Stochastic Partial Update LMS Algorithm

Mahesh Godavarti and Alfred O. Hero-III

October 23, 2001

Submitted to *IEEE Transactions on Signal Processing*, October 2001

Abstract

Partial updating of LMS filter coefficients is an effective method for reducing the computational load and the power consumption in adaptive filter implementations. Several algorithms have been proposed in the literature based on partial updating. Unfortunately, it has been observed that these algorithms don't have good convergence properties in practice. In particular, there generally exist signals for which these algorithms stagnate or diverge. In this paper, we propose a new algorithm, called the Stochastic Partial Update LMS (SPU-LMS) algorithm which attempts to remedy some of the drawbacks of existing algorithms. The SPU-LMS algorithm differs from the existing algorithms in that the subsets to be updated are chosen in a random manner at each iteration. We derive conditions for filter stability, convergence rate, and steady state mean-square error for the proposed algorithm and show that SPU-LMS suffers no loss in steady state performance when compared to the regular (full-update) LMS algorithm.

Keywords: partial update LMS algorithms, random updates, sequential algorithm, periodic algorithm.

Corresponding author:

Mahesh godavarti
Altra Broadband, Inc.
Irvine Technology Center
16275 Laguna Canyon Rd
Suite 150
Irvine, CA 92618.
Tel: (949) 341-0106 x230.
Fax: (949) 341-0226.
e-mail: godavarti@altrabroadband.com.

[1]This work was performed while Mahesh Godavarti was a Ph.D. candidate at the University of Michigan, Ann Arbor under the supervision of Prof. Alfred O. Hero-III and was partially supported by the Army Research Office under grant ARO DAAH04-96-1-0337. Parts of the work were presented at SAM 2000, Boston MA

1 Introduction

The LMS algorithm is a popular algorithm for adaptation of weights in the field of adaptive beamforming using antenna arrays. This has application in many areas including interference cancellation, space time modulation and coding, signal copy in surveillance and wireless communications. Some of the applications like echo cancellation and channel equalization require a large number of filter coefficients and hence the coefficient updates might prove too expensive for mobile units with limited processing power. Therefore, partial updating of the LMS adaptive filter has been proposed to reduce these per-iteration computational costs [9, 10, 15] of the algorithm.

Two types of partial update LMS algorithms are prevalent in the literature and have been described in [7]. They are referred to as the "Periodic LMS algorithm" and the "Sequential LMS algorithm". To reduce computation by a factor of P, the Periodic LMS algorithm (P-LMS) updates all the filter coefficients every P^{th} iteration instead of every iteration. The Sequential LMS (S-LMS) algorithm updates only a fraction of coefficients every iteration. Another variant referred to as "Max Partial Update LMS algorithm" (Max PU-LMS) has been proposed in [5, 6] and [1]. In this algorithm, the subset of coefficients to be updated is dependent on the input signal. The subset is chosen so as to minimize the increase in the mean squared error due to partial as opposed to full updating. The input signals multiplying each coefficient are ordered according to their magnitude and the coefficients corresponding to the largest $\frac{1}{P}$ of input signals are chosen for update in an iteration. Some analysis of this algorithm has been performed in [6] for the special case of $P = 1$ but, analysis for more general cases still needs to be completed.

In [11], the authors have analysed the convergence of S-LMS for stationary and cyclo-stationary signals. For stationary signals, it was shown that for any arbitary sequence of updates, S-LMS converges in the mean if LMS converges. For a class of cyclo-stationary signals, for the case of even-odd updates, it was shown that the convergence in the mean conditions on μ were much stricter than that for regular LMS. However, as will be shown in this paper there exist signals for which there is no region of μ for which S-LMS will converge.

The important characteristic of S-LMS and P-LMS is that the coefficients to be updated at an iteration are pre-determined. It is this characteristic which renders P-LMS and S-LMS unstable for certain signals and which makes random coefficient updating attractive.

The algorithm proposed in the paper is similar to S-LMS except that the subset of the filter coefficients that are updated each iteration is selected at random. The algorithm, referred to as Stochastic Partial Update LMS algorithm (SPU-LMS), involves selection of a subset of size $\frac{N}{P}$ coefficients out of P possible subsets from a fixed partition of the N coefficients in the weight vector. For example, filter coefficients can be partitioned into even and odd subsets and either even or odd coefficients are randomly selected to be updated in each iteration. In this paper we derive conditions on the step-size parameter which ensures convergence in the mean and the mean square sense for stationary signals, for deterministic signals and the general case of mixing signals.

The contributions of this paper can be summarized as follows:

- We demonstrate signal scenarios for which the prevalent partial update algorithms do not converge.
- We propose a new algorithm called the Stochastic Partial Update LMS Algorithm (SPU-LMS) based on random updating of filter coefficients that ensures convergence of filter coefficients.
- We derive stability conditions for SPU-LMS for stationary signal scenarios and demonstrate that the steady state performance of the new algorithm is as good as that of the regular LMS algorithm.
- We derive the persistence of excitation condition for the convergence of SPU-LMS for the case of deterministic signals and show that this condition is same as that of the regular LMS algorithm.
- For the general case of mixing signals we show that the stability conditions for SPU-LMS are same as that of LMS. We extend the analysis of [2] to SPU-LMS and use the results to show that SPU-LMS does not suffer a degradation in steady state performance as compared to LMS even when we relax the assumptions made for the performance analysis of the algorithm for stationary signals.
- We demonstrate through different examples that for non-stationary signal scenarios (echo cancellation in telephone networks, digital communication systems) partial updating using P-LMS and S-LMS should not be employed as these are not guaranteed to converge. SPU-LMS is a better choice because of its convergence properties.

The organization of the paper is as follows. First, a brief description of the algorithm is given in Section 2 followed by analysis of the stochastic partial update algorithm for stationary stochastic signals in Section 2.1, for deterministic signals in Section 2.2 and for mixing signals in Section 3. Section 4 discusses the advantage of the new algorithm over the existing Partial Update LMS algorithms. This is followed by Section 4.1 where verification of theoretical analysis in Section 2.1 of the new algorithm is carried out via simulations and examples are given to illustrate the usefulness of SPU-LMS. Sections 4.2 and 4.3 apply the analysis in Section 3 to separate signal scenarios for comparing the steady state performance of LMS and SPU-LMS. Finally conclusions and directions for future work are indicated in Section 5.

2 The Stochastic PU LMS Algorithm

For description purposes we will assume that the filter coefficients can be divided into P mutually exclusive subsets of equal size, i.e. the filter length N is a multiple of P. For convenience, define the index set $S = \{1, 2, \ldots, N\}$. Partition S into P mutually exclusive subsets of equal size, $S_1,\ S_2, \ldots, S_P$.

Let $W_k = [w_{k,1},\ w_{k,2}, \ldots,\ w_{k,N}]^T$ be the column weight vector at iteration k of the LMS algorithm. Let X_k be as defined in section 1. Define I_i by zeroing out the j^{th} row of the identity matrix I if $j \notin S_i$. In that case, $I_i X_k$ will have precisely $\frac{N}{P}$ non-zero entries. Let the sentence "choosing S_i at iteration k" stand to mean "choosing the weights with their indices in S_i for update at iteration k".

The SPU-LMS algorithm is described as follows. At a given iteration, k, one of the sets S_i, $i = 1, \ldots, P$, is chosen at random from $\{S_1,\ S_2, \ldots,\ S_P\}$ with probability $\frac{1}{P}$ and the update is performed. i.e.

$$w_{k+1,j} = \begin{cases} w_{k,j} + \mu e_k^* x_{k,j} & \text{if } j \in S_i \\ w_{k,j} & \text{otherwise} \end{cases} \tag{1}$$

where $e_k = d_k - W_k^H X_k$. The above update equation can be written in a more compact form in the following manner

$$W_{k+1} = W_k + \mu e_k^* I_i X_k \tag{2}$$

where I_i now is a randomly chosen matrix.

2.1 Analysis: Stationary Stochastic Signals

In this setting the offline problem is to choose an optimal W such that

$$\begin{aligned} \xi^{(W)} &= E\left[(d_k - y_k)(d_k - y_k)^*\right] \\ &= E\left[(d_k - W^H X_k)(d_k - W^H X_k)^*\right] \end{aligned}$$

is minimized, where a^* denotes the complex conjugate of a. The solution to this problem is given by

$$W_{opt} = R^{-1} r \tag{3}$$

where $R = E[X_k X_k^H]$ and $r = E[d_k^* X_k]$. The minimum attainable mean square error $\xi^{(W)}$ is given by

$$\xi_{min} = E[d_k d_k^*] - r^H R^{-1} r.$$

For the following analysis, we assume that the desired signal, d_k satisfies the following relation[2] [7]

$$d_k = W_{opt}^H X_k + n_k \tag{4}$$

where X_k is a zero mean complex circular Gaussian[3] random vector and n_k is a zero mean circular complex Gaussian (not necessarily white) noise, with variance ξ_{min}, uncorrelated with X_k.

[2]Note: the model assumed for d_k is same as assuming d_k and X_k are jointly Gaussian sequences. Under this assumption d_k can be written as $d_k = W_{opt}^H X_k + m_k$, where W_{opt} is as in (3) and $m_k = d_k - W_{opt}^H X_k$. Since $E[m_k X_k] = E[X_k d_k] - E[X_k X_k^H] W_{opt} = 0$ and m_k and X_k are jointly Gaussian we conclude that m_k and X_k are independent of each other which is same as model (4).

[3]A complex circular Gaussian random vector consists of Gaussian random variables whose marginal densities depend only on their magnitudes. For more information see [16, p. 198] or [14]

We also make the usual independence assumption used in the analysis of standard LMS [3]. We assume that X_k is a Gaussian random vector and that X_k is independent of X_j for $j < k$. We also assume that I_i and X_k are mutually independent.

For convergence-in-mean analysis we obtain the following update equation conditioned on a choice of S_i.

$$E[V_{k+1}|S_i] = (I - \mu I_i R)E[V_k|S_i]$$

which after averaging over all choices of S_i gives

$$E[V_{k+1}] = (I - \frac{\mu}{P}R)E[V_k]. \tag{5}$$

To obtain the above equation we have made use of the fact that the choice of S_i is independent of V_k and X_k. Therefore, μ has to satisfy $0 < \mu < \frac{2P}{\lambda_{max}}$ to guarantee convergence in mean.

For convergence-in-mean square we are interested in the convergence of $E[e_k e_k^*]$. Under the independence assumptions we obtain $E[e_k e_k^*] = \xi_{min} + \mathrm{tr}\{RE[V_k V_k^H]\}$ where ξ_{min} is as defined earlier.

We have followed the procedure of [13] for our mean-square analysis. First, conditioned on a choice of S_i, the evolution equation of interest for $\mathrm{tr}\{RE[V_k V_k^H]\}$ is given by

$$\begin{aligned} RE[V_{k+1}V_{k+1}^H|S_i] &= RE[V_k V_k^H|S_i] - 2\mu RI_i RE[V_k V_k^H|S_i] + \\ &\quad \mu^2 I_i R I_i E[X_k X_k^H A_k X_k X_k^H|S_i] + \mu^2 \xi_{min} RI_i RI_i \end{aligned}$$

where $A_k = E[V_k V_k^H]$. For simplicity, consider the case of block diagonal R satisfying $\sum_{i=1}^P I_i R I_i = R$. Then, we obtain the final equation of interest for convergence-in-mean square to be

$$G_{k+1} = (I - \frac{2\mu}{P}\Lambda + \frac{\mu^2}{P}\Lambda^2 + \frac{\mu^2}{P}\Lambda^2 \mathbf{1}\mathbf{1}^T)G_k + \frac{\mu^2}{P}\xi_{min}\Lambda^2 \mathbf{1} \tag{6}$$

where G_k is a vector of diagonal elements of $\Lambda E[U_k U_k^H]$ where $U_k = QV_k$ with Q such that $QRQ^H = \Lambda$. It is easy to obtain the following necessary and sufficient conditions (see Appendix A) for convergence of the SPU-LMS algorithm

$$\begin{gathered} 0 < \mu < \frac{2}{\lambda_{max}} \\ \eta(\mu) \stackrel{\text{def}}{=} \sum_{i=1}^N \frac{\mu\lambda_i}{2-\mu\lambda_i} < 1 \end{gathered} \tag{7}$$

which is independent of P and identical to that of LMS.

