

New Sequential Partial Update Switch-Mode Noise-Constrained NLMS Adaptive Filtering Algorithms

S. C. Chan, Y. J. Chu, Z. G. Zhang and K. M. Tsui

Electrical and Electronic Engineering Department, The University of Hong Kong, Hong Kong, P. R. China
E-mail: {scchan; yjchu; zgzhang; kmtsui}@eee.hku.hk

Abstract—The sequential partial update LMS (S-LMS)-based algorithms are efficient adaptive filtering algorithms for reducing the high arithmetic complexity in acoustic and related applications. A limitation of the algorithms is the degraded convergence speed. In this paper, a new family of sequential partial update switch-mode noise-constrained NLMS (S-SNC-NLMS) algorithms is proposed. These algorithms use a new variable step-size (VSS) method to increase the convergence speed of the traditional partial update algorithms while achieving the same steady-state excess mean square error (EMSE). It employs a maximum step-size to improve the initial convergence and exploits the prior knowledge of the additive noise variance as in the noise-constrained (NC) approach near convergence. The mean and mean square convergence behaviors of these new switch mode algorithms are studied to characterize its convergence condition and steady-state EMSE. Based on the theoretical results, an automatic threshold selection method for mode switching is also developed. Computer simulations are conducted to verify the theoretical results and effectiveness of the proposed algorithms.

I. INTRODUCTION

Adaptive filters are frequently employed in applications such as system identification and related problems, in which the statistics of the underlying signals are either unknown *a priori*, or slowly-varying. The well known least mean square (LMS) [1] algorithm and normalized LMS (NLMS) algorithm [2] are widely used in these fields due to their numerical stability and computational simplicity. However, they can become computationally demanding for some applications such as acoustic signal processing, where high-order adaptive filters are usually needed. To solve this problem, many techniques have been proposed to reduce their computational complexity. The partial update (PU) algorithms [3], [4], which update a portion of the filter coefficients at each iteration, is attractive for hardware implementation at the cost of reduced convergence rate. On the other hand, an efficient technique to increase the convergence speed with only a modest increase in complexity is to employ a variable step-size (VSS). The VSS algorithms aim to employ large step-sizes to speed up the convergence rate initially and gradually decrease the step-size in order to achieve a low excess mean square error (EMSE). This is often accomplished by varying the step-size values according to a certain measure of the convergence status [5-9].

In this paper, a new family of sequential partial update switch-mode noise-constrained NLMS (S-SNC-NLMS) algorithm is proposed. It exploits prior knowledge of the

additive noise variance as in the NCLMS approach [7] to achieve simultaneously fast convergence speed and low steady-state error. The mean convergence behavior of the proposed algorithm is investigated and it suggests that improved performance can be obtained if the maximum step-size is employed at initial convergence and during tracking. On the other hand, the noise constrained (NC) adaptation is more suitable to be used near convergence to reduce the steady-state misadjustment. Hence, a new switch-mode adaptation scheme, which employs a maximum step-size mode (MSM) during initial convergence and tracking together with a noise-constrained mode (NCM) near convergence, is proposed to simultaneously improve the convergence speed and steady-state performance of the PU algorithms. The mean squares convergence behavior of this new S-SNC-NLMS algorithm is also studied so as to characterize its steady-state EMSE. Based on the theoretical results, a new and automatic threshold selection method for mode switching is developed. General recommendations for choosing other parameters are also proposed to facilitate engineers in applying this algorithm to achieve a desired EMSE or accuracy in practice. Simulation results show that the S-SNC-NLMS algorithm has better performance than other PU algorithms tested and the validity of the theoretical analysis is also verified. The concept is also applicable to other LMS-based algorithms.

The rest of the paper is organized as follows: in Section II, the sequential PU algorithms are reviewed and the S-SNC-NLMS algorithm is derived. In Section III, the mean and mean square convergence performance of the proposed algorithm are presented. In Section IV, simulation results and comparisons with other conventional algorithms are presented. Finally, conclusion is drawn in Section V.

