

# Robust Set-Membership Affine-Projection Adaptive-Filtering Algorithm

Md. Zulfiqar Ali Bhotto, *Student Member, IEEE*, and Andreas Antoniou, *Life Fellow, IEEE*

**Abstract**—An improved set-membership affine-projection (AP) adaptive-filtering algorithm is proposed. The new algorithm uses two error bounds that are estimated during the learning phase and by this means significantly reduced steady-state misalignment is achieved as compared to those in the conventional AP and set-membership AP algorithms while achieving similar convergence speed and re-adaptation capability. In addition, the proposed algorithm offers robust performance with respect to the error bound, projection order, impulsive-noise interference, and in tracking abrupt changes in the underlying system. These features of the proposed algorithm are demonstrated through extensive simulation results in system-identification and echo-cancellation applications.

**Index Terms**—Affine-projection (AP) algorithms, robust adaptive-filtering algorithms, set-membership algorithms.

## I. INTRODUCTION

AN important class of adaptive-filtering algorithms is the class of affine-projection (AP) algorithms. These algorithms offer superior convergence performance relative to least-mean squares (LMS) algorithms, especially for correlated input signals, although some increase in the computational effort is involved [1], [2]. The basic AP algorithm was introduced in [3] and, subsequently, several variants of this algorithm have been developed by a number of researchers for different scenarios such as the regularized AP algorithm [4], the partial-rank AP algorithm [5], and the simplified set-membership AP (SSMAP) algorithm described in [6].<sup>1</sup> The set-membership binormalized data-reusing LMS algorithm in [7] is an alternative implementation of the SSMAP algorithm in [6] with a projection order of two. Performance analyses of the AP and SSMAP algorithms are presented in [8] and [9], respectively. The analysis presented in [8] shows that the convergence speed of the AP algorithm increases as the number of projections, also known as the projection order, is increased, at the expense of increased steady-

state misalignment. The same conclusion was also drawn for the SSMAP algorithm in [9]. However, by using a variable step size, the SSMAP algorithm yields reduced steady-state misalignment relative to that of the AP algorithm for the same projection order [6]. Some other variants of the SSMAP algorithm are described in [10], [11]; the algorithm in [10] yields improved convergence speed for sparse system-identification applications and the algorithm in [11] yields a slightly improved steady-state misalignment as compared to the SSMAP algorithm in [9]. The prespecified error bound in the SSMAP algorithm is usually chosen as  $\sqrt{5}\sigma_v$ , where  $\sigma_v^2$  is the variance of the measurement noise, in order to achieve a good balance between convergence speed and computational effort [6], [9]–[11]. In practice, it may not be possible to accurately specify the error bound in the SSMAP algorithm. In addition, as for the AP algorithm, the performance of the SSMAP algorithm is affected by outliers in the error signal samples that can be brought about by impulsive-noise interference. Several recent algorithms of the LMS and least-squares families that are robust with respect to outliers are compared in [12].

In this paper, we propose a new SMAP adaptation algorithm that is robust to outliers. The proposed algorithm uses two error bounds one of which is used to achieve faster convergence and the other is used to suppress impulsive-noise interference. By this means, reduced steady-state misalignment is achieved relative to that in the conventional AP and SSMAP algorithms, especially when the projection order is increased. Both of the error bounds are estimated by using the power of the error signal during the learning phase and, consequently, the proposed algorithm is robust in the sense that: (1) its performance remains largely insensitive to outliers brought about by impulsive noise; (2) the sensitivity of the steady-state misalignment on the projection order is significantly reduced; (3) its re-adaptation capability is preserved; and (4) the sensitivity of the convergence performance on the proper choice of error bound is significantly reduced. The performance of the proposed algorithm is tested in system-identification and echo-cancellation applications for different noise levels and projection orders, and is compared with that of known competing algorithms of the AP and SMAP families. The results obtained demonstrate the superior performance of the proposed algorithm.

The paper is organized as follows. Section II provides a brief review of SM adaptive filtering. In Section III the SSMAP adaptation algorithm is discussed. The proposed robust SMAP algorithm is then described in Section IV. A steady-state analysis of the proposed algorithm is carried out in Section V. Simulation results are presented in Section VI and conclusions are drawn in Section VII.

Manuscript received March 06, 2011; revised June 15, 2011 and September 20, 2011; accepted September 22, 2011. Date of publication October 10, 2011; date of current version December 16, 2011. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Roberto Lopez-Valcarce. This work was supported by the Natural Sciences and Engineering Research Council of Canada.

The authors are with the Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC V8W 3P6, Canada (e-mail: zbhotto@ece.uvic.ca; aantoniou@ieee.org).

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/TSP.2011.2170980

<sup>1</sup>This algorithm was referred to as the SMAP algorithm in [6] but was later referred to as the “simplified” SMAP algorithm in [9]. For the sake of consistency with the more recent publication on this algorithm, namely, [9], we will refer to this algorithm as the simplified SMAP (SSMAP) algorithm hereafter.

The symbols used for the  $L_\infty$  and  $L_2$  norms are  $\|\cdot\|_\infty$  and  $\|\cdot\|_2$  whereas  $|\cdot|$  and  $E[\cdot]$  denote absolute value and expectation, respectively. Symbol  $\mathcal{R}$  denotes the real space and  $\mathcal{R}_+$  denotes the space of nonnegative real quantities.

## II. SET-MEMBERSHIP ADAPTIVE FILTERING

Conventional SM adaptive-filtering schemes estimate the weight vector  $\mathbf{w} \in \mathcal{R}^{M \times 1}$  that would cause the magnitude of the estimation error

$$e = d - \mathbf{w}^T \mathbf{x} \quad (1)$$

where  $\mathbf{x} \in \mathcal{R}^{M \times 1}$  and  $d \in \mathcal{R}$  are the input signal vector and desired signal, respectively, to be less than or equal to a prespecified bound  $\gamma \in \mathcal{R}_+$  for all possible input-desired signal pairs  $(\mathbf{x}, d)$ . The set of all possible input-desired signal pairs  $(\mathbf{x}, d)$  is commonly referred to as the *data space* and is denoted as  $S$ . The estimation error based on the SM adaptive-filtering criterion must satisfy

$$|e|^2 \leq \gamma^2 \quad \forall (\mathbf{x}, d) \in S. \quad (2)$$

The set of all possible vectors  $\mathbf{w}$  that satisfy (2) whenever  $(\mathbf{x}, d) \in S$ , designated as  $\Theta$ , is referred to as the *feasibility* or *solution set* and can be expressed as

$$\Theta = \cap_{(\mathbf{x}, d) \in S} \{\mathbf{w} \in \mathcal{R}^M : |d - \mathbf{w}^T \mathbf{x}| \leq \gamma\}. \quad (3)$$

If the adaptive filter is trained with  $k$  input-desired data pairs  $\{\mathbf{x}_i, d_i\}_{i=1}^k$ , then the set containing all vectors  $\mathbf{w}$  for which the associated output error at iteration  $k$  is consistent with (2) is called the *constraint* or *observation set*. It is given by

$$H_k = \{\mathbf{w} \in \mathcal{R}^M : |d_k - \mathbf{w}^T \mathbf{x}_k| \leq \gamma\}. \quad (4)$$

The intersection of the constraint sets  $H_k$  over all iterations  $i = 1, 2, \dots, k$  is called the *exact membership set* and is given by

$$\Psi_k = \cap_{i=1}^k H_i. \quad (5)$$

Evidently, the feasibility set  $\Theta$  is a subset of the exact membership set  $\Psi_k$  in any given iteration.

