

A Family of Variable Step-Size Affine Projection Adaptive Filtering Algorithms with Selective Regressors and Selective Partial Updates

M. Shams Esfand Abadi^{1,*}, V. Mehrdad¹ and A. Gholipour¹

Abstract. This paper presents a family of Variable Step-Size (VSS) Affine Projection (AP) adaptive filtering algorithms with Selective Partial Updates (SPU) and Selective Regressors (SR). The presented algorithms have good convergence speed, low steady state Mean Square Error (MSE), and low computational complexity features. The stability bounds of the family of SPU-APA, SR-APA and SPU-SR-APA are analyzed, based on the energy conservation arguments. This analysis does not need to assume a Gaussian or white distribution for the regressors. We demonstrate the good performance of the proposed algorithms through simulations in system identification and acoustic echo cancellation scenarios.

Keyword: Adaptive filter; Affine projection; Selective partial update; Selective regressor; Stability bound; Variable step-size.

INTRODUCTION

Adaptive filtering has been, and still is, an area of active research that plays an active role in an ever increasing number of applications, such as noise cancellation, channel estimation, channel equalization and acoustic echo cancellation [1,2]. The Least Mean Square (LMS) and its normalized version (NLMS) are the workhorses of adaptive filtering. In the presence of colored input signals, the LMS and NLMS algorithms have extremely slow convergence rates. To solve this problem, a number of adaptive filtering structures, based on affine subspace projections [3-5], data reusing adaptive algorithms [6-8], Block adaptive filters [2] and multirate techniques, have been proposed in the literature [9-11]. In all these algorithms, the selected fixed step-size can change the convergence and the steadystate mean square error. It is well known that the steady-state Mean Square Error (MSE) decreases when the step-size decreases, while the convergence speed increases when the step-size increases. By optimally selecting the step-size during the adaptation, we can

obtain both fast convergence rates and low steady-state mean square errors [12-15]. These selections are based on various criteria. In [12], squared instantaneous errors were used. To improve noise immunity under Gaussian noise, the squared autocorrelation of errors at adjacent times was used in [14], and in [15], the fourthorder cumulant of instantaneous error was adopted. Important examples of two new Variable Step-Size (VSS) versions of the NLMS and the APA algorithm (APA) can be found in [16]. In [16], the norm of the projected weighted error vector is used as a criterion to determine how close the adaptive filter is to optimum performance.

Another feature that should be noticed in VSS adaptive filter algorithms is computational complexity. Several adaptive filters with fixed step-size, such as the adaptive filter algorithms with selective partial updates, have been proposed to reduce the computational complexity. These algorithms update only a subset of the filter coefficients in each time iteration. The Max-NLMS [17], the MMax-NLMS [18,19], the variants of the selective partial update Normalized Least Mean Square algorithms (SPU-NLMS) [20-22], and SPU Affine Projection (SPU-AP) algorithm [21] are important examples of this family of adaptive filter algorithms. Recently an affine projection adaptive filtering algorithm with Selective Regressors (SR) was proposed

^{1.} Department of Electrical Engineering, Shahid Rajaee Teacher Training University, Tehran, P.O. Box 16785-163, Iran.

 $^{*. \} Corresponding \ author. \ E-mail: \ mshams@srttu.edu$

Received 12 August 2009; received in revised form 12 January 2010; accepted 30 January 2010

in [23]. This paper presents a novel affine projection algorithm which reduces computational complexity by optimally selecting a subset of input regressors at every iteration.

In this paper, we extend the approach in [16] to establish the family of VSS-APA. Also, to reduce the computational complexity, the family of VSS-SPU-APA and VSS-SR-APA is established. By combining the SPU and SR approaches, the family of VSS-SPU-SR-APA is established. The stability of these VSS adaptive algorithms is analyzed based on the energy conservation relations. This analysis does not need to assume a Gaussian or white distribution for the regressors.

What we propose in this paper can be summarized as follows:

- The establishment of the family VSS-AP algorithms.
- The establishment of the family of VSS-SPU-AP algorithms.
- The establishment of the family of VSS-SR-AP algorithms.
- The establishment of the family of VSS-SPU-SR-AP algorithms.
- Mean-square stability analysis of the family of AP, and SR-AP, SPU-AP, and SPU-SR-AP algorithms.
- Demonstrating the presented algorithms in system identification and acoustic echo cancellation scenarios.

We have organized our paper as follows. First, the NLMS and SPU-NLMS algorithms will be briefly reviewed. Then, the family of APA, SR-APA and SPU-APA is presented and the family of variable stepsize adaptive filters is established. Following that, the computational complexity of the VSS adaptive filters is discussed. Finally, before concluding the paper, we demonstrate the usefulness of these algorithms by presenting several experimental results.

Throughout the paper, the following notations are adopted:

.	Euclidean norm of a vector,
$\ .\ ^2$	squared Euclidean norm of a vector,
$\ \mathbf{t}\ _{\boldsymbol{\Sigma}}^2$	Σ -weighted Euclidean norm of a column
	vector \mathbf{t} , defined as $\mathbf{t}^T \mathbf{\Sigma} \mathbf{t}$,
$\operatorname{Tr}(.)$	trace of a matrix,
$(.)^{T}$	transpose of a vector or a matrix,
$\mathbf{A}\otimes\mathbf{B}$	Kronecher product of matrices A and B,
$\operatorname{vec}(\mathbf{T})$	creating an $M^2 \times 1$ column vector t
	through stacking the columns of the
	$M \times M$ matrix \mathbf{T} ,
vec(t)	creating an $M \times M$ matrix T form the
× /	$M^2 \times 1$ column vector t .

 λ_{\max} the largest eigenvalue of a matrix,

M. Shams Esfand Abadi, V. Mehrdad and A. Gholipour

 \Re^+ the set of positive real numbers,

 $E\{.\}$ expectation operator.

BACKGROUND ON NLMS, AND SPU-NLMS ALGORITHMS

Figure 1 shows a typical adaptive filter setup, where x(n), d(n) and e(n) are the input, the desired and output error signals, respectively. Here, $\mathbf{h}(n)$ is the $M \times 1$ column vector of filter coefficients at iteration n. The desired signal is assumed to conform to the following linear data model:

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

where $\mathbf{x}(n) = [x(n), x(n-1), \dots, x(n-M+1)]^T$ are the input signal regressors, v(n) is the measurement noise, assumed to be zero mean, white, Gaussian, and independent of $\mathbf{x}(n)$, and \mathbf{h}_t is the unknown filter vector.

It is well known that the NLMS algorithm can be derived from the solution of the following optimization problem:

$$\min_{\mathbf{h}(n+1)} \|\mathbf{h}(n+1) - \mathbf{h}(n)\|^2,$$
(2)

subject to:

$$d(n) = \mathbf{x}^T(n)\mathbf{h}(n+1). \tag{3}$$

Using the method of Lagrange multipliers to solve this optimization problem leads to the following recursion:

$$\mathbf{h}(n+1) = \mathbf{h}(n) + \frac{\mu}{\|\mathbf{x}(n)\|^2} \mathbf{x}(n) e(n), \tag{4}$$

where $e(n) = d(n) - \mathbf{x}^T(n)\mathbf{h}(n)$, and μ is the stepsize that determines the convergence speed and Excess MSE (EMSE).

Now, partition the input signal vector and the vector of filter coefficients into B blocks, each of length L (note that B = M/L and is an integer), which are defined as:

$$\mathbf{x}(n) = [\mathbf{x}_1^T(n), \mathbf{x}_2^T(n), \cdots, \mathbf{x}_B^T(n)]^T,$$
(5)

$$\mathbf{h}(n) = [\mathbf{h}_1^T(n), \mathbf{h}_2^T(n), \cdots, \mathbf{h}_B^T(n)]^T.$$
(6)



Figure 1. A typical adaptive filter setup.

The SPU-NLMS algorithm for a single block update at every iteration minimizes the following optimization problem:

$$\min_{\mathbf{h}_j(n+1)} \|\mathbf{h}_j(n+1) - \mathbf{h}_j(n)\|^2,$$
(7)

subject to Equation 3, where j denotes the index of the block that should be updated [21]. Again by using the method of Lagrange multipliers, the update equation for SPU-NLMS is given by:

$$\mathbf{h}_j(n+1) = \mathbf{h}_j(n) + \frac{\mu}{\|\mathbf{x}_j(n)\|^2} \mathbf{x}_j(n) e(n), \tag{8}$$

where $j = \arg \max \|\mathbf{x}_i(n)\|^2$ for $1 \le i \le B$.

FAMILY OF AP, SR-APA AND SPU-APA

In this section, the family of APA, SR-APA and SPU-APA is presented.

Family of Affine Projection Algorithms (APA)

Now, define the $M \times K$ matrix of the input signal as:

$$\mathbf{X}(n) = [\mathbf{x}(n), \mathbf{x}(n-D), \cdots, \mathbf{x}(n-(K-1)D)], \quad (9)$$

and the $K \times 1$ vector of desired signal as:

$$\mathbf{d}(n) = [d(n), d(n-D), \cdots, d(n-(K-1)D)]^{T}, \quad (10)$$

where K is a positive integer (usually, but not necessarily $K \leq M$), and D is the positive integer parameter $(D \geq 1)$ that can increase the separation and, consequently, reduce the correlation among the regressors in $\mathbf{X}(n)$.

The family of APA can be established by minimizing Equation 2 but subject to $\mathbf{d}(n) = \mathbf{X}^T(n)\mathbf{h}(n)$. Again, by using the method of Lagrange multipliers, the filter vector update equation for the family of APA is given by:

$$\mathbf{h}(n+1) = \mathbf{h}(n) + \mu \mathbf{X}(n) \mathbf{W}(n) \mathbf{e}(n), \tag{11}$$

where $\mathbf{e}(n)$ is the output error vector, which is defined as:

$$\mathbf{e}(n) = \mathbf{d}(n) - \mathbf{X}^{T}(n)\mathbf{h}(n), \qquad (12)$$

and the matrix $\mathbf{W}(n)$ is obtained from Table 1 (in Table 1, ϵ is the regularization parameter, and **I** is the identity matrix). The NLMS, ϵ -NLMS, standard version of the APA, the Binormalized Data-Reusing LMS (BNDR-LMS) [7], the Regularized APA (R-APA) [24] and the NLMS with orthogonal correction factors (NLMS-OCF) [25] are established from Equation 11. From this equation the Partial Rank Algorithm (PRA) [26] can also be established when the adaptation of the filter coefficients is performed only once every K iterations.

