

# Study of the General Kalman Filter for Echo Cancellation

Constantin Paleologu, *Member, IEEE*, Jacob Benesty, and Silviu Ciochină, *Member, IEEE*

**Abstract**—The Kalman filter is a very interesting signal processing tool, which is widely used in many practical applications. In this paper, we study the Kalman filter in the context of echo cancellation. The contribution of this work is threefold. First, we derive a different 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. Second, we show how this general Kalman filter (GKF) is connected with some of the most popular adaptive filters for echo cancellation, i.e., the normalized least-mean-square (NLMS) algorithm, the affine projection algorithm (APA) and its proportionate version (PAPA). Third, a simplified Kalman filter is developed in order to reduce the computational load of the GKF; this algorithm behaves like a variable step-size adaptive filter. Simulation results indicate the good performance of the proposed algorithms, which can be attractive choices for echo cancellation.

**Index Terms**—Echo cancellation, Kalman filter, adaptive filters, recursive least-squares (RLS) algorithm, affine projection algorithm (APA), proportionate APA (PAPA), normalized least-mean-square (NLMS) algorithm.

## I. INTRODUCTION

IN 1960, R. E. Kalman proposed “a new approach to linear filtering and prediction problems” [1]. Fundamentally, the Kalman filter 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. Following [1], the Kalman filter and different versions of it have been involved in a wide range of applications [2].

The echo cancellation application [3] is one of the most challenging system identification problems. In both network and acoustic scenarios, 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 corrupted by different types of “noise” (e.g., the background noise

and the near-end speech). Despite its appealing performance, the Kalman filter has been avoided in this context.

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]–[7], [8]. Most of this work has been focusing on the development of an efficient frequency-domain Kalman filter that has the potential to outperform the classical frequency-domain adaptive filters. For example, in [4] a stochastic state-space model of the acoustic echo path has been formulated in the frequency domain, showing that the Kalman filter is related to the frequency-domain adaptive filters. The idea was further extended in [6], where the near-end disturbance has been included in the estimation framework, thus resulting a robust state-space frequency-domain adaptive filter. Very recently, a similar approach has been applied to multichannel non-linear acoustic echo cancellation [8]. The motivation behind the frequency-domain approach is mainly related with complexity and convergence features. In this context, other related works can be found in [9], [10], [11]. Another obvious possibility is to implement the time-domain Kalman filter in subbands.

In this paper, we study the time-domain Kalman filter in the context of echo cancellation. First, we derive 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 relation between the Kalman filter and the recursive least-squares (RLS) algorithm is well established; see [12] for example. However, the most popular adaptive filters for echo cancellation are the normalized least-mean-square (NLMS) algorithm, the affine projection algorithm (APA) and its proportionate version (PAPA) [13]. We also show how all these algorithms can be obtained as good approximations of the proposed general Kalman filter (GKF). Furthermore, we develop a simplified version of the GKF, which behaves like a variable step-size (VSS) adaptive filter. Due to its convergence features and moderate computational complexity, this simplified algorithm could represent an attractive choice for real-world applications.

The rest of the paper is organized as follows. In Sections II and III, we present the signal model for echo cancellation and the state variable model, respectively. The proposed GKF is developed in Section IV. The connections between the GKF and different other algorithms are explained in Sections V and VI. The simplified version of the GKF is developed in Section VII. Some practical considerations are provided in Section VIII. A detailed experimental study on these algorithms is conducted in Section IX. Finally, Section X concludes this work and outlines several perspectives.

Manuscript received October 04, 2012; revised January 25, 2013; accepted January 27, 2013. Date of publication February 07, 2013; date of current version April 25, 2013. 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. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Woon-Seng Gan.

C. Paleologu and S. Ciochină are with the Telecommunications Department, University Politehnica of Bucharest, Bucharest 060042, Romania (e-mail: constantin.paleologu@upb.ro; silviu.ciochina@upb.ro).

J. Benesty is with INRS-EMT, University of Quebec, Montreal, QC H5A 1K6, Canada (e-mail: benesty@emt.inrs.ca).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TASL.2013.2245654

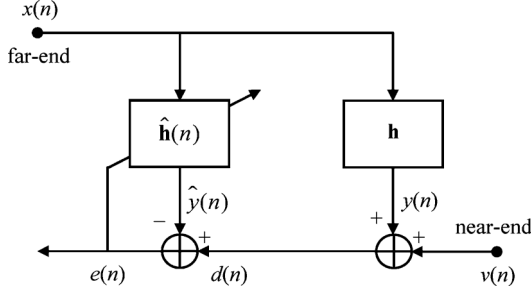


Fig. 1. General configuration for echo cancellation.

## II. SIGNAL MODEL FOR ECHO CANCELLATION

In the context of echo cancellation (Fig. 1), the microphone or desired signal at the discrete-time index  $n$  is

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

where

$$\mathbf{x}(n) = [x(n) \ x(n-1) \ \cdots \ x(n-L+1)]^T \quad (2)$$

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,

$$\mathbf{h} = [h_0 \ h_1 \ \cdots \ h_{L-1}]^T \quad (3)$$

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  $\sigma_v^2 = E[v^2(n)]$ , 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 [3], [13].

Then, our objective is to estimate or identify  $\mathbf{h}$  with an adaptive filter:

$$\hat{\mathbf{h}}(n) = [\hat{h}_0(n) \ \hat{h}_1(n) \ \cdots \ \hat{h}_{L-1}(n)]^T, \quad (4)$$

in such a way that for a reasonable value of  $n$ , we have for the (normalized) misalignment:

$$\frac{\|\hat{\mathbf{h}}(n) - \mathbf{h}\|_2^2}{\|\mathbf{h}\|_2^2} \leq \iota, \quad (5)$$

where  $\iota$  is a predetermined small positive number and  $\|\cdot\|_2$  is the  $\ell_2$  norm.

## III. STATE VARIABLE MODEL

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

$$\begin{aligned} \mathbf{d}(n) &= [d(n) \ d(n-1) \ \cdots \ d(n-P+1)]^T \\ &= \mathbf{y}(n) + \mathbf{v}(n), \end{aligned} \quad (6)$$

where

$$\mathbf{y}(n) = \mathbf{X}^T(n)\mathbf{h}(n) \quad (7)$$

is the echo signal vector of length  $P$ ,

$$\mathbf{X}(n) = [\mathbf{x}(n) \ \mathbf{x}(n-1) \ \cdots \ \mathbf{x}(n-P+1)] \quad (8)$$

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 (6) 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.,

$$\mathbf{h}(n) = \mathbf{h}(n-1) + \mathbf{w}(n), \quad (9)$$

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:

$$\mathbf{h}(n) = \mathbf{h}(n-1) + \mathbf{w}(n), \quad (10)$$

$$\mathbf{d}(n) = \mathbf{X}^T(n)\mathbf{h}(n) + \mathbf{v}(n), \quad (11)$$

our objective is to find the optimal recursive estimator of  $\mathbf{h}(n)$  denoted by  $\hat{\mathbf{h}}(n)$ . 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 good 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. Therefore, there is always a compromise between good tracking and low misalignment. The way we select this parameter in practice will be explained in Section VIII. As simulations will show, this simplified model is very satisfactory for the echo cancellation problem.

From (11), we can define the (nonstationary) echo-to-noise ratio (ENR) as

$$\text{ENR}(n) = \frac{\text{tr}[\mathbf{X}^T(n)\mathbf{R}_h(n)\mathbf{X}(n)]}{P\sigma_v^2}, \quad (12)$$

where  $\text{tr}[\cdot]$  denotes the trace of a square matrix and  $\mathbf{R}_h(n) = E[\mathbf{h}(n)\mathbf{h}^T(n)]$ . The ENR coincides with the classical signal-to-noise ratio (SNR) metric.