## III. SSMAP ALGORITHM

The SSMAP algorithm in [6] performs weight adaptation at iteration  $k$  in such a way as to ensure that the updated weight vector belongs to the exact membership set  $\Psi_k^L$ , i.e.,  $\mathbf{w}_k \in \Psi_k^L$ . Whenever the weight vector  $\mathbf{w}_{k-1}$  is not a member of  $\Psi_k^L$ , an update is performed by solving the optimization problem

$$\begin{aligned} & \text{minimize } \|\mathbf{w}_k - \mathbf{w}_{k-1}\|^2 \\ & \text{subject to: } \mathbf{d}_k - \mathbf{X}_k^T \mathbf{w}_k = \mathbf{g}_k \end{aligned} \quad (6)$$

where  $\mathbf{d}_k \in \mathcal{R}^{L \times 1}$  is the desired signal vector,  $\mathbf{g}_k \in \mathcal{R}^{L \times 1}$  is the error-bound vector,  $\mathbf{X}_k \in \mathcal{R}^{M \times L}$  is the input signal matrix,

i.e.,  $\mathbf{X}_k = [\mathbf{x}_k \mathbf{x}_{k-1} \dots \mathbf{x}_{k-L+1}]$ . The solution of the problem in (6) results in the weight-vector update formula

$$\mathbf{w}_k = \begin{cases} \mathbf{w}_{k-1} + \mathbf{X}_k \left( \mathbf{X}_k^T \mathbf{X}_k \right)^{-1} (\mathbf{e}_k - \mathbf{g}_k) & \text{if } |e_k| > \gamma_k \\ \mathbf{w}_{k-1} & \text{otherwise} \end{cases} \quad (7)$$

where  $e_k = d_k - \mathbf{x}_k^T \mathbf{w}_{k-1}$  is the *a priori* error,  $\mathbf{e}_k = [e_k \ \epsilon_{k-1} \dots \epsilon_{k-L+1}]^T$  and  $\epsilon_{k-i} = d_{k-i} - \mathbf{x}_{k-i}^T \mathbf{w}_{k-1}$  is the *a posteriori* error at iteration  $k-1$ . The update formula in (7) forms the basis of set-membership AP algorithms. As can be seen,  $\mathbf{g}_k$  in (7) needs to be specified and in [6] it is chosen as  $\mathbf{g}_k = [\gamma_k \text{sign}(e_k) \ \epsilon_{k-1} \dots \epsilon_{k-L+1}]^T$  in which case the update formula becomes

$$\mathbf{w}_k = \mathbf{w}_{k-1} + \alpha_k \mathbf{X}_k \left( \mathbf{X}_k^T \mathbf{X}_k \right)^{-1} e_k \mathbf{u}_1 \quad (8)$$

where

$$\alpha_k = \begin{cases} 1 - \gamma/|e_k| & \text{if } |e_k| > \gamma_k \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

and  $\mathbf{u}_1 = [1 \ 0 \dots 0]^T$  [6]. Using projection order  $L=2$  and the closed-form inverse of  $\mathbf{X}_k^T \mathbf{X}_k$  in (8), the weight-vector update formula becomes identical with that of the SM binormalized data-reusing LMS-II adaptation algorithm in [7].

## IV. PROPOSED ROBUST SET-MEMBERSHIP AP ALGORITHM

In the proposed robust SMAP (RSMAP) algorithm, the error-bound vector  $\mathbf{g}_k$  is chosen as

$$\mathbf{g}_k^T = \gamma_k \begin{bmatrix} \frac{e_k}{|e_k|} & \frac{\epsilon_{k-1}}{|\epsilon_{k-1}|} & \dots & \frac{\epsilon_{k-L+1}}{|\epsilon_{k-L+1}|} \end{bmatrix}. \quad (10)$$

For this error bound, if  $|e_k| > g_k(1)$ , we would obtain  $|\epsilon_{k-i}| > g_k(i)$  for  $i = 2, \dots, L$  where  $g_k(i)$  denotes the  $i$ th element of vector  $\mathbf{g}_k$ . The use of (10) in the update formula in (7) results in the weight-vector update formula

$$\mathbf{w}_k = \mathbf{w}_{k-1} + \alpha_k \mathbf{X}_k \left( \mathbf{X}_k^T \mathbf{X}_k \right)^{-1} \mathbf{e}_k \quad (11)$$

where  $\alpha_k$  is defined in (9). The conventional AP algorithm in [3] uses the weight-vector update formula

$$\mathbf{w}_k = \mathbf{w}_{k-1} + \mu \mathbf{X}_k \left( \mathbf{X}_k^T \mathbf{X}_k \right)^{-1} \mathbf{e}_k \quad (12)$$

where  $\mu$  is a fixed step size and  $0 < \mu < 1$ . On the basis of (11) and (12), the proposed RSMAP algorithm could be considered to be a variable-step-size AP algorithm in which the step size  $\alpha_k$  can vary in the range zero to unity.

Two versions of the proposed RSMAP algorithm are possible, one with fixed and the other with variable threshold, as detailed below. These will be referred to as the RSMAP1 and RSMAP2 algorithms, respectively.

### A. RSMAP Algorithm With Fixed Threshold

In the proposed RSMAP algorithm with fixed threshold, namely, the RSMAP1 algorithm, factor  $\gamma_k$  in the error-bound vector in (10) is chosen as

$$\gamma_k = \begin{cases} \|e_k\|_\infty - \nu\theta_k & \text{if } \|e_k\|_\infty > \theta_k \\ \gamma_c & \text{otherwise} \end{cases} \quad (13)$$

where  $0 < \nu \ll 1$ ,  $\gamma_c = \sqrt{5\sigma_v^2}$  [6], and  $\theta_k$  is chosen to be less than  $\gamma_c$  in order to ensure that the algorithm would work with error bound  $\|e_k\|_\infty - \nu\theta_k$  during steady state and also for an impulsive-noise corrupted  $e_k$ . Extensive simulations have shown that a suitable value for  $\theta_k$  can be obtained as  $\theta_k = Q\sigma_{1,k}$  where  $1.86 \leq Q \leq 1.98$ . The value  $Q = 1.88$  was found to lead to improved computational efficiency. The variance of the error signal  $\sigma_{1,k}^2$  is obtained as

$$C_k = \text{median}(c_k) \quad (14)$$

$$\sigma_{1,k}^2 = \lambda\sigma_{1,k-1}^2 + (1 - \lambda)C_k \quad (15)$$

where  $c_k = [e_k^2 + \epsilon \ e_{k-1}^2 + \epsilon \ \dots \ e_{k-P+1}^2 + \epsilon] \in \mathcal{R}^{1 \times P}$  and  $\epsilon$  is a very small positive scalar of the order of  $10^{-12}$  whereas  $\lambda = 1 - \frac{1}{(c_1 M)}$  is the *forgetting factor* and  $c_1$  is an integer. A zero vector can initially be assigned to vector  $c_k$ . With  $\sigma_{1,k}$  known,  $\theta_k$  is obtained and thus  $\gamma_k$  is evaluated using (13). In turn,  $\gamma_k$  is used in (9) to obtain  $\alpha_k$  which is then used in (11) to obtain the weight vector  $w_k$ .