We use the integrated MSE difference $J = \sum_{k=0}^{\infty}[\xi_k - \xi_\infty]$ introduced in [8] as a measure of the convergence rate and $M(\mu) = \frac{\xi_\infty - \xi_{min}}{\xi_{min}}$ as a measure of misadjustment. The misadjustment factor is simply (see Appendix C)

$$M(\mu) = \frac{\eta(\mu)}{1-\eta(\mu)} \tag{8}$$

which is the same as that of the standard LMS. Thus, we conclude that random update of subsets has no effect on the final excess mean-squared error.

Finally, it is straightforward to show (see Appendix B) the integrated MSE difference is

$$J = P \ \mathrm{tr}\{[2\mu\Lambda - \mu^2\Lambda^2 - \mu^2\Lambda^2\mathbf{1}\mathbf{1}^T]^{-1}(G_0 - G_\infty)\} \tag{9}$$

which is P times the quantity obtained for standard LMS algorithm. Therefore, we conclude that for block diagonal R, random updating slows down convergence by a factor of P without affecting the misadjustment. Furthermore, it can be easily verified that a much simpler condition $0 < \mu < \frac{1}{\mathrm{tr}\{R\}}$, is a sufficient region for convergence of SPU-LMS and the standard LMS algorithm.

2.2 Analysis: Deterministic Signals

Here we followed the analysis for LMS with real signals given in [17, pp. 140–143]. This analysis can be easily extended to SPU-LMS with complex signals which we present here. We assume that the input signal X_k is bounded, that is $\sup_k(X_k^H X_k) \le B < \infty$ and that the desired signal d_k follows the model

$$d_k = W_{opt}^H X_k$$

which is different from (4) in that we assume that there is no noise present at the output.

Define $V_k = W_k - W_{opt}$ and $e_k = d_k - W_k^H X_k$.

Lemma 1 *If $\mu < 2/B$ then $\overline{e_k^2} \to 0$ as $k \to \infty$. Here, $\overline{\{\cdot\}}$ indicates statistical expectation over all possible choices of S_i, where each S_i is chosen uniformly from $\{S_1, \ldots, S_P\}$.*

Proof: See Appendix D

Theorem 1 *If $\mu < 2/B$ and the signal satisfies the following persistence of excitation condition: For all k, there exist $K < \infty$, $\alpha_1 > 0$ and $\alpha_2 > 0$ such that*

$$\alpha_1 I < \sum_{i=k}^{k+K} X_i X_i^H < \alpha_2 I \tag{10}$$

then $\overline{V_k}^H \overline{V_k} \to 0$ and $\overline{V_k^H V_k} \to 0$ exponentially fast.

Proof: See Appendix D

Condition (10) is identical to the persistence of excitation condition for standard LMS. Therefore, the sufficient condition for exponential stability of LMS is enough to guarantee exponential stability of SPU-LMS.

3 Analysis of SPU-LMS: Mixing Signals

In this section, we analytically compare the performance of LMS and SPU-LMS in terms of stability and misconvergence when the independent snapshots assumption is invalid. For this we employ the theory developed in [12] and [2]. Even though the theory developed is for the case of real random variables it can easily be adapted to the case of complex circular random variables.

In this section, results for stability and performance for the case of SPU-LMS are developed for describing the performance hit taken when going from LMS to SPU-LMS. One of the important results obtained is that for stability LMS and SPU-LMS have the same necessary and sufficient conditions. The theory used for stability analysis and performance analysis follows along [12] and [2], respectively.

3.1 Stability Analysis

Notations are the same as those used in [12]. $\|A\|_p$ is used to denote the L_p-norm of a random matrix A given as $\|A\|_p \stackrel{\text{def}}{=} \{E\|A\|^p\|^{1/p}$ for $p \geq 1$ where $\|A\| \stackrel{\text{def}}{=} \{\sum_{i,j} |a|_{ij}^2\}^{1/2}$ is the Euclidean norm of the matrix A. Note that in [12], $\|A\| \stackrel{\text{def}}{=} \{\lambda_{max}(AA^H)\}^{1/2}$. Since the two norms are related by a constant the results in [12] could as well have been stated with the definition used here. Our definition is identical to the norm defined in [2].

A process X_k is said to be ϕ-mixing if there is a function $\phi(m)$ such that $\phi(m) \to 0$ as $m \to \infty$ and

$$\sup_{A \in \mathcal{M}_{-\infty}^{k}(X), B \in \mathcal{M}_{k+m}^{\infty}(X)} |P(B|A) - P(B)| \leq \phi(m), \forall m \geq 0, k \in (-\infty, \infty)$$

where $\mathcal{M}_i^j(X)$, $-\infty \leq i \leq j \leq \infty$ is the σ-algebra generated by $\{X_k\}$, $i \leq k \leq j$

For any random matrix sequence $F = \{F_k\}$, define $\mathcal{S}_p(\alpha, \mu^*)$ for $\mu* > 0$ and $0 < \alpha < 1/\mu^*$ by

$$\begin{aligned} \mathcal{S}_p(\alpha, \mu^*) \quad = \quad & \left\{ F : \left\| \prod_{j=i+1}^{k} (I - \mu F_j) \right\|_p \leq K_{\alpha,\mu^*}(F)(1-\mu\alpha)^{k-i} \right. \\ & \left. \forall \mu \in (0, \mu^*], \forall k \geq i \geq 0 \right\} \end{aligned}$$

$\mathcal{S}_p(\alpha, \mu^*)$ is the family of L_p-stable random matrices.

Similarly, the averaged exponentially stable family is defined as $\mathcal{S}(\alpha, \mu^*)$ for $\mu* > 0$ and $0 < \alpha < 1/\mu^*$ by

$$\begin{aligned} \mathcal{S}(\alpha, \mu^*) \quad = \quad & \left\{ F : \left\| \prod_{j=i+1}^{k} (I - \mu E[F_j]) \right\|_p \leq K_{\alpha,\mu^*}(E[F])(1-\mu\alpha)^{k-i} \right. \\ & \left. \forall \mu \in (0, \mu^*], \forall k \geq i \geq 0 \right\}. \end{aligned} \tag{11}$$

We also define $\mathcal{S}_p$ and $\mathcal{S}$ as $\mathcal{S}_p \stackrel{\text{def}}{=} \cup_{\mu^*\in(0,1)} \cup_{\alpha\in(0,1/\mu^*)} \mathcal{S}_p(\alpha,\mu^*)$ and $\mathcal{S} \stackrel{\text{def}}{=} \cup_{\mu^*\in(0,1)} \cup_{\alpha\in(0,1/\mu^*)} \mathcal{S}(\alpha,\mu^*)$.

Let X_k be the input signal vector generated from the following process

$$X_k = \sum_{j=-\infty}^{\infty} A(k,j)\epsilon_{k-j} + \psi_k \tag{12}$$

with $\sum_{j=-\infty}^{\infty} \sup_k \|A(k,j)\| < \infty$. $\{\psi_k\}$ is a d-dimensional deterministic process, and $\{\epsilon_k\}$ is a general m-dimensional ϕ-mixing sequence. The weighting matrices $A(k,j) \in \mathcal{R}^{d\times m}$ are assumed to be deterministic.

Define the index set $S = \{1, 2, \ldots, N\}$. Partition S into P mutually exclusive subsets of equal size, $S_1,\ S_2, \ldots, S_P$. Define $\mathcal{I}_i$ by zeroing out the j^{th} row of the identity matrix I if $j \notin S_i$. Let I_j be a sequence of i.i.d $d \times d$ masking matrices chosen with equal probability from $\mathcal{I}_i$, $i = 1, \ldots, P$.

Then, we have the following theorem which is similar to Theorem 2 in [12].

Theorem 2 *Let X_k be as defined above with $\{\epsilon_k\}$ a ϕ-mixing sequence such that it satisfies for any $n \geq 1$ and any increasing integer sequence $j_1 < j_2 < \ldots < j_n$*

$$E\left[\exp\left(\alpha \sum_{i=1}^{n} \|\epsilon_{j_i}\|^2\right)\right] \leq M \exp(Kn) \tag{13}$$

where α, M, and K are positive constants. Then for any $p \geq 1$, there exist constants $\mu* > 0$, $M > 0$, and $\alpha \in (0,1)$ such that for all $\mu \in (0, \mu*]$ and for all $t \geq k \geq 0$

$$\left[E\left\|\prod_{j=k+1}^{t} (I - \mu I_j X_j X_j^H)\right\|^p\right]^{1/p} \leq M(1-\mu\alpha)^{t-k}$$

if and only if there exists an integer $h > 0$ and a constant $\delta > 0$ such that for all $k \geq 0$

$$\sum_{i=k+1}^{k+h} E[X_i X_i^H] \geq \delta I. \tag{14}$$

Proof: The proof is just a slightly modified version of the proof of Theorem 2 derived in Section IV of [12, pp. 766-769]. The modification takes into account that F_k in [12] is $F_k = X_k X_k^H$ whereas it is $F_k = I_k X_k X_k^H$ in the present context. □

Note that the LMS algorithm has the same necessary and sufficient conditions for convergence (Theorem 2 in [12]). Therefore, the necessary and sufficient conditions for convergence of SPU-LMS are same as that of LMS. Also note that, Theorem 2 in [12] can be stated as a corollary to Theorem 2 by setting $I_j = I$ for all j.

3.2 Performance Analysis

For performance analysis, we assume that

$$d_k = X_k^H W_{opt,k} + n_k$$

$W_{opt,k}$ varies as follows $W_{opt,k+1} - W_{opt,k} = w_{k+1}$, where w_{k+1} is the lag noise. Then for LMS we can write the evolution equation for the tracking error $V_k \stackrel{\text{def}}{=} W_k - W_{opt,k}$ as

$$V_{k+1} = (I - \mu X_k X_k^H)V_k + \mu X_k n_k - w_{k+1}$$

and for SPU-LMS the corresponding equation can be written as

$$V_{k+1} = (I - \mu I_k X_k X_k^H)V_k + \mu X_k n_k - w_{k+1}$$

Now, V_{k+1} can be decomposed [2] as $V_{k+1} = {}^uV_k + \mu^n V_k + {}^wV_k$ where

$$\begin{aligned} {}^uV_{k+1} &= (I - \mu P_k X_k X_k^H)^u V_k, \quad {}^uV_0 = V_0 = -W_{opt,0} \\ {}^nV_{k+1} &= (I - \mu P_k X_k X_k^H)^n V_k + P_k X_k n_k, \quad {}^nV_0 = 0 \\ {}^wV_{k+1} &= (I - \mu P_k X_k X_k^H)^w V_k - w_{k+1}, \quad {}^nV_0 = 0 \end{aligned}$$

where $P_k = I$ for LMS and $P_k = I_k$ for SPU-LMS. $\{{}^uV_k\}$ denotes the unforced term, reflecting the way the successive estimates of the filter coefficients forget the initial conditions. $\{{}^nV_k\}$ accounts for the errors introduced by the measurement noise, n_k and $\{{}^vV_k\}$ accounts for the errors associated with the lag-noise $\{w_k\}$.

In general ${}^nV_k$ and ${}^wV_k$ obey the following inhomogeneous equation

$$\delta_{k+1} = (I - \mu F_k)\delta_k + \xi_k, \quad \delta_0 = 0$$

δ_k can be represent by a set of recursive equations as follows

$$\delta_k = J_k^{(0)} + J_k^{(1)} + \ldots + J_k^{(n)} + H_k^{(n)}$$

where the processes $J_k^{(r)}, 0 \le r < n$ and $H_k^{(n)}$ are described by

$$\begin{aligned} J_{k+1}^{(0)} &= (I - \mu \bar{F}_k)J_k^{(0)} + \xi_k; J_0^{(0)} = 0 \\ J_{k+1}^{(r)} &= (I - \mu \bar{F}_k)J_k^{(r)} + \mu Z_k J_k^{(r-1)}; \quad J_k^{(r)} = 0, 0 \le k < r \\ H_{k+1}^{(n)} &= (I - \mu F_k)H_k^{(n)} + \mu Z_k J_k^{(n)}; \quad H_k^{(n)} = 0, 0 \le k < n \end{aligned}$$

where $Z_k = F_k - \bar{F}_k$ and $\bar{F}_k$ is an appropriate deterministic process, usually chosen as $\bar{F}_k = E[F_k]$. In [2] under appropriate conditions it was shown that there exists some constant $C < \infty$ and $\mu_0 > 0$ such that for all $0 < \mu \le \mu_0$, we have

$$\sup_{k \ge 0} \|H_k^{(n)}\|_p \le C\mu^{n/2}.$$

Now, we modify the definition of weak dependence as given in [2] for circular complex random variables. The theory developed in [2] can be easily adapted for circular random variables using this definition. Let

$q \geq 1$ and $X = \{X_n\}_{n\geq 0}$ be a $(l \times 1)$ matrix valued process. Let $\beta = (\beta(r))_{r\in N}$ be a sequence of positive numbers decreasing to zero at infinity. The complex process $X = \{X_n\}_{n\geq 0}$ is said to be (δ, q)-weak dependent if there exist finite constants $C = \{C_1, \ldots, C_q\}$, such that for any $1 \leq m < s \leq q$ and m-tuple $k_1, \ldots, k_m$ and any $(s-m)$-tuple $k_{m+1}, \ldots, k_s$, with $k_1 \leq \ldots \leq k_m < k_m + r \leq k_{m+1} \leq \ldots \leq k_s$, it holds that

$$\begin{aligned}\sup_{1\leq i_1,\ldots,i_s\leq l, f_{k_1,i_1}, f_{k_2,i_2}\cdots f_{k_m,i_m}} & \Big|\mathrm{cov}\Big(f_{k_1,i_1}(\tilde{X}_{k_1,i_1})\cdot\ldots\cdot f_{k_m,i_m}(\tilde{X}_{k_m,i_m}), \\ & f_{k_{m+1},i_{m+1}}(\tilde{X}_{k_{m+1},i_{m+1}})\cdot\ldots\cdot f_{k_s,i_s}(\tilde{X}_{k_s,i_s})\Big)\Big| \quad \leq C_s\beta(r)\end{aligned}$$

where $\tilde{X}_{n,i}$ denotes the i-th component of $X_n - E(X_n)$ and the set of functions $f_{n,i}()$ that the sup is being taken over are given by $f_{n,i}(\tilde{X}_{n,i}) = \tilde{X}_{n,i}$ and $f_{n,i}(\tilde{X}_{n,i}) = \tilde{X}^*_{n,i}$.