II. THE S-SNC-NLMS ALGORITHM

A. Review of the S-NLMS Algorithm

Consider the identification of a linear time-invariant (LTI) finite duration impulse response (FIR) system with an impulse response coefficient vector \mathbf{w}^* with L taps by an adaptive filter with weight vector of the same length. The unknown system and adaptive filter are both excited by an input $x(n)$. The measured output of the system is $d(n)$, which is assumed to be corrupted by a zero-mean white Gaussian noise $\eta(n)$ with variance σ_η^2 . $d(n)$ is applied to the desired input of the additive filter. Hence

$$d(n) = (\mathbf{w}^*)^T \mathbf{x}(n) + \eta(n), \quad (1)$$

where $\mathbf{x}(n) = [x(n), x(n-1), \dots, x(n-L+1)]^T$ is the input signal vector. For the sake of presentation, we collectively call the conventional sequential PU (S-), the sequential block (SB-) and the stochastic partial update (SPU-) NLMS algorithms as the S-NLMS family of algorithms. The general update equations can be written as

$$e(n) = d(n) - \mathbf{x}^T(n)\mathbf{w}(n), \quad (2)$$

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu \frac{\mathbf{S}_x(n)\mathbf{x}(n)e(n)}{\varepsilon + \mathbf{x}^T(n)\mathbf{x}(n)}, \quad (3)$$

where $\mathbf{w}(n) = [w_0(n), w_1(n), \dots, w_{L-1}(n)]^T$ is the coefficient vector of the adaptive filter, μ is a constant step-size parameter controlling the convergence rate and steady-state EMSE of the algorithm, ε is a small positive value used to avoid division by zero, and $\mathbf{S}_x(n) = \text{diag}\{s_1(n), \dots, s_L(n)\}$ is the selection matrix, where $s_i(n) \in \{0, 1\}$, $i = 1, 2, \dots, L$. At time instant n , when $s_i(n)$ is equal to one, the corresponding element $w_i(n)$ in $\mathbf{w}(n)$ will be updated. When $\mathbf{S}_x(n) = \mathbf{I}$ is an identity matrix, (2) will reduce to the conventional NLMS algorithm. In the PU algorithms, $\mathbf{w}(n)$ is divided into C non-overlapping groups which are updated sequentially and C is called the decimation factor. Each column of $\mathbf{S}_x(n)$ contains P ($= L/C$) equally spaced 1's and 0's elsewhere and consecutive columns are obtained shifting cyclically. Consequently, only P coefficients need to be updated per iteration.

B. The S-SNC-NLMS Algorithm

It can be seen that the PU algorithms reduce considerably the computational complexity at the expense of a slower convergence speed. In order to compensate for the reduced convergence speed, we propose to adopt the following NC-based variable step-size update [7] in the S-NLMS algorithms

$$\mu(n) = \alpha(1 + \gamma\lambda(n)), \quad (4)$$

$$\begin{aligned} \lambda(n+1) &= \lambda(n) + \beta(\frac{1}{2}(e^2(n) - \sigma_\eta^2) - \lambda(n)) \\ &= (1 - \beta)\lambda(n) + \frac{1}{2}\beta\hat{J}(n), \end{aligned} \quad (5)$$

where α, β, γ are constant parameters and $\hat{J}(n) = e^2(n) - \sigma_\eta^2$ is the instantaneous EMSE. It can be seen that the variable step-size $\mu(n)$ has large values when the algorithm is far from convergence in order to speed up the convergence rate since the convergence measure $J(n)$ is comparatively large. $\mu(n)$ is then gradually decreased to achieve a low steady-state EMSE.

As suggested in [7], after fixing the nominal step-size α , γ should be chosen as a value as large as possible to obtain a fast convergence speed, while β should be chosen as a small value to achieve a desired EMSE. However, the values of γ and α are still constrained so that the step-size and hence the convergence speed will be significantly limited. From the mean convergence analysis, to be presented in Section III, we

found that the mean weight error vector will converge faster if a maximum possible step-size is employed. On the other hand, the NC adaptation is very useful when the adaptive filter nearly converges to the designed steady-state EMSE.