## IV. 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 [14]:

$$\begin{aligned} \hat{\mathbf{h}}(n) &= \hat{\mathbf{h}}(n-1) + \mathbf{K}(n)[\mathbf{d}(n) - \mathbf{X}^T(n)\hat{\mathbf{h}}(n-1)] \\ &= \hat{\mathbf{h}}(n-1) + \mathbf{K}(n)\mathbf{e}(n), \end{aligned} \quad (13)$$

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

$$\begin{aligned}\mathbf{e}(n) &= \mathbf{d}(n) - \hat{\mathbf{y}}(n) \\ &= \mathbf{d}(n) - \mathbf{X}^T(n)\hat{\mathbf{h}}(n-1)\end{aligned}\quad (14)$$

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

$$\begin{aligned}\boldsymbol{\epsilon}(n) &= \mathbf{d}(n) - \mathbf{X}^T(n)\hat{\mathbf{h}}(n) \\ &= \mathbf{X}^T(n)\boldsymbol{\mu}(n) + \mathbf{v}(n),\end{aligned}\quad (15)$$

where

$$\boldsymbol{\mu}(n) = \mathbf{h}(n) - \hat{\mathbf{h}}(n) \quad (16)$$

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

$$\mathbf{R}_{\boldsymbol{\mu}}(n) = E[\boldsymbol{\mu}(n)\boldsymbol{\mu}^T(n)]. \quad (17)$$

We can also define the a priori misalignment as

$$\begin{aligned}\mathbf{m}(n) &= \mathbf{h}(n) - \hat{\mathbf{h}}(n-1) \\ &= \boldsymbol{\mu}(n-1) + \mathbf{w}(n),\end{aligned}\quad (18)$$

for which its correlation matrix is

$$\begin{aligned}\mathbf{R}_{\mathbf{m}}(n) &= E[\mathbf{m}(n)\mathbf{m}^T(n)] \\ &= \mathbf{R}_{\boldsymbol{\mu}}(n-1) + \sigma_w^2(n)\mathbf{I}_L.\end{aligned}\quad (19)$$

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

$$\mathbf{e}(n) = \mathbf{X}^T(n)\mathbf{m}(n) + \mathbf{v}(n). \quad (20)$$

The Kalman gain matrix is obtained by minimizing the criterion:

$$J(n) = \frac{1}{L} \text{tr}[\mathbf{R}_{\boldsymbol{\mu}}(n)] \quad (21)$$

with respect to  $\mathbf{K}(n)$ . We easily find that

$$\mathbf{K}(n) = \mathbf{R}_{\mathbf{m}}(n)\mathbf{X}(n) [\mathbf{X}^T(n)\mathbf{R}_{\mathbf{m}}(n)\mathbf{X}(n) + \sigma_v^2(n)\mathbf{I}_P]^{-1}, \quad (22)$$

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

$$\mathbf{R}_{\boldsymbol{\mu}}(n) = [\mathbf{I}_L - \mathbf{K}(n)\mathbf{X}^T(n)]\mathbf{R}_{\mathbf{m}}(n). \quad (23)$$

It is of interest to observe that the correlation matrix of the a priori error signal vector is

$$\mathbf{R}_{\mathbf{e}}(n) = \mathbf{X}^T(n)\mathbf{R}_{\mathbf{m}}(n)\mathbf{X}(n) + \sigma_v^2(n)\mathbf{I}_P, \quad (24)$$

which inverse appears explicitly in the Kalman gain matrix. In the same way, the correlation matrix of the a posteriori error signal vector is

$$\mathbf{R}_{\boldsymbol{\epsilon}}(n) = \mathbf{X}^T(n)\mathbf{R}_{\boldsymbol{\mu}}(n)\mathbf{X}(n) + \sigma_v^2(n)\mathbf{I}_P. \quad (25)$$

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

$$\mathbf{R}_{\mathbf{m}}(n) = \mathbf{R}_{\boldsymbol{\mu}}(n-1) + \sigma_w^2(n)\mathbf{I}_L, \quad (26)$$

$$\mathbf{R}_{\mathbf{e}}(n) = \mathbf{X}^T(n)\mathbf{R}_{\mathbf{m}}(n)\mathbf{X}(n) + \sigma_v^2(n)\mathbf{I}_P, \quad (27)$$

$$\mathbf{K}(n) = \mathbf{R}_{\mathbf{m}}(n)\mathbf{X}(n)\mathbf{R}_{\mathbf{e}}^{-1}(n), \quad (28)$$

$$\mathbf{e}(n) = \mathbf{d}(n) - \mathbf{X}^T(n)\hat{\mathbf{h}}(n-1), \quad (29)$$

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \mathbf{K}(n)\mathbf{e}(n), \quad (30)$$

$$\mathbf{R}_{\boldsymbol{\mu}}(n) = [\mathbf{I}_L - \mathbf{K}(n)\mathbf{X}^T(n)]\mathbf{R}_{\mathbf{m}}(n). \quad (31)$$

The initialization is  $\hat{\mathbf{h}}(n) = \mathbf{0}$  and  $\mathbf{R}_{\boldsymbol{\mu}}(0) = \varepsilon\mathbf{I}_L$ , where  $\varepsilon$  is a small positive constant.

Substituting (30) into (15), we find an interesting relationship between the a posteriori and a priori error signal vectors:

$$\boldsymbol{\epsilon}(n) = \sigma_v^2(n)\mathbf{R}_{\mathbf{e}}^{-1}(n)\mathbf{e}(n), \quad (32)$$

which implies that the update equation can also be expressed as

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \mathbf{K}(n)\boldsymbol{\epsilon}(n), \quad (33)$$

where

$$\mathbf{K}(n) = \frac{1}{\sigma_v^2(n)}\mathbf{R}_{\mathbf{m}}(n)\mathbf{X}(n) \quad (34)$$

is the a posteriori Kalman gain matrix.

An objective measure to assess the echo cancellation by the Kalman filter is the echo-return loss enhancement (ERLE) defined as [15]

$$\begin{aligned}\text{ERLE}(n) &= \frac{E[\mathbf{y}^T(n)\mathbf{y}(n)]}{E\{[\mathbf{y}(n) - \hat{\mathbf{y}}(n)]^T[\mathbf{y}(n) - \hat{\mathbf{y}}(n)]\}} \\ &= \frac{\text{tr}[\mathbf{X}^T(n)\mathbf{R}_{\mathbf{h}}(n)\mathbf{X}(n)]}{\text{tr}[\mathbf{X}^T(n)\mathbf{R}_{\mathbf{m}}(n)\mathbf{X}(n)]} \\ &= \text{ENR}(n) \cdot \frac{P\sigma_v^2(n)}{\text{tr}[\mathbf{R}_{\mathbf{e}}(n)] - P\sigma_v^2(n)}.\end{aligned}\quad (35)$$

At infinity, we have

$$\lim_{n \rightarrow \infty} \text{tr}[\mathbf{R}_{\boldsymbol{\mu}}(n)] = L\sigma_{\boldsymbol{\mu}}^2(n), \quad (36)$$

$$\lim_{n \rightarrow \infty} \text{tr}[\mathbf{R}_{\mathbf{m}}(n)] = L[\sigma_{\boldsymbol{\mu}}^2(n) + \sigma_w^2(n)], \quad (37)$$

where  $\sigma_{\boldsymbol{\mu}}^2(n)$  is a small positive number to which all diagonal elements of  $\mathbf{R}_{\boldsymbol{\mu}}(n)$  converge. Therefore, the normalized misalignment as defined in Section II should be after convergence:

$$\frac{\|\hat{\mathbf{h}}(n) - \mathbf{h}\|_2^2}{\|\mathbf{h}\|_2^2} \approx \frac{L[\sigma_{\boldsymbol{\mu}}^2(n) + \sigma_w^2(n)]}{\|\mathbf{h}\|_2^2}, \quad (38)$$

where  $\mathbf{h}$  is the impulse response that we try to identify. It is instructive to observe how the final misalignment is determined by the values of  $\sigma_w^2(n)$ . Also, when  $\sigma_w^2(n) = 0$ , we have

$$\lim_{n \rightarrow \infty} \mathbf{R}_{\boldsymbol{\mu}}(n) = \mathbf{0}, \quad (39)$$

$$\lim_{n \rightarrow \infty} \mathbf{K}(n) = \mathbf{0}, \quad (40)$$

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 (30) to stay “alert” to any possible random changes of the echo path.

## V. APA AND PAPA

In this part, we show how the APA and PAPA are actually good approximations of the GKF derived in the previous section.

In the first scenario, we assume that the GKF has started to converge. In this case,  $\mathbf{R}_m(n)$  tends to become a diagonal matrix with all its elements equal to a small positive number,  $\sigma_m^2(n)$ ; so we can make the approximation:

$$\mathbf{R}_m(n) \approx \sigma_m^2(n) \mathbf{I}_L. \quad (41)$$

As a result, the Kalman gain matrix simplifies to

$$\mathbf{K}(n) \approx \mathbf{K}_{\text{APA}}(n) = \mathbf{X}(n)[\mathbf{X}^T(n)\mathbf{X}(n) + \delta(n)\mathbf{I}_P]^{-1}, \quad (42)$$

where

$$\delta(n) = \frac{\sigma_v^2}{\sigma_m^2(n)} \quad (43)$$

can be seen as a variable regularization parameter. We deduce that the GKF simplifies to the APA:

$$\mathbf{K}_{\text{APA}}(n) = \mathbf{X}(n)[\mathbf{X}^T(n)\mathbf{X}(n) + \delta(n)\mathbf{I}_P]^{-1}, \quad (44)$$

$$\mathbf{e}(n) = \mathbf{d}(n) - \mathbf{X}^T(n)\hat{\mathbf{h}}(n-1), \quad (45)$$

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \mathbf{K}_{\text{APA}}(n)\mathbf{e}(n). \quad (46)$$

Simulations confirm that with these parameters, the APA behaves the same way as the GKF at the convergence.