Define $\mathcal{N}(p)$ from [2] as follows

$$\begin{aligned}\mathcal{N}(p) = \quad & \Big\{\epsilon : \Big\|\textstyle\sum_{k=s}^t D_k\epsilon_k\Big\|_p \leq \rho_p(\epsilon)\left(\sum_{k=s}^t |D_k|^2\right)^{1/2} \ \forall 0 \leq s \leq t \\ & \text{and } \forall D = \{D_k\}_{k\in N}(q \times l) \text{ deterministic matrices } \}\end{aligned}$$

where $\rho_p(\epsilon)$ is a constant depending only on the process ϵ and the number p.

F_k can be written as $F_k = P_k X_k X_k^H$ where $P_k = I$ for LMS and $P_k = I_k$ for SPU-LMS. It is assumed that the following hold true for F_k. For some $r, q \in N$, $\mu_0 > 0$ and $0 < \alpha < 1/\mu_0$

- **F1**(r, α, μ_0): $\{F_k\}_{k\geq 0}$ is in $\mathcal{S}(r, \alpha, \mu_0)$ that is $\{F_k\}$ is L_r-exponentially stable.
- **F2**(α, μ_0): $\{E[F_k]\}_{k\geq 0}$ is in $\mathcal{S}(\alpha, \mu_0)$, that is $\{E[F_k]\}_{k\geq 0}$ is averaged exponentially stable.

Conditions **F3** and **F4** stated below are trivially satisfied for $P_k = I$ and $P_k = I_k$.

- **F3**(q, μ_0): $\sup_{k\in N}\sup_{\mu\in(0,\mu_0]}\|P_k\|_q < \infty$ and $\sup_{k\in N}\sup_{\mu\in(0,\mu_0]}|E[P_k]| < \infty$
- **F4**(q, μ_0): $\sup_{k\in N}\sup_{\mu\in(0,\mu_0]}\mu^{-1/2}\|P_k - E[P_k]\|_q < \infty$

The excitation sequence $\xi = \{\xi_k\|_{k\geq 0}$ [2] is assumed to be decomposed as $\xi_k = M_k\epsilon_k$ where the processes $M = \{M_k\}_{k\geq 0}$ is a $d \times l$ matrix valued process and $\epsilon = \{\epsilon_k\}_{k\geq 0}$ is a $(l \times 1)$ vector-valued process that verifies the following assumptions

- **EXC1**: $\{M_k\}_{k\in Z}$ is $\mathcal{M}_0^k(X)$-adapted[4] and $\mathcal{M}_0^k(\epsilon)$ and $\mathcal{M}_0^k(X)$ are independent.
- **EXC2**(r, μ_0): $\sup_{\mu\in(0,\mu_0]}\sup_{k\geq 0}\|M_k\|_r < \infty$, $(r > 0, \mu_0 > 0)$
- **EXC3**(p, μ_0): $\epsilon = \{\epsilon_k\}_{k\in N}$ belongs to $\mathcal{N}(p)$, $(p > 0, \mu_0 > 0)$

[4] A sequence of random variables, X_i is called adapted with respect to a sequence of σ-fields $\mathcal{F}_i$ if X_i is $\mathcal{F}_i$ measurable [4].

The following theorems from [2] are relevant.

Theorem 3 (Theorem 1 in [2]) *Let $n \in N$ and let $q \geq p \geq 2$. Assume* **EXC1**, **EXC2**$(pq/(q-p), \mu_0)$ *and* **EXC3**(p, μ_0). *For $a, b, \alpha > 0$, $a^{-1} + b^{-1} = 1$, and some $\mu_0 > 0$, assume in addition* **F2**(α, μ_0), **F4**(aqn, μ_0) *and*

- $\{G_k\}_{k\geq 0}$ *is $(\beta, (q+2)n)$ weakly dependent and* $\sum (r+1)^{((q+2)n/2)-1}\beta(r) < \infty$
- $\sup_{k\geq 0} \|G_k\|_{bqn} < \infty$

Then, there exists a constant $K < \infty$ (depending on $\beta(k)$, $k \geq 0$ and on the numerical constants $p, q, n, q, b, \mu_0, \alpha$ but not otherwise on $\{X_k\}$, $\{\epsilon_k\}$ or on μ), such that for all $0 < \mu \leq \mu_0$, for all $0 \leq r \leq n$

$$\sup_{s\geq 1} \|J_s^{(r)}\|_p \leq K \rho_p(\epsilon) \sup_{k\geq 0} \|M_k\|_{pq/(q-p)} \mu^{(r-1)/2}.$$

Theorem 4 (Theorem 2 in [2]) *Let $p \geq 2$ and let $a, b, c > 0$ such that $1/a + 1/b + 1/c = 1/p$. Let $n \in N$. Assume* **F1**(a, α, μ_0) *and*

- $\sup_{s\geq 0} \|Z_s\|_b < \infty$
- $\sup_{s\geq 0} \|J_s^{(n+1)}\|_c < \infty$

Then there exists a constant $K' < \infty$ (depending on the numerical constants a, b, c, α, μ_0, n but not on the process $\{\epsilon_k\}$ or on the stepsize parameter μ), such that for all $0 < \mu \leq \mu_0$,

$$\sup_{s\geq 0} \|H_s^{(n)}\|_p \leq K' \sup_{s\geq 0} \|J_s^{(n+1)}\|_c.$$

We next show that if LMS satisfies the assumptions above (assumptions in section 3.2 in [2]) then so does SPU-LMS. Conditions **F1** and **F2** follow directly from Theorem 2. It is easy to see that **F3** and **F4** hold easily for LMS and SPU-LMS.

Lemma 2 *The constant K in Theorem 3 calculated for LMS can also be used for SPU-LMS.*

Proof: Here all that is needed to be shown is that if LMS satisfies the conditions **(EXC1)**, **(EXC2)** and **(EXC3)** then so does SPU-LMS. Moreover, the upper bounds on the norms for LMS are also upper bounds for SPU-LMS. That easily follows because $M_k^{LMS} = X_k$ whereas $M_k^{SPU-LMS} = I_k X_k$ and $\|I_k\| \leq 1$ for any norm $\|\cdot\|$. $\square$

Lemma 3 *The constant K' in Theorem 4 calculated for LMS can also be used for SPU-LMS.*

Proof: First we show that if for LMS $\sup_{s\geq 0}\|Z_s\|_b<\infty$ then so it is for SPU-LMS. First, note that for LMS we can write $Z_s^{LMS}=X_sX_s^H-E[X_sX_s^H]$ whereas for SPU-LMS

$$\begin{aligned} Z_s^{SPU-LMS} &= I_sX_sX_s^H-\frac{1}{P}E[X_sX_s^H] \\ &= I_sX_sX_s^H-I_sE[X_sX_s^H]+(I_s-\frac{1}{P}I)E[X_sX_s^H] \end{aligned}$$

That means $\|Z_s^{SPU-LMS}\|_b\leq\|I_s\|_b\|Z_s^{LMS}\|_b+\|I_s-\frac{1}{P}I\|_b\|E[X_sX_s^H]\|_b$. Therefore, since $\sup_{s\geq 0}\|_bE[X_sX_s^H]\|_b<\infty$ and $\sup_{s\geq 0}\|Z_s^{LMS}\|_b<\infty$ we have

$$\sup_s\|Z_s^{SPU-LMS}\|_b<\infty.$$

Since all conditions for Theorem 2 have been satisfied by SPU-LMS in a similar manner the constant obtained is also the same. $\square$

The results in this section are an extension of analysis in [2] to SPU-LMS. The results enable us to predict the steady state behaviour of SPU-LMS without the standard independent snapshots assumption employed in Section 2.1. Moreoever, the two lemmas in this section state that the error terms for LMS and SPU-LMS are bounded above by the same constants. These results are very useful for comparison of steady state errors of SPU-LMS and LMS in the sense that the error terms are of the same magnitude. This will become evident in Section 4.2 and Section 4.3 where we compare the steady state performance of the two algorithms for two different scenarios.

4 Discussion and Examples

It is useful to compare the performance of the new algorithm to those of the existing algorithms namely the peridic Partial Update LMS Algorithm (P-LMS) and the sequential Partial Update LMS Algorithm (S-LMS).

For P-LMS, the update equation can be written as follows

$$W_{k+P}=W_k+\mu e_k^*X_k$$

For the Sequential LMS algorithm the update equation is same as (2) except that the choice of I_i is no longer random. The sequence of I_i as k progresses is pre-determined and fixed.

For the P-LMS algorithm, using the method of analysis described in [13] we conclude that the conditions for convergence are identical to standard LMS. That is (7) holds also for P-LMS. Also, the misadjustment factor remains the same. The only difference between LMS and P-LMS is that the measure J for P-LMS is P times that of LMS. Therefore, we see that the behavior of SPU-LMS and P-LMS algorithms is very similar for stationary signals.

The difference between P-LMS and SPU-LMS becomes evident from the persistence of excitation condition for deterministic signals. From [7] we conclude that the persistence of excitation condition for P-LMS is stricter than that for SPU-LMS. For mixing signals, the persistence of excitation condition can similarly be shown to be stricter than that of SPU-LMS. In fact, in the next section we construct signals for which P-LMS is guaranteed not to converge whereas SPU-LMS will converge.

The convergence of Sequential LMS algorithm has been analyzed using the small μ assumption in [7]. Theoretical results for this algorithm without this assumption is an open problem. [11] analyses this algorithm for its convergence in the mean properties, but so far mean square convergence for stationary, non-stationary and mixing signals is still untacked. However, we show through simulation examples that this algorithm diverges for certain signals and therefore should be employed with caution.

4.1 Numerical Examples

In the first two examples, we simulated an m-element uniform linear antenna array operating in a multiple signal environment. Let A_i denote the response of the array to the i^{th} plane wave signal:

$$A_i = [e^{-\mathbf{j}(\frac{m}{2}-\tilde{m})\omega_i} \; e^{-\mathbf{j}(\frac{m}{2}-1-\tilde{m})\omega_i} \; \ldots \; e^{\mathbf{j}(\frac{m}{2}-1-\tilde{m})\omega_i} \; e^{\mathbf{j}(\frac{m}{2}-\tilde{m})\omega_i}]^T$$

where $\tilde{m} = (m+1)/2$ and $\omega_i = \frac{2\pi D \sin\theta_i}{\lambda}$, $i = 1, \ldots, M$. θ_i is the broadside angle of the i^{th} signal, D is the inter-element spacing between the antenna elements and λ is the common wavelength of the narrowband signals in the same units as D and $\frac{2\pi D}{\lambda} = 2$. The array output at the k^{th} snapshot is given by $X_k = \sum_{i=1}^{M} A_i s_{k,i} + n_k$ where M denotes the number of signals, the sequence $\{s_{k,i}\}$ the amplitude of the i^{th} signal and n_k the noise present at the array output at the k^{th} snapshot. The objective, in both the examples, is to maximize the SNR at the output of the beamformer. Since the signal amplitudes are random the objective translates to obtaining the best estimate of $s_{k,1}$, the amplitude of the desired signal, in the MMSE sense. Therefore, the desired signal is chosen as $d_k = s_{k,1}$.

