KALMAN FILTER BASED STEREO SYSTEM IDENTIFICATION WITH AUTO- AND CROSS-DECORRELATION

Stefan Kühl, Christiane Antweiler, Tobias Hübschen, Peter Jax

Institute of Communication Systems (IKS)
RWTH Aachen University, Germany

ABSTRACT

In stereo or multichannel system identification, the most critical problems regarding online identification, e.g., for acoustic echo control, are the correlation properties of the excitation signals of the different audio channels. In this paper the impact of both the auto- and cross-correlation properties is considered and investigated. A new system combining appropriate decorrelation techniques with a Kalman filter driven adaptation algorithm in the frequency domain is presented. For the auto-decorrelation a new structure is proposed where the signals in the adaptation paths are decorrelated via linear prediction without affecting the acoustic signals. A small non-linearity is added into each channel for the reduction of the cross-correlation between channels. The performance evaluation clearly shows the influence of the different countermeasures and the effectiveness of the combined approach.

Index Terms— Stereo system identification, Kalman filter, linear prediction, decorrelation, acoustic echo cancellation.

1. INTRODUCTION

In stereo system identification, in the past much attention has been given to the so-called non-uniqueness problem [1–3], i.e., the system identification cannot provide a unique transmission path solution if the stereo audio signals originate from the same source. As solution to high cross-correlation between the two incoming audio channels, numerous approaches have been proposed to mitigate this non-uniqueness problem [4–8]. One proven approach is to add a moderate non-linearity to the signals improving the system identification with minimal introduced distortion. For the adaptation algorithm itself, cost-efficient solutions like multichannel LMS algorithms exist. To tackle the problem of colored signal input, algorithms such as RLS, APA variants [9–11] and in the recent years Kalman filter based algorithms either in the time [12,13] or the frequency domain [14–16] have been proposed. In [17], however, it has been shown that in the single channel case also sophisticated algorithms such as the Kalman filter based algorithm suffer to a certain extent from autocorrelated signal input and that decorrelation techniques improve identification. Consequently, it can be expected that in the stereo case autocorrelation *as well as* cross-correlation between the channels affect the identification.

The main novelty in this paper is the usage of auto- *and* cross-decorrelation methods in combination. In a first step, we extend the concept of the Kalman filter based algorithm with linear prediction techniques to a stereo system. In a second step, a small non-linearity is applied to both input signals aiming at the non-uniqueness problem. For this new approach the impact of the two combined decorrelation methods is investigated in terms of a stereo echo cancellation application.

The paper is structured as follows. In Sec. 1.1 the underlying system model is introduced and the frequency domain Kalman filter reviewed. Afterwards in Sec. 2, principle experiments regarding the impact of autocorrelation and cross-correlation on stereo system identification motivate the implementation and integration of appropriate decorrelation techniques into one system. A performance evaluation in terms of various simulations concludes the paper.

1.1. System Model

Figure 1 shows the block diagram of the stereo system identification approach using Kalman filtering. The excitation signals for the first and second channel and the measurement noise are denoted by $x_1(i)$, $x_2(i)$ and $s(i)$, respectively, with the time index i. The impulse responses of the two unknown acoustic paths are denoted by the vectors $\mathbf{w}_1(i)$ and $\mathbf{w}_2(i)$. Throughout this paper, vectors and matrices are denoted by boldface letters in contrast to scalars. Assuming a finite length L of the impulse responses, they can be represented by the vectors

$$\mathbf{w}_\nu(i) = \big(w_{\nu,1}(i), w_{\nu,2}(i), \dots, w_{\nu,L}(i)\big)^T, \tag{1}$$

with $(\cdot)^T$ denoting the transpose of the vector while $\nu \in \{1, 2\}$ marks the channel number. Thus, the inner product $d_\nu(i) = \mathbf{x}_\nu^T(i)\mathbf{w}_\nu(i)$ with the excitation vectors

$$\mathbf{x}_\nu(i) = \big(x_\nu(i), x_\nu(i-1), \dots, x_\nu(i-L+1)\big)^T \tag{2}$$

refers to the system response for each channel. The measured signal, consisting of the two system responses $d_\nu(i)$ and the measurement noise $s(i)$, is denoted by $y(i)$. For the estimation of $\mathbf{w}_\nu(i)$, two adaptive filters $\hat{\mathbf{w}}_\nu(i)$ of length L are used. The resulting error signal $e(i) = y(i) - \hat{d}_1(i) - \hat{d}_2(i)$ as well as the excitation signals $x_\nu(i)$ serve as inputs for the identification algorithm, marked by the Kalman block in Fig.1.

