

# IMPROVED KALMAN GAIN COMPUTATION FOR MULTICHANNEL FREQUENCY-DOMAIN ADAPTIVE FILTERING AND APPLICATION TO ACOUSTIC ECHO CANCELLATION

*Herbert Buchner and Walter Kellermann*

Telecommunications Laboratory, University of Erlangen-Nuremberg  
Cauerstr. 7, D-91058 Erlangen, Germany  
{buchner, wk}@LNT.de

## ABSTRACT

For multichannel adaptive filtering applications like acoustic echo cancellation, sophisticated, but efficient algorithms are necessary that take both auto- and cross-correlations of the excitation signals into account. Based on a recently published frequency-domain framework, we examine in this contribution the computation of the so-called frequency-domain Kalman gain, which is crucial in the multichannel case to meet the above requirements for correlated input signals. We present an extended scheme with increased computational efficiency for a higher number of input channels. Robust convergence behaviour during perturbation of the desired signal is ensured by a new dynamical regularization. Application of the new algorithm to multichannel acoustic echo cancellation with real-world signals achieves improved performance.

## 1. INTRODUCTION

Multichannel adaptive filtering for applications of high-order FIR filters excited with several highly auto- and cross-correlated input signals is a demanding problem in signal processing [1, 2].

A typical example with these requirements is multichannel acoustic echo cancellation (M-C AEC) as a multiple-input system identification problem. M-C AEC is crucial for hands-free and full-duplex communication systems with multichannel sound reproduction (e.g., stereo or 5.1 channel - surround systems). Applications can be found, e.g., in home entertainment, virtual reality (e.g. games, simulations, training), or advanced teleconferencing and teleteaching.

However, due to the often very ill-conditioned nature of the underlying least-squares problem in the multichannel case [1], even for only two input channels (i.e. stereo), most adaptive algorithms are very complex and/or show poor convergence behaviour. Only few algorithms provide acceptable properties in a multi-loudspeaker environment for real-world conditions.

Frequency-domain adaptive filtering (FDAF) is well known for its very low complexity and increased convergence speed when properly designed. However, only recently a rigorous derivation of FDAF allowing a straightforward and powerful generalization to the multi-channel case has been presented [4].

In the following, we address important issues for the applicability of this scheme with increased number of input channels. The main focus of this contribution is the efficient calculation of the frequency-domain Kalman gain and associated regularization.

All considerations in this contribution are referring to only one system output and desired signal, respectively (e.g., microphone in

the receiving room), but can be easily and efficiently generalized to multichannel output (e.g., sound recording as discussed in [5]).

## 2. MULTICHANNEL FREQUENCY-DOMAIN ADAPTIVE FILTERING

The high computational efficiency of single-channel frequency-domain adaptive filtering (FDAF), originally derived from the basic LMS algorithm [3] and further optimized in [6] has been known for many years. However, direct application to the multichannel case usually does not lead to satisfying results, because the LMS algorithm does not take the correlations of the input signals into account. In [4], a rigorous derivation of FDAF based on a weighted least-squares criterion in the frequency domain has been shown.

### 2.1. Basic concept

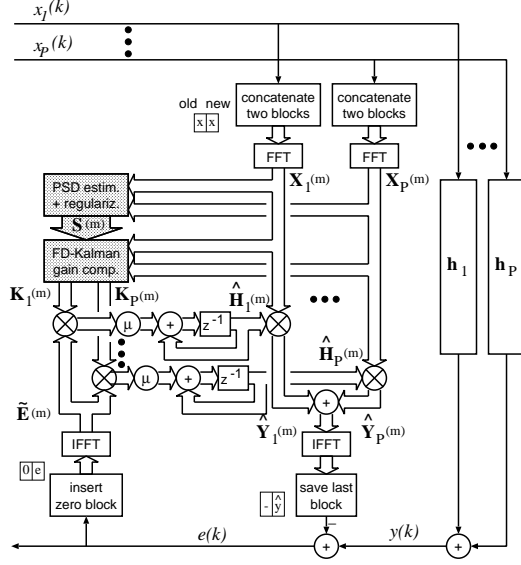
The two advantages of this structure, low complexity and fast convergence, are based on linear filtering via fast convolution as well as the approximate diagonalization of Toeplitz matrices by the Discrete Fourier Transform (DFT) in the unconstrained version of the algorithm. Fig. 1 shows an overview of the scheme in a system identification setup as used for M-C AEC. In the following we assume a block size equal to the filter length  $L$ . Generalization to a partitioned version is straightforward. For the fast convolution we apply here the overlap-save (OLS) method [1]. According to this method, we first transform two concatenated blocks (length  $2L$ ) of each input signal  $x_i$ ,  $i = 1, \dots, P$  to the discrete Fourier domain,

