the illustration in the second column of Fig. 2 shows the case $N = L$, i.e., $K = 1$ for $P = 2$.

4 Spatial Transformation

For the rigorous derivation of the WDAF we start again with the MIMO formulation of the error equation (6),

$$\mathbf{E}(m) = \mathbf{Y}(m) - \mathbf{W}_{N \times 2N}^{01} \mathbf{F}_{2N}^{-1} \underline{\mathbf{X}}(m) \underline{\hat{\mathbf{H}}}'. \quad (20)$$

To additionally include a spatial transform, we now reorder the elements of the *blockwise diagonal* matrix $\underline{\mathbf{X}}(m)$, so that we obtain (for $N = L$ or for each partition, respectively) a *block-diagonal* matrix $\overline{\underline{\mathbf{X}}}(m)$, as illustrated in Fig. 4. In other words, we combine the MIMO relations for each DFT bin into sub-matrices using a suitable permutation matrix $\mathbf{A}$, i.e.,

$$\underline{\mathbf{X}}(m) = \overline{\underline{\mathbf{X}}}(m) \mathbf{A}^T. \quad (21)$$

Plugging this relation into (20), we obtain

$$\mathbf{E}(m) = \mathbf{Y}(m) - \mathbf{W}_{N \times 2N}^{01} \mathbf{F}_{2N}^{-1} \overline{\underline{\mathbf{X}}}(m) \overline{\underline{\hat{\mathbf{H}}}}', \quad (22)$$
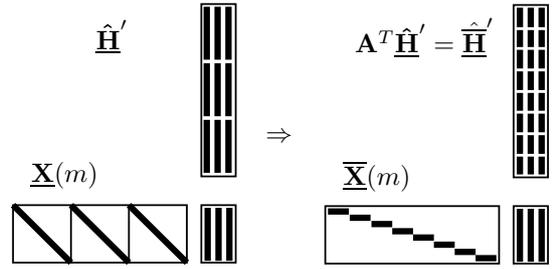

Figure 4: Illustration ($N = L$, $P = Q = 3$) of the reordering of matrix elements using the permutation matrix $\mathbf{A}$ according to (21).

where we defined

$$\overline{\underline{\hat{\mathbf{H}}}}' := \mathbf{A}^T \underline{\hat{\mathbf{H}}}' = \mathbf{A}^T \mathbf{G}_{2LP \times LP}^{10} \underline{\hat{\mathbf{H}}}. \quad (23)$$

The next step is to introduce the actual spatial transformation using the (block-diagonal) unitary matrices $\overline{\mathbf{C}}_1'$ and $\overline{\mathbf{C}}_2$, so that

$$\overline{\underline{\hat{\mathbf{H}}}}' = \overline{\mathbf{C}}_1' \underline{\hat{\mathbf{H}}}' \overline{\mathbf{C}}_2^H. \quad (24)$$

This generic transformation is motivated by the singular value decomposition in the *eigenspace adaptive filtering* (EAF) approach [19] in which ideally the (yet unknown) MIMO system is diagonalized. In this paper, double underlined symbols denote spatio-temporally transformed quantities, or *wave-domain* quantities. Incorporating this transformation into (22) immediately leads to

$$\mathbf{E}(m) = \mathbf{Y}(m) - \mathbf{W}_{N \times 2N}^{01} \mathbf{F}_{2N}^{-1} \underline{\mathbf{X}}(m) \overline{\underline{\hat{\mathbf{H}}}}' \overline{\mathbf{C}}_2^H, \quad (25)$$

where we defined

$$\underline{\mathbf{X}}(m) := \overline{\underline{\mathbf{X}}}(m) \overline{\mathbf{C}}_1' = \underline{\mathbf{X}}(m) \mathbf{A} \overline{\mathbf{C}}_1'. \quad (26)$$

Note that (25) is the wave-domain formulation of the frequency-domain expression (20). Moreover, although the transformation approach in (24) was originally system-based, it is interesting to see from (26) that $\overline{\mathbf{C}}_1'$ in fact acts as a spatial transformation of the input signals. Similarly, since $\overline{\mathbf{C}}_2^H \overline{\mathbf{C}}_2 = \mathbf{I}$, we can show via multiplication of $\overline{\mathbf{C}}_2$ on both sides of (25), that $\overline{\mathbf{C}}_2$ acts as a spatial transformation of the output signals $\mathbf{Y}(m)$ and $\mathbf{E}(m)$.