Example 1: In the first example (Figure 1), the array has 4 elements and a single planar waveform with amplitude, $s_{k,1}$ propagates across the array from direction angle, $\theta_1 = \frac{\pi}{2}$. The amplitude sequence $\{s_{k,1}\}$ is a binary phase shifty keying (BPSK) signal with period four taking values on $\{-1, 1\}$ with equal probability. The additive noise n_k is circular Gaussian with variance 0.25 and mean 0. In all the simulations for SPU-LMS, P-LMS, and S-LMS the number of subsets for partial updating, P was chosen to be 4. It can be easily determined from (7) that for Gaussian and independent signals the necessary and sufficient condition for convergence of LMS and SPU-LMS is $\mu < 0.67$. Figure 2 shows representative trajectories of the empirical mean-squared error for LMS, SPU-LMS, P-LMS and S-LMS algorithms averaged over 100 trials for $\mu = 0.6$ and $\mu = 1.0$. All algorithms were found to be stable for the BPSK signals even for μ values greater than 0.67. It was only as μ approached 1 that divergent behavior was observed. As expected, LMS and SPU-LMS

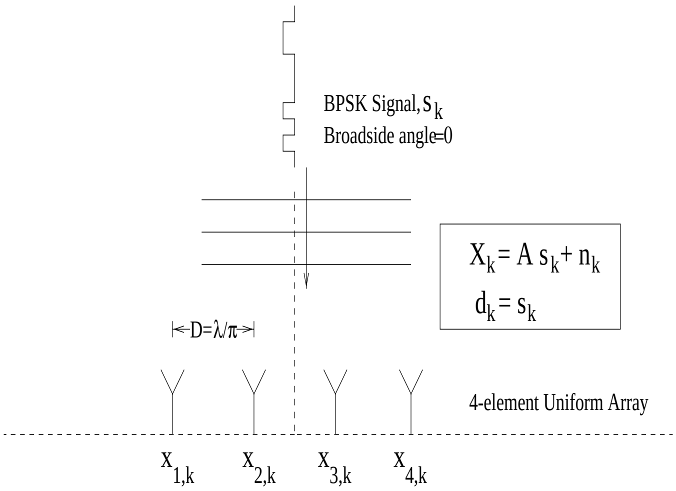

Figure 1: Signal Scenario for Example 1

were observed to have similar μ regions of convergence. It is also clear from Figure 2, that as, expected SPU-LMS, P-LMS, and S-LMS take roughly 4 times longer to converge than LMS.

Example 2: In the second example, we consider an 8-element uniform linear antenna array with one signal of interest propagating at angle θ_1 and 3 interferers propagating at angles θ_i, $i = 2, 3, 4$. The array noise n_k is again mean 0 circular Gaussian but with variance 0.001. We generated signals, such that $s_{k,1}$ is stationary and $s_{k,i}$, $i = 2, 3, 4$ are cyclostationary with period four, which make both S-LMS and P-LMS non-convergent. All the signals were chosen to be independent from time instant to time instant. First, we found signals for which S-LMS doesn't converge by the following procedure. Make the small μ approximation $I - \mu \sum_{i=1}^{P} I_i E[X_{k+i} X_{k+i}^H]$ to the transition matrix $\prod_{i=1}^{P} (I - \mu I_i E[X_{k+i} X_{k+i}])$ and generate sequences $s_{k,i}$, $i = 1, 2, 3, 4$ such that $\sum_{i=1}^{P} I_i E[X_{k+i} X_{k+i}^H]$ has roots in the negative left half plane. This ensures that $I - \mu \sum_{i=1}^{P} I_i E[X_{k+i} X_{k+i}^H]$ has roots outside the unit circle. The sequences found in this manner were then verified to cause the roots to lie outside the unit circle for all μ. One such set of signals found was: $s_{k,1}$ is equal to a BPSK signal with period one taking values in $\{-1, 1\}$ with equal probability. The interferers, $s_{k,i}$, $i = 2, 3, 4$ are cyclostationary BPSK type signals taking values in $\{-1, 1\}$ with the restriction that $s_{k,2} = 0$ if $k \% 4 \neq 1$, $s_{k,3} = 0$ if $k \% 4 \neq 2$ and $s_{k,4} = 0$ if $k \% 4 \neq 3$. Here $a \% b$ stands for a modulo b. θ_i, $i = 1, 2, 3, 4$ are chosen such that $\theta_1 = 1.0388$, $\theta_2 = 0.0737$, $\theta_3 = 1.0750$ and $\theta_4 = 1.1410$. These signals render the S-LMS algorithm unstable for all μ.

The P-LMS algorithm also fails to converge for the signal set described above irrespective of μ and the choice of θ_1, θ_2, θ_3, and θ_4. Since P-LMS updates the coefficients every 4^{th} iteration it sees at most one of the three interfering signals throughout all its updates and hence can place a null at atmost one signal incidence angle θ_i. Figure 4 shows the envelopes of the e_k^2 trajectories of S-LMS and P-LMS for the signals given above with the representative value $\mu = 0.03$. As can be seen P-LMS fails to converge whereas S-LMS shows divergent behavior. SPU-LMS and LMS were observed to converge for the signal set described above

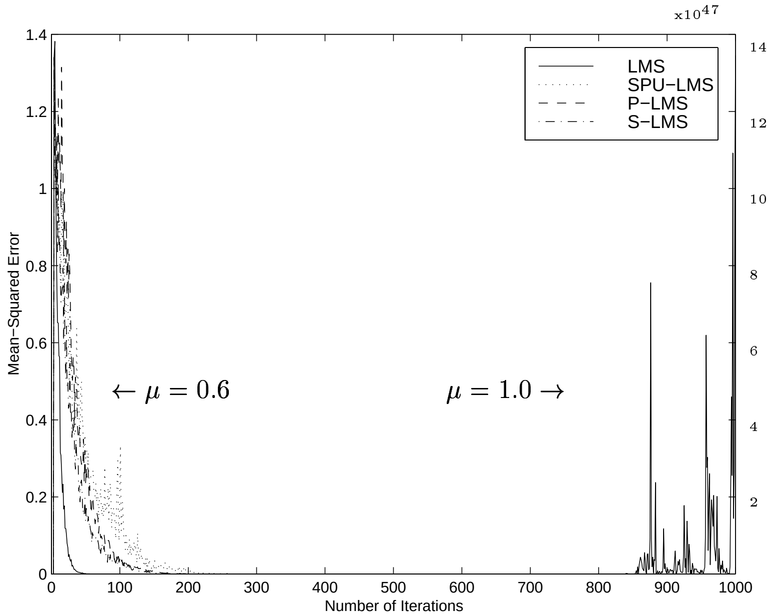

Figure 2: Trajectories of MSE for Example 1

when $\mu = 0.03$.

Example 3: In the third example, we consider a 4-tap filter ($N = 4$) with a time series input, that is $X_k = [x_k \ x_{k-1} \ x_{k-2} \ x_{k-3}]$. The input, the filter coefficients and the desired output are all real valued. x_k is given by $s_k + n_k$ where s_k is a BPSK signal with a symbol duration of 4 time intervals and n_k is a zero mean Gaussian noise with variance 0.01. The desired output d_k, is given by $d_k = W_{opt}^T S_k$ where $S_k = [s_k \ s_{k-1} \ s_{k-2} \ s_{k-3}]$ and $W_{opt} = [1 \ 2 \ 3 \ 4]^T$. The update is such that one coefficient is updated per iteration, i.e. $P = 4$. For this signal it was verified that S-LMS diverges for all values of μ. Figures 5 and 6 show the trajectory of mean-squared error for LMS, SPU-LMS, P-LMS and S-LMS for a representative value of $\mu = 0.05$, respectively. As can be seen P-LMS and S-LMS fail to converge whereas LMS and SPU-LMS do.

4.2 I.i.d Gaussian Input Sequence

In this section, we assume that $X_k = [x_k \ x_{k-1} \ \ldots \ x_{k-N+1}]^T$ where N is the length of the vector X_k. $\{x_k\}$ is a sequence of zero mean i.i.d Gaussian random variables. We assume that $w_k = 0$ for all $k \geq 0$. In that case

$$V_{k+1} = (I - \mu P_k X_k X_k^H) V_k + X_k n_k \quad V_0 = -W_{opt,0} = W_{opt}$$

where for LMS we have $P_k = I$ and $P_k = I_k$ in case of SPU-LMS. We assume n_k is a white i.i.d. Gaussian noise with variance σ_v^2. We see that since the conditions (13) and (14) are satisfied for theorem 2 both

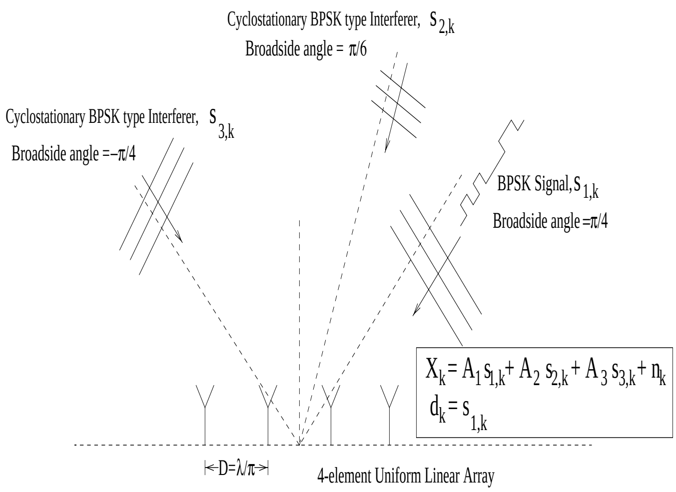

Figure 3: Signal Scenario for Example 2

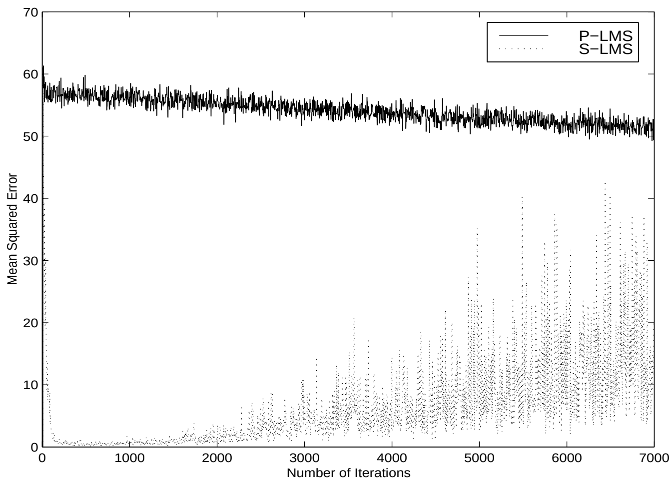

Figure 4: Trajectories of MSE for Example 2

LMS and SPU-LMS are exponentially stable. In fact both have the same α exponent of decay. Therefore, conditions **F1** and **F2** are satisfied.