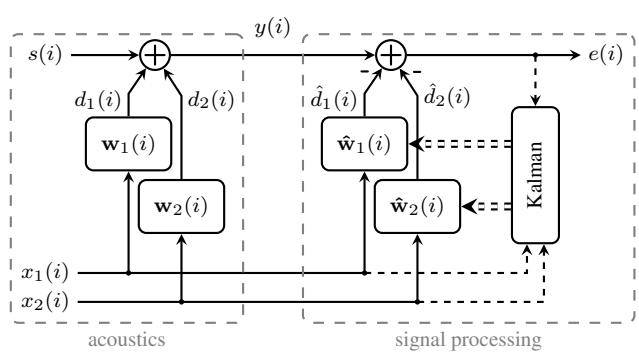

Fig. 1. Stereo system identification with Kalman filter

Within our system, the *fully diagonalized* Kalman filter adaptation in the frequency domain with an overlap-save framework is used [16, 18, 19] and extended to the two-channel problem. The processing is performed blockwise in frames of length M with frameshift R and frame index k. In this context, the vectors

$$\mathbf{x}_{\nu,M}(k) = \big(x_\nu(kR-M+1), x_\nu(kR-M+2), \dots, x_\nu(kR)\big)^T$$
$$\mathbf{y}_R(k) = \big(y(kR-R+1), y(kR-R+2), \dots, y(kR)\big)^T, \quad (3)$$

the Fourier-matrix $\mathbf{F}_M$ of size $M \times M$, and the zero-padding matrix $\mathbf{Q}_R = \big(\mathbf{0}_{M-R} \quad \mathbf{I}_R\big)^T$ are defined. $\mathbf{I}_R$ is the identity-matrix of size $R \times R$ and $\mathbf{0}_{M-R}$ the zero-matrix of size $R \times (M-R)$. In the following, $(\cdot)^H$ denotes the Hermitian and $(\cdot)^{-1}$ the inverse of a matrix. With these definitions, the *fully diagonalized* Kalman equations can be extended to the stereo case according to

$$\mathbf{K}_\nu(k) = \frac{R}{M}\mathbf{P}_\nu\mathbf{X}_\nu^H(k)\mathbf{D}^{-1}(k) \quad (4a)$$
$$\hat{\mathbf{W}}_\nu^+(k) = \hat{\mathbf{W}}_\nu(k) + \mathbf{K}_\nu(k)\mathbf{E}(k) \quad (4b)$$
$$\mathbf{P}_\nu^+(k) = \Big(\mathbf{I}_M - \frac{R}{M}\mathbf{K}_\nu(k)\mathbf{X}_\nu(k)\Big)\mathbf{P}_\nu(k) \quad (4c)$$
$$\hat{\mathbf{W}}_\nu(k+1) = A \cdot \hat{\mathbf{W}}_\nu^+(k) \quad (4d)$$
$$\mathbf{P}_\nu(k+1) = A^2 \cdot \mathbf{P}_\nu^+(k) + \mathbf{\Psi}_{\Delta\Delta_\nu}(k)\,, \quad (4e)$$

with $\mathbf{X}_\nu(k) = \mathrm{diag}\{\mathbf{F}_M \cdot \mathbf{x}_{\nu,M}(k)\}$ and forgetting factor A. The entities

$$\mathbf{D}(k) = \frac{R}{M}\sum_{\nu=1}^{2}\mathbf{X}_\nu(k)\mathbf{P}_\nu(k)\mathbf{X}_\nu^H(k) + \mathbf{\Psi}_{ss}(k) \quad (5)$$

$$\mathbf{E}(k) = \mathbf{Y}(k) - \mathbf{F}_M\mathbf{Q}\mathbf{Q}_R^H\mathbf{F}_M^{-1}\sum_{\nu=1}^{2}\mathbf{X}_\nu(k)\hat{\mathbf{W}}_\nu(k) \quad (6)$$

