EUSIPCO 2013 1569726733

ROBUST GENERAL KALMAN FILTER FOR ECHO CANCELLATION

Constantin Paleologu[†]*, Jacob Benesty*[‡]*, Silviu Ciochină*[†]*, and Victor Popescu*[†]

[†] University Politehnica of Bucharest, Romania, e-mail: {pale, silviu, vpopescu}@comm.pub.ro
[‡] INRS-EMT, University of Quebec, Montreal, Canada, e-mail: benesty@emt.inrs.ca

ABSTRACT

The Kalman filter is widely used in many practical applications. In this paper, we motivate the use of a more general form of the Kalman filter in the context of echo cancellation. This algorithm is derived by considering, at each iteration, a block of time samples instead of one time sample as it is the case in the conventional approach. The variance of the near-end signal explicitly appears within the algorithm, which allows us to better control it. Using a proper estimation of this parameter, the general Kalman filter is robust to large near-end signals (like double-talk). Simulation results indicate the appealing performance of the proposed algorithm.

***Index Terms*—** Echo cancellation, Kalman filter, adaptive filters, recursive least-squares (RLS) algorithm.

1. INTRODUCTION

The echo cancellation application [1] can be formulated as a system identification problem. In this context, the main goal is to estimate an unknown system, i.e., the echo path, from a set of noisy observations, i.e., the microphone (reference) signal that contains the echo and the near-end signal (e.g., the background noise and the near-end speech).

Basically, the Kalman filter [2] estimates a set of unknown variables based on a set of (noisy) observations acquired over time. This algorithm recursively provides an optimal estimate of these variables, which is based on the Bayesian approach. The Kalman filter and different versions of it have been involved in a wide range of applications [3]. However, despite its appealing performance, the Kalman filter has been avoided in the context of echo cancellation.

The most important study of the Kalman filter in the context of echo cancellation, is due to G. Enzner and his co-authors [4], [5], [6], [7]. Most of this work has been focusing on the development of efficient frequency-domain Kalman filters that have the potential to outperform the classical frequency-domain adaptive filters. The motivation behind the frequency-domain approach is mainly related to the complexity issue. Also, other related works can be found in [8], [9], [10].

This work was supported by the UEFISCDI Romania under Grant PN-II-RU-TE no. 7/5.08.2010 and Grant PN-II-ID-PCE-2011-3-0097.

In this work, we present another form of the Kalman filter by considering, at each iteration, a block of time samples instead of one time sample as it is the case in the conventional approach. The variance of the near-end signal explicitly appears within the algorithm. This allows us to better control the filter, in terms of its robustness to near-end signal variations. Simulations show that the proposed general Kalman filter is robust to large near-end signals, like double-talk.

2. STATE VARIABLE MODEL FOR ECHO CANCELLATION

In the context of echo cancellation, the microphone or desired signal at the discrete-time index n is

$$d(n) = \mathbf{x}^T(n)\mathbf{h} + v(n) = y(n) + v(n), \tag{1}$$

where $\mathbf{x}(n) = \begin{bmatrix} x(n) & x(n-1) & \cdots & x(n-L+1) \end{bmatrix}^T$ is a vector containing the L most recent time samples of the input (loudspeaker) signal $x(n)$, superscript ^{T} denotes transpose of a vector or a matrix, $\mathbf{h} = \begin{bmatrix} h_0 & h_1 & \cdots & h_{L-1} \end{bmatrix}^T$ is the impulse response (of length L) of the system (from the loudspeaker to the microphone) that we need to identify, and $v(n)$ is a zero-mean stationary white Gaussian noise signal. The variance of this additive noise is $\sigma_v^2 = E\left[v^2(n)\right]$, where $E[\cdot]$ denotes mathematical expectation. The signal $y(n)$ is called the echo in the context of echo cancellation that we want to cancel with an adaptive filter [1], [11].