### B. RSMAP Algorithm With Variable Threshold

The RSMAP algorithm with variable threshold, namely, the RSMAP2 algorithm, uses all the equations of the RSMAP1 algorithm, i.e., (13) to (15), along with

$$\eta_k = \beta\eta_{k-1} + (1 - \beta) \min \left( \eta_{k-1}, \frac{|d_k^2 - y_k^2|}{d_k^2} \right) \quad (16)$$

$$\sigma_{2,k}^2 = \lambda\sigma_{2,k-1}^2 + (1 - \lambda) \min (\sigma_{2,k-1}^2, \sigma_{1,k}^2) \quad (17)$$

where  $\eta_k$  is a parameter used to control the switching between error bounds  $\gamma_c$  and  $\|e_k\|_\infty - \nu\theta_k$ ,  $\sigma_{2,k}$  is the variance of the error signal,  $\beta = 1 - \frac{1}{(c_2 M)}$  is the *forgetting factor* used to obtain the threshold

$$\gamma_c = \{ \gamma_{c,0}^2 + \Upsilon [1 + \text{sign}(1 - \eta_k)] \sigma_{2,k}^2 \}^{1/2} \quad (18)$$

where  $\Upsilon$  is a tuning constant and  $\gamma_{c,0}^2$  is an approximate estimate of the noise variance  $\sigma_v^2$ . If no information is available about the noise variance, a zero value can be assigned to  $\gamma_{c,0}^2$ . In effect, the RSMAP2 algorithm uses (14) and (15) to obtain  $\sigma_{1,k}$  and, in turn, evaluates  $\theta_k$  as  $\theta_k = 1.88\sigma_{1,k}$ ; it then uses (16)–(18) to obtain  $\gamma_c$  and (13) to obtain  $\gamma_k$ . With  $\gamma_k$  known, (9) yields  $\alpha_k$  which is then used in (11) to obtain the new weight vector  $w_k$ . The estimators in (13)–(18) are similar to the estimators we have used in the robust quasi-Newton adaptation algorithm we proposed in [13]. Note that the estimate  $\gamma_c$  in (18) is different from that used in the SMLMS family in [14] and [15].

### C. Discussion

We have used the median of the squared error-signal samples over a finite window for the evaluation of  $\sigma_{1,k}$  in (15) in order to render the estimate of  $\sigma_{1,k}$  robust with respect to outliers for both the RSMAP1 and RSMAP2 algorithms. On the other hand, we have used  $\sigma_{2,k}$  to estimate the error bound  $\gamma_c$  for the RSMAP2 algorithm. In order to achieve robustness with respect to sudden system changes, the estimator of  $\sigma_{2,k}$  uses the minimum of the previous and current values of  $\sigma_{1,k}$ . Eq. (16) controls how long  $\gamma_c$  in (18) would work with  $\gamma_{c,0}$ .

As reported in [9], with a small  $\gamma$  the SSMAP algorithm yields faster convergence and increased steady-state misalignment; on the other hand, with a large  $\gamma$  it yields slower convergence and reduced steady-state misalignment. Under these circumstances, optimal performance can be achieved by using a small error bound during transience and a large error bound during steady state. In the proposed RSMAP algorithm, optimal performance is achieved by choosing the initial values of  $\sigma_{1,k}$ ,  $\sigma_{2,k}$ , and  $\eta_k$  to be large. In such a situation, during transience the algorithm would work with error bound  $\gamma_k = \gamma_c$  which is significantly smaller than  $\|e_k\|_\infty - \nu\theta_k$  and, as a result, the transient state would decay at a fast rate. On the other hand, during steady state the algorithm would work with error bound  $\gamma_k = \|e_k\|_\infty - \nu\theta_k$  as  $\theta_k < \gamma_c$  during steady-state which would yield reduced steady-state misalignment. In addition, this choice of  $\gamma_k$  would work in the occurrence of impulsive noise to yield robust performance with respect to outliers. In the presence of a sudden system disturbance,  $\theta_k$  would tend to grow and, therefore, error bound  $\gamma_k = \gamma_c$  would come into play. As a result, the re-adaptation capability of the proposed algorithm would be retained. Although a rough choice of the initial values of  $\sigma_{1,k}$ ,  $\sigma_{2,k}$ , and  $\eta_k$  would work, simulation results have shown that the choices  $\sigma_{1,0} = \frac{20E_1}{\sigma_v^2}$ ,  $\sigma_{2,0} = \frac{20E_2}{\sigma_v^2}$ , and  $\eta_0 = \frac{20E_3}{\sigma_v^2}$ , where  $\hat{\sigma}_v^2$  is a crude approximation of the noise variance, give good results. Parameters  $E_1, E_2, E_3, c_1, c_2$  are integers which can be chosen in the range 1 to 8 based on simulation results. The proposed algorithms are more robust with respect to the choice of tuning parameters for medium to high signal-to-noise ratios (SNRs) as compared to low SNRs. Parameter  $\Upsilon$  in the RSMAP2 algorithm is a constant chosen to be 2.5. The regularization matrix  $\delta \mathbf{I}$ , where  $\mathbf{I}$  is the  $L \times L$  identity matrix and  $\delta$  is a small constant, is added to the correlation matrix  $\mathbf{X}_k^T \mathbf{X}_k$  in (8), (11), and (12) to assure its invertibility if the signal becomes ill-conditioned. The value of  $\delta$  should be very small, of the order of  $10^{-6}$ , so that it does not noticeably influence the behavior of the algorithm.

The steady-state performance of the proposed algorithm is not influenced by the tuning parameters as it is determined by parameter  $\nu$ . The convergence speed could become sensitive to changes in the tuning parameters especially for low SNRs and unknown systems of high order. However, no such problems have been experienced for high SNRs and unknown systems of low to moderate orders. For unknown systems of low order, say, 7 to 15, of the type commonly used in communication systems, the tuning parameters are easy to adjust and the performance of the adaptive filter is quite robust even for relatively low SNRs.

Since  $\sigma_v$  is prespecified in the RSMAP1 algorithm, it can be chosen to achieve a good balance between convergence speed and computational savings in the RSMAP1 algorithm. On the other hand, the RSMAP2 algorithm can be used in applications where  $\gamma_c$  cannot be prespecified. In most other situations, the two versions of the proposed algorithm offer similar performance.

## V. STEADY-STATE ANALYSIS OF RSMAP ALGORITHM

In this section, we provide a steady-state analysis of the proposed RSMAP algorithm. The analysis is based on the framework of an energy-conservation relation described in [8], [16] which was used to analyze several adaptation algorithms, for example, in [17]–[19]. Since the formula in (11) is essentially the same as that in (12), the steady-state analysis of the proposed RSMAP algorithm can be carried out by replacing  $\mu$  in (12) by  $\alpha_k$  given by (9) and then proceeding as in [8]. The update equation in (7) can be expressed in terms of the weight-error vector  $\hat{\mathbf{w}}_k = \mathbf{w}_o - \mathbf{w}_k$  as

$$\hat{\mathbf{w}}_k = \hat{\mathbf{w}}_{k-1} - \alpha_k \mathbf{X}_k \left( \delta \mathbf{I} + \mathbf{X}_k^T \mathbf{X}_k \right)^{-1} \mathbf{e}_k \quad (19)$$

where  $\mathbf{w}_o$  denotes the weight vector of the unknown system and  $\delta \mathbf{I}$  is the regularization matrix. Premultiplying both sides in (19) by the input signal matrix, we obtain

$$\mathbf{X}_k^T \hat{\mathbf{w}}_k = \mathbf{X}_k^T \hat{\mathbf{w}}_{k-1} - \alpha_k \mathbf{X}_k^T \mathbf{X}_k \left( \delta \mathbf{I} + \mathbf{X}_k^T \mathbf{X}_k \right)^{-1} \mathbf{e}_k. \quad (20)$$

Using the definition of the noise-free *a posteriori* error given by

$$\hat{\mathbf{e}}_k = \mathbf{X}_k^T \hat{\mathbf{w}}_k \quad (21)$$

and the noise-free *a priori* error given by

$$\hat{\mathbf{e}}_k = \mathbf{X}_k^T \hat{\mathbf{w}}_{k-1} \quad (22)$$

in (20), we obtain

$$\hat{\mathbf{e}}_k = \hat{\mathbf{e}}_k - \alpha_k \mathbf{X}_k^T \mathbf{X}_k \left( \delta \mathbf{I} + \mathbf{X}_k^T \mathbf{X}_k \right)^{-1} \mathbf{e}_k. \quad (23)$$

Hence, we can write

$$\left( \delta \mathbf{I} + \mathbf{X}_k^T \mathbf{X}_k \right)^{-1} \mathbf{e}_k = \frac{1}{\alpha_k} \left( \mathbf{X}_k^T \mathbf{X}_k \right)^{-1} (\hat{\mathbf{e}}_k - \hat{\mathbf{e}}_k) \quad (24)$$