are the same for both channels and, thus, are responsible for a coupled adaptation. In the overlap-save framework, the estimates $\hat{\mathbf{W}}_\nu(k)$ are constrained by zeroing of the last $R-1$ values in the time domain to avoid cyclic artifacts. The $M \times M$ matrices $\mathbf{P}_\nu(k)$ represent estimates of the auto-covariance matrix of the estimation error in each channel. Due to the fact that the cross-covariance of the estimation errors of each channel vanishes with the convergence of the algorithm, the cross terms are omitted. $\mathbf{\Psi}_{ss}(k)$ and $\mathbf{\Psi}_{\Delta\Delta_\nu}(k)$ denote the covariance matrices of the measurement noise at the microphone and the process noise in each channel, respectively, which can be approximated by diagonalized estimates.

2. CORRELATION

The influence of the autocorrelation and cross-correlation of the excitation signals on the Kalman filter based adaptation can be visualized with a simple experiment for the single and the two-channel case each excited either by correlated (speech) and uncorrelated (white noise) signal input. For the speech input a stereo signal, measured in a sound proof booth with two microphones placed closely together, was used. For the white noise input two channels of uncorrelated noise were generated. At the microphone we added white measurement noise with an SNR of 40 dB. For the simulation, slowly time variant impulse responses measured in an empty sound proof booth ($T_{60} \approx 100-200$ ms) according to [20, 21] of length 192 were used, allowing for reproducible simulations and objective assessment in case of time variant acoustic room impulse responses. The sampling rate is 8 kHz and the forgetting factor A of the Kalman algorithm

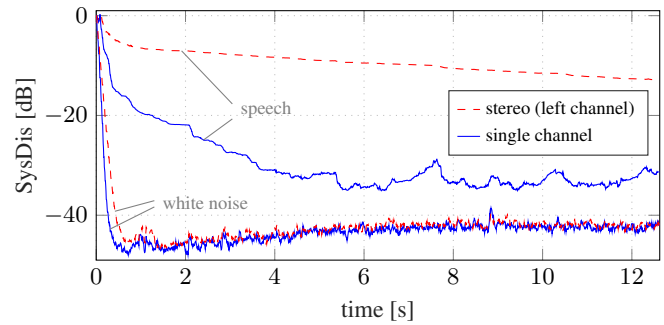

Fig. 2. Comparison of single and dual channel adaption. ($x_\nu(i)$: speech or white noise, $s(i)$: white noise, $A = 0.99995$, SNR $= 40$ dB, $\mathbf{w}_{\nu,\mathrm{slow}}(i)$)

was set to 0.99995. Figure 2 illustrates the results in terms of the logarithmic system distance

$$\mathrm{SysDis}(i)\,[\mathrm{dB}] = 10\lg\left(\frac{\|\mathbf{w}(i)-\hat{\mathbf{w}}(i)\|^2}{\|\mathbf{w}(i)\|^2}\right). \quad (7)$$

For the ease of clarity, for the stereo case the system distance regarding the left channel only is shown.

In the single channel case, the identification obviously converges much faster with white noise excitation, which additionally yields better steady state values compared to the speech excitation. As described in [17], this is due to the autocorrelation properties of the speech signal. It has been shown that even the Kalman filter based algorithm (in the frequency domain) suffers from auto-correlated signal input. In the stereo case, the identification process for speech excitation fails more or less in providing the unique echo path solution. In contrast, for a white noise stereo excitation almost the same excellent results as in the single channel case are obtained. Only a slight degradation in terms of convergence speed can be observed, which is to be expected due to the doubled number of coefficients to be identified. So besides the autocorrelation, also the cross-correlation between the two excitation signals $x_1(i)$ and $x_2(i)$ affects the identification performance, known as non-uniqueness problem. Thus, for the stereo problem countermeasures for the decorrelation of both, autocorrelation and cross-correlation, are needed.

2.1. Autocorrelation Countermeasures

In this section a method for auto-decorrelation based on linear prediction (LP) is presented. It is mainly based on the approach for the single channel case [17] and is extended accordingly to the stereo problem. Note, that the later description and derivation follows the equations in [17] with additional indexing of the channels and coupling of the adaptation of the two channels in terms of $\mathbf{D}(k)$ and $\mathbf{E}(k)$.

For the derivation of the new system, we start with the introduction of a *reference* structure as given in Fig. 3. All input signals, including the measurement noise, are passed through linear prediction filters such that the system is excited with the decorrelated signals $x_{\nu,e}(i)$, marked by an additional index e.