The objective is to estimate or identify $\mathbf{h}$ with an adaptive filter, $\widehat{\mathbf{h}}(n) = \begin{bmatrix} \widehat{h}_0(n) & \widehat{h}_1(n) & \cdots & \widehat{h}_{L-1}(n) \end{bmatrix}^T$, in such a way that for a reasonable value of n, we have for the (normalized) misalignment:

$$\frac{\left\|\widehat{\mathbf{h}}(n) - \mathbf{h}\right\|_2^2}{\|\mathbf{h}\|_2^2} \leq \iota, \tag{2}$$

where ι is a predetermined small positive number and $\|\cdot\|_2$ is the ℓ_2 norm.

We are now going to model the system impulse response as a state equation. By considering the P most recent time samples of the microphone signal, (1) can be expressed as

$$\begin{aligned}\mathbf{d}(n) &= \left[\begin{array}{cccc} d(n) & d(n-1) & \cdots & d(n-P+1) \end{array}\right]^T \\ &= \mathbf{y}(n) + \mathbf{v}(n), \end{aligned} \quad (3)$$

where

$$\mathbf{y}(n) = \mathbf{X}^T(n)\mathbf{h}(n) \quad (4)$$

is the echo signal vector of length P,

$$\mathbf{X}(n) = \left[\begin{array}{cccc} \mathbf{x}(n) & \mathbf{x}(n-1) & \cdots & \mathbf{x}(n-P+1) \end{array}\right] \quad (5)$$

is the input signal matrix of size $L \times P$, and the noise signal vector, $\mathbf{v}(n)$, is defined similarly to $\mathbf{d}(n)$. In our context, $\mathbf{X}^T(n)$ is the measurement matrix and $x(n)$ is considered as deterministic. Expression (3) is called the observation equation. We assume that $\mathbf{h}(n)$ is a zero-mean random vector, which follows a simplified first-order Markov model, i.e.,

$$\mathbf{h}(n) = \mathbf{h}(n-1) + \mathbf{w}(n), \quad (6)$$

where $\mathbf{w}(n)$ is a zero-mean white Gaussian noise signal vector, which is uncorrelated with $\mathbf{h}(n-1)$ and $\mathbf{v}(n)$. The correlation matrix of $\mathbf{w}(n)$ is assumed to be $\mathbf{R_w}(n) = \sigma_w^2(n)\mathbf{I}_L$, where $\mathbf{I}_L$ is the $L \times L$ identity matrix. The variance, $\sigma_w^2(n)$, captures the uncertainties in $\mathbf{h}(n)$. Expression (6) is called the state equation.

Now, the echo cancellation problem may be restated as follows. Given the two fundamental equations:

$$\begin{aligned}\mathbf{h}(n) &= \mathbf{h}(n-1) + \mathbf{w}(n), \\ \mathbf{d}(n) &= \mathbf{X}^T(n)\mathbf{h}(n) + \mathbf{v}(n), \end{aligned} \quad (7)$$

our objective is to find the optimal recursive estimate of $\mathbf{h}(n)$ denoted by $\widehat{\mathbf{h}}(n)$. In the context of echo cancellation, the values of $\sigma_w^2(n)$ play a major role in the performance of the estimator. Indeed, small values of $\sigma_w^2(n)$ imply a low misalignment but a poor tracking; while large values of $\sigma_w^2(n)$ (meaning that the uncertainties in the echo path are high) imply a good tracking but a high misalignment. In other words, the values of $\sigma_w^2(n)$ highly determine the tracking abilities and the convergence of the Kalman filter to be derived.