Substituting  $\left( \delta \mathbf{I} + \mathbf{X}_k^T \mathbf{X}_k \right)^{-1} \mathbf{e}_k$  into (19), we get

$$\hat{\mathbf{w}}_k + \mathbf{X}_k \left( \mathbf{X}_k^T \mathbf{X}_k \right)^{-1} \hat{\mathbf{e}}_k = \hat{\mathbf{w}}_{k-1} + \mathbf{X}_k \left( \mathbf{X}_k^T \mathbf{X}_k \right)^{-1} \hat{\mathbf{e}}_k \quad (25)$$

By taking the square of the  $L_2$  norm on both sides of the above equation, after some manipulation we obtain the following energy conservation relation [8]:

$$\|\hat{\mathbf{w}}_k\|^2 + \hat{\mathbf{e}}_k^T \left( \mathbf{X}_k^T \mathbf{X}_k \right)^{-1} \hat{\mathbf{e}}_k = \|\hat{\mathbf{w}}_{k-1}\|^2 + \hat{\mathbf{e}}_k^T \left( \mathbf{X}_k^T \mathbf{X}_k \right)^{-1} \hat{\mathbf{e}}_k. \quad (26)$$

Taking the expectation of both sides in (26), we obtain

$$E \left[ \|\hat{\mathbf{w}}_k\|^2 \right] + E \left[ \hat{\mathbf{e}}_k^T \left( \mathbf{X}_k^T \mathbf{X}_k \right)^{-1} \hat{\mathbf{e}}_k \right] = E \left[ \|\hat{\mathbf{w}}_{k-1}\|^2 \right] + E \left[ \hat{\mathbf{e}}_k^T \left( \mathbf{X}_k^T \mathbf{X}_k \right)^{-1} \hat{\mathbf{e}}_k \right]. \quad (27)$$

At steady state, we have  $E[\|\hat{\mathbf{w}}_k\|^2] = E[\|\hat{\mathbf{w}}_{k-1}\|^2]$  and hence the above relation assumes the form

$$E \left[ \hat{\mathbf{e}}_k^T \left( \mathbf{X}_k^T \mathbf{X}_k \right)^{-1} \hat{\mathbf{e}}_k \right] = E \left[ \hat{\mathbf{e}}_k^T \left( \mathbf{X}_k^T \mathbf{X}_k \right)^{-1} \hat{\mathbf{e}}_k \right]. \quad (28)$$

Substituting  $\hat{\mathbf{e}}_k$  in (23) into (28), we obtain

$$\begin{aligned} & E \left[ \hat{\mathbf{e}}_k^T \left( \mathbf{X}_k^T \mathbf{X}_k \right)^{-1} \hat{\mathbf{e}}_k \right] \\ &= E \left[ \hat{\mathbf{e}}_k^T \left( \mathbf{X}_k^T \mathbf{X}_k \right)^{-1} \hat{\mathbf{e}}_k^T - \alpha_k \hat{\mathbf{e}}_k^T \left( \delta \mathbf{I} + \mathbf{X}_k^T \mathbf{X}_k \right)^{-1} \mathbf{e}_k \right. \\ &\quad \left. - \alpha_k \mathbf{e}_k^T \left( \delta \mathbf{I} + \mathbf{X}_k^T \mathbf{X}_k \right)^{-1} \hat{\mathbf{e}}_k \right. \\ &\quad \left. + \alpha_k^2 \mathbf{e}_k^T \left( \delta \mathbf{I} + \mathbf{X}_k^T \mathbf{X}_k \right)^{-1} \left( \mathbf{X}_k^T \mathbf{X}_k \right) \right. \\ &\quad \left. \times \left( \delta \mathbf{I} + \mathbf{X}_k^T \mathbf{X}_k \right)^{-1} \mathbf{e}_k \right]. \quad (29) \end{aligned}$$

Using the definition of the input-signal correlation matrix, i.e.,  $\mathbf{R}_k = \mathbf{X}_k^T \mathbf{X}_k$ , and assuming that

$$\mathbf{S}_k = \left( \delta \mathbf{I} + \mathbf{X}_k^T \mathbf{X}_k \right)^{-1} \approx \mathbf{R}_k^{-1} \quad (30)$$

for  $\delta \ll 1$ , (29) can be simplified to

$$E \left[ \alpha_k^2 \mathbf{e}_k^T \mathbf{T}_k \mathbf{e}_k \right] = E \left[ \alpha_k \hat{\mathbf{e}}_k^T \mathbf{S}_k \mathbf{e}_k \right] + E \left[ \alpha_k \mathbf{e}_k^T \mathbf{S}_k \hat{\mathbf{e}}_k \right] \quad (31)$$

where  $\mathbf{T}_k = \mathbf{S}_k \mathbf{R}_k \mathbf{S}_k$ . At steady state  $\|\mathbf{e}_k\|_\infty \approx |e_k|$  and  $\frac{|e_k|}{\sigma_{1,k}} \rightarrow 1$ . Since  $\gamma_k = \|\mathbf{e}_k\|_\infty - \nu \theta_k$  at steady state, then for a small value of  $\nu$  (9) yields

$$\alpha_k \approx 1.88\nu. \quad (32)$$

Thus by using a deterministic value for  $\alpha_k$ , (31) simplifies to

$$\alpha_k E \left[ \hat{\mathbf{e}}_k^T \mathbf{T}_k \mathbf{e}_k \right] = E \left[ \hat{\mathbf{e}}_k^T \mathbf{S}_k \mathbf{e}_k \right] + E \left[ \mathbf{e}_k^T \mathbf{S}_k \hat{\mathbf{e}}_k \right]. \quad (33)$$

The noisy and noise-free *a priori* error vectors  $\mathbf{e}_k$  and  $\hat{\mathbf{e}}_k$ , respectively, are interrelated by

$$\mathbf{e}_k = \hat{\mathbf{e}}_k + \mathbf{v}_k \quad (34)$$

where  $\mathbf{v}_k = [v_k \ v_{k-1} \ \cdots \ v_{k-L+1}]^T$  is the measurement noise vector. Now substituting (34) into (33), we obtain

$$\alpha_k E \left[ (\hat{\mathbf{e}}_k + \mathbf{v}_k)^T \mathbf{T}_k (\hat{\mathbf{e}}_k + \mathbf{v}_k) \right] = 2E \left[ \hat{\mathbf{e}}_k^T \mathbf{S}_k (\hat{\mathbf{e}}_k + \mathbf{v}_k) \right]. \quad (35)$$

Considering the noise signal  $\mathbf{v}_k$  to be white as well as statistically independent of the input signal and neglecting the dependency of  $\hat{\mathbf{e}}_k$  on  $\mathbf{v}_k$ , the expression in (35) can be further simplified to

$$\alpha_k E \left[ \hat{\mathbf{e}}_k^T \mathbf{T}_k \hat{\mathbf{e}}_k \right] + \alpha_k E \left[ \mathbf{v}_k^T \mathbf{T}_k \mathbf{v}_k \right] = 2E \left[ \hat{\mathbf{e}}_k^T \mathbf{S}_k \hat{\mathbf{e}}_k \right]. \quad (36)$$

By using the definition in (30), we obtain  $\mathbf{T}_k \approx \mathbf{S}_k$ ; with this modification applying the trace operation to both sides of (36), we obtain

$$E \left\{ \text{tr} \left[ \hat{\mathbf{e}}_k \hat{\mathbf{e}}_k^T \mathbf{S}_k \right] \right\} = \frac{\alpha_k}{2 - \alpha_k} E \left\{ \text{tr} \left[ \mathbf{v}_k \mathbf{v}_k^T \mathbf{S}_k \right] \right\}. \quad (37)$$