We rewrite $V_k = J_k^{(0)} + J_k^{(1)} + J_k^{(2)} + H_k^{(2)}$. Choosing $\bar{F}_k = E[F_k]$ we have $E[P_k X_k X_k^H] = \sigma^2 I$ in the case of LMS and $\frac{1}{P}\sigma^2 I$ in the case of SPU-LMS. By Theorems 3 and 4 and Lemmas 2 and 3 we can upperbound both $|J_k^{(2)}|$ and $|H_k^{(2)}|$ by exactly the same constants for LMS and SPU-LMS. In particular, there exists some constant $C < \infty$ such that for all $\mu \in (0, \mu_0]$, we have

$$\sup_{t\geq 0}\left|E[J_t^{(1)}(J_t^{(2)} + H_t^{(2)})^H]\right| \leq C\|X_0\|_{r(r+\delta)/\delta}\rho_r^2(v)\mu^{1/2}$$

$$\sup_{t\geq 0}\left|E[J_t^{(0)} H_t^{(2)}]\right| \leq C\rho_r(v)\|X_0\|_{r(r+\delta)/\delta}\mu^{1/2}.$$

Next, for LMS we concentrate on

$$J_{k+1}^{(0)} = (1-\mu\sigma^2)J_k^{(0)} + X_k n_k$$

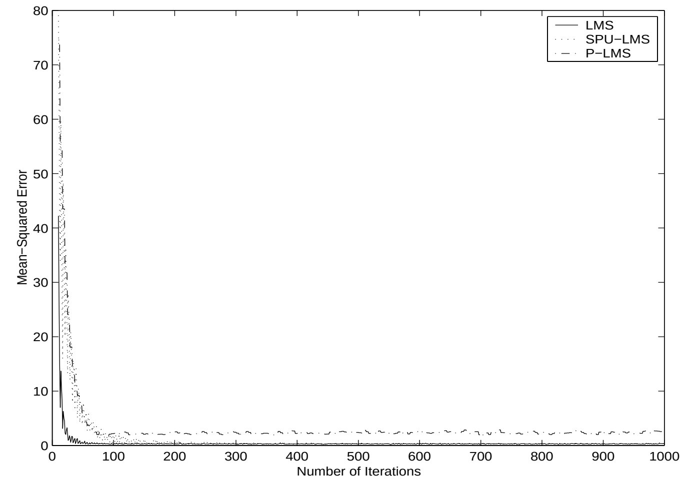

Figure 5: Trajectories of MSE for LMS, SPU-LMS and P-LMS for Example 3

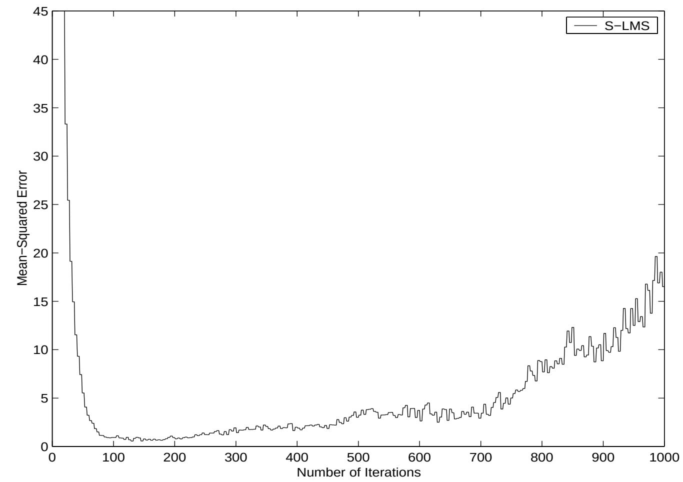

Figure 6: Trajectories of MSE for S-LMS for Example 3

$$J_{k+1}^{(1)} = (1-\mu\sigma^2)J_k^{(1)} + \mu(\sigma^2 I - X_k X_k^H)J_k^{(0)}$$

and for SPU-LMS we concentrate on

$$\begin{aligned} J_{k+1}^{(0)} &= (1-\frac{\mu}{P}\sigma^2)J_k^{(0)} + I_k X_k n_k \\ J_{k+1}^{(1)} &= (1-\frac{\mu}{P}\sigma^2)J_k^{(1)} + \mu(\frac{\sigma^2}{P}I - I_k X_k X_k^H)J_k^{(0)}. \end{aligned}$$

Solving (see Appendix E), we obtain for LMS

$$\begin{aligned} \lim_{k\to\infty} E[J_k^{(0)}(J_k^{(0)})^H] &= \frac{\sigma_v^2}{\mu(2-\mu\sigma^2)}I \\ \lim_{k\to\infty} E[J_k^{(0)}(J_k^{(1)})^H] &= 0 \\ \lim_{k\to\infty} E[J_k^{(0)}(J_k^{(2)})^H] &= 0 \\ \lim_{k\to\infty} E[J_k^{(1)}(J_k^{(1)})^H] &= \frac{N\sigma^2\sigma_v^2}{(2-\mu\sigma^2)^2}I \end{aligned}$$

$$= \frac{N\sigma^2\sigma_v^2}{4} I + O(\mu) I$$

which yields $\lim_{k\to\infty} E[V_k V_k^H] = \frac{\sigma_v^2}{2\mu} I + \frac{N\sigma^2\sigma_v^2}{4} I + O(\mu^{1/2}) I$ and for SPU-LMS we obtain

$$\begin{aligned}
\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(0)})^H] &= \frac{\sigma_v^2}{\mu(2-\frac{\mu}{P}\sigma^2)} I \\
\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(1)})^H] &= 0 \\
\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(2)})^H] &= 0 \\
\lim_{k\to\infty} E[J_k^{(1)}(J_k^{(1)})^H] &= \frac{\frac{(N+1)P-1}{P}\sigma^2\sigma_v^2}{(2-\frac{\mu}{P}\sigma^2)^2} I \\
&= \frac{\frac{(N+1)P-1}{P}\sigma^2\sigma_v^2}{4} I + O(\mu) I
\end{aligned}$$

which yields $\lim_{k\to\infty} E[V_k V_k^H] = \frac{\sigma_v^2}{2\mu} I + \frac{\frac{(N+1)P-1}{P}\sigma^2\sigma_v^2}{4} I + O(\mu^{1/2}) I$. Therefore, we see that SPU-LMS is marginally worse than LMS in terms of misadjustment.

4.3 Temporally Correlated Spatially Uncorrelated Array Output

In this section we consider X_k given by

$$X_k = \kappa X_{k-1} + \sqrt{1-\kappa^2} U_k$$

where U_k is a vector of circular Gaussian random variables with unit variance. Similar to section 4.2, we rewrite $V_k = J_k^{(0)} + J_k^{(1)} + J_k^{(2)} + H_k^{(2)}$. Since, we have chosen $\bar{F}_k = E[F_k]$ we have $E[P_k X_k X_k^H] = I$ in the case of LMS and $\frac{1}{P} I$ in the case of SPU-LMS. Again, conditions **F1** and **F2** are satisfied because of Theorem 2. By [2] and Lemmas 1 and 2 we can upperbound both $J_k^{(2)}$ and $H_k^{(2)}$ by exactly the same constants for LMS and SPU-LMS. By Theorems 3 and 4 and Lemmas 2 and 3 we have that there exists some constant $C < \infty$ such that for all $\mu \in (0, \mu_0]$, we have

$$\begin{aligned}
\sup_{t\geq 0} \left| E[J_t^{(1)}(J_t^{(2)} + H_t^{(2)})^H] \right| &\leq C \|X_0\|_{r(r+\delta)/\delta} \rho_r^2(v) \mu^{1/2} \\
\sup_{t\geq 0} \left| E[J_t^{(0)} H_t^{(2)}] \right| &\leq C \rho_r(v) \|X_0\|_{r(r+\delta)/\delta} \mu^{1/2}.
\end{aligned}$$

Next, for LMS we concentrate on

$$\begin{aligned}
J_{k+1}^{(0)} &= (1-\mu) J_k^{(0)} + X_k n_k \\
J_{k+1}^{(1)} &= (1-\mu) J_k^{(1)} + \mu(I - X_k X_k^H) J_k^{(0)}
\end{aligned}$$

and for SPU-LMS we concentrate on

$$\begin{aligned}
J_{k+1}^{(0)} &= (1-\frac{\mu}{P}) J_k^{(0)} + I_k X_k n_k \\
J_{k+1}^{(1)} &= (1-\frac{\mu}{P}) J_k^{(1)} + \mu(\frac{1}{P} I - I_k X_k X_k^H) J_k^{(0)}.
\end{aligned}$$

Solving (see Appendix F), we obtain for LMS

$$\begin{aligned}
\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(0)})^H] &= \frac{\sigma_v^2}{\mu(2-\mu)} I \\
\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(1)})^H] &= -\frac{\kappa^2\sigma_v^2 N}{2(1-\kappa^2)} I + O(\mu) I \\
\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(2)})^H] &= \frac{\kappa^2\sigma_v^2 N}{4(1-\kappa^2)} I + O(\mu) I \\
\lim_{k\to\infty} E[J_k^{(1)}(J_k^{(1)})^H] &= \frac{(1+\kappa^2)\sigma_v^2 N}{4(1-\kappa^2)} I + O(\mu) I
\end{aligned}$$

which leads to $\lim_{k\to\infty} E[V_k V_k^H] = \frac{\sigma_v^2}{2\mu} I + \frac{N\sigma_v^2}{4} I + O(\mu^{1/2}) I$ and for SPU-LMS we obtain

$$\begin{aligned}
\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(0)})^H] &= \frac{\sigma_v^2}{\mu(2-\frac{\mu}{P})} I \\
\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(1)})^H] &= -\frac{\kappa^2\sigma_v^2 N}{2(1-\kappa^2)P} I + O(\mu) I \\
\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(2)})^H] &= \frac{\kappa^2\sigma_v^2 N}{4(1-\kappa^2)P} I + O(\mu) I \\
\lim_{k\to\infty} E[J_k^{(1)}(J_k^{(1)})^H] &= \frac{\sigma_v^2}{4}[\frac{N}{P}\frac{1+\kappa^2}{1-\kappa^2} + (N+1)\frac{P-1}{P}] I + O(\mu) I
\end{aligned}$$

which leads to $\lim_{k\to\infty} E[V_k V_k^H] = \frac{\sigma_v^2}{2\mu} I + \frac{\sigma^2}{4}[N+1-\frac{1}{P}] I + O(\mu^{1/2}) I$. Again, SPU-LMS is marginally worse than LMS in terms of misadjustment.

5 Conclusion

We have proposed a new algorithm based on randomization of filter coefficient subsets for partial updating of filter coefficients. The conditions on step-size for convergence-in-mean and mean-square were shown to be equivalent to those of standard LMS. It was verified by theory and by simulation that LMS and SPU-LMS have similar regions of convergence. We also have shown that the Stochastic Partial Update LMS algorithm has the same performance as the Periodic LMS algorithm for stationary signals but, can have superior performance for some non-stationary signals. We also demonstrated that the randomization of filter coefficient updates does not increase the final steady state error as compared to the regular LMS algorithm.

The idea of random choice of subsets proposed in the paper can be extended to include arbitrary subsets of size $\frac{N}{P}$ and not just subsets from a particular partition. No special advantage is immediately evident from this extension though. In future work, we will analyze the algorithm in the time-series setting, that is, without the independent snapshots assumption.

The Max PU-LMS described in Section 1 is similar SPU-LMS in the sense that the coefficient subset chosen to be updated at an iteration are also random. However, update equations (5) and (6) are not valid for Max PU-LMS as we can no longer assume that X_k and I_i are independent since the coefficients to be updated in an iteration explicitly depend on X_k. The analysis of this algorithm is an open problem.

APPENDICES

A Derivation of Stability Condition (7)

We will follow the Z-transform method of [13]. Let $\tilde{\xi}(z)$ donate the Z-transform of ξ_k and $\tilde{G}_i(z)$ donate the Z-transform of the i^{th} component of G_k. Then we have the following

$$\begin{aligned}
\tilde{\xi}(z) &= \xi_{min}\frac{1}{1-z^{-1}} + \sum_{i=1}^{N}\tilde{G}_i(z) \\
\tilde{G}_i(z) &= (1 - \frac{2\mu}{P}\lambda_i + \frac{2\mu^2}{P}\lambda_i^2)\tilde{G}_i(z) + \frac{\mu^2}{P}\lambda_i^2 z^{-1}\tilde{\xi}(z) + G_i(0)
\end{aligned}$$

which leads to

$$\tilde{\xi}(z) = \frac{\xi_{min}\frac{1}{1-z^{-1}} + \sum_{i=1}^{N}\frac{G_i(0)}{1-z^{-1}(1-\frac{2\mu}{P}\lambda_i+\frac{2\mu^2}{P}\lambda_i^2)}}{1 - \sum_{i=1}^{N}\frac{\frac{\mu^2}{P}\lambda_i^2 z^{-1}}{1-z^{-1}(1-\frac{2\mu}{P}\lambda_i+\frac{2\mu^2}{P}\lambda_i^2)}} \tag{15}$$

and

$$\tilde{G}_i(z) = \frac{1}{D(z)}\frac{\frac{\mu^2}{P}\lambda_i^2 N(z) + G_i(0)D(z)}{1 - z^{-1}(1 - \frac{2\mu}{P}\lambda_i + \frac{2\mu^2}{P}\lambda_i^2)} \tag{16}$$

where $N(z)$ and $D(z)$ denote the numerator and the denominator in (15). Therefore, the condition for stability is that the roots of

$$z - (1 - \frac{2\mu}{P}\lambda_i + \frac{2\mu^2}{P}\lambda_i^2) = 0$$

for $i = 1, \ldots, N$ and

$$\prod_{i=1}^{N}\left[z - (1 - \frac{2\mu}{P}\lambda_i + \frac{2\mu^2}{P}\lambda_i^2)\right] - \sum_{i=1}^{N}\frac{\mu^2}{P}\lambda_i^2\prod_{k\neq i}\left[z - (1 - \frac{2\mu}{P}\lambda_k + \frac{2\mu^2}{P}\lambda_k^2)\right] = 0$$

should lie within the unit circle.