The linear prediction analysis is performed only on $x_1(i)$. The resulting prediction filter of degree P, represented by its impulse response $\mathbf{a}_{LP}(i) = (a_0(i), a_1(i), \dots, a_P(i))^T$, is applied to all input signals. This leads to an (optimal) decorrelation of $x_1(i)$ and, in case of stereo signals originated from the same source, also to a decorrelation of $x_2(i)$. Note that different individual filters for the two channels are not applicable due to the required retransformation of $e_e(i)$. The

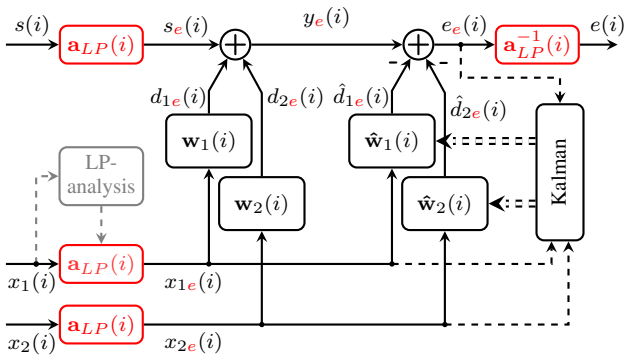

Fig. 3. Stereo system identification with Kalman filter with linear prediction techniques in a *reference* structure.

reference structure, only introduced for a better understanding of the new system to be derived, is obviously not implementable, because the residual signals would need to be transmitted over the acoustic channel and the measurement noise $s(i)$ is not separately available.

In order to develop a realizable structure, we aim to shift the decorrelation filters into the adaptation paths such that a decoupling of transmission and adaptation can be achieved. As one part of this process, the upper LP-filter in the "$s(i)$ path" is system theoretically shifted over the summation points and swapped with the filters $\mathbf{w}_\nu(i)$ and $\hat{\mathbf{w}}_\nu(i)$. This introduces an error due to the time variance of the involved filters. In case of swapping with $\mathbf{w}_\nu(i)$, the time variance is relatively small and the introduced error can be neglected in practice. The error caused by the time variance of $\hat{\mathbf{w}}_\nu(i)$, however, is significant and has to be compensated for by the introduction of a so-called *refiltering stage* [17]. The refiltering eliminates the time dependency on former LP-coefficients $\mathbf{a}_{LP}$ and former filter coefficients $\hat{\mathbf{w}}_\nu$ such that the inputs for the adaptation process depend only on the current sets $\mathbf{a}_{LP}(i)$ and $\hat{\mathbf{w}}_\nu(i)$. In this way the involved filters appear to be time invariant for the current adaptation step and can be swapped without causing an error. The corresponding blockdiagram of the realizable new system is depicted in Fig. 4.

For the refiltering of the excitation signals $x_\nu(i)$, the matrices

$$\mathbf{x}_{\nu,\text{states}}(k) = \begin{pmatrix} x_\nu(kR-M+1) & \cdots & x_\nu(kR-M-P+1) \\ \vdots & & \vdots \\ x_\nu(kR) & \cdots & x_\nu(kR-P) \end{pmatrix}$$

which contain the filter states of the decorrelation filters $\mathbf{a}_{LP}(kR)$ of each channel for the last M time instances, are defined. The decorrelated and refiltered $x_\nu(i)$ can now be formulated in the frequency domain according to

$$\mathbf{X}_{\nu,e}^r(k) = \text{diag}\left\{\mathbf{F}_M \cdot \mathbf{x}_{\nu,\text{states}}(k) \cdot \mathbf{a}_{LP}(kR)\right\}\,, \tag{8}$$

which depends only on the current set of coefficients $\mathbf{a}_{LP}(kR)$. The corresponding decorrelated and refiltered error signal in the frequency domain can then be calculated via

$$\mathbf{E}_e^r(k) = \mathbf{F}_M\mathbf{Q}_R\mathbf{Q}_R^H\mathbf{F}_M^{-1}\mathbf{A}_{LP}(k)\mathbf{F}_M\mathbf{Q}_{R+P}\Big(\mathbf{y}_{R+P}(k) - \mathbf{Q}_{R+P}^H\mathbf{F}_M^{-1}\sum_{\nu=1}^{2}\mathbf{X}_\nu(k)\hat{\mathbf{W}}_\nu(k)\Big) \tag{9}$$

with $\mathbf{y}_{R+P}(k)$ and $\mathbf{Q}_{R+P}$ defined analogically to (3) and $\mathbf{Q}_R$, respectively. The matrix