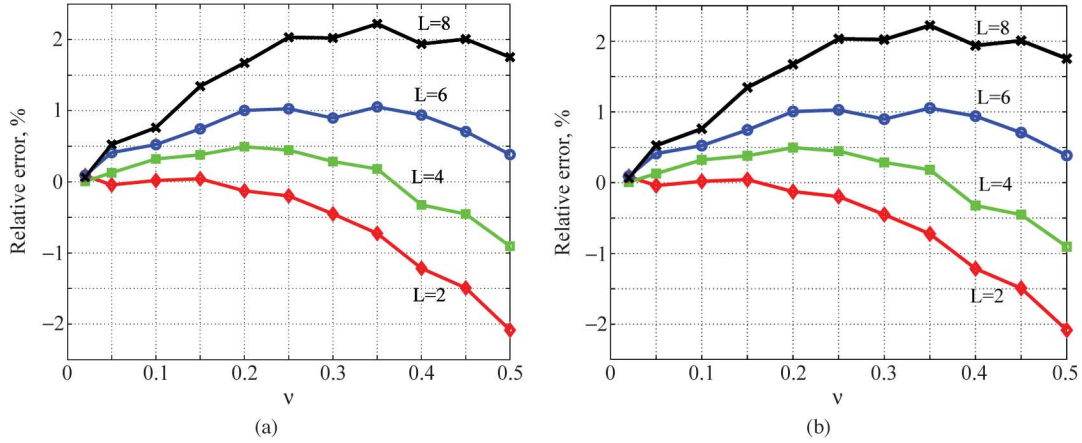


Fig. 1. Relative error between expected and experimental MSE: (a) RSMAP algorithm with fixed threshold. (b) RSMAP algorithm with variable threshold.

Now assuming that  $\mathbf{S}_k$  is statistically independent of both the *a priori* error at steady state and the noise signal, we obtain

$$\text{tr} \left\{ E \left[ \hat{\mathbf{e}}_k \hat{\mathbf{e}}_k^T \right] E \left[ \mathbf{S}_k \right] \right\} = \frac{\alpha_k}{2 - \alpha_k} \text{tr} \left\{ E \left[ \mathbf{v}_k \mathbf{v}_k^T \right] E \left[ \mathbf{S}_k \right] \right\}. \quad (38)$$

This result is used in the next subsection to derive an expression for the excess MSE in the RSMAP algorithm.

#### A. Excess MSE for RSMAP Algorithm

In order to derive an expression for the excess MSE in the proposed RSMAP algorithm, we have to deduce an expression for the noise-free error covariance  $E[\hat{\mathbf{e}}_k \hat{\mathbf{e}}_k^T]$ . The update equation in (11) during steady-state can be written in terms of (34) as

$$\hat{\mathbf{e}}_k = \hat{\mathbf{e}}_k - \alpha_k \mathbf{e}_k = (1 - \alpha_k) \hat{\mathbf{e}}_k - \alpha_k \mathbf{v}_k. \quad (39)$$

Now following the steps in [8], we obtain

$$E \left[ \hat{\mathbf{e}}_k \hat{\mathbf{e}}_k^T \right] = E \left[ \hat{\mathbf{e}}_k^2 \right] \mathbf{U}_1 + \alpha_k^2 \sigma_v^2 \mathbf{U}_2 \quad (40)$$

where  $\mathbf{U}_1$  and  $\mathbf{U}_2$  are diagonal matrices given by

$$\mathbf{U}_1 = \begin{bmatrix} 1 & & & \\ & (1 - \alpha_k)^2 & & \\ & & \ddots & \\ & & & (1 - \alpha_k)^{2(L-1)} \end{bmatrix} \quad (41)$$

and

$$\mathbf{U}_2 = \begin{bmatrix} 0 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 + \dots + (1 - \alpha_k)^{2L} \end{bmatrix} \quad (42)$$

respectively. Since the step size  $\alpha_k$  is very small at steady state, we obtain  $\alpha_k^2 \sigma_v^2 \mathbf{U}_2 \approx \mathbf{0}$  and  $\mathbf{U}_1 \approx \mathbf{I}$ . Consequently, (40) assumes the form

$$E \left[ \hat{\mathbf{e}}_k \hat{\mathbf{e}}_k^T \right] = E \left[ \hat{\mathbf{e}}_k^2 \right] \mathbf{U}_1. \quad (43)$$

Using (43) in (38), we obtain

$$E \left[ \hat{\mathbf{e}}_k^2 \right] = \frac{\alpha_k}{2 - \alpha_k} \sigma_v^2 \frac{\text{tr} \{ E[\mathbf{S}_k] \}}{\text{tr} \{ \mathbf{U}_1 E[\mathbf{S}_k] \}}. \quad (44)$$

For a sufficiently small  $\alpha_k$ , we can assume that  $\mathbf{U}_1 \approx \mathbf{I}$  and in this case (44) yields the excess MSE as

$$\text{EMSE} = \frac{\alpha_k}{2 - \alpha_k} \sigma_v^2 \quad (45)$$

As can be seen, the excess MSE is independent of the projection order.

#### B. Verification of EMSE

In this subsection, we examine the accuracy of the expression for the EMSE given in (45) in a system-identification application. The unknown system in this experiment was a sixteenth-order FIR filter with a cut-off frequency of 0.3, which was designed using MATLAB command  $\mathbf{w}_{\text{opt}} = \text{fir1}(M - 1, 0.3)$ . The weight vector of the unknown system was normalized to have unity power. An IIR filter with a pole at 0.95 was used to produce a colored input signal from a zero-mean white Gaussian noise signal with a variance of unity. The desired signal was contaminated with a zero-mean white Gaussian noise signal with variance  $\sigma_v^2 = 10^{-4}$ . The expected steady-state MSE was obtained as  $\text{MSE} = \sigma_v^2 + \text{EMSE}$  where EMSE is given in (45). The parameters for the proposed algorithms were set to  $c_1 = 1$ ,  $\mathbf{w}_0 = \mathbf{0}$ ,  $E_1 = E_2 = E_3 = 1$ , and  $\gamma_c = \sqrt{5\sigma_v^2}$  for all  $L$  and  $c_2 = 2$  and 3 for  $L = [2, 4]$  and  $L = [6, 8]$ , respectively. The relative error between the expected and experimental MSE defined as

$$\text{Relative error} = \frac{\text{Expected MSE} - \text{Experimental MSE}}{\text{Expected MSE}} 100 \quad (46)$$

with fixed and variable thresholds for different values of  $L$  and  $\nu$  is plotted in Fig. 1(a), (b).

As can be seen in Fig. 1(a) and (b), (46) provides a fairly accurate estimate of the relative MSE error for values of  $\nu$  in the range  $\nu = (0, 0.5]$ . It should be mentioned that a reduced  $\nu$  does not affect the convergence speed or the re-adaptation capability

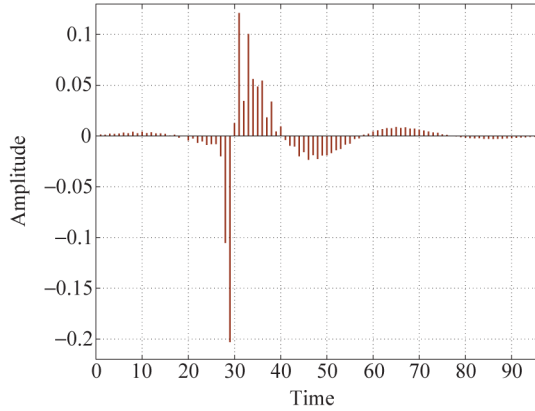


Fig. 2. Impulse response of network echo path in first experiment.

of the adaptive filter and at the same time for values of  $\nu$  in the range  $(0, 0.2]$ , the relative error is quite small for different values of  $L$  as can be seen in Fig. 1(a), (b).

The above analysis can be used to select appropriate values for parameters  $\nu$  and  $L$ .