 Table 1. Family of affine projection adaptive filter algorithms.

Algorithm	K	D	$\mathbf{W}(n)$
NLMS	K = 1	D = 1	$\frac{1}{\ \mathbf{x}(n)\ ^2}$
$\epsilon ext{-NLMS}$	K = 1	D = 1	$\frac{1}{\epsilon + \ \mathbf{x}(n)\ ^2}$
APA	$K \leq M$	D = 1	$(\mathbf{X}^T(n)\mathbf{X}(n))^{-1}$
BNDR-LMS	K = 2	D = 1	$(\mathbf{X}^T(n)\mathbf{X}(n))^{-1}$
R-APA	$K \leq M$	D = 1	$(\epsilon \mathbf{I} + \mathbf{X}^T(n)\mathbf{X}(n))^{-1}$
NLMS-OCF	$K \leq M$	$D \ge 1$	$(\mathbf{X}^T(n)\mathbf{X}(n))^{-1}$

Family of Selective Regressor APA (SR-APA)

In [23], another novel affine projection algorithm with Selective Regressors (SR), called (SR-APA) was presented. In this section, we extend this approach to present the family of SR-APA. The SR-APA minimizes Relation 2 subject to:

$$\mathbf{d}_G(n) = \mathbf{X}_G^T(n)\mathbf{h}(n),\tag{13}$$

where $G = \{i_1, i_2, \cdots, i_P\}$ denote the P subset (subset with P member) of the set $\{0, 1, \cdots, K-1\}$,

$$\mathbf{X}_G(n) = [\mathbf{x}(n-i_1D), \mathbf{x}(n-i_2D), \cdots, \mathbf{x}(n-i_PD)], (14)$$

is the $M \times P$ matrix of the input signal and:

$$\mathbf{d}_{G}(n) = [d(n-i_{1}D), d(n-i_{2}D), \cdots, d(n-i_{P}D)]^{T},$$
(15)

is the $P \times 1$ vector of the desired signal. Using the method of Lagrange multipliers to solve this optimization problem leads to the following update equation:

$$\mathbf{h}(n+1) = \mathbf{h}(n) + \mu \mathbf{X}_G(n) (\mathbf{X}_G^T(n) \mathbf{X}_G(n))^{-1} \mathbf{e}_G(n), \quad (16)$$

where:

$$\mathbf{e}_G(n) = \mathbf{d}_G(n) - \mathbf{X}_G^T(n)\mathbf{h}(n).$$
(17)

The indices of G are obtained by the following procedure:

1. Compute the following values for $0 \le i \le K - 1$:

$$\frac{e^2(n-iD)}{\|\mathbf{x}(n-iD)\|^2},$$
(18)

where $\mathbf{e}(n) = [e(n), e(n-D), \cdots, e(n-(K-1)D)]^T$.

2. The indices of G correspond to P largest values of Equation 18.

Setting D = 1 leads to SR-APA presented in [23]. Furthermore, from Equation 16, the family of SR-APA, such as SR-BNDR-LMS, SR-NLMS-OCF adaptive algorithms will be established.

Equation 16 can also be represented as:

 $\mathbf{h}(n+1) = \mathbf{h}(n)$

+
$$\mu \mathbf{X}(n) \mathbf{B}(n) (\mathbf{B}^{T}(n) \mathbf{X}^{T}(n) \mathbf{X}(n) \mathbf{B}(n))^{-1} \mathbf{B}^{T}(n) \mathbf{e}(n),$$
(19)

where $\mathbf{B}(n) = [\mathbf{1}_{i_1}, \mathbf{1}_{i_2}, \cdots, \mathbf{1}_{i_P}]$ is the $K \times P$ matrix and $\mathbf{1}_{i_p} = [0, \cdots, 0, 1, 0, \cdots, 0]^T$ is the $K \times 1$ vector with the element 1 in the position i_p .

Family of Selective Partial Update APA (SPU-APA)

The SPU-APA solves the following optimization problem:

$$\min_{\mathbf{h}_{F}(n+1)} \|\mathbf{h}_{F}(n+1) - \mathbf{h}_{F}(n)\|^{2},$$
(20)

subject to $\mathbf{d}(n) = \mathbf{X}^T(n)\mathbf{h}(n)$ where $F = \{j_1, j_2, \cdots, j_S\}$ denote the indices of the S blocks out of B blocks that should be updated at every adaptation. Again, by using the Lagrange multipliers approach, the filter vector update equation is given by:

$$\mathbf{h}_F(n+1) = \mathbf{h}_F(n) + \mu \mathbf{X}_F(n) (\mathbf{X}_F^T(n) \mathbf{X}_F(n))^{-1} \mathbf{e}(n),$$
(21)

where:

$$\mathbf{X}_F(n) = [\mathbf{X}_{j_1}^T(n), \mathbf{X}_{j_2}^T(n), \cdots, \mathbf{X}_{j_S}^T(n)]^T,$$
(22)

is the $SL \times K$ matrix and:

$$\mathbf{X}_{i}(n) = [\mathbf{x}_{i}(n), \mathbf{x}_{i}(n-D), \cdots, \mathbf{x}_{i}(n-(K-1)D)],$$
(23)

is the $L \times K$ matrix. The indices of F are obtained by the following procedure:

1. Compute the following values for $1 \le i \le B$:

$$\operatorname{Tr}(\mathbf{X}_{i}^{T}(n)\mathbf{X}_{i}(n)) \tag{24}$$

2. The indices of F correspond to S largest values of Relation 24.

By setting D = 1, the SPU-APA in [21] can be derived from Equation 14. Furthermore, from Equation 21, the new SPU adaptive algorithms such as SPU-BNDR-LMS and SPU-NLMS-OCF will be established. Also, the SPU-PRA can be established when the adaptation of the filter coefficients is performed only once every Kiterations. Equation 21 can be represented in the form of a full update equation as:

$$\mathbf{h}(n+1) = \mathbf{h}(n)$$

+ $\mu \mathbf{A}(n) \mathbf{X}(n) (\mathbf{X}^T(n) \mathbf{A}(n) \mathbf{X}(n))^{-1} \mathbf{e}(n),$ (25)

where the $\mathbf{A}(n)$ matrix is the $M \times M$ diagonal matrix with **1** and **0** blocks, each of length L on the diagonal, and the positions of 1's on the diagonal determine which coefficients should be updated in each iteration. The positions of **1** blocks (S blocks and $S \leq B$) on the diagonal of the $\mathbf{A}(n)$ matrix for each iteration in the family of SPU-APA are determined by the indices of F.

FAMILY OF VSS-SR-APA, VSS-SPU-APA AND VSS-SPU-SR-APA

In this section, we present the family of VSS-SR-APA, VSS-SPU-APA and VSS-SPU-SR-APA.

Family of VSS-SR-APA

We now proceed by determining the optimum stepsize, $\mu^{o}(n)$, instead of using μ in the VSS version of Equation 16. The latter equation can be stated in terms of weight error vector, $\tilde{\mathbf{h}}(n) = \mathbf{h}_{t} - \mathbf{h}(n)$, as:

$$\tilde{\mathbf{h}}(n+1) = \tilde{\mathbf{h}}(n) - \mu \mathbf{X}_G(n) (\mathbf{X}_G^T(n) \mathbf{X}_G(n))^{-1} \mathbf{e}_G(n),$$
(26)

Taking the squared Euclidean norm and expectations from both sides of Equation 26,

$$E\{\|\tilde{\mathbf{h}}(n+1)\|^{2}\} = E\{\|\tilde{\mathbf{h}}(n)\|^{2}\}$$
$$+ \mu^{2} E\{\mathbf{e}_{G}^{T}(n)(\mathbf{X}_{G}^{T}(n)\mathbf{X}_{G}(n))^{-1}\mathbf{e}_{G}(n)\}$$
$$- 2\mu E\{\mathbf{e}_{G}^{T}(n)(\mathbf{X}_{G}^{T}(n)\mathbf{X}_{G}(n))^{-1}\mathbf{X}_{G}^{T}(n)\tilde{\mathbf{h}}(n)\},$$
(27)

Equation 27 can be represented in the form of Equation 44:

$$E\{\|\tilde{\mathbf{h}}(n+1)\|^2\} = E\{\|\tilde{\mathbf{h}}(n)\|^2\} - \Delta\mu,$$
(28)

where $\Delta \mu$ is:

$$\Delta \mu = -\mu^2 E\{\mathbf{e}_G^T(n)(\mathbf{X}_G^T(n)\mathbf{X}_G(n))^{-1}\mathbf{e}_G(n)\}$$
$$+2\mu E\{\mathbf{e}_G^T(n)(\mathbf{X}_G^T(n)\mathbf{X}_G(n))^{-1}\mathbf{X}_G^T(n)\tilde{\mathbf{h}}(n)\}.$$
(29)

If $\Delta \mu$ is maximized, then Mmean-Square Deviation (MSD) will undergo the largest decrease from iteration n to iteration n + 1. The optimum step-size will be found with derivation of $\Delta \mu$, with respect to μ , $d\Delta \mu/d\mu = 0$,

$$\mu^{\circ}(n) = \frac{E\{\mathbf{e}_{G}^{T}(n)(\mathbf{X}_{G}^{T}(n)\mathbf{X}_{G}(n))^{-1}\mathbf{X}_{G}^{T}(n)\tilde{\mathbf{h}}(n)\}}{E\{\mathbf{e}_{G}^{T}(n)(\mathbf{X}_{G}^{T}(n)\mathbf{X}_{G}(n))^{-1}\mathbf{e}_{G}(n)\}}.$$
(30)