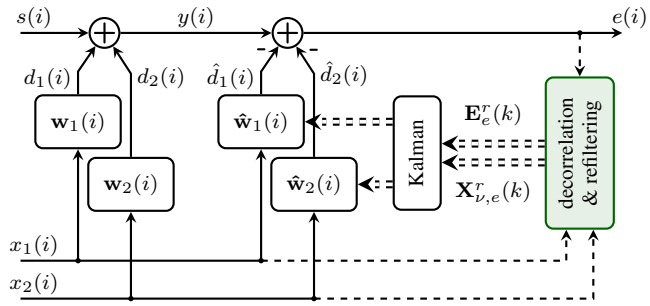

Fig. 4. Stereo system identification with Kalman filter with linear prediction techniques and refiltering stage.

$$\mathbf{A}_{LP}(k) = \text{diag}\left\{\mathbf{F}_M\begin{pmatrix}\mathbf{a}_{LP}(kR)\\ \mathbf{0}_{M-P-1}\end{pmatrix}\right\} \tag{10}$$

represents the impulse response of the decorrelation filter in the frequency domain, with $\mathbf{0}_{M-P-1}$ as a column vector containing $(M-P-1)$ zeros. The term in brackets in (9) has to yield $R+P$ valid values in contrast to R values for the ordinary Kalman algorithm in (6) to allow for R cyclic-free samples resulting from the afterwards performed decorrelation by the frequency domain multiplication with $\mathbf{A}_{LP}(k)$. If (8) and (9) are used instead of $\mathbf{X}_\nu(k)$ and $\mathbf{E}(k)$, respectively, the *fully diagonalized* stereo Kalman equations with linear prediction

$$\mathbf{K}_\nu(k) = \frac{R}{M}\mathbf{P}_\nu\mathbf{X}_{\nu,e}^{r\,H}(k)\mathbf{D}_e^{-1}(k) \tag{11a}$$

$$\hat{\mathbf{W}}_\nu^+(k) = \hat{\mathbf{W}}_\nu(k) + \mathbf{K}_\nu(k)\mathbf{E}_e^r(k) \tag{11b}$$

$$\mathbf{P}_\nu^+(k) = \Big(\mathbf{I}_M - \frac{R}{M}\mathbf{K}_\nu(k)\mathbf{X}_{\nu,e}^r(k)\Big)\mathbf{P}_\nu(k) \tag{11c}$$

$$\hat{\mathbf{W}}_\nu(k+1) = A\cdot\hat{\mathbf{W}}_\nu^+(k) \tag{11d}$$

$$\mathbf{P}_\nu(k+1) = A^2\cdot\mathbf{P}_\nu^+(k) + \mathbf{\Psi}_{\Delta\Delta_\nu}(k)\,, \tag{11e}$$

with

$$\mathbf{D}_e(k) = \frac{R}{M}\sum_{\nu=1}^{2}\mathbf{X}_{\nu,e}^r(k)\mathbf{P}_\nu(k)\mathbf{X}_{\nu,e}^{r\,H}(k) + \mathbf{\Psi}_{ss,e}(k) \tag{12}$$

can be derived.

The prediction analysis is performed every R samples on the last M samples of $x_1(i)$ to ensure that the prediction is performed on the relevant samples. As positive side effect, the prediction causes no algorithmic delay and is adapted for each frame.

2.2. Cross-Correlation Countermeasures

In the past, many different approaches like, e.g., the introduction of non-linearities [2, 22, 23], allpass filtering [4, 6, 24], exploiting psychoacoustical effects [7, 8, 25] or decorrelation in sub-bands [26–28] have been proposed to reduce the cross-correlation among the channels. All these approaches have in common that they modify the acoustically transmitted signal. Therefore, the tradeoff between tolerable distortion and amount of decorrelation has to be chosen carefully. In this paper a non-linearity with a half-wave rectifier (HWR) is applied, which achieves a good compromise between good decorrelation performance and little distortion because the shape of the non-linearity is adapted to the input signal. The resulting cross-decorrelated signals are calculated by

$$x_{\nu,\text{hwr}}(i) = x_\nu(i) + \frac{\alpha}{2}\Big(x_\nu(i) + (-1)^{\nu+1}|x_\nu(i)|\Big). \tag{13}$$