## VI. SIMULATION RESULT

In this section, we compare the performance of the two versions of the RSMAP algorithm with that of some known AP and SMAP algorithms through extensive simulations. Two applications, namely, a system-identification and an echo-cancellation application were considered.

### A. System Identification Application

In the first set of experimental results, the performance of the RSMAP1 and RSMAP2 algorithms is compared with that of the conventional AP, SSMAP, and SM proportionate AP (SMPAP) algorithms reported in [3], [6], and [10], respectively, for a system-identification application. A network echo path of length  $M = 96$  with the impulse response illustrated in Fig. 2 was chosen as the unknown system.

A colored input signal for the adaptive filter was generated by filtering a zero-mean white Gaussian noise signal with power 10 by using an IIR filter with transfer function [20]

$$H(z) = \frac{z^2 + 0.5z + 0.81}{z^2 - 0.59z + 0.4}$$

The impulsive noise was modeled as  $\phi_k = \omega_k \chi_k$  where  $\omega_k$  is a Bernoulli process with  $P(\omega_k = 1) = p = 0.01$  and  $\chi_k$  is a zero-mean Gaussian signal with variance  $\sigma_\chi^2 = 10^4 \sigma_y^2$  where  $\sigma_y^2$  is the power of the uncorrupted output signal [21]. The weight vector was initialized as  $\mathbf{w} = \mathbf{0}$  in all algorithms and all experiments. In addition, the error bounds,  $\gamma$  for the SSMAP and SMPAP algorithms and  $\gamma_c$  for the RSMAP1 algorithm, were set to  $\sqrt{5\sigma_v^2}$  in all experiments. Unless otherwise stated, the parameters of the RSMAP algorithms were chosen as  $c_1 = 1$ ,  $P = 15$ ,  $\nu = 0.05$ ,  $c_2 = 1$ ,  $E_1 = 2$ ,  $E_2 = 2$ , and  $E_3 = 2$ . The measurement noise,  $v_k$ , added to the desired signal was also a Gaussian noise signal with a zero mean and variance  $\sigma_v^2 = 10^{-3}$  which corresponds to an SNR of 30 dB. The impulse response of the echo path was multiplied by  $-1$  at iteration  $10^4$  to investigate the re-adaptation capability of the algorithms. Impulsive

noise was added to the desired signal at iterations  $0.5 \times 10^4$  and  $1.5 \times 10^4$ . The learning curves obtained averaged over 1000 independent trials for  $L = 8$  are illustrated in Fig. 3(a). As can be seen, the proposed RSMAP algorithms yield significantly reduced steady-state misalignment and improved robustness with respect to impulsive noise as compared to the other algorithms without compromising the initial speed of convergence. The evolution of the averaged step size  $\alpha_k$  for the SSMAP, SMPAP, and the proposed algorithms is illustrated in Fig. 3(b). As can be seen, the proposed algorithms yield lower values of  $\alpha_k$  during steady state, namely, 0.004, compared to 0.0834 and 0.0823 in the SSMAP and SMPAP algorithms, respectively. As a result, a much reduced steady-state misalignment is achieved in the proposed algorithms. In addition, since  $\alpha_k$  returns back to high values during sudden system changes, the proposed algorithms preserve their re-adaptation capability.

To examine the effect of projection order on the performance of the proposed algorithms, we repeated the same experiment with the variance of the measurement noise changed to  $\sigma_v^2 = 10^{-6}$ , which corresponds to an SNR of 60 dB, for values of  $L = 2, 4, 6$ , and 8. In this experiment, the parameters of the RSMAP algorithms were chosen as  $E_1 = E_2 = E_3 = 1$  and impulsive noise was added to the desired signal at iteration  $2.5 \times 10^4$ . Parameters  $c_1$  and  $c_2$  were both set to 6, 3, 2, and 1 for the simulations of Fig. 3(c)–(f), respectively. The learning curves shown were obtained using 1000 independent trials. As can be seen, as the projection order is increased, the convergence speed of the AP and SSMAP algorithms is increased at the cost of increased steady-state misalignment, which is consistent with the analysis presented in [8] and [9]. On the other hand, the proposed algorithm yields significantly reduced steady-state misalignment relative to that in the AP and SSMAP algorithms especially when the projection order is large. The numerical values of the steady-state MSE in dB for the AP, SSMAP, RSMAP1, and RSMAP2 algorithms for  $L = 2$  and  $L = 8$  are given in Table I. As can be seen, the use of the proposed algorithms leads to a reduction in the steady-state misalignment in the range of 4.0 to 7.9 dB relative to the AP algorithm and 0.4 to 6.8 dB relative to the SSMAP algorithm depending on the value of  $L$ . In addition, the proposed algorithms offer improved robustness with respect to impulsive noise.

### B. Echo Cancellation Application

In the second set of experimental results, the performance of the RSMAP algorithms is compared with that of the competing algorithms for the case of an echo-cancellation application. In such an application, the length of the acoustic echo path would be significantly larger than the length of the network echo path and the adaptive filter would require thousands of taps to successfully cancel the acoustic echo signal. The impulse response of the acoustic echo path assumed is illustrated in Fig. 4(a). Only the first 512 samples were used to produce the echo signal. An IIR filter with transfer function

$$H(z) = \frac{1}{z^4 - 0.95z^3 - 0.19z^2 - 0.09z + 0.5}$$

was used to bandlimit a zero-mean white Gaussian noise signal with variance one which could serve as the input signal for the