Obviously, depending on the choice of these transformations, the generic algorithm to be presented in the following section includes various important special cases: For $\overline{\mathbf{C}}_1' = \mathbf{I}$, $\overline{\mathbf{C}}_2 = \mathbf{I}$ we obtain the MC

FDAF (and all its variants). The ideal EAF follows if $\overline{\mathbf{C}}_1'$ and $\overline{\mathbf{C}}_2$ are composed of the left and right singular vectors of the MIMO system. For the practical WDAF, the transformation matrices represent the wave transform as outlined, e.g., in [10].

Finally, before we present the generic WDAF algorithm, $\hat{\underline{\mathbf{H}}}'$ in (25) has to be expressed by $\hat{\underline{\mathbf{H}}}$, i.e., without zero-padding (matrix $\mathbf{G}_{2LP \times LP}^{10}$ in (23)) for the coefficient optimization. Similar to [9], we then obtain both the constrained and unconstrained WDAF as special cases. From (23), (24), and the unitarity of the transform matrices, we immediately obtain for the zero-padded versions of the system matrices the relation

$$\hat{\underline{\mathbf{H}}}' = \overline{\mathbf{C}}_1'^H \mathbf{A}^T \mathbf{G}_{2LP \times LP}^{10} \hat{\mathbf{H}} \overline{\mathbf{C}}_2. \qquad (27)$$

On the other hand, analogously to (23) and (24) we write for the versions without zero padding (notation without the prime)

$$\hat{\underline{\mathbf{H}}} = \overline{\mathbf{C}}_1 \hat{\mathbf{H}} \overline{\mathbf{C}}_2^H = \mathbf{B}^T \hat{\underline{\mathbf{H}}}, \qquad (28)$$

where $\mathbf{B}$ again denotes a suitable permutation matrix. Introducing (27) and (28) into (25) immediately leads to the wave-domain counterpart of (6):

$$\mathbf{E}(m) = \mathbf{Y}(m) - \mathbf{W}_{N \times 2N}^{01}$$
$$\cdot \mathbf{F}_{2N}^{-1} \underline{\mathbf{X}}(m) \underbrace{\overline{\mathbf{C}}_1'^H \mathbf{A}^T \mathbf{G}_{2LP \times LP}^{10} \mathbf{B} \overline{\mathbf{C}}_1}_{=: \tilde{\mathbf{G}}_{2LP \times LP}^{10}} \hat{\mathbf{H}} \overline{\mathbf{C}}_2^H. \qquad (29)$$

5 Generic WDAF Algorithm

5.1 Normal Equation

Having derived the broadband error signal (29), we now minimize the optimization criterion (8) with respect to the MIMO filter coefficients in the wave-domain. Note that due to the unitarity of $\overline{\mathbf{C}}_1'$ and $\overline{\mathbf{C}}_2$, the time-frequency equivalence is assured by Parseval's theorem,

$$\text{tr}[\mathbf{E}^H \mathbf{E}] = \text{tr}[\underline{\mathbf{E}}^H \underline{\mathbf{E}}] = \text{tr}[\hat{\underline{\mathbf{E}}}^H \hat{\underline{\mathbf{E}}}], \qquad (30)$$

where

$$\hat{\underline{\mathbf{E}}} := \underline{\mathbf{E}} \overline{\mathbf{C}}_2 := \mathbf{F}_N \mathbf{E} \overline{\mathbf{C}}_2. \qquad (31)$$