$$\mathbf{X}_i(m) = \text{diag}\{\mathbf{F}[x_i(m\frac{L}{\alpha} - L + 1) \cdots x_i(m\frac{L}{\alpha} + L)]^T\}. \quad (1)$$

$\mathbf{F}$  denotes the DFT matrix,  $L$  is the modeling filter length,  $m$  is the block index over time, and  $1 \leq \alpha < L$  denotes a factor capturing the overlap of successive blocks which balances the computational complexity versus the number of iterations (leading to faster convergence) [7].

The estimated channel output signals  $\hat{\mathbf{Y}}_i(m)$  in the DFT domain are then generated by multiplying the weights  $\hat{\mathbf{H}}_i(m)$  with the elements of the diagonal matrix  $\mathbf{X}_i(m)$ . The block of residual errors (including an OLS constraint [4] with the data window  $\mathbf{W} = \text{diag}\{\mathbf{0}_{1 \times L} \quad \mathbf{1}_{1 \times L}\}$ , where  $\mathbf{0}_{1 \times L}$  and  $\mathbf{1}_{1 \times L}$  denote length  $L$  row vectors consisting of zeros and ones, respectively) becomes

$$\tilde{\mathbf{e}}(m) = [\mathbf{0}_{1 \times L} \quad y(m\frac{L}{\alpha} + 1) \cdots y(m\frac{L}{\alpha} + L)]^T -$$



**Fig. 1.** Unconstrained multichannel frequency-domain adaptive filtering in a system identification setup

$$-\mathbf{W}\mathbf{F}^{-1}\sum_{i=1}^P\mathbf{X}_i(m)\hat{\mathbf{H}}_i(m). \quad (2)$$

This vector serves both as output signal (last  $\frac{L}{\alpha}$  elements of  $\tilde{\mathbf{e}}(m)$ ) and as feedback for the next adaptation step in order to identify and track the impulse responses for the models. Using the transformed error vectors  $\tilde{\mathbf{E}}(m) = \mathbf{F}\tilde{\mathbf{e}}(m)$ , the filter weights are adapted using a recursive least-squares error criterion in the frequency domain [4]. This leads to the following update equation for  $\hat{\mathbf{H}} = [\hat{\mathbf{H}}_1^T \hat{\mathbf{H}}_2^T \cdots \hat{\mathbf{H}}_P^T]^T$ :

$$\hat{\mathbf{H}}(m+1) = \hat{\mathbf{H}}(m) + \mu\mathbf{S}^{-1}(m)\mathbf{X}^H(m)\tilde{\mathbf{E}}(m), \quad (3)$$

where

$$\mathbf{X} = [\mathbf{X}_1 \mathbf{X}_2 \cdots \mathbf{X}_P] \quad (4)$$

and

$$\mathbf{S}(m) = (1-\lambda)\sum_{q=0}^m \lambda^{m-q}\mathbf{X}^H(q)\mathbf{X}(q) \quad (5)$$

in the *unconstrained* version [4] of the algorithm (the constraint removed here is different from the one in [6]).  $\lambda$  is a forgetting factor close to, but less than one. Eqs. (3) and (5) are a good approximation to the well-known recursive least-squares (RLS) algorithm in the time domain (e.g. [1, 3]) which provides optimum convergence speed at very high cost. The product