3. GENERAL KALMAN FILTER

From the simplified (with respect to the state equation) model presented in the previous section, we can derive the Kalman filter. It is well known that, in the context of the linear sequential Bayesian approach, the optimum estimate of the state vector, $\mathbf{h}(n)$, has the form [12]:

$$\begin{aligned}\widehat{\mathbf{h}}(n) &= \widehat{\mathbf{h}}(n-1) + \mathbf{K}(n)\left[\mathbf{d}(n) - \mathbf{X}^T(n)\widehat{\mathbf{h}}(n-1)\right] \\ &= \widehat{\mathbf{h}}(n-1) + \mathbf{K}(n)\mathbf{e}(n), \end{aligned} \quad (8)$$

where $\mathbf{K}(n)$ is the Kalman gain matrix and

$$\begin{aligned}\mathbf{e}(n) &= \mathbf{d}(n) - \widehat{\mathbf{y}}(n) \\ &= \mathbf{d}(n) - \mathbf{X}^T(n)\widehat{\mathbf{h}}(n-1) \end{aligned} \quad (9)$$

is the a priori error signal vector between the microphone signal vector and the estimate of the echo signal vector. The a posteriori error signal vector is defined as

$$\begin{aligned}\boldsymbol{\epsilon}(n) &= \mathbf{d}(n) - \mathbf{X}^T(n)\widehat{\mathbf{h}}(n) \\ &= \mathbf{X}^T(n)\boldsymbol{\mu}(n) + \mathbf{v}(n), \end{aligned} \quad (10)$$

where

$$\boldsymbol{\mu}(n) = \mathbf{h}(n) - \widehat{\mathbf{h}}(n) \quad (11)$$

is the state estimation error or a posteriori misalignment. The correlation matrix of $\boldsymbol{\mu}(n)$ is

$$\mathbf{R}_{\boldsymbol{\mu}}(n) = E\left[\boldsymbol{\mu}(n)\boldsymbol{\mu}^T(n)\right]. \quad (12)$$

We can also define the a priori misalignment as

$$\begin{aligned}\mathbf{m}(n) &= \mathbf{h}(n) - \widehat{\mathbf{h}}(n-1) \\ &= \boldsymbol{\mu}(n-1) + \mathbf{w}(n), \end{aligned} \quad (13)$$

for which its correlation matrix is

$$\begin{aligned}\mathbf{R_m}(n) &= E\left[\mathbf{m}(n)\mathbf{m}^T(n)\right] \\ &= \mathbf{R}_{\boldsymbol{\mu}}(n-1) + \sigma_w^2(n)\mathbf{I}_L. \end{aligned} \quad (14)$$

It is clear that the a priori misalignment appears in the a priori error signal vector as

$$\mathbf{e}(n) = \mathbf{X}^T(n)\mathbf{m}(n) + \mathbf{v}(n). \quad (15)$$

The Kalman gain matrix is obtained by minimizing the criterion:

$$J(n) = \frac{1}{L}\mathrm{tr}\left[\mathbf{R}_{\boldsymbol{\mu}}(n)\right] \quad (16)$$

with respect to $\mathbf{K}(n)$, where $\mathrm{tr}\left[\cdot\right]$ denotes the trace of a square matrix. We easily find that

$$\mathbf{K}(n) = \mathbf{R_m}(n)\mathbf{X}(n)\left[\mathbf{X}^T(n)\mathbf{R_m}(n)\mathbf{X}(n) + \sigma_v^2\mathbf{I}_P\right]^{-1}, \quad (17)$$

where $\mathbf{I}_P$ is the $P \times P$ identity matrix and

$$\mathbf{R}_{\boldsymbol{\mu}}(n) = \left[\mathbf{I}_L - \mathbf{K}(n)\mathbf{X}^T(n)\right]\mathbf{R_m}(n). \quad (18)$$

The correlation matrix of the a priori error signal vector is

$$\mathbf{R_e}(n) = \mathbf{X}^T(n)\mathbf{R_m}(n)\mathbf{X}(n) + \sigma_v^2\mathbf{I}_P, \quad (19)$$

which inverse appears explicitly in the Kalman gain matrix.