As we mentioned, we assumed a linear model for the desired signal, d(n), which we can also express as:

$$\mathbf{d}(n) = \mathbf{X}^{T}(n)\mathbf{h}_{t} + \mathbf{v}(n), \qquad (31)$$

A Family of Variable Step-Size Affine Projection

where $\mathbf{v}(n) = [v(n), v(n-D), \cdots, v(n-(K-1)D^T)$ is measurement noise vector assumed to be zero mean, white, Gaussian, and independent of the input signal matrix $\mathbf{X}(n)$. Since $\mathbf{e}_G(n) = \mathbf{d}_G(n) - \mathbf{X}_G^T(n)\mathbf{h}(n)$ and using Equation 31, we obtain:

$$\mathbf{e}_G(n) = \mathbf{X}_G^T(n)\tilde{\mathbf{h}}(n) + \mathbf{v}_G(n).$$
(32)

where:

$$\mathbf{v}_G(n) = [v(n-i_1D), v(n-i_2D), \cdots, v(n-i_PD)]^T.$$

Using the previous assumptions for the noise sequence and neglecting the dependency of $\tilde{\mathbf{h}}(n)$ on the past noises, we establish the following two sub equations from the two parts of Equation 30:

• PART I:

$$E\{\mathbf{e}_{G}^{T}(n)(\mathbf{X}_{G}^{T}(n)\mathbf{X}_{G}(n))^{-1}\mathbf{X}_{G}^{T}(n)\tilde{\mathbf{h}}(n)\}$$

$$= E\{(\tilde{\mathbf{h}}^{T}(n)\mathbf{X}_{G}(n)$$

$$+ \mathbf{v}_{G}^{T}(n))((\mathbf{X}_{G}^{T}(n)\mathbf{X}_{G}(n))^{-1}\mathbf{X}_{G}^{T}(n)\tilde{\mathbf{h}}(n))\}$$

$$= E\{\tilde{\mathbf{h}}^{T}(n)\mathbf{X}_{G}(n)(\mathbf{X}_{G}^{T}(n)\mathbf{X}_{G}(n))^{-1}\mathbf{X}_{G}^{T}(n)\tilde{\mathbf{h}}(n)\},$$
(33)

• PART II:

$$E\{\mathbf{e}_{G}^{T}(n)(\mathbf{X}_{G}^{T}(n)\mathbf{X}_{G}(n))^{-1}\mathbf{e}_{G}(n)\}$$

$$=E\{\tilde{\mathbf{h}}^{T}(n)\mathbf{X}_{G}(n)(\mathbf{X}_{G}^{T}(n)\mathbf{X}_{G}(n))^{-1}\mathbf{X}_{G}^{T}(n)\tilde{\mathbf{h}}(n)\}$$

$$+E\{\mathbf{v}_{G}^{T}(n)(\mathbf{X}_{G}^{T}(n)\mathbf{X}_{G}(n))^{-1}\mathbf{v}_{G}(n)\}$$

$$=E\{\tilde{\mathbf{h}}^{T}(n)\mathbf{X}_{G}(n)(\mathbf{X}_{G}^{T}(n)\mathbf{X}_{G}(n))^{-1}\mathbf{X}_{G}^{T}(n)\tilde{\mathbf{h}}(n)\}$$

$$+\sigma_{v}^{2}\mathrm{Tr}(E\{(\mathbf{X}_{G}^{T}(n)\mathbf{X}_{G}(n))^{-1}\}).$$
(34)

Finally, the optimum size in Equation 30 becomes:

 $\mu^{\,o}\,(\,n\,) \,\,=\,\,$

$$\frac{E\{\tilde{\mathbf{h}}^{T}(n)\mathbf{X}_{G}(n)(\mathbf{X}_{G}^{T}(n)\mathbf{X}_{G}(n))^{-1}\mathbf{X}_{G}^{T}(n)\tilde{\mathbf{h}}(n)\}}{E\{\tilde{\mathbf{h}}^{T}(n)\mathbf{X}_{G}(n)(\mathbf{X}_{G}^{T}(n)\mathbf{X}_{G}(n))^{-1}\mathbf{X}_{G}^{T}(n)\tilde{\mathbf{h}}(n)\} + \sigma_{v}^{2}\mathrm{Tr}(E\{(\mathbf{X}_{G}^{T}(n)\mathbf{X}_{G}(n))^{-1}\})}}$$

$$(35)$$

Substituting $\mu^{\circ}(n)$ instead of μ in Equation 16, the generic variable step-size update equation for SR-APA will be established. From Equation 35, the optimum step-size in the family of SR-APA can be stated as:

$$\mu^{o}(n) = \frac{E\{\|\mathbf{q}_{G}(n)\|^{2}\}}{E\{\|\mathbf{q}_{G}(n)\|^{2}\} + \sigma_{v}^{2} \operatorname{Tr}(E\{(\mathbf{X}_{G}^{T}(n)\mathbf{X}_{G}(n))^{-1}\})}, (36)$$

where:

$$\mathbf{q}_G(n) = \mathbf{X}_G(n) (\mathbf{X}_G^T(n) \mathbf{X}_G(n))^{-1} \mathbf{X}_G^T(n) \tilde{\mathbf{h}}(n).$$
(37)

Since, from Equation 32, $\mathbf{X}_{G}^{T}(n)\tilde{\mathbf{h}}(n) = \mathbf{e}_{G}(n) - \mathbf{v}_{G}(n)$, the expectation of $\mathbf{q}_{G}(n)$ can be stated as:

$$E\{\mathbf{q}_G(n)\} = E\{\mathbf{X}_G(n)(\mathbf{X}_G^T(n)\mathbf{X}_G(n))^{-1}\mathbf{e}_G(n)\}.$$
 (38)

Now, we can estimate $\mathbf{q}_G(n)$ from the following recursion:

$$\hat{\mathbf{q}}_G(n) = \beta \hat{\mathbf{q}}_G(n-1) + (1-\beta) \mathbf{X}_G(n) (\mathbf{X}_G^T(n) \mathbf{X}_G(n))^{-1} \mathbf{e}_G(n), \quad (39)$$

where β ($0 \leq \beta < 1$) is the smoothing factor. Using $\|\hat{\mathbf{q}}_G(n)\|^2$ instead of $E\{\|\mathbf{q}_G(n)\|^2\}$ in Equation 36, the update equation for the family of VSS-SR-APA is given by:

$$\mathbf{h}(n+1) = \mathbf{h}(n)$$
$$+ \mu(n)\mathbf{X}_G(n)(\mathbf{X}_G^T(n)\mathbf{X}_G(n))^{-1}\mathbf{e}_G(n), \quad (40)$$

where:

$$\mu(n) = \mu_{\max} \cdot \frac{\|\hat{\mathbf{q}}_G(n)\|^2}{\|\hat{\mathbf{q}}_G(n)\|^2 + \Psi}.$$
(41)

Also, $\Psi = \sigma_v^2 \operatorname{Tr}(E\{(\mathbf{X}_G^T(n)\mathbf{X}_G(n))^{-1}\})$ which can be approximated as P/SNR (the details of this approximation is given in Appendix A). The step-size changes with the $\|\hat{\mathbf{q}}_G(n)\|^2$, and the constant Ψ . Also, μ_{\max} should be selected in the stability bound to guarantee the stability (in Appendix B, the stability of the family of APA has been discussed in details). From Equation 40, the VSS-SR-PRA will also be established when the adaptation of the filter coefficients is performed only once every K iterations.

Family of VSS-SPU-APA

The same as in the previous subsection, we again proceed by determining the optimum step-size, $\mu^o(n)$, instead of using μ in the VSS version of Equation 21. This equation can be stated in terms of weight error vector, $\mathbf{\tilde{h}}_F(n) = \mathbf{h}_{t,F} - \mathbf{h}_F(n)$, where $\mathbf{h}_{t,F}$ is the partial unknown true filter vector, as:

$$\tilde{\mathbf{h}}_{F}(n+1) = \tilde{\mathbf{h}}_{F}(n) - \mu \mathbf{X}_{F}(n) (\mathbf{X}_{F}^{T}(n) \mathbf{X}_{F}(n))^{-1} \mathbf{e}(n),$$
(42)

Taking the squared Euclidean norm and expectations from both sides of Equation 42:

$$E\{\|\tilde{\mathbf{h}}_{F}(n+1)\|^{2}\} = E\{\|\tilde{\mathbf{h}}_{F}(n)\|^{2}\}$$
$$+ \mu^{2}E\{\mathbf{e}^{T}(n)(\mathbf{X}_{F}^{T}(n)\mathbf{X}_{F}(n))^{-1}\mathbf{e}(n)\}$$
$$- 2\mu E\{\mathbf{e}^{T}(n)(\mathbf{X}_{F}^{T}(n)\mathbf{X}_{F}(n))^{-1}\mathbf{X}_{F}^{T}(n)\tilde{\mathbf{h}}_{F}(n)\},$$
(43)

Equation 43 can be represented in the following form:

$$E\{\|\tilde{\mathbf{h}}_{F}(n+1)\|^{2}\} = E\{\|\tilde{\mathbf{h}}_{F}(n)\|^{2}\} - \Delta\mu,$$
(44)

where $\Delta \mu$ is given by:

$$\Delta \mu = -\mu^2 E\{\mathbf{e}^T(n) (\mathbf{X}_F^T(n) \mathbf{X}_F(n))^{-1} \mathbf{e}(n)\}$$

+ $2\mu E\{\mathbf{e}^T(n) (\mathbf{X}_F^T(n) \mathbf{X}_F(n))^{-1} \mathbf{X}_F^T(n) \tilde{\mathbf{h}}_F(n)\},$
(45)

If $\Delta \mu$ is maximized, then mean-square deviation (MSD) will undergo the largest decrease from iteration n to iteration n + 1. The optimum step-size will be found with a derivation of $\Delta \mu$, with respect to μ , $d\Delta\mu/d\mu = 0,$

$$\mu^{o}(n) = \frac{E\{\mathbf{e}^{T}(n)(\mathbf{X}_{F}^{T}(n)\mathbf{X}_{F}(n))^{-1}\mathbf{X}_{F}^{T}(n)\tilde{\mathbf{h}}_{F}(n)\}}{E\{\mathbf{e}^{T}(n)(\mathbf{X}_{F}^{T}(n)\mathbf{X}_{F}(n))^{-1}\mathbf{e}(n)\}}$$
(46)

Now, by using the following approximation, $\mathbf{X}^T(n)\mathbf{h}(n)$ $\approx \mathbf{X}_{F}^{T}(n)\mathbf{h}_{F}(n)$, and neglecting the dependency of $\tilde{\mathbf{h}}_F(n)$ on past noises, the optimum size in Equation 46 becomes:

 $\mu^{o}(n) =$

$$\frac{E\{\tilde{\mathbf{h}}_{F}^{T}(n)\mathbf{X}_{F}(n)(\mathbf{X}_{F}^{T}(n)\mathbf{X}_{F}(n))^{-1}\mathbf{X}_{F}^{T}(n)\tilde{\mathbf{h}}_{F}(n)\}}{E\{\tilde{\mathbf{h}}_{F}^{T}(n)\mathbf{X}_{F}(n)(\mathbf{X}_{F}^{T}(n)\mathbf{X}_{F}(n))^{-1}\mathbf{X}_{F}^{T}(n)\tilde{\mathbf{h}}_{F}(n)\} + \sigma_{v}^{2}\operatorname{Tr}(E\{(\mathbf{X}_{F}^{T}(n)\mathbf{X}_{F}(n))^{-1}\})} \cdot (47)$$

Now, by defining:

$$\mathbf{q}_F(n) = \mathbf{X}_F(n) (\mathbf{X}_F^T(n) \mathbf{X}_F(n))^{-1} \mathbf{X}_F^T(n) \tilde{\mathbf{h}}_F(n), \quad (48)$$

the optimum variable step-size can be approximated as:

$$\mu(n) = \mu_{\max} \cdot \frac{\|\hat{\mathbf{q}}_F(n)\|^2}{\|\hat{\mathbf{q}}_F(n)\|^2 + \mathbf{\Upsilon}},\tag{49}$$

where $\Upsilon = \sigma_v^2 \operatorname{Tr}(E\{(\mathbf{X}_F^T(n)\mathbf{X}_F(n))^{-1}\})$, and we can estimate $\mathbf{q}_F(n)$ with the following recursion:

$$\hat{\mathbf{q}}_F(n) = \beta \hat{\mathbf{q}}_F(n-1) + (1-\beta) \mathbf{X}_F(n) (\mathbf{X}_F^T(n) \mathbf{X}_F(n))^{-1} \mathbf{e}(n).$$
(50)

Therefore, the updated equation for the family of VSS-SPU-APA is established as:

 $\mathbf{h}_F(n+1) = \mathbf{h}_F(n)$

+
$$\mu(n)\mathbf{X}_F(n)(\mathbf{X}_F^T(n)\mathbf{X}_F(n))^{-1}\mathbf{e}(n),$$
 (51)

where $\mu(n)$ is obtained from Equation 49. In this equation, Υ can be approximated as K/SNR (the details of this approximation is given in Appendix A). The step-size changes with $\|\hat{\mathbf{q}}_F(n)\|^2$, and $\boldsymbol{\Upsilon}$. Again, $\mu_{\rm max}$ should be selected in the stability bound to guarantee the stability (in Appendix B, the stability of the family of SPU-APA has been discussed in details). From Equation 51, the VSS-SPU-PRA will also be established when the adaptation of the filter coefficients is performed only once every K iterations. Table 2 summarizes the presented algorithms.

Family of VSS-SPU-SR-APA

We can combine the VSS-SPU and VSS-SR approaches in affine projection adaptive filter algorithms to establish the family of VSS-SPU-SR-APA. Defining $SL \times P$ input signal matrix through:

$$\mathbf{X}_{F,G}(n) = \begin{pmatrix} \mathbf{x}_{j_1}(n-i_1D) & \mathbf{x}_{j_1}(n-i_2D) & \cdots & \mathbf{x}_{j_1}(n-i_PD) \\ \mathbf{x}_{j_2}(n-i_1D) & \mathbf{x}_{j_2}(n-i_2D) & \cdots & \mathbf{x}_{j_2}(n-i_PD) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_{j_S}(n-i_1D) & \mathbf{x}_{j_S}(n-i_2D) & \cdots & \mathbf{x}_{j_S}(n-i_PD) \end{pmatrix},$$
(52)

	$\mathbf{X}_G(n) = [\mathbf{x}(n-i_1D), \mathbf{x}(n-i_2D), \cdots, \mathbf{x}(n-i_PD)]$
	$\mathbf{d}_G(n) = [d(n-i_1D), d(n-i_2D), \cdots, d(n-i_PD)]^T$
VSS-SR-APA	$\mathbf{e}_G(n) = \mathbf{d}_G(n) - \mathbf{X}_G^T(n)\mathbf{h}(n)$
	$\hat{\mathbf{q}}_G(n) = \beta \hat{\mathbf{q}}_G(n-1) + (1-\beta) \mathbf{X}_G(n) (\mathbf{X}_G^T(n) \mathbf{X}_G(n))^{-1} \mathbf{e}_G(n)$
	$\mu(n) = \mu_{\max} \cdot \frac{\ \hat{\mathbf{q}}_G(n)\ ^2}{\ \hat{\mathbf{q}}_G(n)\ ^2 + \Psi}, \ \Psi \approx P/\text{SNR}$
	$\mathbf{h}(n+1) = \mathbf{h}(n) + \mu(n) \mathbf{X}_G(n) (\mathbf{X}_G^T(n) \mathbf{X}_G(n))^{-1} \mathbf{e}_G(n),$
	$\mathbf{X}_F(n) = [\mathbf{X}_{j_1}^T(n), \mathbf{X}_{j_2}^T(n), \cdots, \mathbf{X}_{j_S}^T(n)]^T$
	$\mathbf{d}(n) = [d(n), d(n-D), \cdots, d(n-(K-1)D)]^{T}$
VSS-SPU-APA	$\mathbf{e}(n) = \mathbf{d}(n) - \mathbf{X}^T(n)\mathbf{h}(n)$
	$\hat{\mathbf{q}}_F(n) = \beta \hat{\mathbf{q}}_F(n-1) + (1-\beta) \mathbf{X}_F(n) (\mathbf{X}_F^T(n) \mathbf{X}_F(n))^{-1} \mathbf{e}(n)$
	$ \mu(n) = \mu_{\max} \cdot \frac{\ \hat{\mathbf{q}}_F(n)\ ^2}{\ \hat{\mathbf{q}}_F(n)\ ^2 + \Upsilon}, \Upsilon \approx K/\text{SNR} $
	$\mathbf{h}_F(n+1) = \mathbf{h}_F(n) + \mu(n)\mathbf{X}_F(n)(\mathbf{X}_F^T(n)\mathbf{X}_F(n))^{-1}\mathbf{e}(n)$

Table 2. VSS adaptive filter algorithms.

N7

the update equation for the family of VSS-SPU-SR-APA is given by:

$$\mathbf{h}_{F}(n+1) = \mathbf{h}_{F}(n)$$
$$+ \mu(n) \mathbf{X}_{F,G}(n) (\mathbf{X}_{F,G}^{T}(n) \mathbf{X}_{F,G}(n))^{-1} \mathbf{e}_{G}(n), \quad (53)$$

where:

$$\mu(n) = \mu_{\max} \cdot \frac{\|\hat{\mathbf{q}}_{F,G}(n)\|^2}{\|\hat{\mathbf{q}}_{F,G}(n)\|^2 + \boldsymbol{\Theta}}.$$
(54)

The positive constant, $\Theta = \sigma_v^2$ Tr $(E\{(\mathbf{X}_{F,G}^T(n) | \mathbf{X}_{F,G}(n))^{-1}\})$, can be approximated as P/SNR, and we can estimate $\mathbf{q}_{F,G}(n)$ with the following recursion:

$$\hat{\mathbf{q}}_{F,G}(n) = \beta \hat{\mathbf{q}}_{F,G}(n-1) + (1-\beta) \mathbf{X}_{F,G}(n) (\mathbf{X}_{F,G}^T(n) \mathbf{X}_{F,G}(n))^{-1} \mathbf{e}_G(n).$$
(55)

COMPUTATIONAL COMPLEXITY

The computational complexity of the VSS adaptive algorithms has been given in Table 3. The computational complexity of the APA is from [5]. The SPU-APA needs $(K^2 + 2K)SL + K^3 + K^2$ multiplications and 1 division. This algorithm needs $B \log_2 S + O(B)$ comparisons when using the heapsort algorithm [27]. The SR-APA needs $(P^2 + 2P)M + P^3 + \tilde{P}^2$ multiplications and K divisions. This algorithm needs (K - P)M +K + 1 additional multiplications and $K \log_2 P + O(K)$ comparisons. Comparing the updated equation for APA and VSS-APA shows that the VSS-APA needs M additional multiplications due to variable stepsize. In VSS-SPU-APA, the additional multiplication is SL. Also, this algorithm needs $B \log_2 S + O(B)$ comparisons. It is obvious that the computational complexity of VSS-SPU-APA is lower than VSS-APA. The number of reductions in multiplication for VSS-SPU-APA is $(M - SL)(K^2 + 2K + 1)$, which is large in some applications such as networks and acoustic echo cancellations. Also, the computational complexity of VSS-PRA and VSS-SPU-PRA is reduced by the factor of K, because the adaptation of the filter coefficients is performed only once every K iterations. In VSS-SR-APA, the additional multiplication is M compared with SR-APA. Also this algorithm needs $K \log_2 P + O(K)$ comparisons. It is obvious that the computational complexity of VSS-SR-APA is lower than VSS-APA. The computational complexity of VSS-PRA and VSS-SR-PRA is reduced by the factor of K, because the adaptation of the filter coefficients is performed only once every K iterations. The VSS-SPU-SR-APA needs $(P^2 + 2P)SL + P^3 + P^2$ multiplications, which is lower than VSS-SR-APA and VSS-SPU-APA. This algorithm needs also K + 1 divisions, SL + (K - P)M + K +1 additional multiplications and $K \log_2 P + O(K) + O(K)$ $B \log_2 S + O(B)$ comparisons.

SIMULATION RESULTS

We justified the performance of the proposed algorithm by carrying out computer simulations in system identification and line echo cancellation scenarios.