$$\mathbf{K}(m) := \mathbf{S}^{-1}(m)\mathbf{X}^H(m) \quad (6)$$

in Eq. (3) is called here the *frequency-domain Kalman gain* in analogy to the RLS algorithm [3]. In our case, however, the matrices  $\mathbf{S}$  to be inverted and  $\mathbf{X}$  are block diagonal with blocks of size  $2L \times 2L$  (in the unconstrained version), so that each submatrix  $\mathbf{S}_{i,j}$  ( $i, j = 1, \dots, P$ ) representing a cross power spectrum is diagonal.

$\mathbf{K}(m)$  after Eq. (6) is the solution of a  $P \times P$  system of linear equations of block matrices. This allows decomposition of Eq. (3) into  $P$  single-channel update equations (for vectors of length  $2L$ )

$$\hat{\mathbf{H}}_i(m+1) = \hat{\mathbf{H}}_i(m) + \mu\mathbf{K}_i(m)\tilde{\mathbf{E}}(m). \quad (7)$$

with modified Kalman gains  $\mathbf{K}_i(m)$  taking the cross-correlations between the channels into account. These decomposed update equations can then be calculated element-wise and the (cross) power spectra in Eq. (5) are estimated recursively:

$$\mathbf{S}_{i,j}(m) = \lambda\mathbf{S}_{i,j}(m-1) + (1-\lambda)\mathbf{X}_i^*(m)\mathbf{X}_j(m), \quad (8)$$

where  $\mathbf{S}_{j,i}(\cdot) = \mathbf{S}_{i,j}^*(\cdot)$ . For the two-channel case, the update equations can be very easily written in an explicit form [1], e.g. for the first channel, we have

$$\begin{aligned} \hat{\mathbf{H}}_1(m+1) &= \hat{\mathbf{H}}_1(m) + \mu\mathbf{S}_1^{-1}[\mathbf{X}_1^*(m) - \\ &\quad - \mathbf{S}_{1,2}\mathbf{S}_{2,2}^{-1}\mathbf{X}_2^*(m)]\tilde{\mathbf{E}}(m), \\ \mathbf{S}_i(m) &= \mathbf{S}_{i,i}(m)[\mathbf{I}_{2L \times 2L} - \\ &\quad - \mathbf{S}_{1,2}^*(m)\mathbf{S}_{1,2}(m)\{\mathbf{S}_{1,1}(m)\mathbf{S}_{2,2}(m)\}^{-1}]. \end{aligned} \quad (9)$$

## 2.2. Dynamical Regularization

In most practical scenarios, the desired signal  $y(k)$  in Fig. 1 is disturbed, e.g. by some acoustic background noise. According to perturbation theory [8], the parameter estimation (i.e. misalignment in M-C AEC [1]) is very sensitive in poorly excited frequency bins. For robust adaptation the power spectral densities  $\mathbf{S}_{i,i}$  are replaced by regularized versions according to  $\tilde{\mathbf{S}}_{i,i} = \mathbf{S}_{i,i} + \text{diag}\{\delta_i\}$  prior to inversion in Eq. (9). The basic feature of the regularization is a compromise between fidelity to data and fidelity to some prior information about the solution [9], which increases the robustness, but leads to biased solutions. Therefore, we propose here a novel *bin-selective dynamical regularization vector*

$$\delta_i(m) = \delta_{max} [e^{-S_{i,i}^{(0)}(m)/S_0} \cdots e^{-S_{i,i}^{(2L-1)}(m)/S_0}]^T \quad (10)$$

with two scalar parameters  $\delta_{max}$  and  $S_0$ .  $S_{i,i}^{(\nu)}$  denotes the  $\nu$ -th frequency component ( $\nu = 0, \dots, 2L-1$ ) on the main diagonal of  $\mathbf{S}_{i,i}$ . Note that for implementation,  $e$  in Eq. (10) may be replaced by a basis 2 and modified  $S_0$ .