Calculation of the derivative (e.g., [20]) in

$$\nabla J(m, \hat{\underline{\mathbf{H}}}) = 2 \frac{\partial}{\partial \hat{\underline{\mathbf{H}}}^*} J(m, \hat{\underline{\mathbf{H}}}) = \mathbf{0} \qquad (32)$$

leads to the *normal equation* in the wave domain

$$\mathbf{T}_{\mathbf{xx}}(m) \hat{\underline{\mathbf{H}}}(m) = \mathbf{T}_{\mathbf{xy}}(m), \qquad (33)$$

where

$$\mathbf{T}_{\mathbf{xx}}(m) = \lambda \mathbf{T}_{\mathbf{xx}}(m-1) + (1-\lambda) \left(\tilde{\mathbf{G}}_{2LP \times LP}^{10} \right)^H \cdot$$
$$\cdot \underline{\mathbf{X}}^H(m) \mathbf{G}_{2N \times 2N}^{01} \underline{\mathbf{X}}(m) \tilde{\mathbf{G}}_{2LP \times LP}^{10}, \quad (34)$$

$$\mathbf{T}_{\mathbf{xy}}(m) = \lambda \mathbf{T}_{\mathbf{xy}}(m-1) + (1-\lambda) \left(\tilde{\mathbf{G}}_{2LP \times LP}^{10} \right)^H \cdot$$
$$\cdot \underline{\mathbf{X}}^H(m) \left(\mathbf{G}_{N \times 2N}^{01} \right)^H \underline{\mathbf{Y}}(m) \qquad (35)$$

and

$$\mathbf{G}_{2N \times 2N}^{01} = \left(\mathbf{G}_{N \times 2N}^{01} \right)^H \mathbf{G}_{N \times 2N}^{01}, \qquad (36)$$
$$\mathbf{G}_{N \times 2N}^{01} = \mathbf{F}_{2N} \mathbf{W}_{N \times 2N}^{01} \mathbf{F}_{2N}^{-1}, \qquad (37)$$
$$\underline{\mathbf{Y}} = \mathbf{F}_N \mathbf{Y} \overline{\mathbf{C}}_2. \qquad (38)$$

If the multichannel input signal is well-conditioned, matrix $\mathbf{T}_{\mathbf{xx}}(m)$ is nonsingular. In this case, the normal equation has a unique solution which is the optimum Wiener solution.

5.2 Adaptation Algorithm

The different formulations for filter adaptation discussed below, i.e., recursive updates of $\hat{\underline{\mathbf{H}}}(m)$, are all derived directly from the normal equation (33) and associated equations (34) and (35).

Here, we replace $\mathbf{T}_{\mathbf{xy}}(m)$ and $\mathbf{T}_{\mathbf{xy}}(m-1)$ in the recursive equation (35) by formulating (33) in terms of block time indices m and $m-1$, respectively. We then eliminate $\mathbf{T}_{\mathbf{xx}}(m-1)$ from the resulting equation using (34). Reintroducing the multichannel error signal (29), we obtain an exact recursive solution of (33) by the following adaptation algorithm:

$$\hat{\underline{\mathbf{E}}}(m) = \underline{\mathbf{Y}}(m) - \mathbf{G}_{N \times 2N}^{01} \underline{\mathbf{X}}(m) \tilde{\mathbf{G}}_{2LP \times LP}^{10} \hat{\underline{\mathbf{H}}}(m-1) \qquad (39)$$

$$\hat{\underline{\mathbf{H}}}(m) = \hat{\underline{\mathbf{H}}}(m-1) + (1-\lambda) \mathbf{T}_{\mathbf{xx}}^{-1}(m) (\tilde{\mathbf{G}}_{2LP \times LP}^{10})^H$$
$$\cdot \underline{\mathbf{X}}^H(m) (\mathbf{G}_{N \times 2N}^{01})^H \underline{\mathbf{E}}(m). \qquad (40)$$