Consequently, the following equations summarize the general Kalman filter (GKF) [10]:

$$\begin{aligned} \mathbf{R_m}(n) &= \mathbf{R}_{\boldsymbol{\mu}}(n-1) + \sigma_w^2(n)\mathbf{I}_L, && (20) \\ \mathbf{R_e}(n) &= \mathbf{X}^T(n)\mathbf{R_m}(n)\mathbf{X}(n) + \sigma_v^2\mathbf{I}_P, && (21) \\ \mathbf{K}(n) &= \mathbf{R_m}(n)\mathbf{X}(n)\mathbf{R_e}^{-1}(n), && (22) \\ \mathbf{e}(n) &= \mathbf{d}(n) - \mathbf{X}^T(n)\widehat{\mathbf{h}}(n-1), && (23) \\ \widehat{\mathbf{h}}(n) &= \widehat{\mathbf{h}}(n-1) + \mathbf{K}(n)\mathbf{e}(n), && (24) \\ \mathbf{R}_{\boldsymbol{\mu}}(n) &= \left[\mathbf{I}_L - \mathbf{K}(n)\mathbf{X}^T(n)\right]\mathbf{R_m}(n). && (25) \end{aligned}$$

The initialization is $\widehat{\mathbf{h}}(n) = \mathbf{0}$ and $\mathbf{R}_{\boldsymbol{\mu}}(0) = \varepsilon\mathbf{I}_L$, where ε is a small positive constant. When $\sigma_w^2(n) = 0$, we have

$$\begin{aligned} \lim_{n\to\infty} \mathbf{R}_{\boldsymbol{\mu}}(n) &= \mathbf{0}, && (26) \\ \lim_{n\to\infty} \mathbf{K}(n) &= \mathbf{0}, && (27) \end{aligned}$$

and, obviously, the Kalman filter will never be able to track the changes in $\mathbf{h}(n)$. On the other hand, for large values of $\sigma_w^2(n)$, the Kalman gain matrix never goes to zero, which allows the update equation (24) to stay "alert" to any possible random changes of the echo path.

One very important particular case of the GKF is when $P = 1$. Consequently, we get the well-known Kalman filter [13]:

$$\begin{aligned} \mathbf{R_m}(n) &= \mathbf{R}_{\boldsymbol{\mu}}(n-1) + \sigma_w^2(n)\mathbf{I}_L, && (28) \\ \sigma_e^2(n) &= \mathbf{x}^T(n)\mathbf{R_m}(n)\mathbf{x}(n) + \sigma_v^2, && (29) \\ \mathbf{k}(n) &= \frac{1}{\sigma_e^2(n)}\mathbf{R_m}(n)\mathbf{x}(n), && (30) \\ e(n) &= d(n) - \mathbf{x}^T(n)\widehat{\mathbf{h}}(n-1), && (31) \\ \widehat{\mathbf{h}}(n) &= \widehat{\mathbf{h}}(n-1) + \mathbf{k}(n)e(n), && (32) \\ \mathbf{R}_{\boldsymbol{\mu}}(n) &= \left[\mathbf{I}_L - \mathbf{k}(n)\mathbf{x}^T(n)\right]\mathbf{R_m}(n). && (33) \end{aligned}$$

This algorithm has striking resemblances with the recursive least-squares (RLS) algorithm [13]. However, contrary to what it may be believed, the two algorithms are very much different and do not behave the same way in practice; this is perhaps the reason why the Kalman filter was not really so deeply studied in echo cancellation except by G. Enzner [5]. We also believe that it is even confusing to try to compare the two algorithms from a theoretical point of view. Nevertheless, there are at least four fundamental differences between these two filters. First, the Kalman filter does not require any matrix inversion, which is not the case for the RLS (whose inverse input signal correlation matrix is implicitly calculated at each iteration). Second, the Kalman filter depends explicitly on the correlation matrix of the misalignment while the RLS adaptive filter depends on the (inverse) correlation matrix of the input signal. Third, the RLS does not depend on the variance of the additive noise. Finally, the RLS does not depend on the uncertainties in $\mathbf{h}(n)$ since it is considered as deterministic in its derivation. The two parameters $\sigma_w^2(n)$ and σ_v^2 in the Kalman filter (for which the RLS does not depend on) allow us to better control it.