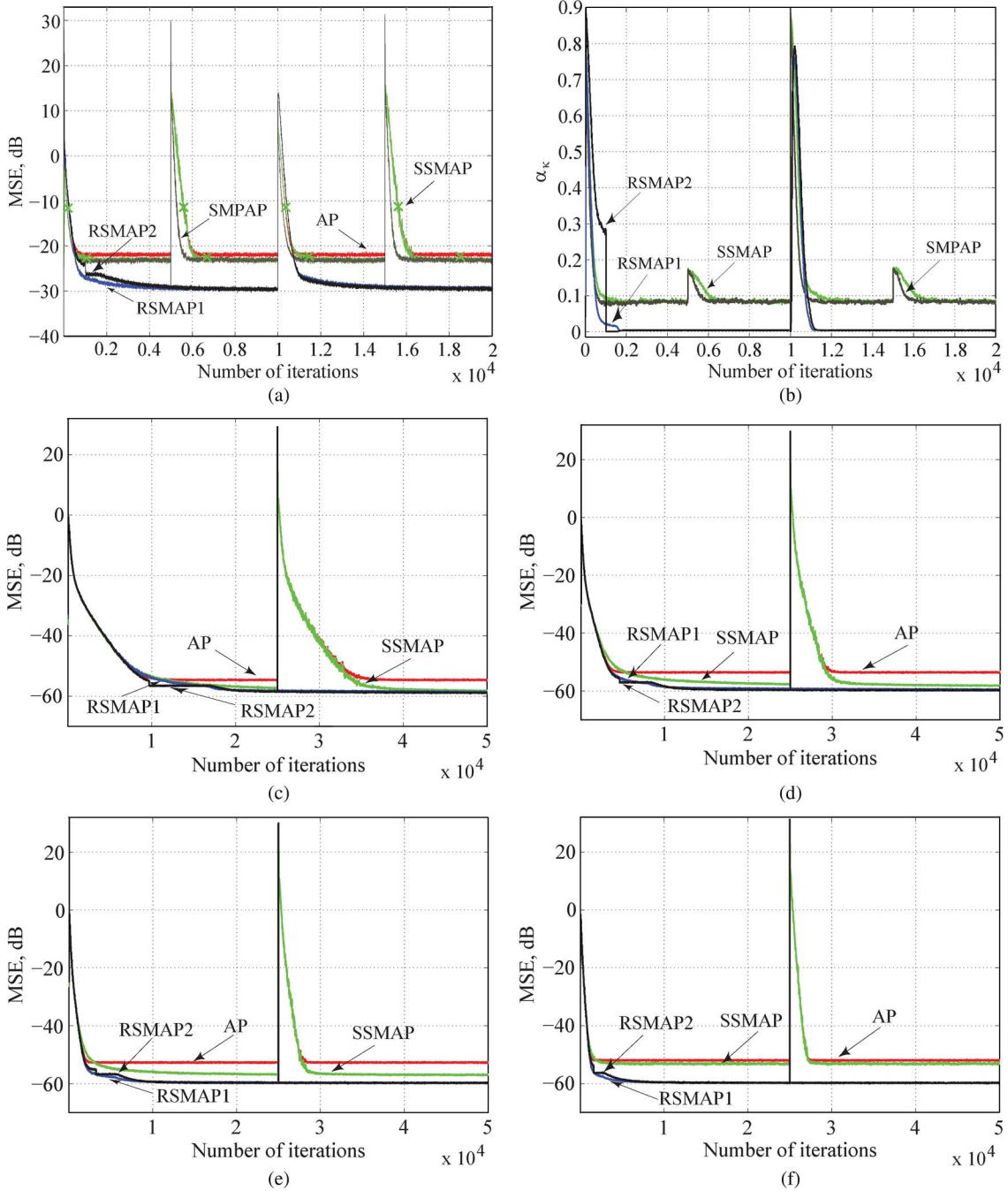


Fig. 3. Simulation results for system-identification application. (a) Learning curves:  $L = 8$ , SNR = 30 dB. (b) Evolution of  $\alpha_k$ :  $L = 8$ , SNR = 30 dB. (c) Learning curves:  $L = 2$ , SNR = 60 dB. (d) Learning curves:  $L = 4$ , SNR = 60 dB. (e) Learning curves:  $L = 6$ , SNR = 60 dB. (f) Learning curves:  $L = 8$ , SNR = 60 dB.

TABLE I  
NUMERICAL VALUES OF THE STEADY-STATE MSE, dB

L	AP	SSMAP	RSMAP1	RSMAP2
2	-54.6	-58.2	-58.6	-58.9
8	-51.9	-53.0	-59.8	-59.8

adaptive filters [6]. The input signal would thus be more correlated than that used in the previous application. The measurement noise was a zero-mean white Gaussian noise signal with variance  $10^{-2}$  with impulsive noise at iterations  $2.5 \times 10^4$  and  $7.5 \times 10^4$ . The impulse response of the echo path was multiplied

by  $-1$  at iteration  $5 \times 10^4$ . The parameters for the proposed RSMAP algorithms were set to  $E_1 = E_2 = E_3 = 5$  and  $L = 8$  in all subsequent experiments. The learning curves obtained by using the AP, SSMAP, SMPAP, SMRAP1, and SMRAP2 algorithms for 1000 independent trials are illustrated in Fig. 4(b). As can be seen, the RSMAP algorithms yield a significant reduction in the steady-state misalignment compared to the other algorithms while retaining their re-adaptation capability and initial speed of convergence. The proposed algorithms also exhibit robust performance with respect to the impulsive noise.

Next we repeated the same experiment with real speech signals as the input signals. Twenty speech signals were recorded

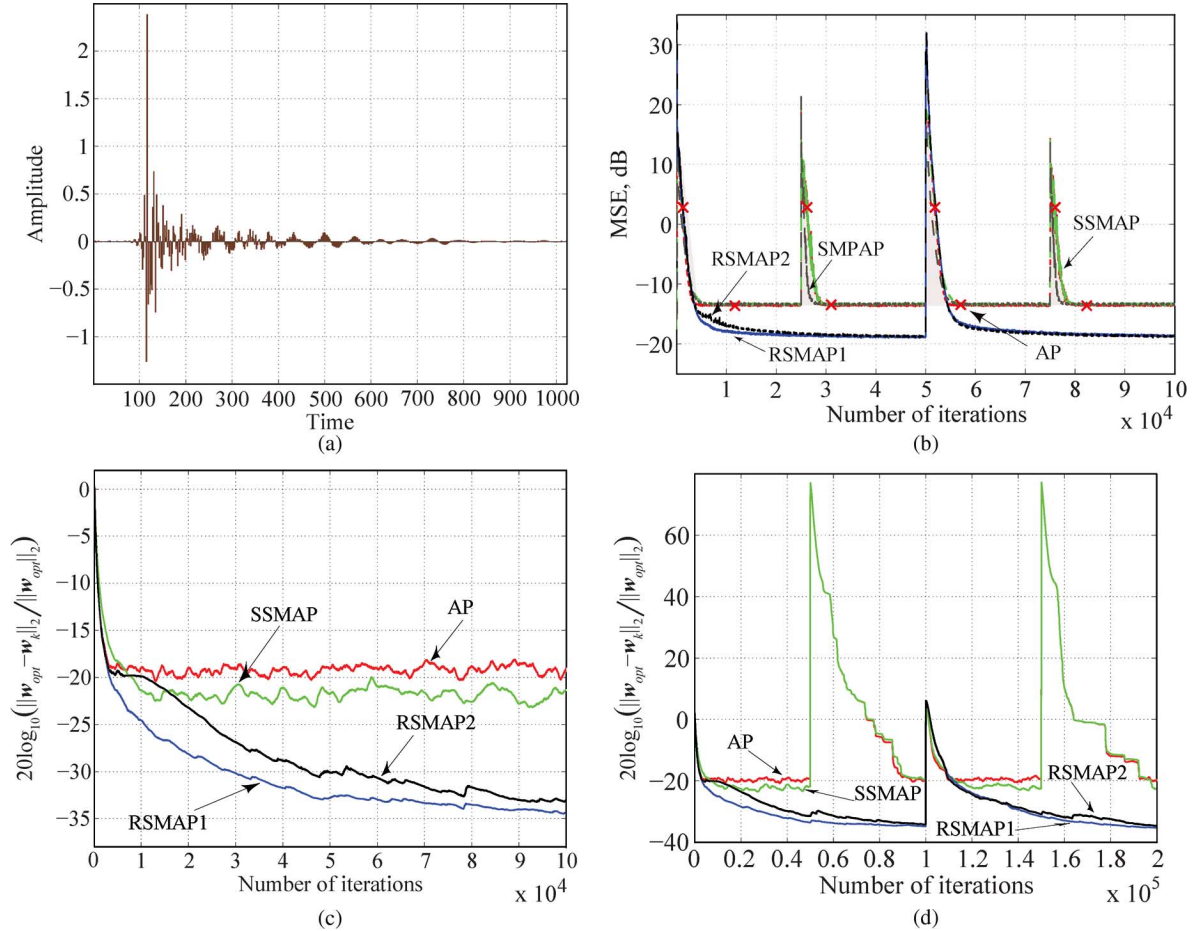


Fig. 4. Simulation results for echo-cancellation application. (a) Impulse response of acoustic echo path. (b) Learning curves: Gaussian signals as input  $P = 15$  ( $\times$  symbols correspond to AP algorithm). (c) Learning curves: speech signals as input  $P = 15$ . (d) Learning curves: speech signals as input  $P = 215$ .

and each was then used as the input signal in a given trial. The desired signal and each speech signal were contaminated with a zero-mean white Gaussian noise signal with a variance of  $10^{-4}$ . In order to obtain a smooth learning curve, the misalignment was evaluated as  $20 \log_{10} \left( \frac{\|w_{\text{opt}} - w_k\|_2}{\|w_{\text{opt}}\|_2} \right)$  where  $w_{\text{opt}}$  is the impulse response shown in Fig. 4(a). The learning curves obtained in 20 independent trials are illustrated in Fig. 4(c) and (d). The plots of Fig. 4(c) correspond to the case where no impulsive noise was added to the desired signal. The plots of Fig. 4(d) correspond to the case where impulsive noise of duration  $100T_s$  was added to the desired signal at iterations  $0.5 \times 10^5$  and  $1.5 \times 10^5$ , where  $T_s$  is the sampling period, and the impulse response of the echo path was multiplied by  $-1$  at iteration  $1.0 \times 10^5$ . As can be seen in Fig. 4(c) and (d), the RSMAP1 and RSMAP2 algorithms yield reduced steady-state misalignment for the same convergence speed as compared to the other algorithms. In addition, the RSMAP1 and RSMAP2 algorithms are more robust with respect to a long burst of impulsive noise and offer similar re-adaptation capability relative to the other algorithms. The RSMAP2 algorithm yields slightly more misalignment compared to the RSMAP1 algorithm due to the fact that the latter algorithm starts with a much lower error bound than the former. The long burst of impulsive noise replicates crosstalk and intersymbol interference situations commonly found in two-way telephone and communication systems.