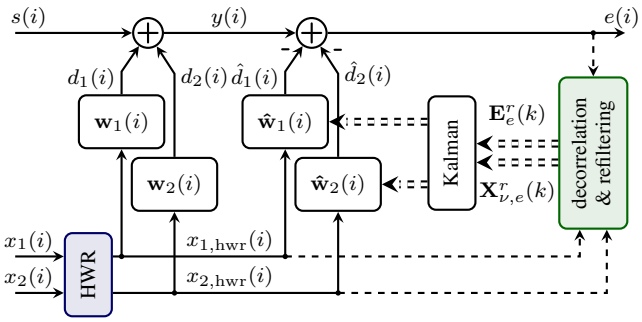

Fig. 5. Stereo system identification with Kalman filter with cross-decorrelation via half-wave rectifier and auto-decorrelation via linear prediction techniques with refiltering stage.

The factor α determines the amount of non-linearity in the signal and has been chosen to a value of 0.3 in all simulations, which leads to non audible distortions.

2.3. Combined System

By combining the decorrelation methods from the last two sections, the system depicted in Fig. 5 can be developed. This system is now capable to perform both, auto- and cross-decorrelation of the excitation signals. The LP-analysis for the prediction coefficients as well as the prediction filtering itself is carried out on the half-wave rectified signals $x_{\nu,\mathrm{hwr}}(i)$. The next interesting question is, how much impact the different countermeasures have.

3. SIMULATION RESULTS

For a performance evaluation, simulations of the new proposed system in an acoustic echo control (AEC) application were carried out for different scenarios. Besides the stationary, slowly time variant impulse responses $\mathbf{w}_{\nu,\mathrm{slow}}(i)$ (see Sec. 2), also strongly time variant impulse responses $\mathbf{w}_{\nu,\mathrm{fast}}(i)$ have been applied. The corresponding measurements for $\mathbf{w}_{\nu,\mathrm{fast}}(i)$ were performed with an object moving between the loudspeakers and microphone in the soundproof booth, see [20, 21] for more details concerning reproducible time variant room impulse response simulations. As a simple time variance indication (TVI) of $\mathbf{w}_{\nu,\mathrm{slow}}(i)$ and $\mathbf{w}_{\nu,\mathrm{fast}}(i)$, respectively, the logarithmic system distance between successive impulse responses of the left channel is calculated. The parameters of the Kalman filter are set to $M = 256$, $R = 64$ and $P = 2$, as numerous simulations verified that in this context a prediction degree of 2 is already sufficient.

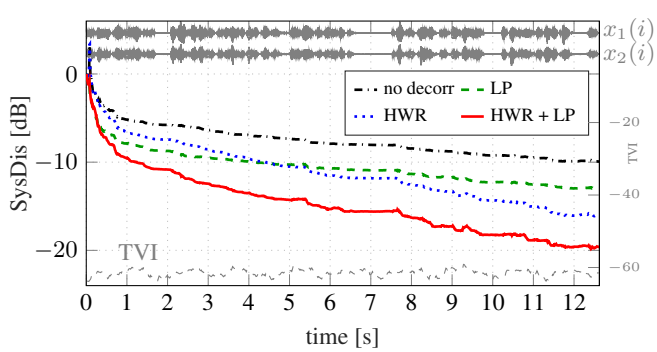

Fig. 6. Comparison of different decorrelation approaches for good conditions. ($x_\nu(i)$: speech, $s(i)$: white noise, $A = 0.99995$, SNR $=$ 30 dB, $\mathbf{w}_{\nu,\mathrm{slow}}(i)$)