This exponential method provides a smooth transition between regularization for low input power and data fidelity whenever the input power is high enough, and yields improved results compared to fixed regularization and to the popular approach of choosing the maximum out of the respective component  $\mathbf{S}_{i,i}^{(\nu)}$  and a fixed threshold  $\delta_{th}$  (Fig. 2). It copes well with asymmetric excitation of the input channels, and most importantly, it can be easily extended for the Kalman gain calculation introduced in the next section.

## 2.3. Efficient calculation of the frequency-domain Kalman gain

The solutions of  $\mathbf{S}(m)\mathbf{K}(m) = \mathbf{X}^H(m)$  (Eq. (6)) for more than two channels may be formulated similarly to the corresponding part of the stereo update equations in Eq. (9) (e.g. using Cramer's rule). For three channels, we have (omitting, for simplicity, the time index  $m$  of all matrices)

$$\begin{aligned} \mathbf{K}_1 &= \mathbf{D}^{-1}[\mathbf{X}_1^*(\mathbf{S}_{2,2}\mathbf{S}_{3,3} - \mathbf{S}_{3,2}\mathbf{S}_{2,3}) - \mathbf{X}_2^*(\mathbf{S}_{1,2}\mathbf{S}_{3,3} - \\ &\quad - \mathbf{S}_{1,3}\mathbf{S}_{3,1}) - \mathbf{X}_3^*(\mathbf{S}_{1,3}\mathbf{S}_{2,2} - \mathbf{S}_{1,2}\mathbf{S}_{2,3})], \\ \mathbf{D} &:= \mathbf{S}_{1,1}(\mathbf{S}_{2,2}\mathbf{S}_{3,3} - \mathbf{S}_{3,2}\mathbf{S}_{2,3}) - \mathbf{S}_{2,1}(\mathbf{S}_{1,2}\mathbf{S}_{3,3} - \\ &\quad - \mathbf{S}_{1,3}\mathbf{S}_{3,1}) - \mathbf{S}_{3,1}(\mathbf{S}_{1,3}\mathbf{S}_{2,2} - \mathbf{S}_{1,2}\mathbf{S}_{2,3}) \end{aligned} \quad (11)$$

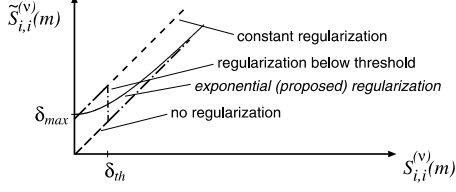


Fig. 2. Different regularization methods (channel  $i$ , bin  $\nu$ )

as the first of the three Kalman gain components with the common factor  $\mathbf{D}$ . The representations of Eqs. (9) and (11) allow an intuitive interpretation: a correction of the interchannel-correlations in  $\mathbf{K}_i$  between  $\mathbf{X}_i^*$  and the other input signals  $\mathbf{X}_j^*$ ,  $j \neq i$ .

However, for a *practical implementation* of a system with  $P > 2$  channels, we propose computationally more efficient methods to calculate Eq. (6) as follows.

Due to the block diagonal structure of Eq. (6), it can be simply decomposed into  $2L$  equations

$$\mathbf{K}^{(\nu)}(m) = (\mathbf{S}^{(\nu)}(m))^{-1} (\mathbf{X}^{(\nu)})^H(m) \quad (12)$$

with (small)  $P \times P$  unitary and positive definite matrices  $\mathbf{S}^{(\nu)}$  for the components  $\nu = 0, \dots, 2L - 1$  on the diagonals. Both  $\mathbf{K}^{(\nu)}$  and  $\mathbf{X}^{(\nu)}$  are *vectors* of length  $P$ . This property has been pointed out independently by [10]. Note that for real input signals  $x_i$  we need to solve Eq. (12) only for  $L + 1$  bins.

A well-known and numerically stable method for this type of problems is the Cholesky decomposition of  $\mathbf{S}^{(\nu)}$  followed by solution via backsubstitution, see [8]. The resulting total complexity for one output value is then

$$O(P \cdot \log(2L)) + O(P^3), \quad (13)$$

where in the two-channel (stereo) case the second term  $O(P^3)$  is much smaller than the share due to the first term.