The matrix  $\mathbf{R}_\mu(n-1)$  depends on the filter weights. In the second scenario, we suggest to approximate this matrix by a diagonal matrix:

$$\begin{aligned} \mathbf{R}_\mu(n-1) &\approx \Psi(n-1) \\ &= \text{diag}[\psi_0(n-1), \dots, \psi_{L-1}(n-1)], \end{aligned} \quad (47)$$

which contains the so-called “gain” (or proportionate) factors; the parameters  $\psi_l(n-1)$  (with  $0 \leq l \leq L-1$ ) can be evaluated in many ways, depending on the proportionate-type algorithm [13]. For example, in the case of the improved PAPA (IPAPA) [16], [17], these parameters are defined as

$$\psi_l(n-1) = \frac{1-\kappa}{2L} + (1+\kappa) \frac{|\hat{h}_l(n-1)|}{2 \sum_{l=0}^{L-1} |\hat{h}_l(n-1)|}, \quad 0 \leq l \leq L-1, \quad (48)$$

where  $\kappa$  ( $-1 \leq \kappa < 1$ ) is a parameter that controls the amount of proportionality [16]. Therefore, the PAPAs are obtained by simply replacing  $\mathbf{R}_\mu(n-1)$  in the GKF by the diagonal matrix

$\Psi(n-1)$ . In this scenario,  $\mathbf{R}_m(n)$  is also a diagonal matrix. The following equations summarize the algorithm:

$$\mathbf{R}_m(n) = \Psi(n-1) + \sigma_w^2(n)\mathbf{I}_L, \quad (49)$$

$$\begin{aligned} \mathbf{K}_{\text{PAPA}}(n) &= \mathbf{R}_m(n)\mathbf{X}(n) \\ &\times [\mathbf{X}^T(n)\mathbf{R}_m(n)\mathbf{X}(n) + \delta'(n)\mathbf{I}_P]^{-1}, \end{aligned} \quad (50)$$

$$\mathbf{e}(n) = \mathbf{d}(n) - \mathbf{X}^T(n)\hat{\mathbf{h}}(n-1), \quad (51)$$

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \mathbf{K}_{\text{PAPA}}(n)\mathbf{e}(n), \quad (52)$$

where  $\delta'(n)$  can be seen as the variable regularization parameter of PAPA. In the case of IPAPA, this regularization parameter should be evaluated as [18], [19]

$$\delta'(n) = \frac{\delta(n)}{L}, \quad (53)$$

where  $\delta(n)$  is defined in (43).

## VI. A PARTICULAR CASE OF THE GKF

One very important particular case of the GKF developed in Section IV is when  $P = 1$ . Consequently, we get the well-known Kalman filter [12]:

$$\mathbf{R}_m(n) = \mathbf{R}_\mu(n-1) + \sigma_w^2(n)\mathbf{I}_L, \quad (54)$$

$$\sigma_e^2(n) = \mathbf{x}^T(n)\mathbf{R}_m(n)\mathbf{x}(n) + \sigma_v^2, \quad (55)$$

$$\mathbf{k}(n) = \frac{1}{\sigma_e^2(n)}\mathbf{R}_m(n)\mathbf{x}(n), \quad (56)$$

$$e(n) = d(n) - \mathbf{x}^T(n)\hat{\mathbf{h}}(n-1), \quad (57)$$

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \mathbf{k}(n)e(n), \quad (58)$$

$$\mathbf{R}_\mu(n) = [\mathbf{I}_L - \mathbf{k}(n)\mathbf{x}^T(n)]\mathbf{R}_m(n). \quad (59)$$

This algorithm has striking resemblances with the classical RLS algorithm [12]. 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. Second, the Kalman filter depends explicitly on the correlation matrix of the misalignment while the RLS adaptive filter depends on the 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  $\sigma_v^2$  in the Kalman filter (for which the RLS does not depend on) allow us to better control it.

After convergence and with the approximation given in (41), we obtain

$$e(n) = d(n) - \mathbf{x}^T(n)\hat{\mathbf{h}}(n-1), \quad (60)$$

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \frac{\mathbf{x}(n)e(n)}{\mathbf{x}^T(n)\mathbf{x}(n) + \delta(n)}, \quad (61)$$

which is the classical NLMS algorithm, where  $\delta(n)$  is defined in (43). Simulations confirm that with these parameters, the NLMS algorithm converges to the same final misalignment as the Kalman filter.

## VII. A SIMPLIFIED GKF

Let us first consider the case  $P = 1$  and assume that the update matrix that appears explicitly in  $\mathbf{R}_\mu(n)$  can be approximated as

$$\mathbf{I}_L - \mathbf{k}(n)\mathbf{x}^T(n) \approx \left[1 - \frac{\mathbf{k}^T(n)\mathbf{x}(n)}{L}\right] \mathbf{I}_L. \quad (62)$$

The previous approximation is reasonable, taking into account that we deal with a diagonal process noise covariance in (9). Indeed, when the filter starts to converge, the matrix  $\mathbf{R}_\mu(n)$  tends to become a diagonal one since the misalignment of the individual coefficients tend to become uncorrelated. As a consequence of (62), the two matrices  $\mathbf{R}_m(n)$  and  $\mathbf{R}_\mu(n)$  become diagonal, i.e.,  $\mathbf{R}_m(n) \approx \sigma_m^2(n)\mathbf{I}_L = r_m(n)\mathbf{I}_L$  and  $\mathbf{R}_\mu(n) \approx \sigma_\mu^2(n)\mathbf{I}_L = r_\mu(n)\mathbf{I}_L$ . Then, it is not hard to deduce that the Kalman filter simplifies to

$$r_m(n) = r_\mu(n-1) + \sigma_w^2(n), \quad (63)$$

$$\delta(n) = \frac{\sigma_v^2}{r_m(n)}, \quad (64)$$

$$e(n) = d(n) - \mathbf{x}^T(n)\hat{\mathbf{h}}(n-1), \quad (65)$$

$$\mathcal{S}_x(n) = \mathbf{x}^T(n)\mathbf{x}(n), \quad (66)$$

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \frac{\mathbf{x}(n)e(n)}{\mathcal{S}_x(n) + \delta(n)}, \quad (67)$$

$$r_\mu(n) = \left\{1 - \frac{\mathcal{S}_x(n)}{L[\mathcal{S}_x(n) + \delta(n)]}\right\} r_m(n). \quad (68)$$

This simplified Kalman filter for  $P = 1$  was first proposed in [7] and named broadband Kalman filter (BKF). With a proper estimation of the parameter  $\sigma_w^2(n)$  (as will be discussed in the next section), this algorithm behaves like a VSS-type NLMS. Simulations prove this statement. This algorithm may look very similar, at first glance, to the one proposed in [20] but the two are very much different. Indeed, in [20], the echo path is considered as time invariant so that no state equation is involved; this is equivalent to taking  $\sigma_w^2(n) = 0$ , which may be a problem in tracking as discussed earlier.

For any value of  $P$ , we can make the approximation:

$$\mathbf{I}_L - \mathbf{K}(n)\mathbf{X}^T(n) \approx \left\{1 - \frac{\text{tr}[\mathbf{K}^T(n)\mathbf{X}(n)]}{PL}\right\} \mathbf{I}_L. \quad (69)$$

As a result, the two matrices  $\mathbf{R}_m(n)$  and  $\mathbf{R}_\mu(n)$  are diagonal and the simplified GKF (SGKF) for any value of  $P$  is

$$r_m(n) = r_\mu(n-1) + \sigma_w^2(n), \quad (70)$$

$$\delta(n) = \frac{\sigma_v^2}{r_m(n)}, \quad (71)$$

$$\mathbf{e}(n) = \mathbf{d}(n) - \mathbf{X}^T(n)\hat{\mathbf{h}}(n-1), \quad (72)$$

$$\mathcal{S}_x(n) = \mathbf{X}^T(n)\mathbf{X}(n), \quad (73)$$

$$\hat{\mathbf{h}}(n) = \hat{\mathbf{h}}(n-1) + \mathbf{X}(n)[\mathcal{S}_x(n) + \delta(n)\mathbf{I}_P]^{-1}\mathbf{e}(n), \quad (74)$$

$$r_\mu(n) = \left\{1 - \frac{\text{tr}\{[\mathcal{S}_x(n) + \delta(n)\mathbf{I}_P]^{-1}\mathcal{S}_x(n)\}}{PL}\right\} \times r_m(n). \quad (75)$$

Also, using a proper evaluation of  $\sigma_w^2(n)$  (see the next section), the SGKF behaves like a VSS-type APA.

## VIII. PRACTICAL CONSIDERATIONS

There are two parameters that need to be set or estimated within the GKF (and the SGKF). The first (and perhaps the most important) one is  $\sigma_w^2(n)$ , which plays a major role in the overall performance of the algorithms, as explained in Sections II and III. Based on the state equation (9) and also considering the contribution of the model's order, we propose to evaluate this parameter as

$$\hat{\sigma}_w^2(n) = \frac{\|\hat{\mathbf{h}}(n) - \hat{\mathbf{h}}(n-1)\|_2^2}{PL}. \quad (76)$$

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 (76) 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  $\hat{\mathbf{h}}(n)$  and  $\hat{\mathbf{h}}(n-1)$  is significant, so that the parameter  $\hat{\sigma}_w^2(n)$  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  $\hat{\mathbf{h}}(n)$  and  $\hat{\mathbf{h}}(n-1)$  reduces, thus leading to small values of  $\hat{\sigma}_w^2(n)$  and, consequently, to a low misalignment.