Because of the above observations and possible advantages, we propose below a novel switch-mode scheme for the variable step-size. It employs

- 1) the **maximum step-size mode (MSM)**, where a designed maximum step-size μ_{\max} is employed to achieve a faster convergence speed during initial convergence or tracking, and
- 2) the **noise constrained mode (NCM)**, where the step-size is adjusted as in the NC algorithms, shown in (4) and (5), to achieve the desired EMSE after the maximum step-size mode is nearly converged.

Consequently, the corresponding updates for the step-size can be summarized by the following equations

$$\mu(n) = \begin{cases} \mu_{\max}, & \bar{\lambda}(n) \geq T, \text{ (MSM)}, \\ \max[\alpha(1 + \gamma\lambda(n)), \alpha(1 + \delta)], & \bar{\lambda}(n) < T, \text{ (NCM)}, \end{cases} \quad (6)$$

$$\bar{\lambda}(n+1) = (1 - \bar{\beta})\bar{\lambda}(n) + \bar{\beta}\hat{J}(n)/2, \quad (7)$$

$$\lambda(n+1) = (1 - \beta)\lambda(n) + \beta\hat{J}(n)/2, \quad (8)$$

where μ_{\max} is the designed maximum step-size. To switch between the two modes, we also employ the noise power estimate as in the NC algorithm to measure the convergence status. However, to achieve a fast switching response, a large value of β , denoted as $\bar{\beta}$, is used to estimate a short-term EMSE $\bar{\lambda}(n)$ as shown in (7). When $\bar{\lambda}(n)$ is larger than a certain threshold T , the MSM is invoked. On the other hand, when $\bar{\lambda}(n)$ is smaller than T , the adaptive filter is close to convergence and the NCM is invoked, where a small β is used to estimate a long-term EMSE $\lambda(n)$ as shown in (8) to reduce the estimation variance. The value of $\lambda(n)$ immediately after mode switching is obtained from $\bar{\lambda}(n)$. Eqns. (2), (3) and (6)-(8) constitute the S-SNC-NLMS algorithm.

In cases of noise variance mismatch, the true noise variance in (7) and (8) has to be replaced by $\hat{\sigma}_\eta^2 = a\sigma_\eta^2$ where a is the mismatch factor.

As mentioned earlier, the key issue with the switch-mode approach is the proper selection of the switching threshold T between the two modes and the other related parameters. In this paper, a novel threshold parameter selection scheme is proposed based on the theoretical analysis proposed recently for the NLMS algorithm in Gaussian noise [10] by the authors. The selection of related parameters will also be discussed shortly after the performance analysis in the next section.

III. PERFORMANCE ANALYSIS

In this section, we analyze the convergence performance of the proposed S-SNC-NLMS algorithm. The following assumptions are made:

- (A1) the step-size $\mu(n)$ is independent of the input and error

sequence;

(A2) the additive noise $\{\eta(n)\}$ is white Gaussian-distributed with zero-mean and variance σ_η^2 , which is uncorrelated with the input signal $\{\mathbf{x}(n)\}$;

(A3) the weight vector $\mathbf{w}(n)$ is independent of $\{\mathbf{x}(n)\}$ and $\{\eta(n)\}$ (the independent assumption);

(A4) $\{\mathbf{x}(n)\}$ is an independent identically distributed (i.i.d) Gaussian random sequence with zero-mean and covariance matrix \mathbf{R}_{xx} .

(A1) is an approximation commonly used in most analysis of VSS LMS algorithms to make it mathematical tractable. (A3) is the independence assumption, which is also commonly used in analyzing the LMS family of algorithms. It is quite accurate for small to medium step sizes.

A. Mean Convergence Analysis

First, let the weight error vector at time n be $\mathbf{v}(n) = \mathbf{w}(n) - \mathbf{w}^*$. By using (2), (3), (6)-(8) and the assumptions above, the evolution equations of the mean weight error vector, mean step-size and mean multiplier $E[\lambda(n)]$ can be derived as

$$\begin{aligned} E[\mathbf{v}(n+1)] &= E[\mathbf{v}(n)] - E[\mu(n)]E[e(n)\mathbf{x}(n)] \\ &= (\mathbf{I} - E[\mu(n)]\bar{\mathbf{R}}_{xx})E[\mathbf{v}(n)], \end{aligned} \quad (9)$$