For a large number of input channels (e.g. loudspeakers for surround sound or sound field synthesis for virtual reality) we introduce a recursive solution of Eq. (12) that jointly estimates the *inverse* power spectra  $(\mathbf{S}^{(\nu)})^{-1}$  (Eq. (8)). Using the matrix-inversion lemma, e.g. [3], we immediately obtain from Eq. (8) an update equation for the *inverse* matrices

$$(\mathbf{S}^{(\nu)}(m))^{-1} = \lambda^{-1} \left[ (\mathbf{S}^{(\nu)}(m-1))^{-1} - \frac{(\mathbf{S}^{(\nu)}(m-1))^{-1} \mathbf{X}^{(\nu)H}(m) \mathbf{X}^{(\nu)}(m) (\mathbf{S}^{(\nu)}(m-1))^{-1}}{\lambda(1-\lambda)^{-1} + \mathbf{X}^{(\nu)}(m) (\mathbf{S}^{(\nu)}(m-1))^{-1} \mathbf{X}^{(\nu)H}(m)} \right]$$

using the decomposition Eq. (12) (making the denominator in Eq. (14) a scalar value).

Introduction of the common vector

$$\mathbf{T}^{(\nu)}(m) = \mathbf{X}^{(\nu)}(m) (\mathbf{S}^{(\nu)}(m-1))^{-1} \quad (14)$$

in the numerator and the denominator leads to

$$(\mathbf{S}^{(\nu)}(m))^{-1} = \lambda^{-1} (\mathbf{S}^{(\nu)}(m-1))^{-1} - \frac{\mathbf{T}^{(\nu)}(m) (\mathbf{T}^{(\nu)}(m))^H}{\lambda^2 (1-\lambda)^{-1} + \mathbf{X}^{(\nu)}(m) \mathbf{T}^{(\nu)}(m)}. \quad (15)$$

The Kalman gain  $\mathbf{K}^{(\nu)}(m) = (\mathbf{S}^{(\nu)})^{-1} (\mathbf{X}^{(\nu)})^H(m)$  can then be efficiently calculated (using Eq. (15)) by

$$\mathbf{K}^{(\nu)}(m) = \mathbf{T}^{(\nu)}(m) \cdot \left[ \lambda^{-1} - \frac{(\mathbf{T}^{(\nu)}(m))^H (\mathbf{X}^{(\nu)}(m))^H}{\lambda^2 (1-\lambda)^{-1} + \mathbf{X}^{(\nu)}(m) \mathbf{T}^{(\nu)}(m)} \right] \quad (16)$$

Again, there are common factors in Eq. (16) (in the scalar expression in the brackets) and Eq. (15).

Note that our approach should not be confused with the classical RLS approach [3] which also makes use of the matrix-inversion lemma. As we apply the lemma independently to small  $P \times P$  systems (Eq. (12)) it is numerically much less critical than in the RLS algorithm. Moreover, there is no analogon to a more efficient *fast RLS* due to the different matrix structures (vector  $\mathbf{X}^{(\nu)}(m)$  does not reflect a tapped delay line).

## 2.4. Dynamical Regularization for Proposed Kalman Gain Approach

Due to the recursion Eq. (15) the regularization according to Eq. (10) is not immediately applicable. Therefore, an equivalent modification directly is applied to the data matrices  $\mathbf{X}^{(\nu)}(m)$  by addition of mutually uncorrelated white noise sequences to each channel and frequency bin, respectively.