In the case of the BKF [7] (which is similar to the SGKF with  $P = 1$ ), the state equation is slightly different as compared to (9), i.e.,

$$\mathbf{h}(n) = a\mathbf{h}(n-1) + \mathbf{w}(n), \quad (77)$$

where  $0 \leq a \leq 1$  is the transition coefficient. This implies that the estimation of  $\sigma_w^2(n)$  becomes [7]

$$\hat{\sigma}_w^2(n) = (1 - a^2) \frac{E[\hat{\mathbf{h}}^T(n)\hat{\mathbf{h}}(n)]}{L}. \quad (78)$$

In general, the value of the parameter  $a$  should be chosen very close to one [7], which could be problematic for the estimator given in (78). From this point of view, the evaluation of  $\hat{\sigma}_w^2(n)$  from (76) should be more reliable in practice.

The second parameter to be found is the noise power,  $\sigma_v^2$ . Usually, it can be estimated during silences of the near-end talker, i.e., in a single-talk scenario [21]. 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 [22] or [23]. For example, assuming that the adaptive filter has converged

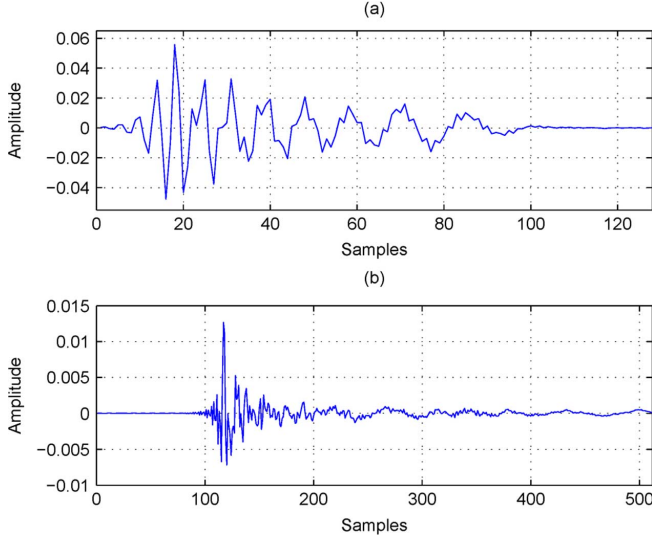


Fig. 2. Impulse responses used in simulations. (a) Network echo path (G168). (b) Acoustic echo path.

to a certain degree, the near-end signal power can be evaluated as [23]

$$\hat{\sigma}_v^2(n) = |\hat{\sigma}_d^2(n) - \hat{\sigma}_y^2(n)|, \quad (79)$$

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

$$\hat{\sigma}_d^2(n) = \beta \hat{\sigma}_d^2(n-1) + (1-\beta) d^2(n), \quad (80)$$

$$\hat{\sigma}_y^2(n) = \beta \hat{\sigma}_y^2(n-1) + (1-\beta) \hat{y}^2(n), \quad (81)$$

where  $\beta = 1 - 1/(KL)$ , with  $K \geq 1$ . The initial values are  $\hat{\sigma}_d^2(0) = 0$  and  $\hat{\sigma}_y^2(0) = 0$ . When using the estimator from (79), we refer to the simplified algorithm as the practical SGKF (PSGKF). However, we should note that different other estimators can be used for the noise power; the analysis of their influence on the algorithms' performance is beyond the scope of this paper.

## IX. EXPERIMENTAL STUDY

In this section, we present an experimental study in order to outline the performance of the proposed algorithms. For the sake of clarity, we divide this section into three parts. The first subsection details the experimental setup. The second subsection is dedicated to the performance of the GKF proposed in Section IV, in conjunction with the benchmarks NLMS algorithm, APA, IPAPA, and RLS algorithm. Finally, the third subsection presents the performance of the simplified algorithm, i.e., SGKF, as compared to the previously proposed BKF [7]; at the end of this last subsection, some results with the PSGKF are also presented.

### A. Experimental Setup

Experiments are performed in the context of both network and acoustic echo cancellation. Two echo paths are used, as shown in Fig. 2. The first one [Fig. 2(a)] is a network echo path

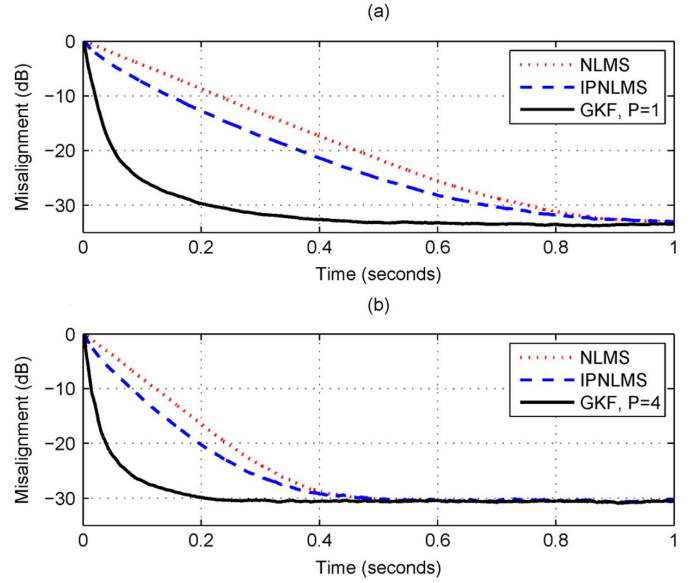


Fig. 3. (a) Misalignment of the NLMS, IPNLMS, and GKF using  $P = 1$ . (b) Misalignment of the APA, IPAPA, and GKF using  $P = 4$ . The input signal is white and Gaussian,  $L = 128$ ,  $\sigma_w^2(n) = 10^{-10}$ , and  $\text{SNR} = 20$  dB.

from G168 Recommendation [24] and it is used in the experiments provided in Subsection IX.B. The second one [Fig. 2(b)] is a measured acoustic echo path, which is used in the simulations presented in Subsection IX.B. All adaptive filters used in the experiments have the same length as the echo paths to be identified, i.e.,  $L = 128$  in Subsection IX.B and  $L = 512$  in Subsection IX.C. The sampling rate is 8 kHz.

The far-end signal (i.e., the input signal) is either a white Gaussian signal or a speech sequence. The output of the echo path is corrupted by an independent white Gaussian noise (i.e., the background noise at the near-end) with 20 dB SNR. Only in one experiment (reported in Fig. 15), a case with a variable SNR is also evaluated. For a fair comparison of the algorithms, we assume that the variance of the noise,  $\sigma_v^2$ , is available in most of the simulations (except in the last experiments with the PSGKF).

Most of the simulations, except for the last one reported in Fig. 16, are performed in the single-talk case. In order to evaluate the tracking capabilities of the algorithms, an echo path change scenario is simulated in some experiments, by shifting the impulse response to the right by 12 samples. The performance measure used in simulations is the normalized misalignment (in dB) evaluated based on (5), i.e.,

$$\text{Mis}(n) = 20 \log_{10} \frac{\|\mathbf{h} - \hat{\mathbf{h}}(n)\|_2}{\|\mathbf{h}\|_2}. \quad (82)$$

In all the experiments, the results are averaged over 20 independent trials.

### B. GKF

The proposed GKF was developed in Section IV. Also, as it was shown in Section V, the APA and PAPA represent good approximations of the GKF. Consequently, with a proper selection of their parameters [see (41), (43), and (53)], all these algorithms

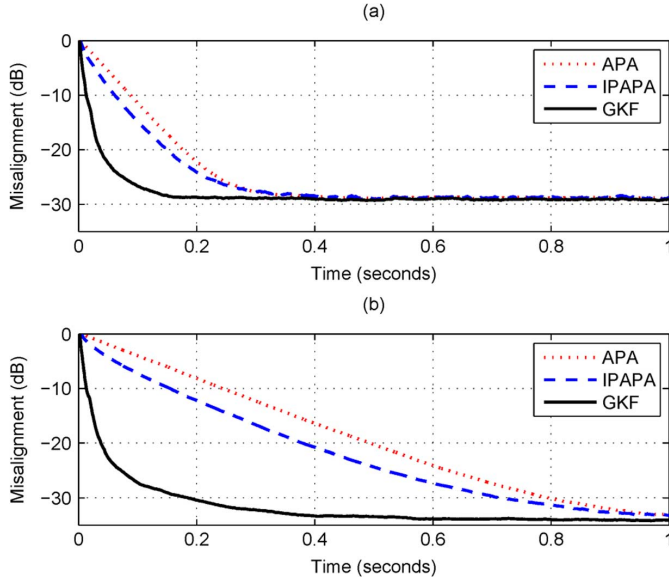


Fig. 4. Misalignment of the APA, IPAPA, and GKF using  $P = 8$ . (a)  $\sigma_w^2(n) = 10^{-10}$ . (b)  $\sigma_w^2(n) = 10^{-11}$ . The input signal is white and Gaussian,  $L = 128$ , and SNR = 20 dB.

behave the same way at the convergence. This aspect is proved in the first two experiments.