Eqs. (34), (39), and (40) form the main equations of the generic adaptive algorithm. In the same way as shown in [9] for the FDAF, and summarized in Appendix A, these equations can be reformulated in a practically more useful form involving only DFTs of length $2N$:

$$\mathbf{T}(m) = \lambda \mathbf{T}(m-1)$$
$$+ (1-\lambda) \underline{\mathbf{X}}^H(m) \mathbf{G}_1 \underline{\mathbf{X}}(m) \qquad (41)$$

$$\underline{\mathbf{K}}(m) = (1-\lambda) \mathbf{T}^{-1}(m) \underline{\mathbf{X}}^H(m) \qquad (42)$$

$$\underline{\mathbf{E}}'(m) = \underline{\mathbf{Y}}'(m) - \mathbf{G}_2 \underline{\mathbf{X}}(m) \hat{\underline{\mathbf{H}}}'(m-1) \quad (43)$$

$$\hat{\underline{\mathbf{H}}}'(m) = \hat{\underline{\mathbf{H}}}'(m-1) + \mathbf{G}_3 \underline{\mathbf{K}}(m) \underline{\mathbf{E}}'(m)$$
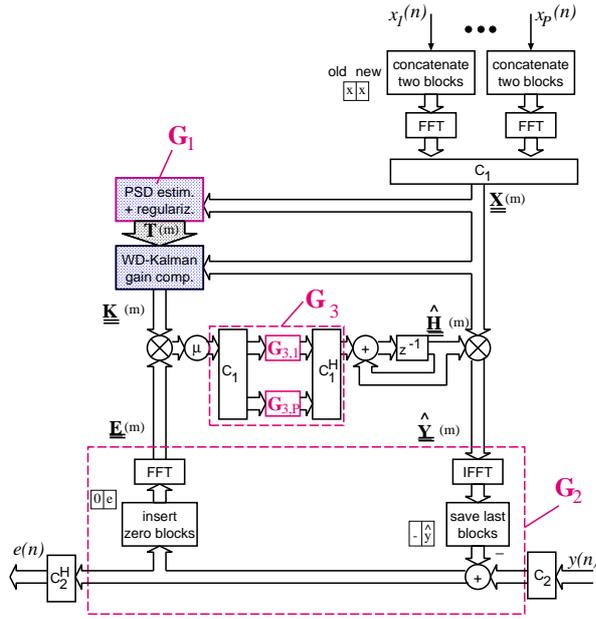$$(44)$$

Figure 5: Principle of WDAF ($N = L$).

with the constraint matrices

$$
\begin{aligned}
\mathbf{G}_1 = \mathbf{G}_2 &= \mathbf{G}_{2N \times 2N}^{01} \\
\mathbf{G}_3 &= \tilde{\mathbf{G}}_{2LP \times 2LP}^{10}, \\
&= \overline{\mathbf{C}}_1^H \mathbf{A}^T \mathbf{G}_{2LP \times 2LP}^{10} \mathbf{A} \overline{\mathbf{C}}_1, \quad (45)
\end{aligned}
$$

and the zero-padded wave-domain quantities

$$
\underline{\mathbf{E}}'(m) = \mathbf{F}_{2N} \left[\mathbf{0}_{Q \times N}, \mathbf{E}^T(m) \right]^T \overline{\mathbf{C}}_2, \quad (46)
$$

$$
\underline{\mathbf{Y}}'(m) = \mathbf{F}_{2N} \left[\mathbf{0}_{Q \times N}, \mathbf{Y}^T(m) \right]^T \overline{\mathbf{C}}_2. \quad (47)
$$

5.3 Discussion

Comparing the generic WDAF algorithm (41)-(44) with the generic FDAF algoritm (10)-(13), we see that the constraints $\mathbf{G}_1$ and $\mathbf{G}_2$ are the same in both cases. This is a very useful result as it allows us to use the same efficient approximation $\mathbf{G}_1 = \mathbf{I}/2$ in both cases (see Sect. 3.1). In contrast, the gradient constraint $\mathbf{G}_3$ differs between FDAF and WDAF. In the WDAF case it includes both spatial and temporal transformations, as illustrated in Fig. 5. Figure 5 also confirms that the WDAF adaptation mechanism simultaneously takes into account all loudspeaker channels and all microphone channels by the transformation matrix $\overline{\mathbf{C}}_2$. This is in contrast to the FDAF and other conventional adaptation algorithms.