The reader should note that (16) should be used to determine the stability of $\tilde{G}_i(z)$ and not

$$\tilde{G}_i(z) = \frac{\frac{\mu^2}{P}\lambda_i^2 z^{-1}\left[\xi_{min}\frac{1}{1-z^{-1}} + \sum_{k\neq i}\tilde{G}_k(z)\right] + G_i(0)}{1 - z^{-1}(1 - \frac{2\mu}{P}\lambda_i + \frac{3\mu^2}{P}\lambda_i^2)}$$

that was used in [13].

Following the rest of the procedure as outlined in [13] exactly, we obtain the conditions for stability to be (7).

B Derivation of expression (9)

Here we follow the procedure in [8]. Assuming G_k converges we have the expression for G_∞ to be

$$G_\infty = P\ [2\mu\Lambda - 2\mu^2\Lambda - \mu^2\Lambda^2\mathbf{1}\mathbf{1}^T]^{-1}\frac{\mu^2}{P}\Lambda^2\mathbf{1}\xi_{min}.$$

Then we have

$$G_{k+1} - G_\infty = F(G_k - G_\infty)$$

where $F = I - \frac{2\mu}{P}\Lambda + \frac{2\mu^2}{P}\Lambda + \frac{\mu^2}{P}\Lambda^2\mathbf{1}\mathbf{1}^T$. Since $\xi_k = \mathrm{tr}\{G_k\}$ we have

$$\begin{aligned}\sum_{k=0}^{\infty}(\xi_k - \xi_\infty) &= \mathrm{tr}\{\sum_{k=0}^{\infty}(G_k - G_\infty)\} \\ &= \mathrm{tr}\{\sum_{k=0}^{\infty}F^k(G_0 - G_\infty)\} \\ &= \mathrm{tr}\{(I-F)^{-1}(G_0 - G_\infty)\}\end{aligned}$$

from which (9) follows.

C Derivation of the misadjustment factor (8)

Here we follow the approach of [13]. The misadjustment numerator and denominator is defined as $M(\mu) = \frac{\xi_\infty - \xi_{min}}{\xi_{min}}$. Since $\xi_\infty = \lim_{z\to 1}(1-z^{-1})\tilde{\xi}(z)$ and the limits of $(1-z^{-1})\xi(z)$ are finite, we have

$$\xi_\infty = \frac{\lim_{z\to 1}\left[\xi_{min} + (1-z^{-1})\sum_{i=1}^{N}\frac{G_i(0)}{1-z^{-1}(1-\frac{2\mu}{P}\lambda_i+\frac{2\mu^2}{P}\lambda_i^2)}\right]}{\lim_{z\to 1}\left[1 - \sum_{i=1}^{N}\frac{\frac{\mu^2}{P}\lambda_i^2 z^{-1}}{1-z^{-1}(1-\frac{2\mu}{P}\lambda_i+\frac{2\mu^2}{P}\lambda_i^2)}\right]},$$

that is

$$\begin{aligned}\xi_\infty &= \frac{\xi_{min}}{1-\frac{1}{2}\sum_{i=1}^{N}\frac{\mu\lambda_i}{1-\mu\lambda_i}} \\ &= \frac{\xi_{min}}{1-\eta(\mu)},\end{aligned}$$

from which (8) follows.

D Proofs of Lemma 1 and Theorem 1

Proof of Lemma 1: First note that $e_k = -V_k^H X_k$. Next, consider the Lyapunov function $\mathcal{L}_{k+1} = \overline{V_{k+1}^H V_{k+1}}$ where $\overline{\{\cdot\}}$ is as defined in Lemma 1. Averaging the following update equation for $V_{k+1}^H V_{k+1}$

$$V_{k+1}^H V_{k+1} = V_k^H V_k - \mu\mathrm{tr}\{V_k V_k^H X_k X_k^H I_i\} - \mu\mathrm{tr}\{V_k V_k^H I_i X_k X_k^H\} + \mu^2\mathrm{tr}\{V_k V_k^H X_k X_k^H I_i X_k X_k^H\}$$

over all possible choices of S_i, $i = 1, \ldots, P$ we obtain

$$\mathcal{L}_{k+1} = \mathcal{L}_k - \frac{\mu}{P}\mathrm{tr}\{\overline{V_k V_k^H} X_k (2 - \mu X_k X_k^H) X_k^H\}.$$

Since $\sup_k (X_k^H X_k) \leq B < \infty$ the matrix $(2I - \mu X_k X_k^H) - (2I - \mu B I)$ is positive definite. Therefore,

$$\mathcal{L}_{k+1} \leq \mathcal{L}_k - \frac{\mu}{P}(2 - \mu B)\mathrm{tr}\{\overline{V_k V_k^H} X_k X_k^H\}.$$

Since $\mu < 2/B$

$$\mathcal{L}_{k+1} \leq \mathcal{L}_k - \mathrm{tr}\{\overline{V_k V_k^H} X_k X_k^H\}$$

Noting that $\overline{e_k^2} = \mathrm{tr}\{\overline{V_k V_k^H} X_k X_k^H\}$ we obtain

$$\mathcal{L}_{k+1} + \sum_{l=0}^{k} \overline{e_k^2} \leq \mathcal{L}_0$$

since $\mathcal{L}_0 < \infty$ we have $\overline{e_k^2} = O(1/k)$ and $\lim_{k\to\infty} \overline{e_k^2} = 0$ □

Before proving Theorem 1 we need Lemmas 4 and 5. We reproduce the proof of Lemma 4 from [17] using our notation because this enables to understand the proof of Lemma 5 better.

Lemma 4 *[17, Lemma 6.1 p. 143-144] Let X_k satisfy the persistence of excitation condition in Theorem 1. let*

$$\Pi_{k,k+D} = \begin{cases} \prod_{l=k}^{k+D}(I - \frac{\mu}{P} X_l X_l^H) & \text{if } D \geq 0 \\ 1 & \text{if } D < 0 \end{cases}$$

and

$$\mathcal{G}_k = \sum_{l=0}^{K} \Pi_{k,k+l-1}^H X_{k+l} X_{k+l}^H \Pi_{k,k+l-1}$$

where K is as defined in Theorem 1 then $\mathcal{G}_k - \eta I$ is a positive definite matrix for some $\eta > 0$ and $\forall k$.

Proof: Proof is by contradiction. Suppose not then for some vector ω such that $\omega^H \omega = 1$ we have $\omega^H \mathcal{G}_k \omega \leq c^2$ where c is any arbitrary positive number.

Then

$$\begin{aligned} &\sum_{l=0}^{K} \omega^H \Pi_{k,k+l-1}^H X_{k+l} X_{k+l}^H \Pi_{k,k+l-1}\omega \leq c^2 \\ \Rightarrow\quad & \omega^H \Pi_{k,k+l-1}^H X_{k+l} X_{k+l}^H \Pi_{k,k+l-1}\omega \leq c^2 \quad \text{for } 0 \leq l \leq K. \end{aligned}$$

Choosing $l = 0$ we obtain $\omega^H X_k X_k^H \omega \leq c^2$ or $\|\omega^H X_k\| \leq c$.

Choosing $l = 1$ we obtain $\|\omega^H (I - \frac{\mu}{P} X_k X_k^H) X_{k+1}\| \leq c$. Therefore,

$$\begin{aligned} \|\omega^H X_{k+1}\| &\leq \|\omega^H (I - \frac{\mu}{P} X_k X_k^H) X_{k+1}\| + \frac{\mu}{P}\|\omega^H X_k\| \|X_k^H X_{k+1}\| \\ &\leq c + \frac{\mu}{P} B c = c(1 + 2/P). \end{aligned}$$

Choosing $l = 2$ we obtain $\|\omega^H(I - \frac{\mu}{P}X_kX_k^H)(I - \frac{\mu}{P}X_{k+1}X_{k+1}^H)X_{k+2}\| \le c$. Therefore,

$$\begin{aligned}
\|\omega^H X_{k+2}\| &\le \|\omega^H(I - \frac{\mu}{P}X_kX_k^H)(I - \frac{\mu}{P}X_{k+1}X_{k+1}^H)X_{k+2}\| + \frac{\mu}{P}\|\omega^H X_kX_k^H X_{k+2}\| \\
&\quad + \frac{\mu}{P}\|\omega^H X_{k+1}X_{k+1}^H X_{k+2}\| + \frac{\mu^2}{P^2}\|\omega^H X_kX_k^H X_{k+1}X_{k+1}^H X_{k+2}\| \\
&\le O(c).
\end{aligned}$$

Proceeding along similar lines we obtain $\|\omega^H X_{k+l}\| \le Lc$ for $l = 0, \ldots, K$ where L is some constant. This implies $\omega^H \sum_{l=k}^{k+K} X_lX_l^H\omega \le (K+1)L^2c^2$. Since c is arbitrary we obtain that $\omega^H \sum_{l=k}^{k+K} X_lX_l^H\omega < \alpha_1$ which is a contradiction. □

Lemma 5 *Let X_k satisfy the persistence of excitation condition in Theorem 1. let*

$$\mathcal{P}_{k,k+D} = \begin{cases} \prod_{l=k}^{k+D}(I - \mu I_l X_l X_l^H) & \text{if } D \ge 0 \\ 1 & \text{if } D < 0 \end{cases}$$

where I_l is the randomly chosen masking matrix and let

$$\Omega_k = \sum_{l=0}^{K} \overline{\Pi_{k,k+l-1}^H X_{k+l}X_{k+l}^H \Pi_{k,k+l-1}}$$

where K is as defined in Theorem 1 and $\overline{\{\cdot\}}$ is the average over randomly chosen I_l then $\Omega_k - \gamma I$ is a positive definite matrix for some $\gamma > 0$ and $\forall k$.

Proof: Proof is by contradiction. Suppose not then for some vector ω such that $\omega^H\omega = 1$ we have $\omega^H\Omega_k\omega \le c^2$ where c is any arbitrary positive number.

Then

$$\begin{aligned}
&\sum_{l=0}^{K} \omega^H\ \overline{\mathcal{P}_{k,k+l-1}^H X_{k+l}X_{k+l}^H \mathcal{P}_{k,k+l-1}}\ \omega \le c^2 \\
\Rightarrow\ &\omega^H\ \overline{\mathcal{P}_{k,k+l-1}^H X_{k+l}X_{k+l}^H \mathcal{P}_{k,k+l-1}}\ \omega \le c^2 \quad \text{for } 0 \le l \le K.
\end{aligned}$$

Choosing $l = 0$ we obtain $\omega^H X_kX_k^H\omega \le c^2$ or $\|\omega^H X_k\| \le c$.

Choosing $l = 1$ we obtain $\omega^H\ \overline{(I - \mu X_kX_k^H I_k)X_{k+1}X_{k+1}^H(I - \mu I_k X_kX_k^H)}\ \omega \le c^2$. Therefore,

$$\begin{aligned}
&\omega^H X_{k+1}X_{k+1}^H\omega - \frac{\mu}{P}\omega^H X_kX_k^H X_{k+1}X_{k+1}^H\omega - \frac{\mu}{P}\omega^H X_{k+1}X_{k+1}^H X_kX_k^H\omega + \\
&\qquad \frac{\mu^2}{P}\omega^H X_kX_k^H \left[\sum_{i=0}^{P} I_i X_{k+1}X_{k+1}^H I_i\right] X_kX_k^H\omega \le c^2.
\end{aligned}$$

Now

$$\begin{aligned}
\|\omega^H X_kX_k^H X_{k+1}X_{k+1}^H\omega\| &\le \|\omega^H X_k\|\|X_k\|\|X_{k+1}^H X_{k+1}\|\|\omega\| \\
&\le cB^{3/2}
\end{aligned}$$

and

$$\|\omega^H X_k X_k^H \left[\sum_{i=0}^{P} I_i X_{k+1} X_{k+1}^H I_i\right] X_k X_k^H \omega\| \leq c^2 P B^2.$$

Therefore, $\omega^H X_{k+1} X_{k+1}^H \omega = O(c)$ which implies $\|\omega^H X_{k+1}\| = O(c^{1/2})$. Proceeding along the same lines we obtain $\|\omega^H X_{k+1}\| = O(c^{1/L})$ for $l = 0, \ldots, K$ for some constant L. This implies $\omega^H \sum_{l=k}^{k+K} X_l X_l^H \omega = O(c^{2/L})$. Since c is arbitrary we obtain that $\omega^H \sum_{l=k}^{k+K} X_l X_l^H \omega < \alpha_1$ which is a contradiction. □

Now, we are ready to Prove Theorem 1.