In Fig. 3(a), the GKF using  $P = 1$  (see the beginning of Section VI) is compared to the NLMS and improved proportionate NLMS (IPNLMS) [16] algorithms; the input signal is a white Gaussian noise. It should be noted that the IPNLMS algorithm is equivalent to the IPAPA with  $P = 1$ . The specific parameter of the IPNLMS is set to  $\kappa = 0$  [see (48)]. The regularization parameters of the NLMS and IPNLMS algorithms are evaluated according to (43) and (53), respectively. For a fair comparison, a constant value of the parameter  $\sigma_w^2(n)$  is considered in this experiment, i.e.,  $\sigma_w^2(n) = 10^{-10}$ . As we can see from Fig. 3(a), all the algorithms converge to the same steady-state misalignment. In terms of the initial convergence rate, the GKF clearly outperforms both NLMS and IPNLMS algorithms. Also, due to the quasi-sparse nature of the impulse response [see Fig. 2(a)], the IPNLMS algorithm converges slightly faster than NLMS.

The same experiment is repeated in Fig. 3(b), but using  $P = 4$ . Therefore, the GKF as described in Section IV is compared now to the APA and IPAPA. The other parameters remain the same as in Fig. 3(a). The conclusions are basically the same, i.e., the APA and IPAPA converge to the same steady-state misalignment as the GKF, but with a slower initial convergence rate. In this context, the influence of the parameter  $\sigma_w^2(n)$  is outlined in Fig. 4, using  $P = 8$  for all the algorithms. The main conclusions remain the same, but it can be noticed that a smaller value of  $\sigma_w^2(n)$  [see Fig. 4(b) as compared to Fig. 4(a)] leads to a slower initial convergence rate but also to a lower steady-state misalignment.

The influence of the parameter  $\sigma_w^2(n)$  on the performance of the GKF is analyzed in Fig. 5. In order to evaluate the tracking capability of the algorithm, an abrupt change of the echo path is considered in the middle of the experiment. The input signal is a white Gaussian noise. The GKF uses  $P = 2$  and different

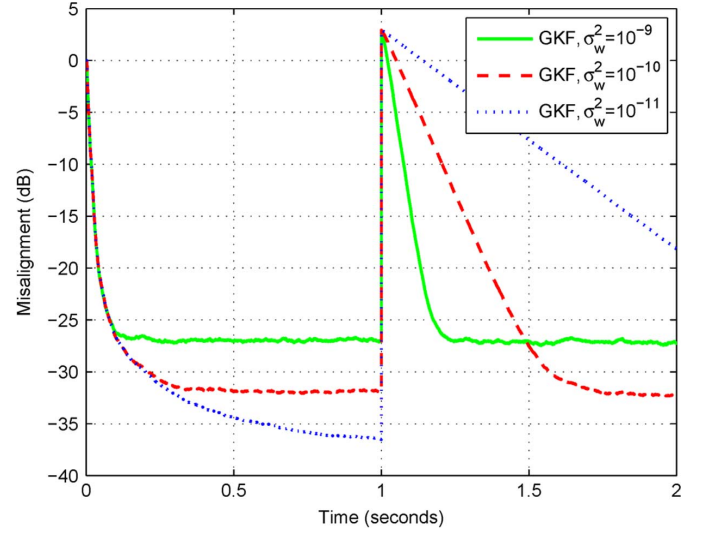


Fig. 5. Misalignment of the GKF using  $P = 2$  and different values of  $\sigma_w^2(n)$ . The input signal is white and Gaussian,  $L = 128$ , and SNR = 20 dB. Echo path changes after 1 second.

constant values of  $\sigma_w^2(n)$ . The conclusions of this experiment support the role of this important parameter, in terms of the compromise between good tracking and low misalignment (as described in Sections II and III). It can be noticed from Fig. 5 that a smaller value of  $\sigma_w^2(n)$  leads to a lower misalignment but a slower tracking (despite the similar initial convergence rate); while a larger value of this parameter increases the tracking capability of the algorithm but also the misalignment level. Concluding, it is not easy to find a proper compromise between these performance criteria when using a constant value of  $\sigma_w^2(n)$ .

A similar experiment is considered in Fig. 6, but using  $\sigma_w^2(n) = 10^{-10}$  and different model orders, i.e.,  $P = 1, 2, 4$ , and 8. It can be noticed that the tracking reaction of the GKF is improved when the value of  $P$  increases. On the other hand, the algorithm achieves a lower steady-state misalignment when the value of  $P$  decreases. Consequently, there is also a compromise between good tracking and low misalignment, which is influenced by the value of  $P$ .

The previous two simulations motivate the use of a variable parameter  $\hat{\sigma}_w^2(n)$ , as estimated in (76). The previous experiment is repeated in Fig. 7, under this new circumstance. According to this figure, it can be noticed that the GKF using the variable  $\hat{\sigma}_w^2(n)$  from (76) compromises much better between the tracking capability and the steady-state misalignment level. It can be noticed that there is a significant performance improvement of the GKF with  $P = 2$  over the case when  $P = 1$ . However, this improvement is not very significant when the value of  $P$  increases, as compared to the case with  $P = 2$ . Due to its specific features, the variable parameter  $\hat{\sigma}_w^2(n)$  from (76) will be used in all the following experiments.

As it was shown in Section VI, the GKF with  $P = 1$  is quite similar to the classical RLS algorithm [12]. However, as it was also explained in the same section, these two algorithms are fundamentally different. In Fig. 8, the GKF with  $P = 1$  and  $P = 2$  is compared to the RLS algorithm using different values of the forgetting factor ( $0 < \lambda \leq 1$ ). This specific parameter of the



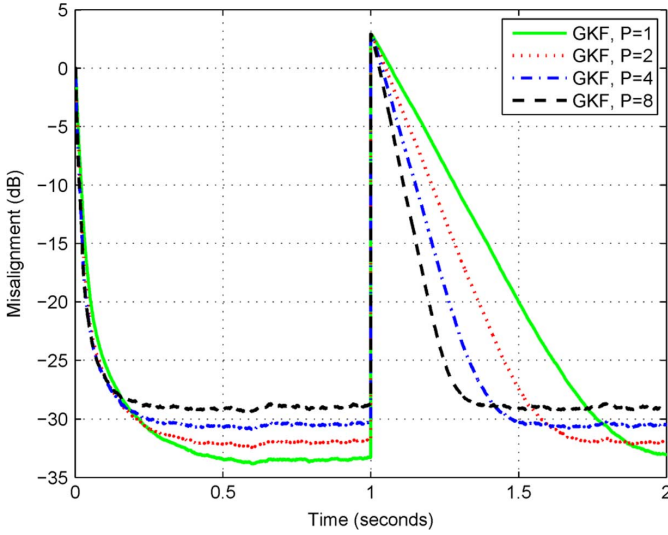


Fig. 6. Misalignment of the GKF using  $\sigma_w^2(n) = 10^{-10}$  and different values of  $P$ . Other conditions as in Fig. 5.

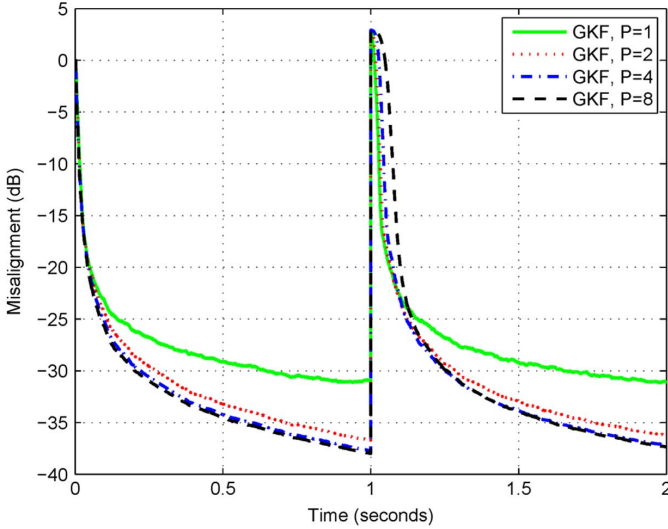


Fig. 7. Misalignment of the GKF using  $\hat{\sigma}_w^2(n)$  from (76) and different values of  $P$ . The input signal is white and Gaussian,  $L = 128$ , and  $\text{SNR} = 20$  dB. Echo path changes after 1 second.