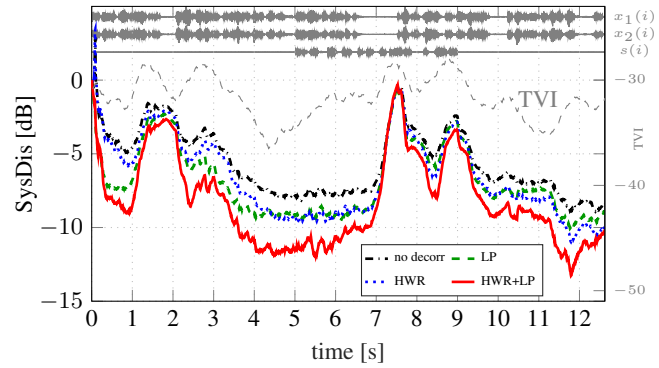

Fig. 7. Comparison of different decorrelation approaches for severe conditions. ($x_\nu(i)$, $s(i)$: speech, $A = 0.999$, SNR $=$ 0 dB in double talk periods, SNR $=$ 40 dB in far-end single talk periods, $\mathbf{w}_{\nu,\mathrm{fast}}(i)$)

Fig. 6 shows the results for a stereo speech excitation and white measurement noise $s(i)$ with an SNR of 30 dB at the microphone and slowly time varying impulse responses $\mathbf{w}_{\nu,\mathrm{slow}}(i)$. This setup represents a realistic "best-case" scenario for far-end single talk. Again only the system distance of the left channel is depicted. Note, that the results of the right channel are very similar. The results in case of no decorrelation, autocorrelation and cross-correlation countermeasures, and their combination are given.

As expected, the Kalman filter without any decorrelating measures (-·-) converges slowly suffering from the high degree of correlation. Comparing the separate usage of auto- and cross-decorrelation via the LP (- -) and HWR (···), respectively, shows that both measures lead to improvements. The improvement of auto-decorrelation has an impact especially in the initial convergence phase, leading to a faster adaptation in the beginning. The cross-decorrelation, however, improves the conditioning regarding the non-uniqueness problem and allows for a better global adaptation. By combining the two methods (—), a significant improvement of the convergence behavior can be achieved. Both countermeasures leverage at different points and complement each other in their effects.

The results of a "worst-case" scenario are shown in Fig. 7 and follow the same principal tendency. Here, impulse responses $\mathbf{w}_{\nu,\mathrm{fast}}(i)$ are used during single and double talk periods. The combination of HWR and LP (—) again yields the best performance. The strong misalignment after 7 s is dominated by the extreme time fluctuation of the impulse responses at that time.

4. CONCLUSION

In this paper a Kalman filter based stereo system identification with auto- and cross-decorrelation is introduced. For the auto-decorrelation we derived a new structure including linear prediction filters in the adaptation paths in order to tackle the conditioning of the estimation problem. For the non-uniqueness problem a small non-linearity is added into each channel reducing the cross-correlation. In summary, both countermeasures contribute to the correlation problem. As a result, the combined approach benefits from auto- and cross-decorrelation leading in our examples to improvements in an order of 10 dB (best-case scenario) and up to 5 dB (worst-case scenario). It is of special interest that the LP filtering does not affect the signal quality of the transmitted signals and that the added non-linearity can be chosen such that it is perceptually not noticeable. Moreover, all presented concepts can be generalized to multichannel systems.

5. REFERENCES

[1] M. M. Sondhi, D. R. Morgan, and J. L. Hall, "Stereophonic acoustic echo cancellation - an overview of the fundamental problem," *IEEE Signal Processing Letters*, vol. 2, no. 8, pp. 148–151, Aug. 1995.

[2] J. Benesty, D. R. Morgan, and M. Sondhi, "A better understanding and an improved solution to the problems of stereophonic acoustic echo cancellation," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Apr. 1997, vol. 1, pp. 303–306.

[3] S. Makino, "Stereophonic acoustic echo cancellation: An overview and recent solutions," *Acoustical Science and Technology*, vol. 22, no. 5, pp. 325–333, Jan. 2001.

[4] M. Ali, "Stereophonic acoustic echo cancellation system using time-varying all-pass filtering for signal decorrelation," in *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 1998, vol. 6, pp. 3689–3692.

[5] Y. Huang and J. Benesty, Eds., *Audio Signal Processing for Next-Generation Multimedia Communication Systems*, Springer US, 2004.

[6] N. Tangsangiumvisai, J. A. Chambers, and A. G. Constantinides, "Higher-order time-varying allpass filters for signal decorrelation in stereophonic acoustic echo cancellation," *Electronics Letters*, vol. 35, no. 1, pp. 88–90, Jan. 1999.