4. PRACTICAL CONSIDERATIONS

There are two parameters that need to be set or estimated within the GKF. The first (and, perhaps, the most important) one is $\sigma_w^2(n)$, which plays a major role in the overall performance of the algorithm, as explained before. Using the ℓ_2 norm in both sides of (6), together with the approximation $\|\mathbf{w}(n)\|_2^2 \approx L\sigma_w^2(n)$ (which is valid when $L \gg 1$), replacing $\mathbf{h}(n)$ by its estimate $\widehat{\mathbf{h}}(n)$, and also considering the contribution of the model's order, we can evaluate

$$\widehat{\sigma}_w^2(n) = \frac{\left\|\widehat{\mathbf{h}}(n) - \widehat{\mathbf{h}}(n-1)\right\|_2^2}{PL}. \quad (34)$$

Let us remember that the value of $\sigma_w^2(n)$ needs to compromise between good tracking and low misalignment. In this sense, it is difficult to obtain a proper compromise by using a constant value of this parameter. Based on this motivation, the estimation from (34) is designed to achieve this goal. When the algorithm starts to converge or when there is an abrupt change of the system (e.g., when the echo path changes), the difference between $\widehat{\mathbf{h}}(n)$ and $\widehat{\mathbf{h}}(n-1)$ is significant, so that the parameter $\widehat{\sigma}_w^2(n)$ takes large values, thus providing fast convergence and tracking. On the other hand, when the algorithm starts to converge to its steady-state, the difference between $\widehat{\mathbf{h}}(n)$ and $\widehat{\mathbf{h}}(n-1)$ reduces, thus leading to small values of $\widehat{\sigma}_w^2(n)$ and, consequently, to a low misalignment.

The second parameter to be found is the noise power, σ_v^2. Usually, it can be estimated during silences of the near-end talker, i.e., in a single-talk scenario [14]. However, this is not always an easy task. The most critical situation in echo cancellation is the double-talk case, when the near-end signal is a combination of background noise and near-end speech. In this scenario, the parameter $\sigma_v^2(n)$ can be estimated as proposed in [15] or [16]. For example, assuming that the adaptive filter has converged to a certain degree, the near-end signal power can be evaluated as [16]

$$\widehat{\sigma}_v^2(n) = \left|\widehat{\sigma}_d^2(n) - \widehat{\sigma}_{\widehat{y}}^2(n)\right|, \quad (35)$$

where $\widehat{\sigma}_d^2(n)$ and $\widehat{\sigma}_{\widehat{y}}^2(n)$ are the power estimates of $d(n)$ and $\widehat{y}(n)$, respectively; $\widehat{y}(n)$ denotes the output of the adaptive filter at time index n. These parameters can be recursively evaluated as

$$\begin{aligned} \widehat{\sigma}_d^2(n) &= \beta\widehat{\sigma}_d^2(n-1) + (1-\beta)d^2(n), && (36) \\ \widehat{\sigma}_{\widehat{y}}^2(n) &= \beta\widehat{\sigma}_{\widehat{y}}^2(n-1) + (1-\beta)\widehat{y}^2(n), && (37) \end{aligned}$$

where $\beta = 1 - 1/(KL)$, with $K \geq 1$. The initial values are $\widehat{\sigma}_d^2(0) = 0$ and $\widehat{\sigma}_{\widehat{y}}^2(0) = 0$.