RLS algorithm is of great importance in practice, since it needs to address the compromise between convergence rate/tracking capabilities on the one hand and misadjustment/stability on the other hand. It can be noticed from Fig. 8 that the GKF using  $P = 1$  compromises much better between the tracking capability and steady-state misalignment level, as compared to the RLS algorithm. Also, the GKF using  $P = 2$  clearly outperforms the RLS algorithm, in terms of both performance criteria. In order to improve the tracking reaction of the RLS algorithm, the value of the forgetting factor should be decreased. However, a small value of  $\lambda$  increases the steady-state misalignment and could affect the stability of the algorithm.

In Fig. 9, the previous experiment is repeated but using a speech sequence as input and changing the echo path at time 2.5 seconds. The results are basically the same, proving that the GKF performs better as compared to the RLS algorithm.

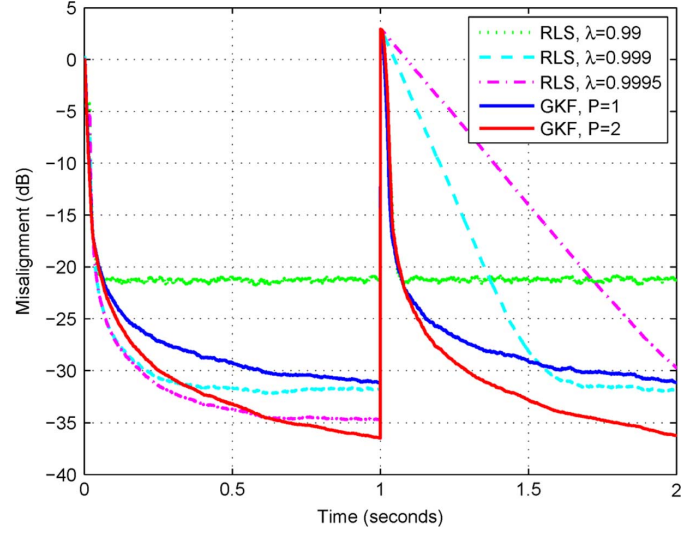


Fig. 8. Misalignment of the RLS algorithm using different values of the forgetting factor  $\lambda$  and misalignment of the GKF using different values of  $P$ . Other conditions as in Fig. 7.

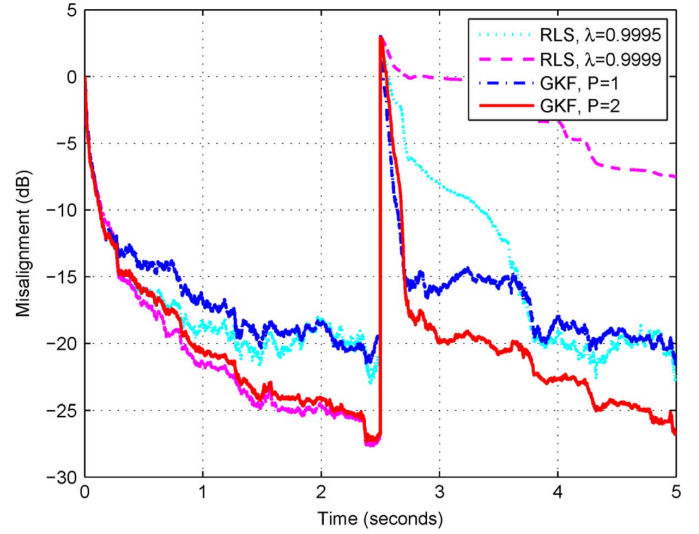


Fig. 9. Misalignment of the RLS algorithm using different values of the forgetting factor  $\lambda$  and misalignment of the GKF using different values of  $P$ . The input signal is speech,  $L = 128$ , and  $\text{SNR} = 20$  dB. Echo path changes after 2.5 seconds.

### C. SGKF

The main motivation behind the development of the SGKF (from Section VII) was to reduce the computational amount of the GKF. When  $P = 1$ , this simplified algorithm is very similar to the BKF proposed in [7]. However, there are two main differences between these two algorithms. First, the BKF uses a transition parameter  $a$  resulted from its state equation [see (77)], which does not appear in the SGKF. Second, the parameter  $\hat{\sigma}_w^2(n)$  is estimated in different ways within the two algorithms, i.e., based on (76) for the SGKF and on (78) for the BKF.

In the first experiment of this subsection, the SGKF using  $P = 1$  (i.e., the algorithm from the beginning of Section VII) is compared to the BKF using different values of the parameter  $a$ . The echo path to be identified is depicted in Fig. 2(b); the abrupt change of its impulse response is introduced in the middle of the simulation. The input signal is a white Gaussian noise. It



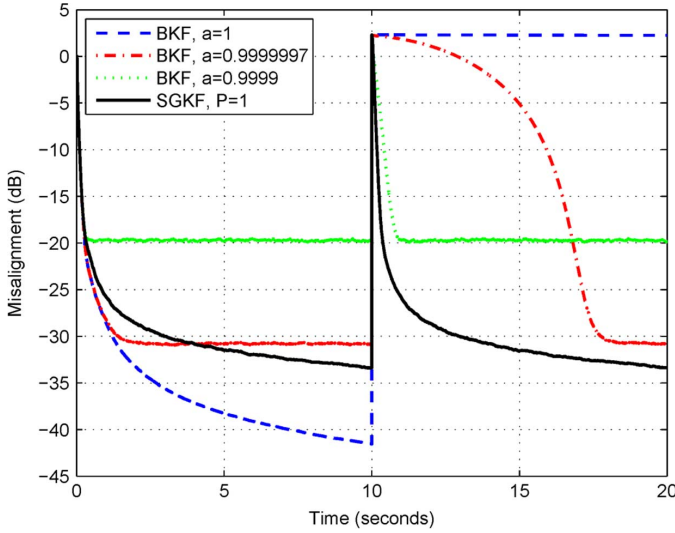


Fig. 10. Misalignment of the BKF using different values of the parameter  $a$  and misalignment of the SGKF using  $P = 1$ . The input signal is white and Gaussian,  $L = 512$ , and  $\text{SNR} = 20$  dB. Echo path changes after 10 seconds.

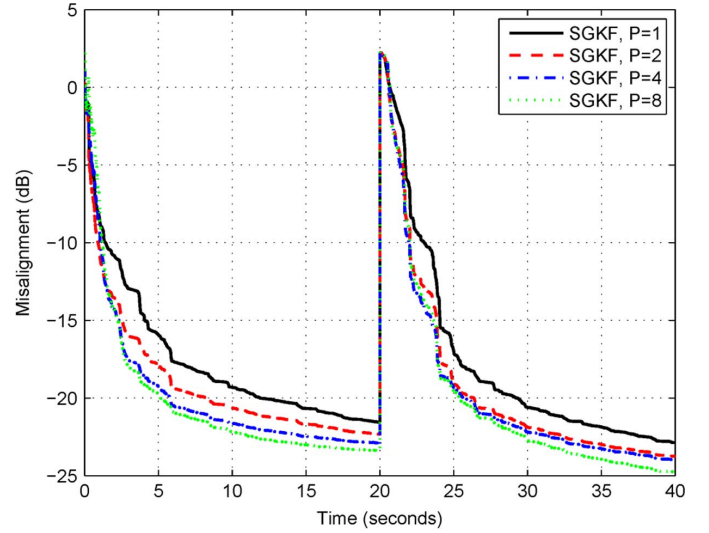


Fig. 12. Misalignment of the SGKF using different values of  $P$ . Other conditions as in Fig. 11.

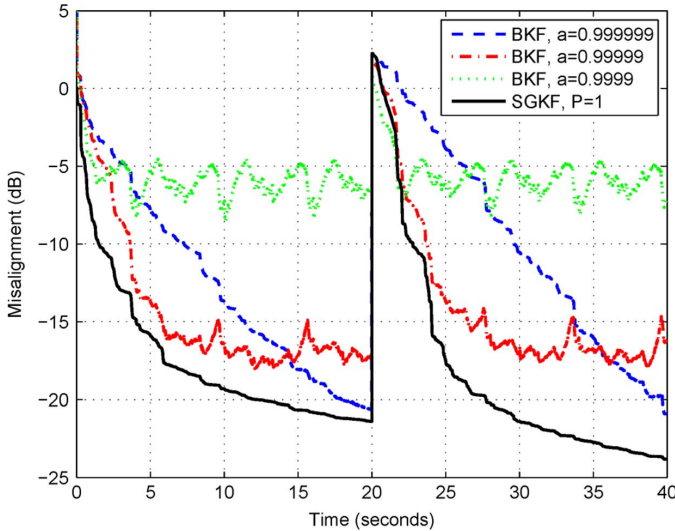


Fig. 11. Misalignment of the BKF using different values of the parameter  $a$  and misalignment of the SGKF using  $P = 1$ . The input signal is speech,  $L = 512$ , and  $\text{SNR} = 20$  dB. Echo path changes after 20 seconds.