Proof of Theorem 1: First, we will prove the convergence of $\overline{V}_k^H \overline{V}_k$. We have $\overline{V}_{k+1} = (I - \frac{\mu}{P} X_k X_k^H)\overline{V}_k$. Proceeding as before, we obtain the following update equation for $\overline{V}_k \overline{V}_k^H$

$$\begin{aligned}
\overline{V}_{k+K+1}^H \overline{V}_{k+K+1} &= \overline{V}_{k+K}^H \overline{V}_{k+K} - 2\frac{\mu}{P}\overline{V}_{k+K}^H X_{k+K} X_{k+K}^H \overline{V}_{k+K} \\
&\quad + \frac{\mu^2}{P^2}\overline{V}_{k+K}^H X_{k+K} X_{k+K}^H X_{k+K} X_{k+K}^H \overline{V}_{k+K} \\
&\leq \overline{V}_{k+K}^H \overline{V}_{k+K} - \frac{\mu}{P}\overline{V}_{k+K}^H X_{k+K} X_{k+K}^H \overline{V}_{k+K}.
\end{aligned}$$

The last step follows from the fact that $\mu < 2/B$. Using the update equation for $\overline{V_k}$ repeatedly, we obtain

$$\overline{V}_{k+K+1}^H \overline{V}_{k+K+1} \quad \leq \quad \overline{V}_k^H \overline{V}_k - \frac{\mu}{P}\overline{V}_k^H \mathcal{G}_k \overline{V}_k.$$

From Lemma 4 we have,

$$\overline{V}_{k+K+1}^H \overline{V}_{k+K+1} \leq (1 - \frac{\mu}{P}\eta)\overline{V}_k^H \overline{V}_k$$

which ensures exponential convergence of $\mathrm{tr}\{\overline{V}_k \overline{V}_k^H\}$.

Next, we prove the convergence of $\overline{V_k^H V_k}$. First, we have the following update equation for $\mathrm{tr}\{\overline{V_k V_k^H}\}$

$$\mathrm{tr}\{\overline{V_{k+K+1} V_{k+K+1}^H}\} \leq \mathrm{tr}\{\overline{V_{k+K} V_{k+K}^H}\} - \frac{\mu}{P}\mathrm{tr}\{X_{k+K} X_{k+K}^H \overline{V_{k+K} V_{k+K}^H}\}. \tag{17}$$

Using (17) and also

$$\overline{V_{k+1} V_{k+1}^H} = \overline{(I - \mu I_k X_k X_k^H)\overline{V_k V_k^H}(I - \mu X_k X_k^H I_k)}$$

repeatedly, we obtain the following update equation

$$\mathrm{tr}\{\overline{V_{k+K+1} V_{k+K+1}^H}\} \leq \mathrm{tr}\{\overline{V_k V_k^H}\} - \mathrm{tr}\{\Omega_k \overline{V_k V_k^H}\}.$$

From Lemma 5 we have

$$\mathrm{tr}\{\overline{V_{k+K+1} V_{k+K+1}^H}\} \leq (1 - \mu\gamma)\mathrm{tr}\{\overline{V_k V_k^H}\}$$

which ensures the exponential convergence of $\mathrm{tr}\{\overline{V_k V_k^H}\}$. □

E Derivation of Expressions in Section 4.2

In this section, we will need the following identity

$$\sum_{s=0}^{\infty} s(1-\alpha\mu)^{2s} = \frac{(1-\alpha\mu)^2}{\alpha^2\mu^2(2-\alpha\mu)^2}.$$

First, we have the following expressions for LMS

$$\begin{aligned}
J^0_{k+1} &= \sum_{s=0}^{k}(1-\mu\sigma^2)^{k-s}X_s n_s \\
J^1_{k+1} &= \mu\sum_{s=0}^{k}(1-\mu\sigma^2)^{k-s-1}D_1(k,s+1)X_s n_s \\
J^2_{k+1} &= \mu^2\sum_{s=0}^{k}(1-\mu\sigma^2)^{k-s-2}D_2(k,s+1)X_s n_s
\end{aligned}$$

where

$$\begin{aligned}
D_1(k,s) &= \sum_{u=s}^{k} Z_u \quad k \geq s \quad D_1(k,s)=0 \quad s>k \\
D_2(k,s) &= \sum_{u=s}^{k} D_1(k,u+1)Z_u
\end{aligned}$$

and $Z_u = E[X_u X_u^H] - X_u X_u^H$.

This leads to

$$\lim_{k\to\infty} E[J^0_{k+1}(J^0_{k+1})^H] = \lim_{k\to\infty} \sigma_v^2 \sum_{s=0}^{k}(1-\mu\sigma^2)^{2(k-s)}E[X_0X_0^H]$$

and we finally obtain

$$\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(0)})^H] = \frac{\sigma_v^2}{\mu(2-\mu\sigma^2)}I.$$

Similarly,

$$\lim_{k\to\infty} E[J^0_{k+1}(J^1_{k+1})^H] = \mu\sigma_v^2\sum_{s=0}^{k}(1-\mu\sigma^2)^{2s-1}E[D_1(s,1)X_0X_0^H].$$

Now, $E[Z_uX_0X_0^H] = E[X_uX_u^H]E[X_0X_0^H] - E[X_uX_u^HX_0X_0^H] = 0$ which gives

$$E[D_1(s,1)X_0X_0^H] = 0.$$

Thus, $\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(1)})^H] = 0$.

Next,

$$\lim_{k\to\infty} E[J^1_{k+1}(J^1_{k+1})^H] = \sigma_v^2\mu^2\sum_{u=0}^{\infty}(1-\mu\sigma^2)^{2u-2}E[D_1(u,1)X_0X_0^HD_1(u,1)^H]$$

$$\begin{aligned}
E[Z_vX_0X_0^HZ_u^H] &= \sigma^6 I - \sigma^2E[X_vX_v^HX_0X_0^H] - \sigma^2E[X_0X_0^HX_uX_u^H] + \\
&\quad E[X_vX_v^HX_0X_0^HX_uX_u^H] \\
&= 0 \quad \text{if } v \neq u \\
&= N\sigma^6 I \quad \text{if } v = u.
\end{aligned}$$

Therefore, $E[D_1(u,1)X_0X_0^HD_1(u,1)^H] = uN\sigma^6I$ and we obtain

$$\lim_{k\to\infty} E[J_k^{(1)}(J_k^{(1)})^H] = \frac{N\sigma^2\sigma_v^2}{(2-\mu\sigma^2)^2}I.$$

Next, we have $\lim_{k\to\infty} E[J_{k+1}^0(J_{k+1}^2)^H] = \mu^2\sigma_v^2\sum_{s=0}^{\infty} E[D_2(s,1)X_0X_0^H]$. Now,

$$\begin{aligned} E[Z_vZ_uX_0X_0^H] &= \sigma^6I - \sigma^2E[X_vX_v^HX_0X_0^H] - \sigma^2E[X_uX_u^HX_0X_0^H] + \\ &\quad E[X_vX_v^HX_uX_u^HX_0X_0^H] \\ &= 0 \ \text{ if } v \neq u. \end{aligned}$$

Therefore, $E[D_2(s,1)X_0X_0^H] = 0$ and consequently $\lim_{k\to\infty} E[J_{k+1}^0(J_{k+1}^2)^H] = 0$.

Second, we have the following expressions for SPULMS

$$\begin{aligned} J_{k+1}^0 &= \sum_{s=0}^{k}(1-\frac{\mu}{P}\sigma^2)^{k-s}I_sX_sn_s \\ J_{k+1}^1 &= \mu\sum_{s=0}^{k}(1-\frac{\mu}{P}\sigma^2)^{k-s-1}D_1(k,s+1)I_sX_sn_s \\ J_{k+1}^2 &= \mu^2\sum_{s=0}^{k}(1-\frac{\mu}{P}\sigma^2)^{k-s-2}D_2(k,s+1)I_sX_sn_s \end{aligned}$$

where

$$\begin{aligned} D_1(k,s) &= \sum_{u=s}^{k} Z_u \quad k \geq s \quad D_1(k,s) = 0 \quad s > k \\ D_2(k,s) &= \sum_{u=s}^{k} D_1(k,u+1)Z_u \end{aligned}$$

and $Z_u = I_uX_uX_u^H - \frac{1}{P}E[X_uX_u^H]$.

This leads to

$$\lim_{k\to\infty} E[J_{k+1}^0(J_{k+1}^0)^H] = \lim_{k\to\infty}\sigma_v^2\sum_{s=0}^{k}(1-\mu\sigma^2)^{2(k-s)}E[I_0X_0X_0^HI_0]$$

and

$$\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(0)})^H] = \frac{\sigma_v^2}{\mu(2-\frac{\mu}{P}\sigma^2)}I.$$

Similarly,

$$\lim_{k\to\infty} E[J_{k+1}^0(J_{k+1}^1)^H] = \mu\sigma_v^2\sum_{s=0}^{k}(1-\mu\sigma^2)^{2s-1}E[D_1(s,1)I_0X_0X_0^HI_0].$$

Now, $E[Z_uI_0X_0X_0^HI_0] = E[I_uX_uX_u^H]E[I_0X_0X_0^HI_0] - E[I_uX_uX_u^HI_0X_0X_0^HI_0] = 0$ which gives

$$E[D_1(s,1)X_0X_0^H] = 0$$

so that $\lim_{k\to\infty} E[J_{k+1}^0(J_{k+1}^1)^H] = 0$.

$$\lim_{k\to\infty} E[J_{k+1}^1 (J_{k+1}^1)^H] = \sigma_v^2 \mu^2 \sum_{u=0}^{\infty} (1-\mu\sigma^2)^{2u-2} E[D_1(u,1) I_0 X_0 X_0^H I_0 D_1(u,1)^H].$$

Furthermore

$$\begin{aligned}
E[Z_v I_0 X_0 X_0 I_0 Z_u^H] &= \sigma^6 I - \sigma^2 E[I_v X_v X_v^H I_0 X_0 X_0^H I_0] - \sigma^2 E[I_0 X_0 X_0^H I_0 X_u X_u^H I_u] \\
&\quad + E[I_v X_v X_v^H I_0 X_0 X_0^H I_0 X_u X_u^H I_u] \\
&= 0 \ \text{ if } v \neq u \\
&= \frac{(N+1)P-1}{P^3} \sigma^6 I \ \text{ if } v = u.
\end{aligned}$$

Therefore, $E[D_1(u,1) X_0 X_0^H D_1(u,1)^H] = u \frac{(N+1)P-1}{P^3} \sigma^6 I$ and we obtain

$$\lim_{k\to\infty} E[J_k^{(1)} (J_k^{(1)})^H] = \frac{\frac{(N+1)P-1}{P} \sigma^2 \sigma_v^2}{(2-\frac{\mu}{P}\sigma^2)^2} I.$$

Finally, we consider $\lim_{k\to\infty} E[J_{k+1}^0 (J_{k+1}^2)^H] = \mu^2 \sigma_v^2 \sum_{s=0}^{\infty} E[D_2(s,1) X_0 X_0^H]$. Now

$$\begin{aligned}
E[Z_v Z_u X_0 X_0^H] &= \sigma^6 I - \sigma^2 E[I_v X_v X_v^H I_0 X_0 X_0^H I_0] - \sigma^2 E[I_u X_u X_u^H I_0 X_0 X_0^H I_0] \\
&\quad + E[I_v X_v X_v^H I_u X_u X_u^H I_0 X_0 X_0^H I_0] \\
&= 0 \ \text{ if } v \neq u.
\end{aligned}$$

Therefore, $E[D_2(s,1) X_0 X_0^H] = 0$ and consequently $\lim_{k\to\infty} E[J_{k+1}^0 (J_{k+1}^2)^H = 0$.

F Derivation of Expressions in Section 4.3

In this section, we will need the following identities

$$\begin{aligned}
\sum_{v,w=1}^{s} a^{2|v-w|} &= \frac{s(1-a^4) - 2a^2 + 2a^{2(s+1)}}{(1-a^2)^2} \\
\sum_{v,w=1}^{s} a^{|v-w|} a^{v+w} &= \frac{a^2}{(1-a^2)^2}[1 + a^2 - (2s+1)a^{2s} + (2s-1)a^{2s+2}] \\
\sum_{s=0}^{\infty} s(1-\alpha\mu)^{2s} &= \frac{(1-\alpha\mu)^2}{\alpha^2\mu^2(2-\alpha\mu)^2}.
\end{aligned}$$

First, we have the following expressions for LMS

$$\begin{aligned}
J_{k+1}^0 &= \sum_{s=0}^{k} (1-\mu)^{k-s} X_s n_s \\
J_{k+1}^1 &= \mu \sum_{s=0}^{k} (1-\mu)^{k-s-1} D_1(k, s+1) X_s n_s \\
J_{k+1}^2 &= \mu^2 \sum_{s=0}^{k} (1-\mu)^{k-s-2} D_2(k, s+1) X_s n_s
\end{aligned}$$

where

$$\begin{aligned}
D_1(k,s) &= \sum_{u=s}^{k} Z_u \quad k \geq s \quad D_1(k,s) = 0 \quad s > k \\
D_2(k,s) &= \sum_{u=s}^{k} D_1(k,u+1) Z_u
\end{aligned}$$

and $Z_u = E[X_u X_u^H] - X_u X_u^H$.