$$E[\mu(n)] = \alpha(1 + \gamma E[\lambda(n)]), \quad (10)$$

$$E[\lambda(n+1)] = (1 - \beta)E[\lambda(n)] + \beta J_b(n), \quad (11)$$

where $\bar{\mathbf{R}}_{xx} = E_{\{\mathbf{x}, \eta\}}[\mathbf{S}_x(n)\mathbf{x}(n)\mathbf{x}^T(n)/(\varepsilon + \mathbf{x}^T(n)\mathbf{x}(n)) | \mathbf{v}]$ with $E_{\{\mathbf{x}, \eta\}}[\cdot]$ denoting the expectation over the sequences $\{\mathbf{x}(n), \eta(n)\}$ conditioned on $\mathbf{v}(n)$. It has been shown in [11] that $\bar{\mathbf{R}}_{xx} = \frac{1}{C}\mathbf{U}\Lambda\mathbf{D}_\Lambda\mathbf{U}^T$, where $\mathbf{U}\Lambda\mathbf{U}^T$ is the eigendecomposition of \mathbf{R}_{xx} for some orthogonal matrix \mathbf{U} , and diagonal matrix Λ containing the eigenvalues of \mathbf{R}_{xx} , and \mathbf{D}_Λ is a diagonal matrix with its i -th diagonal element given by the generalized Abelian integral [10]: $I_i(\Lambda) = \int_0^\infty \exp(-\beta\varepsilon)[\prod_{k=1}^L(2\beta\lambda_k + 1)^{-1/2}] \cdot (2\beta\lambda_i + 1)^{-1} d\beta$. For notational convenience, let $b = a - 1$, and $J_b(n) = E[e^2(n)] - a\sigma_\eta^2 = J(n) - b\sigma_\eta^2$, where $J(n) = E[e^2(n)] - \sigma_\eta^2$. When $b = 0$, there is perfect knowledge of the variance of channel noise, i.e. no noise mismatch exists, and $J_b(n) = J(n) = E[\hat{J}(n)]$.

We shall only focus on the NC adaptation mode, as the maximum step-size mode is equivalent to the S-NLMS algorithms running at a fixed maximum step-size. The latter can be obtained by assuming $\mu(n)$ to be a constant and the details can be found in [12].

Based on (9) and expressing the weight error $\mathbf{v}(n)$ in the canonical coordinate, $\mathbf{V}(n) = \mathbf{U}^T\mathbf{v}(n)$, we get the following difference equation for the i -th element of $E[\mathbf{V}(n)]$

$$E[\mathbf{V}(n)]_i = (1 - \frac{1}{C}E[\mu(n)]\lambda_i I_i(\Lambda))E[\mathbf{V}(n)]_i, \quad (12)$$

where λ_i is the i -th eigenvalue of \mathbf{R}_{xx} . Note, the convergence

rate is slowed down by a factor of C due to the decimation. It can be seen that the mean weight vector of the adaptive filter will converge to the true value if

$$0 < E[\mu(n)] < 2C / \lambda_i I_i(\Lambda). \quad (13)$$

In other words, the maximum step-size can be increased by a factor of C to achieve the same convergence rate as the full update (FU) algorithm for damping down the mean error. If $\mu(n)$ is fixed as in the fixed step-size PU algorithms, the steady-state EMSE will be significantly increased. On the other hand, the use of VSS allows us to achieve both a fast initial convergence speed and a low EMSE.

B. Mean Square Convergence Analysis

To evaluate the mean square behavior, multiplying $\mathbf{v}(n)$ by its transpose and taking expectation on both sides, one gets a difference equation of the weight error covariance matrix $\mathbf{\Xi}(n) = E[\mathbf{v}(n)\mathbf{v}^T(n)]$ as follows

$$\begin{aligned} \mathbf{\Xi}(n+1) &= \mathbf{\Xi}(n) - \frac{1}{C}E[\mu(n)]\bar{\mathbf{R}}_{xx}\mathbf{\Xi}(n) - \frac{1}{C}E[\mu(n)]\mathbf{\Xi}(n)\bar{\mathbf{R}}_{xx} \\ &\quad + E[\mu^2(n)]E_{\{\mathbf{v}\}}[\tilde{\mathbf{s}}_3], \end{aligned} \quad (14)$$

where

$$\tilde{\mathbf{s}}_3 = E_{\{\mathbf{x}, \eta\}}[e(n)^2 \mathbf{S}_x(n)\mathbf{x}(n)\mathbf{x}(n)^T \mathbf{S}_x(n)/(\varepsilon + \mathbf{x}(n)^T \mathbf{x}(n))^2 | \mathbf{v}].$$