Using the modified signal vectors, denoted by

$$\tilde{\mathbf{X}}^{(\nu)}(m) = \mathbf{X}^{(\nu)}(m) + \mathbf{N}^{(\nu)}(m), \quad (17)$$

where  $\mathbf{N}^{(\nu)}(m)$  are the vectors of the white noise signals, we obtain the modified power spectral density matrices

$$\tilde{\mathbf{S}}^{(\nu)}(m) \approx (1-\lambda) \sum_{q=0}^m \lambda^{m-q} \mathbf{X}^{(\nu)H}(q) \mathbf{X}^{(\nu)}(q) + (1-\lambda) \sum_{q=0}^m \lambda^{m-q} \text{diag}\{|N_1^{(\nu)}(q)|^2 \dots |N_P^{(\nu)}(q)|^2\}^T. \quad (18)$$

The diagonal elements of the second term can be interpreted as a bin-selective dynamical regularization vector  $\delta^{(\nu)}(m)$  with elements (for channel  $i$  and bin  $\nu$ )

$$\begin{aligned} \delta_i^{(\nu)}(m) &= (1-\lambda) \sum_{q=0}^m \lambda^{m-q} |N_i^{(\nu)}(q)|^2, \\ &= \lambda \delta_i^{(\nu)}(m-1) + (1-\lambda) |N_i^{(\nu)}(m)|^2. \end{aligned} \quad (19)$$

Thus, in order to update the regularization from  $\delta_i^{(\nu)}(m-1)$  to  $\delta_i^{(\nu)}(m)$  with the appropriate speed (determined by  $\lambda$ ), we need to add noise with power

$$|N_i^{(\nu)}(m)|^2 = \frac{\delta_i^{(\nu)}(m) - \lambda \delta_i^{(\nu)}(m-1)}{1-\lambda}. \quad (20)$$

On the other hand, according to Eq. (10), the regularization should be chosen according to

$$\begin{aligned} \delta_i^{(\nu)}(m) &= \delta_{max} e^{-\frac{S_{i,i}^{(\nu)}(m)}{S_0}} \\ &= \delta_{max} e^{-\frac{\lambda S_{i,i}^{(\nu)}(m-1) + (1-\lambda) |x_i^{(\nu)}(m)|^2}{S_0}}. \end{aligned} \quad (21)$$

Now, unlike other dynamical regularization methods, the exponential regularization allows simple elimination of the elements  $S_{i,i}^{(\nu)}(m-1)$  of the non-inverted matrix (which need not be computed at all), since

$$\begin{aligned} \delta_i^{(\nu)}(m) &= \delta_{max} \left( e^{-\frac{S_{i,i}^{(\nu)}(m-1)}{S_0}} \right)^\lambda e^{-\frac{(1-\lambda)|X_i^{(\nu)}(m)|^2}{S_0}} \\ &= \delta_{max}^{1-\lambda} (\delta_i^{(\nu)}(m-1))^\lambda e^{-\frac{(1-\lambda)|X_i^{(\nu)}(m)|^2}{S_0}}. \end{aligned} \quad (22)$$

Note that the above equations for dynamical regularization can be derived analogously for the well-known single- and multi-channel time-domain RLS algorithm as well.

### 3. APPLICATION TO MULTICHANNEL ACOUSTIC ECHO CANCELLATION

For the simulations, a speech signal (in the transmission room) was convolved by  $P$  different room impulse responses and non-linearly, but inaudibly preprocessed according to [1] ( $P$  different nonlinearities with factor 0.5). The lengths of the receiving room impulse responses were 4096 and the modeling filters were 1024, respectively. A white noise signal for  $SNR = 35dB$  was added to the echo on the microphone. Fig. 3 shows the misalignment and  $ERLE$  (echo return loss enhancement) convergence of the described algorithm (solid) for the multi-channel cases  $P = 2, 3, 4, 5$ , and the basic NLMS [3] (dashed) for comparison. The overlap factor  $\alpha$  was set to 4 for  $P = 2$  and adjusted to 8 for  $P = 3, 4$  and to 16 for  $P = 5$ . Note that for high overlap factors  $\alpha$ , an efficient recursive DFT computation [11] can be used for Eq. (1). Using these parameters, the convergence curves for the different numbers of channels are almost indistinguishable. Fig. 4 compares different regularization methods (white noise distortion as above): no regularization (uppermost curve), constant regularization (dotted), threshold (dashed), exponential with original algorithm (dash-dot), proposed Kalman gain (lower solid line). The complexity for computing the Kalman gains (for one output value  $e(k)$ ) are compared in Fig. 5.