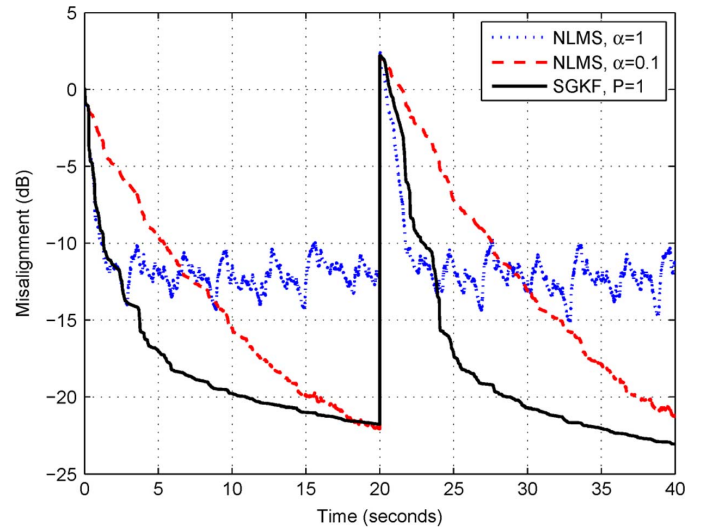


Fig. 13. Misalignment of the NLMS algorithm using different values of the normalized step-size  $\alpha$  and misalignment of the SGKF using  $P = 1$ . Other conditions as in Fig. 11.

can be noticed from Fig. 10 that the BKF obtains a lower misalignment level when the value of parameter  $a$  increases; however, its tracking capability is reduced in this case. On the other hand, the SGKF with  $P = 1$  achieves a proper compromise between these performance criteria. This algorithm tracks faster than the BKF, while also obtaining a reasonable low steady-state misalignment.

The previous experiment is repeated in Fig. 11, but using a speech sequence as input and changing the echo path at time 20 seconds. The conclusions are basically the same. Overall, the SGKF using  $P = 1$  outperforms the BKF in terms of both tracking reaction and steady-state misalignment.

The performance of the SGKF using different values of the model's order (i.e.,  $P = 1, 2, 4$ , and  $8$ ) are presented in Fig. 12, where the input signal is speech. As we can see from this simulation, the overall performance of the SGKF is improved when

the value of  $P$  increases. However, the improvement is not very significant when this value exceeds  $P = 2$ .

As explained in Section VII, the SGKF behaves like a variable step-size adaptive filter. This statement is supported in Fig. 13 (using a speech sequence as input), where the SGKF with  $P = 1$  is compared to the NLMS algorithm using two different values of the normalized step-size. This positive constant  $\alpha$  (usually  $\alpha \leq 1$ ) multiplies the update term of the NLMS algorithm [i.e., the second term in the right-hand side of (61)], in order to achieve a proper compromise between the convergence rate and misadjustment. It can be noticed from Fig. 13 that the convergence rate of the SGKF with  $P = 1$  (and also its tracking reaction) is similar to the NLMS algorithm using the largest normalized step-size. On the other hand, the SGKF obtains a lower steady-state misalignment, specific to the NLMS algorithm with the smaller normalized step-size.

Finally, at the end of this experimental study, we present several results of the PSGKF (see the end of Section VIII).

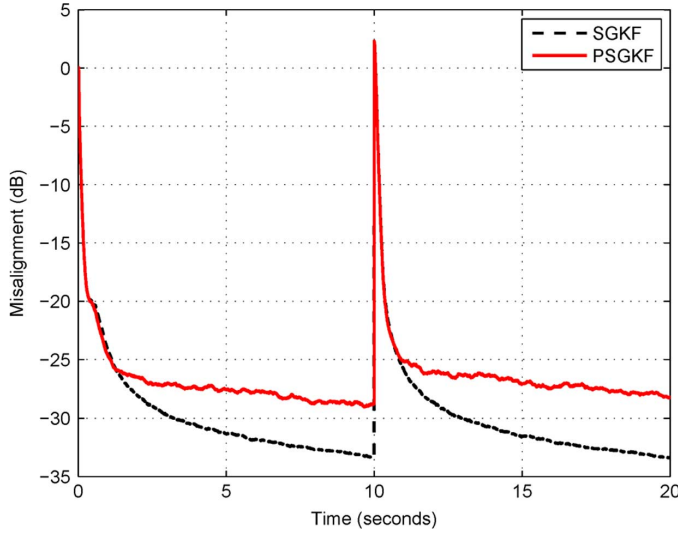


Fig. 14. Misalignment of the SGKF and PSGKF, both using  $P = 1$ . The input signal is white and Gaussian,  $L = 512$ , and  $\text{SNR} = 20$  dB. Echo path changes after 10 seconds.

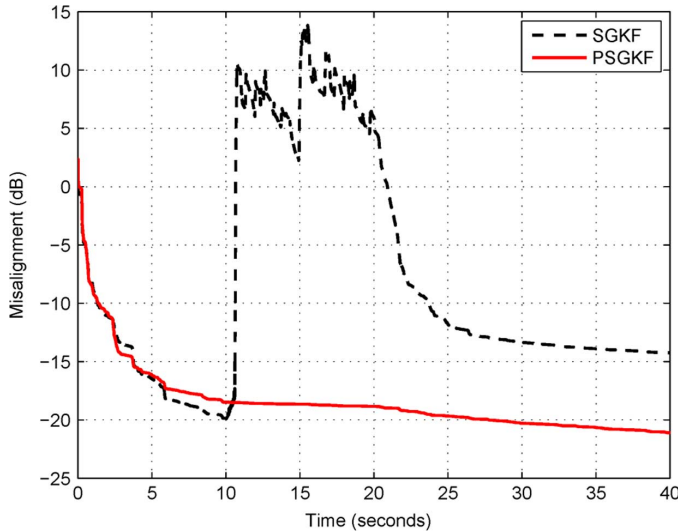


Fig. 15. Misalignment of the SGKF and PSGKF, both using  $P = 1$ . The input signal is speech and  $L = 512$ . The SNR decreases from 20 dB to 10 dB between times 10 and 20 seconds.

This algorithm is basically the same as the SGKF, but using the “practical” estimation of the noise power from (79). As it was mentioned at the end of Section VIII, different other estimation methods can be used to evaluate this parameter. The main feature of the estimator from (79) consists of its robustness to near-end variations (like double-talk).

In Fig. 14, the PSGKF is compared to the SGKF in a tracking situation; both algorithms use  $P = 1$  (this value will be used in all the following experiments). The input signal is a white Gaussian noise and the echo path from Fig. 2(b) changes in the middle of the simulation. The power estimates specific to the PSGKF [see (80) and (81)] are evaluated using  $K = 6$  [21], [23]. As before, it is assumed that the true variance of the background noise ( $\sigma_v^2$ ) is available for the SGKF. It can be noticed from this experiment that the SGKF outperforms PSGKF in terms of the final misalignment level, which is expected. However, the tracking reaction of the two algorithms is very similar.

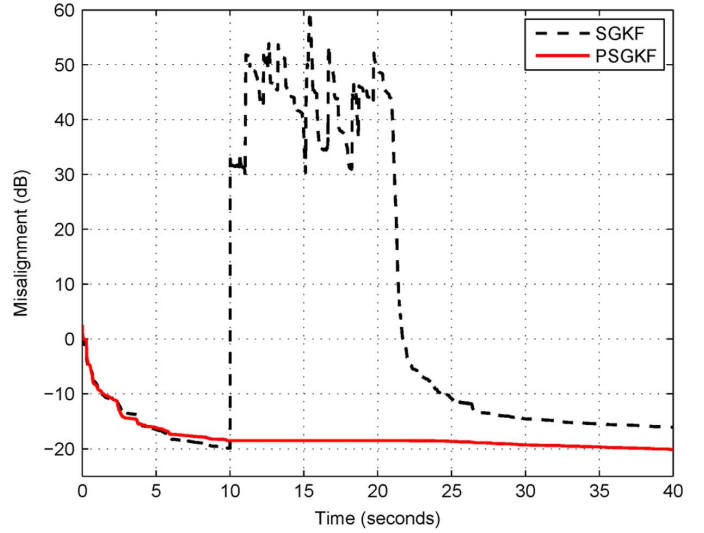


Fig. 16. Misalignment of the SGKF and PSGKF, both using  $P = 1$ . The input signal is speech,  $L = 512$ , and  $\text{SNR} = 20$  dB. The near-end speech (double-talk) appears between times 10 and 20 seconds (without using a DTD).

In the next simulation, a variation of the background noise at the near-end is considered. The input signal is speech and the SNR decreases from 20 dB to 10 dB between times 10 and 20 seconds. Also, we assume that the new value of  $\sigma_v^2$  is not available for the SGKF. It can be noticed from Fig. 15 that the PSGKF is very robust against the background noise variation.

One of the most challenging situation in echo cancellation is the double-talk case. Such a scenario is considered in the last experiment reported in Fig. 16. The near-end speech appears between times 10 and 20 seconds and the algorithms do not use any double-talk detector (DTD); in general, the DTD is designed to stop the adaptation during double-talk periods, in order to avoid the divergence of the algorithm [3]. It can be noticed that the PSGKF is very robust to double-talk, since the parameter from (79) provides an estimate of the near-end signal power (i.e., background noise plus near-end speech).