Using a similar approach as in [11] to evaluate $\tilde{\mathbf{s}}_3$ yields [12]

$$\tilde{\mathbf{s}}_3 = \frac{1}{C}(\mathbf{s}_3 \circ \Omega), \quad (15)$$

where $\mathbf{s}_3 = E_{\{\mathbf{x}, \eta\}}[e(n)^2 \mathbf{x}(n)\mathbf{x}(n)^T / (\varepsilon + \mathbf{x}(n)^T \mathbf{x}(n))^2 | \mathbf{v}]$, “ \circ ” is the element-wise product of matrices, and Ω is a $L \times L$ matrix with L^2/C elements equal to 1 and zero otherwise. For example, in S-LMS/NLMS algorithms [11], $\Omega = (\mathbf{I}_{L/C} \otimes \mathbf{I})$, where “ \otimes ” is the Kronecker product, $\mathbf{I}_{L/C}$ is a $(L/C) \times (L/C)$ matrix with all entries equal to 1 and \mathbf{I} is a $C \times C$ identity matrix.

The expectation \mathbf{s}_3 can be evaluated to be [12]

$$\mathbf{s}_3 = E_{\{\mathbf{v}\}}[\mathbf{I}_1 + \mathbf{I}_{2,3}], \quad (16)$$

where $\mathbf{I}_1 = 2 \int_0^\infty \int_{\beta_1}^\infty \gamma(\beta_2) \mathbf{B} \mathbf{v} \mathbf{v}^T \mathbf{B} d\beta_2 d\beta_1$ and $\mathbf{I}_{2,3} = \int_0^\infty \int_{\beta_1}^\infty \gamma(\beta_2) \sigma_e^2 \cdot \mathbf{B} d\beta_2 d\beta_1$ with $\gamma(\beta) = \exp(-\beta\varepsilon) \prod_{i=1}^L (2\alpha\beta\lambda_i + 1)^{-1/2}$, $\mathbf{B} = (2\alpha\beta\mathbf{I} + \mathbf{R}_{xx}^{-1})^{-1}$ and $\sigma_e^2 = \mathbf{v}^T \mathbf{B} \mathbf{v} + \sigma_\eta^2$. \mathbf{I}_1 can be integrated as $E_{\{\mathbf{v}\}}[\mathbf{I}_1] = \mathbf{U}E[\mathbf{D}_1]\mathbf{U}^T$, where $E[\mathbf{D}_1] = 2\Lambda[(\mathbf{U}\mathbf{\Xi}(n)\mathbf{U}^T) \circ \mathbf{I}(\Lambda)]\Lambda$, $[\mathbf{I}(\Lambda)]_{i,j} = I_j(\Lambda)$, $I_j(\Lambda) = \int_0^\infty \beta \exp(-\beta\varepsilon) [\prod_{k=1}^L (2\beta\lambda_k + 1)^{-1/2}] \cdot (2\beta\lambda_j + 1)^{-1} (2\beta\lambda_j + 1)^{-1} d\beta$. On the other hand, note that $E_{\{\mathbf{v}\}}[\sigma_e^2] = E_{\{\mathbf{v}\}}[\mathbf{v}^T \mathbf{B} \mathbf{v}] + \sigma_\eta^2$ is upper bounded by $E_{\{\mathbf{v}\}}[\mathbf{v}^T \mathbf{R}_{xx} \mathbf{v}] + \sigma_\eta^2 = \sigma_e^2(n) = E[e^2(n)]$, which is the MSE at time instant n . Thus $E_{\{\mathbf{v}\}}[\mathbf{I}_{2,3}] \approx \sigma_e^2(n) \mathbf{U}\Lambda\mathbf{D}'(\Lambda)\mathbf{U}^T$, where $\mathbf{D}'(\Lambda)$ is a diagonal matrix with the i -th diagonal element given by $I'_i(\Lambda) = \int_0^\infty \beta \exp(-\beta\varepsilon) [\prod_{k=1}^L (2\beta\lambda_k + 1)^{-1/2}] (2\beta\lambda_i + 1)^{-1} d\beta$.