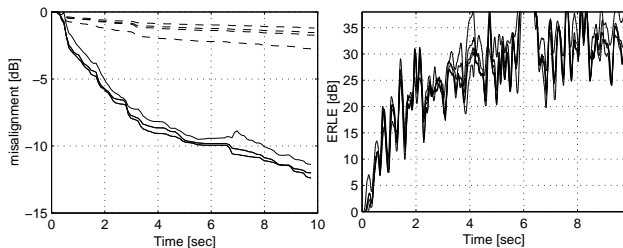


Fig. 3. Convergence for  $P=2,3,4,5$  channels and adjusted  $\alpha$

### 4. CONCLUSIONS

A new recursive method to efficiently calculate the frequency-domain Kalman gain for multichannel frequency-domain adaptive filtering has been presented. The recursive method is based on a decomposition of block diagonal matrices followed by an application of the matrix-inversion lemma. Moreover, we introduced a new dynamical regularization method suitable for the recursive Kalman gain computation. The algorithm is especially interesting for multichannel acoustic echo cancellation. Results for real-world

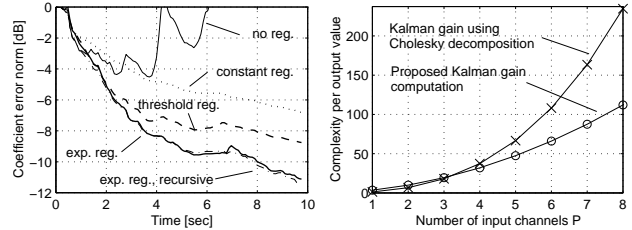


Fig. 4. Comparison of regularization methods,  $P = 5$

Fig. 5. Complexity of Kalman gain

conditions have been shown for more than two reproduction channels. The convergence curves of the system misalignment and the ERLE clearly indicate that the approach considered here is able to cope very well with scenarios such as 5 channel surround sound with highly correlated loudspeaker signals.

### 5. ACKNOWLEDGEMENTS

The authors would like to thank J. Benesty of Bell Laboratories, Lucent Technologies, and D.T.M. Slock of Institut Eurecom for stimulating discussions.

### 6. REFERENCES

- [1] S. L. Gay and J. Benesty (eds.), *Acoustic Signal Processing for Telecommunication*, Kluwer Academic Publishers, 2000.
- [2] M. M. Sondhi and D. R. Morgan, "Stereo Acoustic Echo Cancellation - An Overview of the Fundamental Problem," *IEEE SP Letters*, vol.2, no.8, pp. 148-151, Aug. 1995.
- [3] S. Haykin, *Adaptive Filter Theory*, 3rd ed., Prentice Hall Inc., Englewood Cliffs, NJ, 1996
- [4] J. Benesty and D. R. Morgan, "Frequency-domain adaptive filtering revisited, generalization to the multi-channel case, and application to acoustic echo cancellation," in *Proc. IEEE ICASSP*, Istanbul, Turkey, pp. 789-792, June 2000.
- [5] H. Buchner, W. Herboldt, and W. Kellermann, "An Efficient Combination of Multi-Channel Acoustic Echo Cancellation With a Beamforming Microphone Array," *Proc. Int. Workshop on Hands-Free Speech Communication*, Kyoto, Japan, pp. 55-58, April 2001.
- [6] 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.
- [7] 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.
- [8] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed., Johns Hopkins, Baltimore, MD, 1989.
- [9] V.A. Morozov, *Regularization Methods for Ill-posed Problems*, CRC Press, Boca Raton, FL, 1993.
- [10] J. Benesty, personal communication.
- [11] H. Buchner and W. Kellermann, "Acoustic Echo Cancellation for Two and More Reproduction Channels," *Proc. Int. Workshop on Acoustic Echo and Noise Control*, Darmstadt, Germany, pp. 99-102, Sept. 2001.