[7] L. Romoli, S. Cecchi, P. Peretti, and F. Piazza, "A mixed decorrelation approach for stereo acoustic echo cancellation based on the estimation of the fundamental frequency," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 2, pp. 690–698, Feb. 2012.

[8] L. Romoli, S. Cecchi, and F. Piazza, "A novel decorrelation approach for multichannel system identification," in *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 6652–6656.

[9] S. L. Gay and J. Benesty, Eds., *Acoustic Signal Processing for Telecommunication*, Kluwer Academic Publishers, 2000.

[10] J. Benesty, T. Gänsler, D. R. Morgan, M. Sondhi, and S. L. Gay, *Advances in Network and Acoustic Echo Cancellation*, Digital Signal Processing. Springer-Verlag, 2001.

[11] E. Hänsler and G. Schmidt, *Acoustic Echo and Noise Control – A Practical Approach*, John Wiley & Sons, 2004.

[12] C. Paleologu, J. Benesty, S.L. Grant, and S. Ciochină, "Widely linear general Kalman filter for stereophonic acoustic echo cancellation," *Signal Processing*, vol. 94, pp. 570–575, 2014.

[13] C. Paleologu, J. Benesty, S.L. Grant, and S. Ciochină, "A Kalman filter for stereophonic acoustic echo cancellation," in *Proc. of Asilomar Conference on Signals, Systems and Computers*, Nov. 2014, pp. 1639–1643.

[14] S. Malik and G. Enzner, "Recursive bayesian control of multichannel acoustic echo cancellation," *IEEE Signal Processing Letters*, vol. 18, no. 11, pp. 619–622, Nov. 2011.

[15] S. Malik and G. Enzner, "State-space frequency-domain adaptive filtering for nonlinear acoustic echo cancellation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 7, pp. 2065–2079, Sep. 2012.

[16] S. Malik and J. Benesty, "Variationally diagonalized multichannel state-space frequency-domain adaptive filtering for acoustic echo cancellation," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2013, pp. 595–599.

[17] S. Kühl, C. Antweiler, T. Hübschen, and P. Jax, "Kalman filter based system identification exploiting the decorrelation effects of linear prediction," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Mar. 2017, Accepted for publication.

[18] 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, 2006.

[19] G. Enzner, *A Model-based Optimum Filtering Approach to Acoustic Echo Control: Theory and Practice*, Ph.D. thesis, RWTH Aachen University, 2006.

[20] C. Antweiler and H. G. Symanzik, "Simulation of time variant room impulse responses," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 1995, vol. 5, pp. 3031–3034.

[21] C. Antweiler, A. Telle, and P. Vary, "NLMS-type system identification of MISO systems with shifted perfect sequences," in *Proc. of the International Workshop on Acoustic Echo and Noise Control (IWAENC)*, Seattle, Sep. 2008.

[22] D. R. Morgan, J. L. Hall, and J. Benesty, "Investigation of several types of nonlinearities for use in stereo acoustic echo cancellation," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 6, pp. 686–696, Sep. 2001.

[23] T. Gänsler and J. Benesty, "An adaptive nonlinearity solution to the uniqueness problem of stereophonic echo cancellation," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2002, vol. 2, pp. II–1885–II–1888.

[24] J. M. Valin, "Perceptually-motivated nonlinear channel decorrelation for stereo acoustic echo cancellation," in *Proc. of Joint Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA)*, May 2008, pp. 188–191.

[25] S. Cecchi, L. Romoli, P. Peretti, and F. Piazza, "Low-complexity implementation of a real-time decorrelation algorithm for stereophonic acoustic echo cancellation," *Signal Processing*, vol. 92, no. 11, pp. 2668 – 2675, 2012.

[26] J. Herre, H. Buchner, and W. Kellermann, "Acoustic echo cancellation for surround sound using perceptually motivated convergence enhancement," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Apr. 2007, vol. 1, pp. I–17–I–20.

[27] B. C. Bispo and D. d. S. Freitas, "Hybrid pre-processor based on frequency shifting for stereophonic acoustic echo cancellation," in *Proccedings of European Signal Processing Conference (EUSIPCO)*, Aug. 2012, pp. 2447–2451.

[28] M. L. Valero and E. A. P. Habets, "Insight into a phase modulation technique for signal decorrelation in multi-channel acoustic echo cancellation," in *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 2016, pp. 519–523.