## X. CONCLUSION

In this paper, we have studied the time-domain Kalman filter in the context of echo cancellation. Due to its attractive performance, this optimal filter has the potential to become one of the most interesting choices for this application. A first contribution of this work was to derive a general Kalman filter (namely GKF) based on a different approach, i.e., a block of time samples is considered at each iteration, instead of one time sample (as in the conventional approach). Also, it was proved that the most popular adaptive filters for echo cancellation, i.e., the NLMS algorithm, APA, and PAPA represent good approximations of the GKF. Finally, we have developed a simplified version of the proposed algorithm (namely SGKF), which has a moderate computational complexity and behaves like a variable step-size adaptive filter. Simulation results support the theoretical findings, recommending the proposed solutions for real-world echo cancellation scenarios. Future work will target efficient implementations of these algorithms (e.g., frequency-domain and subbands versions), together with extensions to multichannel and non-linear cases.

## ACKNOWLEDGMENT

The authors would like to thank the Associate Editor and the reviewers for the valuable comments and suggestions.

## REFERENCES

- [1] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng.*, vol. 82, pp. 35–45, Mar. 1960.
- [2] R. Faragher, "Understanding the basis of the Kalman filter via a simple and intuitive derivation," *IEEE Signal Process. Mag.*, vol. 29, no. 5, pp. 128–132, Sep. 2012.
- [3] J. Benesty, T. Gansler, D. R. Morgan, M. M. Sondhi, and S. L. Gay, *Advances in Network and Acoustic Echo Cancellation*. Berlin, Germany: Springer-Verlag, 2001.
- [4] G. Enzner and P. Vary, "Frequency-domain adaptive Kalman filter for acoustic echo control in hands-free telephones," *Signal Process.*, vol. 86, pp. 1140–1156, 2006.
- [5] G. Enzner, "A model-based optimum filtering approach to acoustic echo control: Theory and practice," Ph.D. dissertation, RWTH Aachen, Aachener Beitrage zu digitalen Nachrichtensystemen, Aachen, Germany, Jun. 2006, Vary P. (ed.), Wissenschaftsverlag Mainz.
- [6] S. Malik and G. Enzner, "Model-based vs. traditional frequency-domain adaptive filtering in the presence of continuous double-talk and acoustic echo path variability," in *Proc. IWAENC*, 2008.
- [7] G. Enzner, "Bayesian inference model for applications of time-varying acoustic system identification," in *Proc. EUSIPCO*, 2010, pp. 2126–2130.
- [8] S. Malik and G. Enzner, "State-space frequency-domain adaptive filtering for nonlinear acoustic echo cancellation," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 20, no. 7, pp. 2065–2079, Sep. 2012.
- [9] 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.
- [10] 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.
- [11] T. Yahagi, "An adaptive echo canceller using parallel Kalman filters," in *Proc. IEEE ICASSP*, 1986, pp. 965–968.
- [12] A. H. Sayed and T. Kailath, "A state-space approach to adaptive RLS filtering," *IEEE Signal Process. Mag.*, vol. 11, no. 3, pp. 18–60, Jul. 1994.
- [13] C. Paleologu, J. Benesty, and S. Ciochină, "Sparse adaptive filters for echo cancellation," in *Synthesis Lectures on Speech and Audio Processing*. San Rafael, CA, USA: Morgan & Claypool, 2010.
- [14] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.
- [15] E. Häsler and G. Schmidt, *Acoustic Echo and Noise Control—A Practical Approach*. Hoboken, NJ, USA: Wiley, 2004.
- [16] J. Benesty and S. L. Gay, "An improved PNLMs algorithm," in *Proc. IEEE ICASSP*, 2002, pp. II-1881–II-1884.
- [17] O. Hoshuyama, R. A. Goubran, and A. Sugiyama, "A generalized proportionate variable step-size algorithm for fast changing acoustic environments," in *Proc. IEEE ICASSP*, 2004, pp. IV-161–IV-164.
- [18] J. Benesty, C. Paleologu, and S. Ciochină, "On regularization in adaptive filtering," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 19, no. 6, pp. 1734–1742, Aug. 2011.
- [19] C. Paleologu, J. Benesty, and S. Ciochină, "Regularization of the improved proportionate affine projection algorithm," in *Proc. IEEE ICASSP*, 2012, pp. 169–172.
- [20] D. Lippuner and A. N. Kaelin, "An improved step-size control for LMS filters with correlated input signals," in *Proc. IWAENC*, 2001.
- [21] J. Benesty, H. Rey, L. Rey Vega, and S. Tressens, "A non-parametric VSS NLMS algorithm," *IEEE Signal Process. Lett.*, vol. 13, pp. 581–584, Oct. 2006.
- [22] M. A. Iqbal and S. L. Grant, "Novel variable step size NLMS algorithm for echo cancellation," in *Proc. IEEE ICASSP*, 2008, pp. 241–244.
- [23] C. Paleologu, S. Ciochină, and J. Benesty, "Double-talk robust VSS-NLMS algorithm for under-modeling acoustic echo cancellation," in *Proc. IEEE ICASSP*, 2008, pp. 245–248.
- [24] "Digital network echo cancellers," ITU-T Rec. G.168, 2002.



**Constantin Paleologu** was born in Romania in 1975. In 1998 he received the Masters degree in telecommunications networks from the Faculty of Electronics, Telecommunications, and Information Technology, University Politehnica of Bucharest, Romania. He also received a Masters degree in digital signal processing in 1999, and a Ph.D. degree in adaptive signal processing in 2003, both from the same institution.

During his Ph.D. program (from December 1999 to July 2003), he worked on adaptive filters and echo cancellation. Since October 1998 he has been with the Telecommunications Department, University Politehnica of Bucharest, where he is currently an associate professor. His research interests include adaptive filtering algorithms and acoustic signal processing. He co-authored the books *Sparse Adaptive Filters for Echo Cancellation* (Morgan and Claypool, 2010) and *A Perspective on Stereophonic Acoustic Echo Cancellation* (Springer-Verlag, 2011).

Dr. Paleologu received the "IN HOC SIGNO VINCES" award from the Romanian National Research Council in 2009. In 2010, he received the "IN TEMPORE OPPORTUNO" award from University Politehnica of Bucharest and the "Gheorghe Cartianu Award" from the Romanian Academy.



**Jacob Benesty** was born in 1963. He received a Master degree in microwaves from Pierre & Marie Curie University, France, in 1987, and a Ph.D. degree in control and signal processing from Orsay University, France, in April 1991.

During his Ph.D. (from Nov. 1989 to Apr. 1991), he worked on adaptive filters and fast algorithms at the Centre National d'Etudes des Telecommunications (CNET), Paris, France. From January 1994 to July 1995, he worked at Telecom Paris University on multichannel adaptive filters and acoustic echo cancellation. From October 1995 to May 2003, he was first a Consultant and then a Member of the Technical Staff at Bell Laboratories, Murray Hill, NJ, USA. In May 2003, he joined the University of Quebec, INRS-EMT, in Montreal, Quebec, Canada, as a Professor. His research interests are in signal processing, acoustic signal processing, and multimedia communications. He is the inventor of many important technologies. In particular, he was the lead researcher at Bell Labs who conceived and designed the world-first real-time hands-free full-duplex stereophonic teleconferencing system. Also, he conceived and designed the world-first PC-based multi-party hands-free full-duplex stereo conferencing system over IP networks.

He was the co-chair of the 1999 International Workshop on Acoustic Echo and Noise Control and the general co-chair of the 2009 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. He is the recipient, with Morgan and Sondhi, of the IEEE Signal Processing Society 2001 Best Paper Award. He is the recipient, with Chen, Huang, and Doclo, of the IEEE Signal Processing Society 2008 Best Paper Award. He is also the co-author of a paper for which Huang received the IEEE Signal Processing Society 2002 Young Author Best Paper Award. In 2010, he received the "Gheorghe Cartianu Award" from the Romanian Academy. In 2011, he received the Best Paper Award from the IEEE WASPAA for a paper that he co-authored with Chen.



**Silviu Ciochină** received the Masters degree in electronics and telecommunications in 1971 and the Ph.D. degree in communications in 1978, both from the University Politehnica of Bucharest, Romania.

From 1979 to 1995 was a lecturer at the University Politehnica of Bucharest, Faculty of Electronics, Telecommunications, and Information Technology. Since 1995, he has been a professor at the same faculty. Since 2004, he has been the Head of the Telecommunications Department. His main research interests are in the areas of signal processing and

wireless communications, including adaptive algorithms, spectrum estimation, fast algorithms, channel estimation, multi-antenna systems, and broadband wireless technologies. He co-authored the books *Sparse Adaptive Filters for Echo Cancellation* (Morgan and Claypool, 2010) and *A Perspective on Stereophonic Acoustic Echo Cancellation* (Springer-Verlag, Berlin, 2011).

Dr. Ciochină received the "Traian Vuia" award in 1981 and the "Gheorghe Cartianu" award in 1997, both from the Romanian Academy.