Finally, we can explicitly relate the system-based spatial transformation, as introduced by (24), with the correlation-based spatial transformation illustrated in

the third column of Fig. 2: From (41) and (26) and comparison with (10), we immediately obtain the relation

$$
\mathbf{T}(m) = \overline{\mathbf{C}}_1^H \mathbf{A}^T \mathbf{S}(m) \mathbf{A} \overline{\mathbf{C}}_1. \quad (48)
$$

By a suitable choice of $\overline{\mathbf{C}}_1$ this equation represents a eigenvalue equation for $\mathbf{S}(m)$. It therefore provides a direct link to the (correlation-based) multichannel transform-domain adaptive filtering (MC TDAF) [21]. Note however, that this link between WDAF and MC TDAF only refers to the transformation of the filter *input* signals and the *input* correlation matrix.

6 Conclusions

In this paper, we presented a compact and versatile matrix formulation for Wave-Domain Adaptive Filtering. The WDAF concept has already proven to be a very efficient spatio-temporal generalization of the popular Frequency-Domain Adaptive Filtering. The new formulation of WDAF leads to a whole class of powerful MIMO adaptation algorithms, and it contains various well-known algorithms as special cases. The matrix framework also facilitates the derivation of new improved MIMO algorithms and it gives some important design rules for practical implementations.

A Practical Reformulation (41)-(44) of the Generic WDAF Algorithm

Here, we summarize the steps leading from Eqs. (34), (39), and (40) to the practically more useful form (41)-(44). The advantages of this equivalent formulation are that it involves exclusively DFTs of length $2N$, and that the relation to several known frequency-domain algorithms, such as the conventional MDF can be established.

To begin with, we introduce the zero-padded coefficient matrix $\underline{\hat{\mathbf{H}}}'(m)$. Its relation to the original length-L formulation $\underline{\hat{\mathbf{H}}}(m)$ can be expressed conveniently according to (29) as

$$
\underline{\hat{\mathbf{H}}}'(m) = \tilde{\mathbf{G}}_{2LP \times LP}^{10} \underline{\hat{\mathbf{H}}}(m). \quad (49)
$$

Consequently, the coefficient update (40) is premultiplied by $\tilde{\mathbf{G}}_{2LP \times LP}^{10}$ on both sides so that we obtain

$$
\begin{aligned}
\underline{\hat{\mathbf{H}}}'(m) = &\ \underline{\hat{\mathbf{H}}}'(m-1) + (1 - \lambda) \\
&\cdot \tilde{\mathbf{G}}_{2LP \times LP}^{10} \mathbf{T}_{\mathbf{xx}}^{-1}(m) (\tilde{\mathbf{G}}_{2LP \times LP}^{10})^H \\
&\cdot \underline{\mathbf{X}}^H(m) \underline{\mathbf{E}}'(m).
\end{aligned} \quad (50)
$$

This update equation can be simplified by introducing the matrix $\mathbf{T}(m)$, defined in (41). The relation of $\mathbf{T}(m)$ to $\mathbf{T}_{\mathbf{xx}}(m)$ from (34) is obviously given by

$$
\mathbf{T}_{\mathbf{xx}}(m) = (\tilde{\mathbf{G}}_{2LP \times LP}^{10})^H \mathbf{T}(m) \tilde{\mathbf{G}}_{2LP \times LP}^{10}. \quad (51)
$$

Finally, in order to obtain the coefficient update (44) with (42) from (50), we verify the following relation between the inverses of the two matrices $\mathbf{T}$ and $\mathbf{T_{xx}}$:

$$\tilde{\mathbf{G}}^{10}_{2LP \times 2LP}\mathbf{T}^{-1}(m) = \tilde{\mathbf{G}}^{10}_{2LP \times LP}\mathbf{T}^{-1}_{\mathbf{xx}}(m)(\tilde{\mathbf{G}}^{10}_{2LP \times LP})^H, \tag{52}$$

The verification can be done by post-multiplying both sides of the equation by $\mathbf{T}(m)\tilde{\mathbf{G}}^{10}_{2LP \times LP}$ and noting that $\tilde{\mathbf{G}}^{10}_{2LP \times 2LP}\tilde{\mathbf{G}}^{10}_{2LP \times LP} = \tilde{\mathbf{G}}^{10}_{2LP \times LP}$.