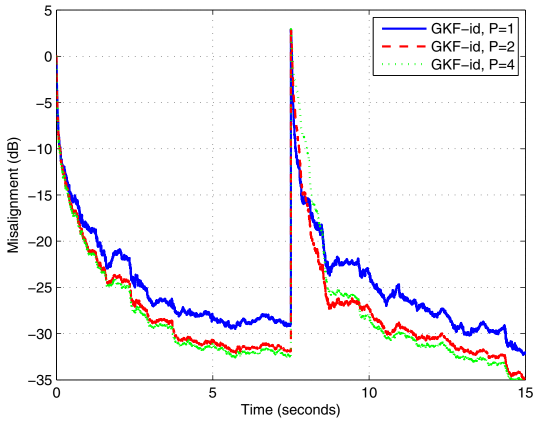

Fig. 1. Misalignment of the GKF-id with different values of P. The input signal is speech, $L = 128$, and SNR = 20 dB. Echo path changes at time 7.5 seconds.

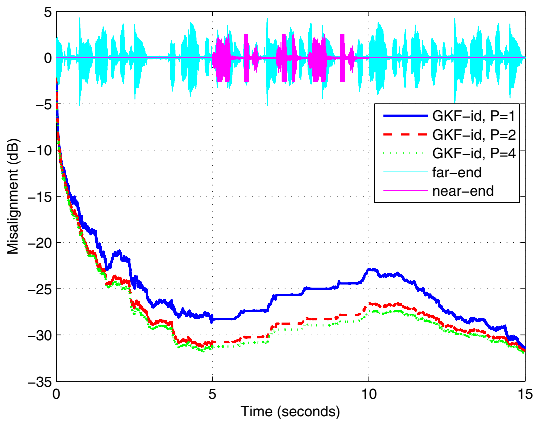

Fig. 2. Misalignment of the GKF-id with different values of P. The input signal is speech, $L = 128$, and SNR = 20 dB. The near-end speech (double-talk) appears at time 5 seconds until time 10 seconds (without using a DTD).

5. SIMULATION RESULTS

Simulations are performed in the context of echo cancellation. The echo path is the fourth impulse response from G168 Recommendation [17] (with 128 taps). The sampling rate is 8 kHz. All adaptive filters used in the experiments have the same length as the echo paths. The far-end signal (i.e., the input signal) is a speech sequence. The output of the echo path is corrupted by an independent white Gaussian noise with 20 dB signal-to-noise ratio (SNR). In order to evaluate the tracking capabilities of the algorithms, an echo path change scenario is simulated in some of the experiments, by shifting the impulse responses to the right by 12 samples. The performance measure used in simulations is the normalized misalignment (in dB) evaluated based on (2).

In the first set of experiments, we evaluate the performance of the GKF in "ideal" conditions, i.e., the near-end signal $v(n)$ is available; its power can be recursively evaluated similar to (36) and (37) as

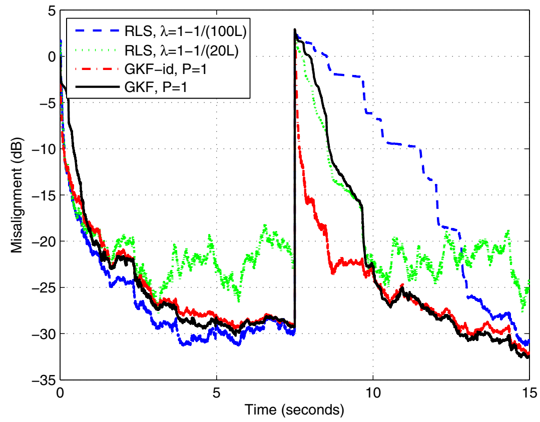

Fig. 3. Misalignment of the RLS algorithm with different values of the forgetting factor λ, and the GKF-id and GKF with $P = 1$. The input signal is speech, $L = 128$, and SNR = 20 dB. Echo path changes at time 7.5 seconds.

$$\widehat{\sigma}_v^2(n) = \beta\widehat{\sigma}_v^2(n-1) + (1-\beta)v^2(n). \quad (38)$$

We refer to this algorithm as the "ideal" GKF (GKF-id). In Fig. 1, the performance of the GKF-id is evaluated for different model orders, i.e., P = 1, 2, and 4. It can be noticed that the GKF-id achieves a good compromise between the tracking capability and the steady-state misalignment level. Also, it can be noticed that there is a significant performance improvement of the GKF-id with $P = 2$ over the case when $P = 1$. This improvement is not very significant when the value of P increases, as compared to the case with $P = 2$.