System Identification

In this experiment, the unknown system has 32 taps and is selected at random. The input signal, x(n), is a first order autoregressive (AR(1)) signal generated according to:

$$x(n) = \rho x(n-1) + w(n),$$
(56)

where w(n) is a zero mean white Gaussian signal. The value of ρ is set to 0.9, generating a highly colored Gaussian signal. The measurement noise, v(n), with $\sigma_v^2 = 10^{-3}$ was added to the noise free desired signal generated through $d(n) = \mathbf{h}_t^T \mathbf{x}(n)$. The adaptive filter and the unknown channel are assumed to have the

Algorithm	Multiplications	Divisions	Additional Multiplications	Comparisons
АРА	$(K^2 + 2K)M + K^3 + K^2$	_	_	_
SPU-APA	$(K^2 + 2K)SL + K^3 + K^2$	_	1	$B\log_2 S + O(B)$
SR-APA	$(P^2 + 2P)M + P^3 + P^2$	K	(K-P)M+K+1	$K \log_2 P + O(K)$
SPU-SR-APA	$(P^2 + 2P)SL + P^3 + P^2$	K	(K-P)M+K+1	$K \log_2 P + O(K) + B \log_2 S + O(B)$
VSS-APA	$(K^2 + 2K)M + K^3 + K^2$	1	M	_
VSS-SPU-APA	$(K^2 + 2K)SL + K^3 + K^2$	1	1 + SL	$B\log_2 S + O(B)$
VSS-SR-APA	$(P^2 + 2P)M + P^3 + P^2$	K+1	M + (K - P)M + K + 1	$K \log_2 P + O(K)$
VSS-SPU-SR-APA	$(P^2 + 2P)SL + P^3 + P^2$	K+1	SL + (K - P)M + K + 1	$K\log_2 P + O(K) + B\log_2 S + O(B)$

Table 3. The computational complexity of the APA, SPU-APA, SR-APA, VSS-APA, VSS-SPU-APA, and VSS-SPU-SR-APA.

Algorithm	$\frac{2}{\lambda_{\max}(E\{\mathrm{D}^T(n)\mathrm{X}^T(n)\})}$	$\frac{1}{\lambda_{\max}(M^{-1}N)}$	$\frac{1}{\max(\lambda(H)\in\Re^+)}$	$\mu_{ m max}$
SR-APA $(K = 4, P = 1)$	3.7009	2.0000	2.8049	2.0000
SR-APA $(K = 4, P = 2)$	3.7030	2.0001	3.0259	2.0001
SR-APA $(K = 4, P = 3)$	3.9993	2.0002	3.2445	2.0002
SR-APA $(K = 4, P = 4)$	3.3778	2.0002	3.6000	2.0002
SPU-APA $(K = 4, B = 4, S = 1)$	4.3020	0.0315	0.5623	0.0315
SPU-APA $(K = 4, B = 4, S = 2)$	3.6542	0.6431	2.7474	0.6431
SPU-APA $(K = 4, B = 4, S = 3)$	3.4910	1.4723	3.2691	1.4723
SPU-APA $(K = 4, B = 4, S = 4)$	3.2658	2.0002	3.4391	2.0002
SPU-SR-APA $(K = 4, B = 4, P = 2, S = 2)$	3.8036	0.8999	2.7722	0.8999
SPU-SR-APA $(K = 4, B = 4, P = 2, S = 3)$	4.0876	1.5911	3.1756	1.5911
SPU-SR-APA $(K = 4, B = 4, P = 2, S = 4)$	4.1587	2.0013	3.1900	2.0013

Table 4. Stability bounds of the SR-AP, SPU-AP and SPU-SR-AP algorithms with different parameters for coloredGaussian input.

same number of taps. The parameters, K, and the number of blocks (B), are set to 4, and different values for S and P are used in simulations. In all simulations, the learning curves are obtained by ensemble averaging over 200 independent trials. Also, $\mu_{\rm max}$ is selected in stability bound to guarantee the stability. Table 4 shows the stability bounds of SR-APA, SPU-APA and SPU-SR-APA for different values of S. These values are obtained from Equations B22 and B24 (Appendix B). Figure 2 shows the simulated steady-state MSE curves of the SPU-AP algorithm as a function of the stepsize for colored Gaussian input. The step-size changes from 0.04 to $\mu_{\rm max}$ for each parameter adjustment. As we can see, the theoretical values for $\mu_{\rm max}$ show the good estimation of the stability bound of SR-AP algorithms. Figure 3 shows the results for the SPU-



Figure 2. Simulated steady-state MSE of SR-APA with K = 4 and P = 1, 2, 3, 4 as a function of the step-size for colored Gaussian input signal.



Figure 3. Simulated steady-state MSE of SPU-APA with K = 4, B = 4 and S = 2, 3, 4 as a function of the step-size for colored Gaussian input signal.

AP algorithm. The parameter, B, was set to 4 and different values for S (2, 3, and 4) were selected. The step-size changes from 0.04 to $\mu_{\rm max}$ for each parameter adjustment. Again, the theoretical values for $\mu_{\rm max}$ show the good estimation of the stability bound of SPU-AP algorithms. In the simulations, $\mu_{\rm max}$ is set to 1 for VSS-SR-APA. In VSS-SPU-APA, for S = 2 and S = 3, $\mu_{\rm max}$ is set to 0.3 and 1, respectively. Also, the constant values of Ψ , Υ and Θ were set to 0.001.

Figure 4 shows the results for VSS-SR-APA. The parameter, P, was set to 2, and different values for μ (0.03, 0.1, 1) were used in SR-APA. This result shows that the VSS-SR-APA has fast convergence speed and low steady-state MSE. Figure 5 compares the learning curves of VSS-APA and VSS-SR-APA for different values of P. As we can see for P =

A Family of Variable Step-Size Affine Projection



Figure 4. Learning curves of SR-APA and VSS-SR-APA for K = 4 and P = 2.



Figure 5. Comparing the learning curves of VSS-APA and VSS-SR-APA with P = 2 and VSS-SR-APA with P = 3.

3, the VSS-SR-APA will be close to the VSS-APA. Furthermore, the computational complexity of VSS-SR-APA is lower than VSS-APA. In Figure 6, we presented the learning curves for PRA and VSS-PRA. Again, VSS-PRA has better performance compared with PRA. Figure 7 shows the results of VSS-SR-PRA for P = 2. The results present a better performance for VSS-SSR-PRA compared with SR-PRA. Figure 8 compares the performance of the VSS-PRA and VSS-SR-PRA for P = 2 and P = 3. For P = 3, the results will be close to VSS-PRA. Also, the computational complexity of VSS-SPU-PRA will be lower than VSS-PRA. In Figure 9, we presented the learning curves of VSS-APA, VSS-PRA, VSS-SR-APA and VSS-SR-PRA. As we can see, the curve (e) in this figure, which is related to VSS-SR-ARA, has close performance to VSS-APA. Also, the curve (b) in this figure, which



Figure 6. Learning curves of PRA and VSS-PRA for K = 4.



Figure 7. Learning curves of SR-PRA and VSS-SR-PRA for K = 4 and P = 2.



Figure 8. Comparing the learning curves of VSS-PRA, VSS-SR-PRA with P = 2 and VSS-SR-PRA with P = 3.

is related to VSS-SR-PRA, with P = 3, has close performance to VSS-SR-PRA. In this algorithm, the input vectors are optimally selected and the adaptation is performed only once every K iterations. Therefore, the reduction of the computational complexity is large in this algorithm. Figure 10 shows the performance of VSS-SR-APA for different values of Ψ . As we can see, this algorithm is not very sensitive to this parameter.

Figure 11 shows the results for VSS-SPU-APA. The parameter, S, was set to 2, and different values for μ (0.03, 0.1, 0.5) were used in SPU-APA. This result shows that the VSS-SPU-APA has fast convergence speed and low steady-state MSE. This fact can be seen in Figure 12 for S = 3. Figure 13 compares the learning curves of VSS-APA and VSS-SPU-APA. As we can see for S = 3, the VSS-SPU-APA will be close to the VSS-



Figure 9. Comparing the learning curves of VSS-APA, VSS-PRA, VSS-SR-APA with P = 2, 3 and VSS-SR-PRA with P = 2, 3.



Figure 10. Learning curves of VSS-SR-APA for different values of Ψ .



Figure 11. Learning curves of SPU-APA and VSS-SPU-APA for K = 4, B = 4 and S = 2.



Figure 12. Learning curves of SPU-APA and VSS-SPU-APA for K = 4, B = 4 and S = 3.



Figure 13. Comparing the learning curves of VSS-APA, VSS-SPU-APA with S = 2 and VSS-SPU-APA with S = 3.

APA. Furthermore, the computational complexity of VSS-SPU-APA is lower than VSS-APA. Figures 14 and 15 show the results of VSS-SPU-PRA for S = 2 and S = 3. The results present better performance for VSS-SPU-PRA compared with SPU-PRA. Figure 16 compares the VSS-PRA and VSS-SPU-PRA for S = 2and S = 3. For S = 3, the results will be close to VSS-PRA. Also, the computational complexity of VSS-SPU-PRA will be lower than VSS-PRA. In Figure 17, we presented the learning curves of VSS-APA, VSS-PRA, VSS-SPU-APA and VSS-SPU-PRA. As we can see, the curve (b) in this figure, which is related to VSS-SPU-PRA, has close performance to VSS-APA. In VSS-SPU-PRA, the filter coefficients are partially updated and this adaptation is performed only once every K iterations. Therefore, the reduction of the computational complexity is large in this algorithm.

Figure 18 shows the learning curves of SPU-SR-



Figure 14. Learning curves of SPU-PRA and VSS-SPU-PRA for K = 4, B = 4 and S = 2.



Figure 15. Learning curves of SPU-PRA and VSS-SPU-PRA for K = 4, B = 4 and S = 3.



Figure 16. Comparing the learning curves of VSS-PRA and VSS-SPU-PRA with S = 2, and VSS-SPU-PRA with S = 3.



Figure 17. Comparing the learning curves of VSS-APA, VSS-PRA and VSS-SPU-APA with S = 2, 3, and VSS-SPU-PRA with S = 2, 3.



Figure 18. Simulated steady-state MSE of SPU-APA with K = 4, B = 4 and S = 2, 3, 4 as a function of the step-size for colored Gaussian input signal.

APA and VSS-SPU-SR-APA with S = 2 and P = 2. This figure shows that VSS-SPU-SR-APA has better performance. This fact can be seen in Figure 19 for P = 3 and S = 3. Figure 20 compares the learning curves of VSS-SPU-SR-APA and VSS-APA. This figure shows that VSS-SPU-SR-APA with S = 3 and P = 3has close performance to VSS-APA.