This leads to

$$\lim_{k\to\infty} E[J_{k+1}^0 (J_{k+1}^0)^H] = \lim_{k\to\infty} \sigma_v^2 \sum_{s=0}^{k} (1-\mu)^{2(k-s)} E[X_0 X_0^H]$$

and as a result

$$\lim_{k\to\infty} E[J_k^{(0)} (J_k^{(0)})^H] = \frac{\sigma_v^2}{\mu(2-\mu)} I.$$

Next,

$$\lim_{k\to\infty} E[J_{k+1}^0 (J_{k+1}^1)^H] = \mu\sigma_v^2 \sum_{s=0}^{k} (1-\mu)^{2s-1} E[D_1(s,1) X_0 X_0^H].$$

Now, $E[Z_u X_0 X_0^H] = E[X_u X_u^H] E[X_0 X_0^H] - E[X_u X_u^H X_0 X_0^H] = -\frac{N}{P^3}\kappa^{2u}$ which gives

$$E[D_1(s,1) X_0 X_0^H] = -\frac{N}{P^3}\kappa^2 \frac{1-\kappa^{2s}}{1-\kappa^2}.$$

Therefore, $\lim_{k\to\infty} E[J_k^{(0)} (J_k^{(1)})^H] = -\frac{\kappa^2 \sigma_v^2 N}{2(1-\kappa^2)} I + O(\mu) I$. Next we consider,

$$\lim_{k\to\infty} E[J_{k+1}^1 (J_{k+1}^1)^H] = \sigma_v^2 \mu^2 \sum_{u=0}^{\infty} (1-\mu)^{2u-2} E[D_1(u,1) X_0 X_0^H D_1(u,1)^H].$$

Note that,

$$\begin{aligned}
E[Z_v X_0 X_0^H Z_u^H] &= I - E[X_v X_v^H X_0 X_0^H] - E[X_0 X_0^H X_u X_u^H] + E[X_v X_v^H X_0 X_0^H X_u X_u^H] \\
&= [(N^2+1)\kappa^{v+u} \kappa^{|v-u|} + N\kappa^{2|v-u|}.
\end{aligned}$$

Therefore,

$$E[D_1(u,1) X_0 X_0^H D_1(u,1)^H] = (N^2+1) \sum_{s=1}^{u} \sum_{t=1}^{u} \kappa^{|v-u|} \kappa^{v+u} + N \sum_{s=1}^{u} \sum_{t=1}^{u} \kappa^{2|v-u|}$$

and consequently

$$\lim_{k\to\infty} E[J_k^{(1)} (J_k^{(1)})^H] = \frac{(1+\kappa^2)\sigma_v^2 N}{4(1-\kappa^2)} I + O(\mu) I.$$

Finally, we have $\lim_{k\to\infty} E[J_{k+1}^0 (J_{k+1}^2)^H] = \mu^2 \sigma_v^2 \sum_{s=0}^{\infty} E[D_2(s,1) X_0 X_0^H]$. Now

$$\begin{aligned}
E[Z_v Z_u X_0 X_0^H] &= I - E[X_v X_v^H X_0 X_0^H] - E[X_u X_u^H X_0 X_0^H] + E[X_v X_v^H X_u X_u^H X_0 X_0^H] \\
&= [(N^2+1)\kappa^{v+u} \kappa^{|v-u|} + N\kappa^{2|v-u|}.
\end{aligned}$$

Therefore,

$$E[D_2(s,1) X_0 X_0^H] = (N^2+1) \sum_{s=1}^{u} \sum_{t=s+1}^{u} \kappa^{|v-u|} \kappa^{v+u} + N \sum_{s=1}^{u} \sum_{t=s+1}^{u} \kappa^{2|v-u|}$$

and

$$\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(2)})^H] = \frac{\kappa^2\sigma_v^2 N}{4(1-\kappa^2)}I + O(\mu)I.$$

Second, we have the following expressions for SPULMS

$$\begin{aligned}
J_{k+1}^0 &= \sum_{s=0}^{k}(1-\frac{\mu}{P})^{k-s} I_s X_s n_s \\
J_{k+1}^1 &= \mu\sum_{s=0}^{k}(1-\frac{\mu}{P})^{k-s-1} D_1(k,s+1) I_s X_s n_s \\
J_{k+1}^2 &= \mu^2\sum_{s=0}^{k}(1-\frac{\mu}{P})^{k-s-2} D_2(k,s+1) I_s X_s n_s
\end{aligned}$$

where

$$\begin{aligned}
D_1(k,s) &= \sum_{u=s}^{k} Z_u \quad k \geq s \quad D_1(k,s) = 0 \quad s > k \\
D_2(k,s) &= \sum_{u=s}^{k} D_1(k,u+1) Z_u
\end{aligned}$$

and $Z_u = I_u X_u X_u^H - \frac{1}{P}E[X_u X_u^H]$.

This leads to

$$\lim_{k\to\infty} E[J_{k+1}^0(J_{k+1}^0)^H] = \lim_{k\to\infty}\sigma_v^2\sum_{s=0}^{k}(1-\mu)^{2(k-s)}E[I_0 X_0 X_0^H I_0]$$

and therefore,

$$\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(0)})^H] = \frac{\sigma_v^2}{\mu(2-\frac{\mu}{P})}I.$$

Next,

$$\lim_{k\to\infty} E[J_{k+1}^0(J_{k+1}^1)^H] = \mu\sigma_v^2\sum_{s=0}^{k}(1-\mu)^{2s-1}E[D_1(s,1)I_0 X_0 X_0^H I_0].$$

Furthermore, $E[Z_u I_0 X_0 X_0^H I_0] = E[I_u X_u X_u^H]E[I_0 X_0 X_0^H I_0] - E[I_u X_u X_u^H I_0 X_0 X_0^H I_0] = -\frac{N}{P^3}\kappa^{2u}$ which gives

$$E[D_1(s,1)X_0 X_0^H] = -\frac{N}{P^3}\kappa^2\frac{1-\kappa^{2s}}{1-\kappa^2}$$

and as a result

$$\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(1)})^H] = -\frac{\kappa^2\sigma_v^2 N}{2(1-\kappa^2)P}I + O(\mu)I.$$

Next, consider

$$\lim_{k\to\infty} E[J_{k+1}^1(J_{k+1}^1)^H] = \sigma_v^2\mu^2\sum_{u=0}^{\infty}(1-\mu)^{2u-2}E[D_1(u,1)I_0 X_0 X_0^H I_0 D_1(u,1)^H].$$

Since,

$$\begin{aligned}
E[Z_v I_0 X_0 X_0 I_0 Z_u^H] &= I - E[I_v X_v X_v^H I_0 X_0 X_0^H I_0] - E[I_0 X_0 X_0^H I_0 X_u X_u^H I_u] \\
&\quad + E[I_v X_v X_v^H I_0 X_0 X_0^H I_0 X_u X_u^H I_u]
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{P^3}[(\frac{N^2}{P}+1)\kappa^{v+u}\kappa^{|v-u|}+N\kappa^{2|v-u|}]I \quad \text{if } v \neq u \\
&= \frac{1}{P^3}[(\frac{N^2}{P}+1)\kappa^{v+u}\kappa^{|v-u|}+N\kappa^{2|v-u|}]I \\
&\quad + \frac{P-1}{P^3}[(N+1)+\frac{N^2+2N+1}{P}\kappa^{2u}]I \quad \text{if } v = u
\end{aligned}
$$

we have $\lim_{k\to\infty} E[J_k^{(1)}(J_k^{(1)})^H] = \frac{\sigma_v^2}{4}[\frac{N}{P}\frac{1+\kappa^2}{1-\kappa^2}+(N+1)\frac{P-1}{P}]I + O(\mu)I$.

Finally, we have $\lim_{k\to\infty} E[J_{k+1}^0(J_{k+1}^2)^H] = \mu^2\sigma_v^2 \sum_{s=0}^{\infty} E[D_2(s,1)X_0X_0^H]$. Furthermore,

$$
\begin{aligned}
E[Z_vZ_uX_0X_0^H] &= \sigma^6 I - \sigma^2 E[I_vX_vX_v^H I_0X_0X_0^H I_0] - \sigma^2 E[I_uX_uX_u^H I_0X_0X_0^H I_0] \\
&\quad + E[I_vX_vX_v^H I_uX_uX_u^H I_0X_0X_0^H I_0] \\
&= \frac{1}{P^3}[(\frac{N^2}{P}+1)\kappa^{v+u}\kappa^{|v-u|}+N\kappa^{2|v-u|}]I \quad \text{if } v \neq u
\end{aligned}
$$

which leads to $\lim_{k\to\infty} E[J_k^{(0)}(J_k^{(2)})^H] = \frac{\kappa^2\sigma_v^2 N}{4(1-\kappa^2)P} I + O(\mu)I$.

References

[1] T. Aboulnasr and K. Mayyas, "Selective coefficient update of gradient-based adaptive algorithms," *Signal Processing*, vol. 47, no. 5, pp. 1421–1424, May 1999.

[2] R. Aguech, E. Moulines, and P. Priouret, "On a perturbation approach for the analysis of stochastic tracking algorithms," *SIAM Journal on Control and Optimization*, vol. 39, no. 3, pp. 872–899, 2000.

[3] T. S. Alexander, *Adaptive Signal Processing: Theory and Applications*, Springer-Verlag and New York, 1986.

[4] P. Billingsley, *Probability and Measure*, Wiley Series in Probability and Mathematical Statistics, 1995.

[5] S. C. Douglas, "Analysis and implementation of the Max-NLMS adaptive filter," *Proceedings of ASIMOLAR-29*, vol. 1, pp. 659l–663, 1996.

[6] S. C. Douglas, "A family of normalized LMS algorithms," *IEEE Signal Processing Letters*, vol. 1, no. 3, pp. 49–51, March 1994.

[7] S. C. Douglas, "Adaptive filters employing partial updates," *IEEE Trans. Acoust., Speech, and Sig. Proc.*, vol. ASSP-44, pp. 209–216, March 1997.

[8] A. Feuer and E. Weinstein, "Convergence analysis of LMS filter with uncorrelated Gaussian data," *IEEE Trans. Acoust., Speech, and Sig. Proc.*, vol. ASSP-33, no. 1, pp. 222–229, Feb. 1985.

[9] M. J. Gingell, B. G. Hay, and L. D. Humphrey, "A block mode update echo canceller using custom LSI," *GLOBECOM Conference Record*, vol. 3, pp. 1394–1397, Nov. 1983.

[10] M. Godavarti, "Implementation of a G.165 line echo canceller on Texas Instruments' TMS320C3x and TMS320C54x chips," *ICSPAT Conference Proceedings*, pp. 65–69, Sep. 1997.

[11] M. Godavarti and A. O. Hero-III, "Analysis of the sequential partial update LMS algorithm," *Submitted to IEEE Trans. on Signal Processing*, 2001.

[12] L. Guo, L. Ljung, and G. J. Wang, "Necessary and sufficient conditions for stability of LMS," *IEEE Trans. Automatic Control*, vol. 42, no. 6, , 1997.

[13] L. L. Horowitz and K. D. Senne, "Performance advantage of complex LMS for controlling narrow-band adaptive arrays," *IEEE Trans. Acoust., Speech, and Sig. Proc.*, vol. ASSP-29, no. 3, pp. 722–736, June 1981.

[14] E. J. Kelly and K. M. Forsythe, "Adaptive detection and parameter estimation for multidimensional signal models," Technical Report 848, M.I.T. Lincoln Laboratory, April, 1989.

[15] D. Messerschmitt, D. Hedberg, C. Cole, A. Haoui, and P. Winship, "Digital voice echo canceller with a TMS32020," *Application report: SPRA129 and Texas Instruments.*

[16] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, McGraw Hill, Inc., 1991.

[17] V. Solo and X. Kong, *Adaptive Signal Processing Algorithms: Stability and Performance*, Prentice-Hall and Englewood Cliffs and NJ, 1995.