In Fig. 2, the performance of the GKF-id is evaluated in a double-talk situation, without using any double-talk detector (DTD). We can notice the robustness of the algorithm for different values of the model order P.

The second set of simulations outlines the performance of the GKF when using the near-end signal power estimate from (35). The model order is set to $P = 1$, in order to make a fair comparison with the RLS algorithm. In Fig. 3 , the GKF and GKF-id are compared with the RLS algorithm using different values of the forgetting factor ($0 < \lambda \leq 1$). This parameter of the RLS algorithm addresses the compromise between convergence rate/tracking capabilities on the one hand and misadjustment/stability on the other hand. As expected, the tracking reaction of the GKF is slower as compared to GKF-id, since (35) is based on the assumption that the adaptive filter has converged to a certain degree (which is biased when the echo path changes). However, the GKF compromises better between the tracking capability and steady-state misalignment level, as compared to the RLS algorithm. To improve the tracking reaction of the RLS algorithm, the value

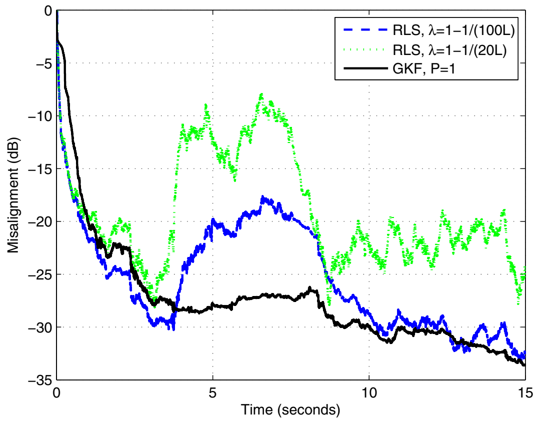

Fig. 4. Misalignment of the RLS algorithm with different values of the forgetting factor λ and the GKF with $P = 1$. The input signal is speech and $L = 128$. The SNR decreases from 20 dB to 10 dB between times 3.75 and 7.5 seconds.

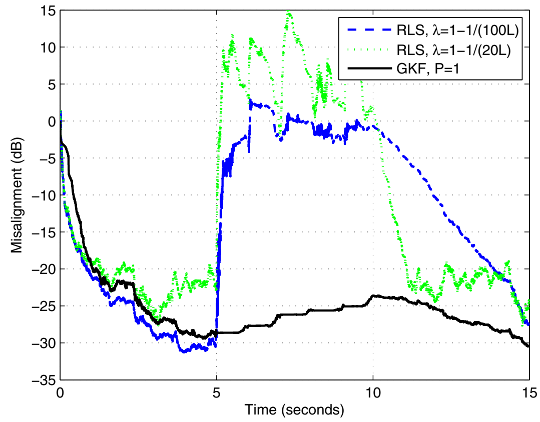

Fig. 5. Misalignment of the RLS algorithm with different values of the forgetting factor λ and the GKF with $P = 1$. The input signal is speech and $L = 128$. The near-end speech (double-talk) appears at time 5 seconds until time 10 seconds (without using a DTD) (see Fig. 2).

of the λ should be decreased, but this increases the misalignment and could affect the stability.

In Fig. 4, a variation of the background noise at the near-end is considered. The SNR decreases from 20 dB to 10 dB between times 3.75 and 7.5 seconds. The GKF is compared with the RLS algorithm using the same values of λ as in the previous simulation. It can be noticed that the GKF is very robust against the background noise variation, outperforming the RLS algorithm.

Finally, a double-talk situation is presented in Fig. 5, considering the same scenario as in Fig. 2. Again, the GKF shows good robustness against the near-end speech, which is not the case for the RLS algorithm.