We have also studied the performance of the presented algorithms for real impulse response systems. Figure 21 shows the impulse response of the car echo path with 256 taps (the impulse response of the car echo path is from [21]). The parameters K and B were set to 4, and the input signal is the same as previous simulations. Figure 22 compares the performance of VSS-APA, VSS-SR-APA, VSS-PRA and VSS-SR-PRA. This figure shows that for P = 3, the convergence



Figure 19. Simulated steady-state MSE of SPU-APA with K = 4, B = 4 and S = 2, 3, 4 as a function of the step-size for colored Gaussian input signal.



Figure 20. Comparing the learning curves of VSS-APA and VSS-SPU-SR-APA with P = 2, S = 2, and VSS-SPU-SR-APA with P = 3, S = 3.



Figure 21. Impulse response of the car echo path.



Figure 22. Comparing the learning curves of VSS-APA, VSS-PRA and VSS-SR-APA with P = 3, and VSS-SR-PRA with P = 3 when the impulse response of the car echo path should be identified.

speed of VSS-SR-APA with P = 3 will be close to the VSS-APA. Figure 23 compares the performance of VSS-APA, VSS-SPU-APA, VSS-PRA and VSS-SPU-PRA. This figure shows that for S = 3, the convergence speed of VSS-SPU-APA will be close to VSS-APA.

Line Echo Cancellation

In communications over phone lines, a signal traveling from a far-end point to a near-end point is usually reflected in the form of an echo at the near-end due to mismatches in circuity. The purpose of a Line Echo Canceller (LEC) is to eliminate the echo from a received signal. In this experiment, the input signal is a speech signal. Also, Figure 24 shows the impulse response sequence of a typical echo path (the impulse response of the line echo path and the input speech signal is from [5], page 347). In this simulation, the length

A Family of Variable Step-Size Affine Projection



Figure 23. Comparing the learning curves of VSS-APA, VSS-PRA and VSS-SPU-APA with S = 3 and VSS-SPU-PRA with S = 3 when the impulse response of the car echo path should be identified.



Figure 24. Impulse response of the line echo path.

of the adaptive filter is 128. Figure 25a shows the farend signal samples. This signal is a synthetic signal that emulates the properties of speech [5]. Figure 25b shows the Echo signal. Figures 26a, b and c show the error signals that are obtained by VSS-SR-APA with P = 2, and P = 3, and VSS-APA. As we can see, by increasing the parameter, P, the error has smaller amplitude. Figures 27a, b and c show the error signals that are obtained by VSS-SPU-APA with S = 2, S = 3and VSS-APA. Again, by increasing the parameter S, the error has smaller amplitude. Figures 28a, b and c show the results for VSS-SPU-S-APA with P = 2, S = 2, and VSS-SPU-SR-APA with P = 3, S = 3 and VSS-APA.

CONCLUSIONS

In this paper, we presented the family of VSS-APA, VSS-SR-APA, VSS-SPU-APA and VSS-SPU-SR-APA.



Figure 25. (a) Far-end signal; (b) Echo signal.



Figure 26. (a) Error obtained by VSS-SPU-APA with P = 2; (b) Error obtained by VSS-SPU-APA with P = 3; (c) Error obtained by VSS-APA.

These algorithms exhibit fast convergence while reducing the steady-state mean square error as compared to the ordinary APA, SR-APA and SPU-APA algorithms. The presented algorithms were also computationally efficient. The stability bounds of these algorithms were analyzed based on energy conservation arguments. We demonstrated the performance of the presented



Figure 27. (a) Error obtained by VSS-SPU-SR-APA with S = 2; (b) Error obtained by VSS-SPU-APA with S = 3; (c) Error obtained by VSS-APA.



Figure 28. (a) Error obtained by VSS-SPU-SR-APA with P = 2, S = 2; (b) Error obtained by VSS-SPU-SR-APA with P = 3, S = 3; (c) Error obtained by VSS-APA.

M. Shams Esfand Abadi, V. Mehrdad and A. Gholipour

VSS adaptive algorithms in system identification and acoustic echo cancellation scenarios.

REFERENCES

- 1. Widrow, B. and Stearns, S.D., *Adaptive Signal Processing*, Englewood Cliffs, NJ, Prentice-Hall (1985).
- Haykin, S., Adaptive Filter Theory, NJ, Prentice-Hall, 4th Ed. (2002).
- Ozeki, K. and Umeda, T. "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties", in *Electron. Commun.* Jpn., 67-A, pp. 19-27 (1984).
- Shin, H.-C. and Sayed, A.H. "Mean-square performance of a family of affine projection algorithms", *IEEE Trans. Signal Processing*, 52(1), pp. 90-102 (Jan. 2004).
- Sayed, A.H., Fundamentals of Adaptive Filtering, Wiley (2003).
- Roy, S. and Shynk, J.J. "Analysis of data-reusing LMS algorithm", in Proc. Midwest Symp. Circuits Syst., pp. 1127-1130 (1989).
- Apolinario, J., Campos, M.L. and Diniz, P.S.R. "Convergence analysis of the binormalized data-reusing LMS algorithm", *IEEE Trans. Signal Processing*, 48(11), pp. 3235-3242 (Nov. 2000).
- Shin, H.-C., Song, W.J. and Sayed, A.H. "Mean-square performance of data-reusing adaptive algorithms", *IEEE Signal Processing Letters*, **12**(12), pp. 851-854 (Dec. 2005).
- Pradhan, S.S. and Reddy, V.E. "A new approach to subband adaptive filtering", *IEEE Trans. Signal Processing*, 47(3), pp. 655-664 (1999).
- de Courville, M. and Duhamel, P. "Adaptive filtering in subbands using a weighted criterion", *IEEE Trans.* Signal Processing, 46(9), pp. 2359-2371 (1998).
- Lee, K.A. and Gan, W.S. "Improving convergence of the NLMS algorithm using constrained subband updates", *IEEE Signal Processing Letters*, **11**, pp. 736-739 (2004).
- Kwong, O.W. and Johnston, E.D. "A variable step size LMS algorithm", *IEEE Trans. Signal Processing*, 40(7), pp. 1633-1642 (1992).
- Mathews, J. and Xie, Z. "Stochastic gradient adaptive filters with gradient adaptive step size", *IEEE Trans.* Signal Processing, 35(6), pp. 450-454 (1993).
- Aboulnasr, T. and Mayyas, K. "A robust variable stepsize lms-type algorithm: Analysis and simulations", *IEEE Trans. Signal Processing*, 45(3), pp. 631-639 (1997).
- Pazaitis, D.I. and Constantinides, A.G. "A novel kurtosis driven variable step-size adaptive algorithm", *IEEE Trans. Signal Processing*, 47(3), pp. 864-872 (1999).

- Shin, H.C., Sayed, A.H. and Song, W.J. "Variable stepsize NLMS and affine projection algorithms", *IEEE* Signal Processing Letters, 11, pp. 132-135 (Feb. 2004).
- Douglas, S.C. "Analysis and implementation of the max-NLMS adaptive filter", in Proc. 29th Asilomar Conf. on Signals, Systems, and Computers, Pacific Grove, CA, pp. 659-663 (Oct. 1995).
- Aboulnasr, T. and Mayyas, K. "Selective coefficient update of gradient-based adaptive algorithms", in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Munich, Germany, pp. 1929-1932 (Apr. 1997).
- Aboulnasr, T. and Mayyas, K. "Complexity reduction of the NLMS algorithm via selective coefficient update", *IEEE Trans. Signal Processing*, 47(5), pp. 1421-1424 (May 1999).
- Schertler, T. "Selective block update NLMS type algorithms", in Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, Seattle, WA, pp. 1717-1720 (May 1998).
- Doğançay, K. and Tanrikulu, O. "Adaptive filtering algorithms with selective partial updates", *IEEE Trans. Circuits, Syst. II: Analogand Digital Signal Processing*, 48(8), pp. 762-769 (Aug. 2001).
- Werner, S., de Campos, M.L.R. and Diniz, P.S.R. "Partial-update NLMS algorithms with data-selective updating", *IEEE Trans. Signal Processing*, 52(4), pp. 938-948 (Apr. 2004).
- Dogancay, K., Partial-Update Adaptive Signal Processing: Design, Analysis and Implementation, Academic Press (2008).
- 24. Hwang, K.Y. and Song, W.J. "An affine projection adaptive filtering algorithm with selective regressors", *IEEE Trans. Circuits, Syst. II: EXPRESS BRIEFS*, 54(1), pp. 43-46 (Jan. 2007).
- 25. Gay, S.L. and Benesty, J., Acoustic Signal Processing for Telecommunication, Boston, MA, Kluwer (2000).
- Sankaran, S.G. and Beex, A.A.L. "Normalized LMS algorithm with orthogonal correction factors", in Proc. Asilomar Conf. on Signals, Systems, and Computers, pp. 1670-1673 (1997).
- Kratzer, S.G. and Morgan, D.R. "The partial-rank algorithm for adaptive beamforming", in *Proc. SPIE Int. Soc. Opt. Eng.*, pp. 9-14 (1985).
- Knuth, D.E., "Sorting and searching", The Art of Computer Programming, 3, 2nd Ed., Reading, MA, Addison-Wesley (1973).
- 29. Moon, T.K. and Sterling, W.C., Mathematical Methods and Algorithms for Signal Processing, Upper Saddle River, Prentice Hall (2000).