References

[1] A.J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," *Journal of the Acoustic Society of America*, vol. 93, no. 5, pp. 2764–2778, May 1993.

[2] M.S. Brandstein and D.B. Ward, *Microphone Arrays,* Springer, 2001.

[3] E. Weinstein, K. Steele, A. Agarwal, and J. Glass, "LOUD: A 1020-node microphone array and acoustic beamformer," in *Proc. Int. Congress on Sound and Vibration*, Cairns, Australia, 2007.

[4] C. Breining et al., "Acoustic Echo Control," *IEEE Signal Processing Magazine*, Vol. 16, No. 4, July 1999, pp. 42-69.

[5] H. Buchner, S. Spors, W. Kellermann, and R. Rabenstein, "Full-Duplex Communication Systems with Loudspeaker Arrays and Microphone Arrays," *Proc. IEEE Int. Conference on Multimedia and Expo (ICME),* Lausanne, Switzerland, Aug. 2002.

[6] 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 Trans. on Speech and Audio Processing*, vol. 6, no.2, March 1998.

[7] S. Haykin, *Adaptive Filter Theory*, 3rd ed., Prentice Hall Inc., Englewood Cliffs, NJ, 1996.

[8] J.J. Shynk, "Frequency-domain and multirate adaptive filtering," *IEEE SP Magazine,* pp. 14-37, Jan. 1992

[9] H. Buchner, J. Benesty, and W. Kellermann, "Multichannel Frequency-Domain Adaptive Algorithms with Application to Acoustic Echo Cancellation," in J.Benesty and Y.Huang (eds.), *Adaptive signal processing: Application to real-world problems*, Springer-Verlag, Berlin/Heidelberg, Jan. 2003.

[10] 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 Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, Montreal, Canada, 2004.

[11] E. Hulsebos, D. de Vries, and E. Bourdillat, "Improved microphone array configurations for auralization of sound fields by Wave Field Synthesis," *110th Convention of the Audio Engineering Society (AES),*

[12] S. Spors, H. Buchner, and R. Rabenstein,"A novel approach to active listening room compensation for wave field synthesis using wave-domain adaptive filtering," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, Montreal, Canada, 2004.

[13] E.R. Ferrara, Jr., "Fast implementation of the LMS adaptive filter," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. ASSP-28, pp. 474-475, Aug. 1980.

[14] D. Mansour and A.H. Gray, "Unconstrained Frequency-Domain Adaptive Filter," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol.30, no.5, Oct. 1982.

[15] J.-S. Soo and K.K. Pang, "Multidelay block frequency domain adaptive filter," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-38, pp. 373-376, Feb. 1990.

[16] H. Buchner, J. Benesty, and W. Kellermann, "An extended multidelay filter: fast low-delay algorithms for very high-order adaptive systems," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 5, pp. 385-388, April 2003.

[17] E. Moulines, O. Ait Amrane, and Y. Grenier, "The generalized multidelay adaptive filter: structure and convergence analysis," *IEEE Trans. Signal Processing,* vol. 43, pp. 14-28, Jan. 1995.

[18] J. Prado and E. Moulines, "Frequency-domain adaptive filtering with applications to acoustic echo cancellation," *Ann. Télécommun.,* vol. 49, pp. 414-428, 1994.

[19] S. Spors, H. Buchner, and R. Rabenstein, "Eigenspace adaptive filtering for efficient pre-equalization of acoustic MIMO systems," *Proc. European Signal Processing Conference (EUSIPCO)*, Florence, Italy, Sept. 2006.

[20] D.A. Harville, *Matrix Algebra From a Statistician's Perspective*, Springer-Verlag, New York, 1997.

[21] S. Spors and H. Buchner, "Multichannel transform domain adaptive filtering: A two stage approach and illustration for acoustic echo cancellation," *Conf. Rec. Intl. Workshop on Acoustic Echo and Noise Control (IWAENC)*, Seattle, WA, USA, Sept. 2008.