## VII. CONCLUSION

A set-membership AP adaptive-filtering algorithm that achieves robustness through the use of two error bounds has been proposed. A steady-state analysis of the proposed algorithm that leads to an approximate formula for the expected MSE has been carried out. This formula can be used to estimate appropriate values for parameters  $\nu$  and  $L$ . On the other hand, extensive simulation results have shown that the guidelines presented for the choice of parameters  $\sigma_{1,0}$ ,  $\sigma_{2,0}$ ,  $\lambda$ ,  $\beta$ ,  $\eta_0$  work quite well in practice. The two versions of the proposed algorithm were applied to system-identification and echo-cancellation applications. The simulation results obtained show that algorithms RSMAP1 and RSMAP2 perform much better than the conventional AP, SSMAP, and SMPAP algorithms in terms of steady-state misalignment and robustness with respect to impulsive noise without compromising the initial convergence speed and re-adaptation capability.

## REFERENCES

- [1] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*, 3rd ed. New York: Springer, 2008.
- [2] A. H. Sayed, *Fundamentals of Adaptive Filtering*. Hoboken, NJ: Wiley, 2003.
- [3] K. Ozeki and T. Umeda, "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties," *Electron. Commun. Japan*, vol. 67-A, pp. 19–27, 1984.



- [4] S. L. Gay and J. Benesty, *Acoustic Signal Processing for Telecommunication*. Boston, MA: Springer, 2000.
- [5] S. G. Kraizer and D. R. Morgan, "The partial-rank algorithm for adaptive beamforming," in *Proc. SPIE Int. Soc. Opt. Eng.*, 1985, vol. 564, pp. 9–14.
- [6] S. Werner and P. S. R. Diniz, "Set-membership affine projection algorithm," *IEEE Signal Process. Lett.*, vol. 8, no. 8, pp. 231–235, Aug. 2001.
- [7] P. S. R. Diniz and S. Werner, "Set-membership binormalized data-reusing LMS algorithms," *IEEE Trans. Signal Process.*, vol. 51, no. 1, pp. 124–134, Jan. 2003.
- [8] H. C. Shin and A. H. Sayed, "Mean-square performance of a family of affine projection algorithms," *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 90–102, Jan. 2004.
- [9] P. S. R. Diniz, "Convergence performance of the simplified set-membership affine projection algorithm," *Circuits, Syst., Signal Process.*, vol. 30, no. 2, pp. 439–462, Nov. 2010.
- [10] S. Werner, J. A. Apolinario, Jr., P. S. R. Diniz, and T. I. Laakso, "Set-membership approach to normalized proportionate adaptation algorithms," in *Proc. Eur. Signal Process. Conf.*, Antalya, Turkey, Sep. 2005.
- [11] S. Werner, P. S. R. Diniz, and J. E. W. Moreira, "Set-membership affine projection algorithm with variable data-reuse factor," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2006, pp. 261–264.
- [12] Y. Zou, S. C. Chan, and T. S. Ng, "A recursive least M-estimate (RLM) adaptive filter for robust filtering in impulsive noise," *IEEE Signal Process. Lett.*, vol. 7, no. 11, pp. 324–326, Nov. 2000.
- [13] M. Z. A. Bhotto and A. Antoniou, "Robust quasi-Newton adaptive filtering algorithms," *IEEE Trans. Circuits Syst. II*, vol. 58, no. 8, pp. 537–541, Aug. 2011.
- [14] L. Guo and Y. F. Huang, "Set membership adaptive filtering with parameter-dependent error bound tuning," in *Proc. Int. Conf. Acoust. Speech, Signal Process.*, 2005, vol. 17, pp. 369–372.
- [15] J. F. Galdino, J. A. Apolinario, Jr., and M. L. R. de Campos, "A set-membership NLMS algorithm with time varying error bound," in *Proc. IEEE Int. Symp. Circuits Syst.*, 2006, pp. 277–280.
- [16] A. H. Sayed and M. Rupp, "Error-energy bounds for adaptive gradient algorithms," *IEEE Trans. Signal Process.*, vol. 44, no. 8, pp. 1982–1989, Aug. 1996.
- [17] N. R. Yousef and A. H. Sayed, "A unified approach to the steady-state and tracking analyses of adaptive filters," *IEEE Trans. Signal Process.*, vol. 49, no. 2, pp. 314–324, Feb. 2001.
- [18] T. Y. Al-Naffouri and A. H. Sayed, "Transient analysis of adaptive filters with error nonlinearities," *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 653–663, Mar. 2003.
- [19] T. Y. Al-Naffouri and A. H. Sayed, "Transient analysis of data-normalized adaptive filters," *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 639–652, Mar. 2003.
- [20] H. C. Shin, W. J. Song, and A. H. Sayed, "Mean-square performance of data-reusing adaptive algorithms," *IEEE Signal Process. Lett.*, vol. 12, no. 12, pp. 851–854, Dec. 2005.
- [21] L. R. Vega, H. Rey, J. Benesty, and S. Tressens, "A fast robust recursive least-squares algorithm," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 1209–1215, Mar. 2009.



**Md. Zulfikar Ali Bhotto** (S'10) received the B.Sc. degree in electrical and electronic engineering from the Department of Electrical and Electronic Engineering, Rajshahi University of Engineering and Technology, Bangladesh, in 2002.

He worked as a lecturer and Assistant Professor with the Rajshahi University of Engineering and Technology from 2003 to 2007. Currently, he is pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, University of Victoria, Victoria, Canada. He was awarded a

University of Victoria fellowship for 2007–2008.



**Andreas Antoniou** (M'69–SM'79–F'82–LF'04) received the B.Sc.(Eng.) and Ph.D. degrees in electrical engineering from the University of London, U.K., in 1963 and 1966, respectively.

He taught at Concordia University, Quebec, Canada, from 1970 to 1983, was the founding Chair of the Department of Electrical and Computer Engineering, University of Victoria, B.C., Canada, from 1983 to 1990, and is now Professor Emeritus. His teaching and research interests are in the area of digital signal processing. He is the author of *Digital*

*Signal Processing: Signals, Systems, and Filters* (New York: McGraw-Hill, 2005) and the coauthor with W.-S. Lu of *Practical Optimization: Algorithms and Engineering Applications* (New York: Springer, 2007).

Dr. Antoniou is a Fellow of the IET. He served as Associate/Chief Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS (CAS) from 1983 to 1987, as a Distinguished Lecturer of the IEEE Signal Processing and the Circuits and Systems Societies during 2003–2004 and 2006–2007, respectively, and as General Chair of the 2004 International Symposium on Circuits and Systems. He was awarded the CAS Golden Jubilee Medal by the IEEE Circuits and Systems Society, the B.C. Science Council Chairman's Award for Career Achievement for 2000, the Doctor Honoris Causa degree by the National Technical University, Athens, Greece, in 2002, the IEEE Circuits and Systems Society Technical Achievement Award for 2005, the 2008 IEEE Canada Outstanding Engineering Educator Silver Medal, the IEEE Circuits and Systems Society Education Award for 2009, and the Craighdarroch Gold Medal for Career Achievement for 2011 from the University of Victoria.