APPENDIX A

Finding an Approximation for Ψ , and Υ

In VSS-SR-APA, positive constant Ψ is related to $\Psi = \sigma_v^2 \operatorname{Tr}(E\{(\mathbf{X}_G^T(n)\mathbf{X}_G(n))^{-1}\})$. This quantity is given by:

$$\Psi = \sigma_v^2 \operatorname{Tr} \left(E \left(\begin{pmatrix} x(n-i_1) & x(n-i_1-1) \\ x(n-i_2) & x(n-i_2-1) \\ \vdots & \vdots \\ x(n-i_P) & x(n-i_2-1) \end{pmatrix} \right) \\ & \cdots & x(n-i_P-M+1) \\ & \ddots & \vdots \\ & \cdots & x(n-i_P-M+1) \end{pmatrix} \\ \times \left(\begin{pmatrix} x(n-i_1) & x(n-i_2) \\ x(n-i_1-1) & x(n-i_2-1) \\ \vdots & \vdots \\ x(n-i_1-M+1) & x(n-i_2-M+1) \\ & \vdots \\ x(n-i_1-M+1) & x(n-i_2-M+1) \end{pmatrix} \right) \\ & \cdots & x(n-i_P-1) \\ & \ddots & \vdots \\ & \cdots & x(n-i_P-M+1) \end{pmatrix} \right)^{-1} \right).$$
(A1)

Similar to [28], by neglecting the off-diagonal elements, Equation 53 can be obtained:



Now applying the expectation and trace operators, we obtain:

$$\Psi = \sigma_v^2 \left(E \left\{ \frac{1}{\|\mathbf{x}(n-i_1)\|^2} \right\} + E \left\{ \frac{1}{\|\mathbf{x}(n-i_2)\|^2} \right\} + \dots + E \left\{ \frac{1}{\|\mathbf{x}(n-i_P)\|^2} \right\}.$$
(A3)

Equation 55 can be stated as:

$$\Psi = P \cdot \sigma_v^2 E\left\{\frac{1}{\|\mathbf{x}(n-i_1)\|^2}\right\}.$$
 (A4)

Now, we obtain that Ψ can be approximated as P/SNR. Therefore, Ψ is inversely proportional to SNR and proportional to P.

In VSS-SPU-APA, parameter $\boldsymbol{\Upsilon}$ is obtained from the following relation:

$$\begin{split} \mathbf{\Upsilon} &= \sigma_v^2 \operatorname{Tr} \left(E \left(\begin{pmatrix} \mathbf{x}_{j_1}^T(n) & \mathbf{x}_{j_2}^T(n) \\ \mathbf{x}_{j_1}^T(n-1) & \mathbf{x}_{j_2}^T(n-1) \\ \vdots & \vdots \\ \mathbf{x}_{j_1}^T(n-K+1) & \mathbf{x}_{j_2}^T(n-K+1) \\ & \cdots & \mathbf{x}_{j_S}^T(n) \\ & \cdots & \mathbf{x}_{j_S}^T(n-1) \\ & \ddots & \vdots \\ & \cdots & \mathbf{x}_{j_S}^T(n-K+1) \\ & \vdots & \vdots \\ \mathbf{x}_{j_S}(n) & \mathbf{x}_{j_S}(n-1) \\ & \vdots & \vdots \\ \mathbf{x}_{j_S}(n) & \mathbf{x}_{j_S}(n-1) \\ & \cdots & \mathbf{x}_{j_S}(n-K+1) \\ & \ddots & \vdots \\ & \cdots & \mathbf{x}_{j_S}(n-K+1) \\ & \ddots & \vdots \\ & \cdots & \mathbf{x}_{j_S}(n-K+1) \\ \end{pmatrix} \right)^{-1} \right), \quad (A5) \end{split}$$

which can be represented as:



By applying the expectation and trace operators, Equation A2 can be stated as:

$$\Upsilon = K \cdot \sigma_v^2 E \left\{ \frac{1}{\|\mathbf{x}_{j_1}(n)\|^2 + \dots + \|\mathbf{x}_{j_S}(n)\|^2} \right\}.$$
 (A7)

Now we obtain that Υ can be approximated as K/SNR.

APPENDIX B

Mean-Square Stability Analysis of the Family of SPU-APA, SR-APA and SPU-SR-APA

Now, we introduce the general filter vector update equation to analyze the mean-square stability of the family of SPU and SR affine projection algorithms. The general filter vector update equation to establish the family of SPU-APA and SR-APA is introduced as:

$$\mathbf{h}(n+1) = \mathbf{h}(n) + \mu \mathbf{C}(n) \mathbf{X}(n) \mathbf{Z}(n) \mathbf{e}(n).$$
(B1)

where $\mathbf{C}(n)$ and $\mathbf{Z}(n)$ matrices are obtained from Table B1. To find the theoretical stability bound, we first study the transientbehavior of the adaptive algorithms. The transient behavior of an adaptive filter algorithm is determined by evolution of the expected squared *a priori* error in time *n*, i.e. $E\{e_a^2(n)\}$, which is:

$$E\{e_a^2(n)\} = E\{\tilde{\mathbf{h}}^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\tilde{\mathbf{h}}(n)\},\tag{B2}$$

where $\mathbf{\tilde{h}}(n) = \mathbf{h}_t - \mathbf{h}(n)$ is the weight-error vector. Employing the common *independence assumption* [2], we have:

$$E\{e_a^2(n)\} = E\{\tilde{\mathbf{h}}^T(n)\mathbf{R}\tilde{\mathbf{h}}(n)\} = E\{\|\tilde{\mathbf{h}}(n)\|_{\mathbf{R}}^2\}, \quad (B3)$$

where the autocorrelation matrix is $\mathbf{R} = E\{\mathbf{x} (n)\mathbf{x}^T(n)\}$. Thus, to obtain the learning curve, we need to find $E\{\|\mathbf{\tilde{h}}(n)\|_{\mathbf{R}}^2\}$ as a function of n. We can recursively obtain $E\{\|\mathbf{\tilde{h}}(n)\|_{\mathbf{\Sigma}}^2\}$, where $\mathbf{\Sigma}$ is a positive definite symmetric matrix whose dimension is commensurate with that of $\mathbf{\tilde{h}}(n)$. If we substitute Equation 1 into Equation 12, the relation between the output estimation error vector, the a priori error vector and the noise vector is:

$$\mathbf{e}(n) = \mathbf{e}_a(n) + \mathbf{v}(n),\tag{B4}$$

where $\mathbf{e}_a(n) = \mathbf{X}^T(n)\tilde{\mathbf{h}}(n)$ is the a priori error vector. The generic weight error vector update equation can be stated as:

$$\tilde{\mathbf{h}}(n+1) = \tilde{\mathbf{h}}(n) - \mu \mathbf{C}(n) \mathbf{X}(n) \mathbf{Z}(n) (\mathbf{X}^T(n) \tilde{\mathbf{h}}(n) + \mathbf{v}(n)).$$
(B5)

By defining $\mathbf{D}(n) = \mathbf{Z}^T(n)\mathbf{X}^T(n)\mathbf{C}^T(n)$, the Σ weighted norm of both sides of Equation B5 is:

$$\begin{split} \|\tilde{\mathbf{h}}(n+1)\|_{\boldsymbol{\Sigma}}^2 &= \|\tilde{\mathbf{h}}(n)\|_{\boldsymbol{\Sigma}'}^2 + \mu^2 \mathbf{v}^T(n) \mathbf{X}^{\boldsymbol{\Sigma}}(n) \mathbf{v}(n) \\ &+ \{ \text{Cross terms involving one instance of} \mathbf{v}(n) \}, \end{split}$$
(B6)

Table B1. Family of SPU, SR, and SPU-SR affine projection algorithms.

Algorithm	K	D	$\mathbf{C}(n)$	$\mathbf{Z}(n)$
SPU-APA	$K \leq M$	D = 1	$\mathbf{A}(n)\mathbf{I}$	$(\mathbf{X}^T(n)\mathbf{A}(n)\mathbf{X}(n))^{-1}$
SPU-BNDR-LMS	K = 2	D = 1	$\mathbf{A}(n)\mathbf{I}$	$(\mathbf{X}^T(n)\mathbf{A}(n)\mathbf{X}(n))^{-1}$
SPU-R-APA	$K \leq M$	D = 1	$\mathbf{A}(n)\mathbf{I}$	$(\epsilon \mathbf{I} + \mathbf{X}^T(n)\mathbf{A}(n)\mathbf{X}(n))^{-1}$
SPU-NLMS-OCF	$K \leq M$	$D \ge 1$	$\mathbf{A}(n)\mathbf{I}$	$(\mathbf{X}^T(n)\mathbf{A}(n)\mathbf{X}(n))^{-1}$
SR-APA	$K \leq M$	D = 1	I	$\mathbf{B}(n)(\mathbf{B}^{T}(n)\mathbf{X}^{T}(n)\mathbf{X}(n)\mathbf{B}(n))^{-1}\mathbf{B}^{T}(n)$
SR-BNDR-LMS	K = 2	D = 1	I	$\mathbf{B}(n)(\mathbf{B}^{T}(n)\mathbf{X}^{T}(n)\mathbf{X}(n)\mathbf{B}(n))^{-1}\mathbf{B}^{T}(n)$
SR-R-APA	$K \leq M$	D = 1	I	$\mathbf{B}(n)(\epsilon \mathbf{I} + \mathbf{B}^{T}(n)\mathbf{X}^{T}(n)\mathbf{X}(n)\mathbf{B}(n))^{-1}\mathbf{B}^{T}(n)$
SR-NLMS-OCF	$K \leq M$	$D \ge 1$	I	$\mathbf{B}(n)(\mathbf{B}^{T}(n)\mathbf{X}^{T}(n)\mathbf{X}(n)\mathbf{B}(n))^{-1}\mathbf{B}^{T}(n)$
SPU-SR-APA	$K \leq M$	D = 1	$\mathbf{A}(n)\mathbf{I}$	$\mathbf{B}(n)(\mathbf{B}^{T}(n)\mathbf{X}^{T}(n)\mathbf{A}(n)\mathbf{X}(n)\mathbf{B}(n))^{-1}\mathbf{B}^{T}(n)$
SPU-SR-BNDR-LMS	K = 2	D = 1	$\mathbf{A}(n)\mathbf{I}$	$\mathbf{B}(n)(\mathbf{B}^{T}(n)\mathbf{X}^{T}(n)\mathbf{A}(n)\mathbf{X}(n)\mathbf{B}(n))^{-1}\mathbf{B}^{T}(n)$
SPU-SR-R-APA	$K \leq M$	D = 1	$\mathbf{A}(n)\mathbf{I}$	$\mathbf{B}(n)(\epsilon \mathbf{I} + \mathbf{B}^{T}(n)\mathbf{X}^{T}(n)\mathbf{A}(n)\mathbf{X}(n)\mathbf{B}(n))^{-1}\mathbf{B}^{T}(n)$
SPU-SB-NLMS-OCF	K < M	D > 1	$\mathbf{A}(n)\mathbf{I}$	$\mathbf{B}(n)(\mathbf{B}^{T}(n)\mathbf{X}^{T}(n)\mathbf{A}(n)\mathbf{X}(n)\mathbf{B}(n))^{-1}\mathbf{B}^{T}(n)$

where:

$$\Sigma' = \Sigma - \mu \Sigma \mathbf{D}^{T}(n) \mathbf{X}^{T}(n) - \mu \mathbf{X}(n) \mathbf{D}(n) + \mu^{2} \mathbf{X}(n) \mathbf{X}^{\Sigma}(n) \mathbf{X}^{T}(n),$$
(B7)

and:

$$\mathbf{X}^{\Sigma}(n) = \mathbf{D}(n)\Sigma\mathbf{D}^{T}(n).$$
(B8)