6. CONCLUSIONS

In this paper, we have motivated the use of the Kalman filter for echo cancellation. The proposed GKF was derived based on a simplified first-order Markov model, which seems to work very well in this context. The variance of the near-end signal naturally appears within the algorithm. As a consequence, the GKF is very robust to large near-end signals, like double-talk. Of course, the main issue remains the computational complexity, which will be addressed in future work by developing fast versions of the GKF.

7. REFERENCES

[1] J. Benesty, T. Gänsler, D. R. Morgan, M. M. Sondhi, and S. L. Gay, *Advances in Network and Acoustic Echo Cancellation*. Berlin, Germany: Springer-Verlag, 2001.

[2] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Engineering*, vol. 82, pp. 35–45, Mar. 1960.

[3] R. Faragher, "Understanding the basis of the Kalman filter via a simple and intuitive derivation," *IEEE Signal Processing Magazine*, vol. 29, pp. 128–132, Sept. 2012.

[4] G. Enzner and P. Vary, "Frequency-domain adaptive Kalman filter for acoustic echo control in hands-free telephones," *Signal Processing*, vol. 86, pp. 1140–1156, 2006.

[5] G. Enzner, *A Model-Based Optimum Filtering Approach to Acoustic Echo Control: Theory and Practice*. Dissertation, RWTH Aachen, Aachener Beitrge zu digitalen Nachrichtensystemen, Vary P. (ed.), Wissenschaftsverlag Mainz, Aachen, June 2006.

[6] G. Enzner, "Bayesian inference model for applications of time-varying acoustic system identification," in *Proc. EUSIPCO*, 2010, pp. 2126–2130.

[7] S. Malik and G. Enzner, "State-space frequency-domain adaptive filtering for nonlinear acoustic echo cancellation," *IEEE Trans. Audio, Speech, Language Processing*, vol. 20, pp. 2065–2079, Sept. 2012.

[8] H. Buchner and W. Kellermann, "Improved Kalman gain computation for multichannel frequency-domain adaptive filtering and application to acoustic echo cancellation," in *Proc. IEEE ICASSP*, 2002, pp. II-1909–II-1912.

[9] K. Helwani, H. Buchner, and S. Spors, "On the robust and efficient computation of the Kalman gain for multichannel adaptive filtering with application to acoustic echo cancellation," in *Proc. ASILOMAR*, 2010, pp. 988–992.

[10] C. Paleologu, J. Benesty, and S. Ciochină, "Study of the general Kalman filter for echo cancellation," *IEEE Trans. Audio, Speech, Language Processing*, vol. 21, pp. 1539–1549 , Aug. 2013.

[11] C. Paleologu, J. Benesty, and S. Ciochină, *Sparse Adaptive Filters for Echo Cancellation*. Morgan & Claypool Publishers, Synthesis Lectures on Speech and Audio Processing, 2010.

[12] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Englewood Cliffs, NJ: Prentice Hall, 1993.

[13] A. H. Sayed and T. Kailath, "A state-space approach to adaptive RLS filtering," *IEEE Signal Processing Magazine*, vol. 11, pp. 18–60, July 1994.

[14] J. Benesty, H. Rey, L. Rey Vega, and S. Tressens, "A non-parametric VSS NLMS algorithm," *IEEE Signal Processing Lett.*, vol. 13, pp. 581–584, Oct. 2006.

[15] M. A. Iqbal and S. L. Grant, "Novel variable step size NLMS algorithm for echo cancellation," in *Proc. IEEE ICASSP*, 2008, pp. 241–244.

[16] C. Paleologu, S. Ciochină, and J. Benesty, "Variable step-size NLMS algorithm for under-modeling acoustic echo cancellation," *IEEE Signal Processing Lett.*, vol. 15, pp. 5–8, 2008.

[17] *Digital Network Echo Cancellers*, ITU-T Rec. G.168, 2002.