Taking the expectation from both sides of Equation B6 yields:

$$\begin{split} E\{\|\tilde{\mathbf{h}}(n+1)\|_{\boldsymbol{\Sigma}}^{2}\} &= E\{\|\tilde{\mathbf{h}}(n)\|_{\boldsymbol{\Sigma}'}^{2}\} \\ &+ \mu^{2} E\{\mathbf{v}^{T}(n)\mathbf{X}^{\boldsymbol{\Sigma}}(n)\mathbf{v}(n)\}. \end{split} \tag{B9}$$

We now obtain the time evolution of the weight-error variance. The expectation of $\|\tilde{\mathbf{h}}(n)\|_{\Sigma'}^2$ is difficult to calculate because of the dependency of Σ' on $\mathbf{C}(n)$, $\mathbf{Z}(n)$, $\mathbf{X}(n)$ and of $\tilde{\mathbf{h}}(n)$ on prior regressors. To solve this problem, we need to use the following independence assumptions [4]:

- 1. $\mathbf{X}(n)$ is an independent and identically distributed sequence matrix. This assumption guarantees that $\tilde{\mathbf{h}}(n)$ is independent of both $\mathbf{\Sigma}'$ and $\mathbf{X}(n)$.
- 2. $\tilde{\mathbf{h}}(n)$ is independent of $\mathbf{D}^T(n)\mathbf{X}^T(n)$.

Using these assumptions, the final results is:

$$E\{\|\tilde{\mathbf{h}}(n+1)\|_{\boldsymbol{\Sigma}}^{2}\} = E\{\|\tilde{\mathbf{h}}(n)\|_{\boldsymbol{\Sigma}'}^{2}\}$$
$$+ \mu^{2} E\{\mathbf{v}^{T}(n)\mathbf{X}^{\boldsymbol{\Sigma}}(n)\mathbf{v}(n)\}, \quad (B10)$$

where:

$$\Sigma' = \Sigma - \mu \Sigma E \{ \mathbf{D}^T(n) \mathbf{X}^T(n) \} - \mu E \{ \mathbf{X}(n) \mathbf{D}(n) \} \Sigma$$
$$+ \mu^2 E \{ \mathbf{X}(n) \mathbf{X}^{\Sigma}(n) \mathbf{X}^T(n) \}.$$
(B11)

Looking only at the second term of the right hand side of Equation B10, we write:

$$E\{\mathbf{v}^{T}(n)\mathbf{X}^{\Sigma}(n)\mathbf{v}(n)\} = E\{\operatorname{Tr}(\mathbf{v}(n)\mathbf{v}^{T}(n)\mathbf{X}^{\Sigma}(n))\}$$
$$= \operatorname{Tr}(E\{\mathbf{v}(n)\mathbf{v}^{T}(n)\}E\{\mathbf{X}^{\Sigma}(n)\}).$$
(B12)

Since $E\{\mathbf{v}(n)\mathbf{v}^{T}(n)\} = \sigma_{v}^{2}\mathbf{I}$, Equation B10 can be stated as:

$$E\{\|\tilde{\mathbf{h}}(n+1)\|_{\Sigma}^{2}\} = E\{\|\tilde{\mathbf{h}}(n)\|_{\Sigma'}^{2}\} + \mu^{2}\sigma_{v}^{2}\mathrm{Tr}(E\{\mathbf{X}^{\Sigma}(n)\}).$$
(B13)

Applying the vec(.) operator [29] on both sides of Equation B11 yields:

$$\operatorname{vec}(\mathbf{\Sigma}') = \operatorname{vec}(\mathbf{\Sigma}) - \mu \operatorname{vec}(\mathbf{\Sigma} E\{\mathbf{D}^{T}(n)\mathbf{X}^{T}(n)\})$$
$$- \mu \operatorname{vec}(E\{\mathbf{X}(n)\mathbf{D}(n)\}\mathbf{\Sigma})$$
$$+ \mu^{2}\operatorname{vec}(E\{\mathbf{X}(n)\mathbf{X}^{\mathbf{\Sigma}}(n)\mathbf{X}^{T}(n)\}).$$
(B14)

Since, in general, $\operatorname{vec}(\mathbf{P}\Sigma\mathbf{Q}) = (\mathbf{Q}^T \otimes \mathbf{P})\operatorname{vec}(\Sigma)$ [29], Equation B14 can be written as:

$$\sigma' = \sigma - \mu(E\{\mathbf{X}(n)\mathbf{D}(n)\} \otimes \mathbf{I}).\sigma$$
$$- \mu(\mathbf{I} \otimes E\{\mathbf{X}(n)\mathbf{D}(n)\}).\sigma$$
$$+ \mu^2(E\{(\mathbf{X}(n)\mathbf{D}(n)) \otimes (\mathbf{X}(n)\mathbf{D}(n))\}).\sigma, \quad (B15)$$

where $\sigma' = \operatorname{vec}(\Sigma')$ and $\sigma = \operatorname{vec}(\Sigma)$. By defining the $M^2 \times M^2$ matrix **G** as:

$$\mathbf{G} = \mathbf{I} - \mu E\{\mathbf{X}(n)\mathbf{D}(n)\} \otimes \mathbf{I} - \mu \mathbf{I} \otimes E\{\mathbf{X}(n)\mathbf{D}(n)\}$$
$$+ \mu^2 E\{(\mathbf{X}(n)\mathbf{D}(n)) \otimes (\mathbf{X}(n)\mathbf{D}(n))\}, \qquad (B16)$$

M. Shams Esfand Abadi, V. Mehrdad and A. Gholipour

Equation B15 becomes:

$$\sigma' = \mathbf{G}.\sigma. \tag{B17}$$

The second term of the right hand side of Equation B13 is:

$$\operatorname{Tr}(E\{\mathbf{X}^{\Sigma}(n)\}) = \operatorname{Tr}(E\{\mathbf{D}^{T}(n)\mathbf{D}(n)\}, \boldsymbol{\Sigma}).$$
(B18)

Defining γ as:

$$\gamma = \operatorname{vec}(E\{\mathbf{D}^T(n)\mathbf{D}(n)\}),\tag{B19}$$

we have:

$$\operatorname{Tr}(E\{\mathbf{D}^{T}(n)\mathbf{D}(n)\},\mathbf{\Sigma}) = \gamma^{T}.\sigma.$$
(B20)

From the above, the recursion of Equation B13 is:

$$E\{\|\tilde{\mathbf{h}}(n+1)\|_{\sigma}^{2}\} = E\{\|\tilde{\mathbf{h}}(n)\|_{\mathbf{G}\sigma}^{2}\} + \mu^{2}\sigma_{v}^{2}\gamma^{T}\sigma.$$
(B21)

The equation is stable if matrix **G** is stable [4]. From Equation B16, we know that $\mathbf{G} = \mathbf{I} - \mu \mathbf{M} + \mu^2 \mathbf{N}$, where $\mathbf{M} = E\{\mathbf{X}(n)\mathbf{D}(n)\} \otimes \mathbf{I} + \mathbf{I} \otimes E\{\mathbf{X}(n)\mathbf{D}(n)\}$, and $\mathbf{N} = E\{(\mathbf{X}(n)\mathbf{D}(n)) \otimes (\mathbf{X}(n)\mathbf{D}(n))\}$. The condition on μ to guarantee the convergence in the mean-square sense of the adaptive algorithms is:

$$0 < \mu < \min\{\frac{1}{\lambda_{\max}(\mathbf{M}^{-1}\mathbf{N})}, \frac{1}{\max(\lambda(\mathbf{H}) \in \Re^+)}\}, (B22)$$

where $\mathbf{H} = \begin{bmatrix} \frac{1}{2}\mathbf{M} & -\frac{1}{2}\mathbf{N} \\ \mathbf{I} & \mathbf{0} \end{bmatrix}$. Taking the expectation from both sides of Equation B5 yields:

$$E\{\tilde{\mathbf{h}}(n+1)\} = [\mathbf{I} - \mu E\{\mathbf{D}^T(n)\mathbf{X}^T(n)\}]E\{\tilde{\mathbf{h}}(n)\}.$$
(B23)

From Equation B23, the convergence to the mean of the adaptive algorithm in Equation B1 is guaranteed for any μ that satisfies:

$$\mu < \frac{2}{\lambda_{\max}(E\{\mathbf{D}^T(n)\mathbf{X}^T(n)\})}.$$
(B24)

BIOGRAPHIES

Mohammad Shams Esfand Abadi was born in Tehran, Iran, in 1978. He received a BS degree in Electrical Engineering from Mazandaran University, in Iran, and his MS degree in Electrical Engineering from Tarbiat Modares University, in Tehran, in 2000 and 2002, respectively. He, subsequently, received his PhD degree in Biomedical Engineering from Tarbiat Modarres University, in Tehran, Iran, in 2007.

Since 2004, he has been with the Department of Electrical Engineering at Shahid Rajaee University, in Tehran, Iran. During the Autumn of 2003, Spring 2005, and again in the Spring of 2007, he was a visiting scholar with the Signal Processing Group at the University of Stavanger, Norway. His research interests include Digital Filter Theory and Adaptive Signal Processing Algorithms.

Vahid Mehrdad was born in Khoramabad, Iran, in 1985. He received a BS degree in Electronic Engineering from Shahid Rajaee Teacher Training University, in Tehran, Iran, in 2007, where he is now a Master's degree student in the Department of Electrical Engineering. His research activities include Adaptive Signal Processing and Adaptive Filtering.

Azizollah Gholipour was born in Behshahr, Mazandaran, Iran, in 1974. He received a BS degree in Electronic Engineering from Mohajer Institute of Technology, in Isfahan, Iran, in 1999. He is presently a Master's degree student in the Department of Electrical Engineering at the Shahid Rajaee Teacher Training University, in Tehran, Iran. His research activities include Adaptive Signal Processing and Adaptive Filtering.

98