Using $\bar{\mathbf{R}}_{xx} = \frac{1}{C}\mathbf{U}\Lambda\mathbf{D}_\Lambda\mathbf{U}^T$ and the above approximation, Eqn. (14) can be simplified to

$$\mathbf{\Xi}(n+1) \approx \mathbf{\Xi}(n) - \frac{1}{C}E[\mu(n)]\mathbf{U}\Lambda\mathbf{D}_\Lambda\mathbf{U}^T\mathbf{\Xi}(n)$$

$$\begin{aligned}
& -\frac{1}{c} E[\mu(n)] \mathbf{E}(n) \mathbf{U} \Lambda \mathbf{D}'_{\Lambda} \mathbf{U}^T \\
& + \frac{1}{c} E[\mu^2(n)] \{2\mathbf{U} \Lambda [(\mathbf{U} \mathbf{E}(n) \mathbf{U}^T) \circ \mathbf{I}(\Lambda)] \Lambda \mathbf{U}^T \\
& + \sigma_e^2(n) (\mathbf{U} \Lambda \mathbf{D}'(\Lambda) \mathbf{U}^T)\} \circ \Omega. \quad (17)
\end{aligned}$$

By pre- and post-multiplying (17) by \mathbf{U}^T and \mathbf{U} , respectively, it can be further simplified to

$$\begin{aligned}
\Phi(n+1) & \approx \Phi(n) - \frac{1}{c} E[\mu(n)] \Lambda \mathbf{D}'_{\Lambda} \Phi(n) \\
& - \frac{1}{c} E[\mu(n)] \Phi(n) \Lambda \mathbf{D}'_{\Lambda} \\
& + \frac{1}{c} E[\mu^2(n)] \mathbf{U}^T \{[2\mathbf{U} \Lambda (\Phi(n) \circ \mathbf{I}(\Lambda))] \Lambda \mathbf{U}^T \\
& + \sigma_e^2(n) (\mathbf{U} \Lambda \mathbf{D}'(\Lambda) \mathbf{U}^T)\} \circ \Omega \mathbf{U}. \quad (18)
\end{aligned}$$

It can be shown that a sufficient condition for convergence is

$$\frac{E[\mu^2(n)]}{E[\mu(n)]} \leq \frac{2}{2\text{Tr}(\mathbf{I}(\Lambda) \Lambda) \text{Tr}(\mathbf{D}'_{\Lambda}^{-1}) + \text{Tr}([\mathbf{U} \Lambda \mathbf{I}'(\Lambda) \mathbf{U}^T] \circ \Omega) \mathbf{U} \mathbf{D}'_{\Lambda}^{-1} \mathbf{U}^T}. \quad (19)$$

To evaluate the steady-state EMSE, $E[\mu^2(n)]$ and $E[\lambda^2(n)]$ are evaluated from (10) and (11) as follows

$$E[\mu^2(n)] = \alpha^2 (1 + 2E[\lambda(n)] + E[\lambda^2(n)]), \quad (20)$$

$$\begin{aligned}
E[\lambda^2(n+1)] & = (1-\beta)^2 E[\lambda^2(n)] + \beta(1-\beta) E[\lambda(n)] (J(n) \\
& - b\sigma_{\eta}^2) + \frac{\beta}{4} (E[(\mathbf{v}(n)^T \mathbf{x}(n))^4] + b_1 \sigma_{\eta}^2 J(n) + b_2 \sigma_{\eta}^4), \quad (21)
\end{aligned}$$

where $b_1 = 2(2-b)$ and $b_2 = 1+(1+b)^2$. With the use of MSM and the switching threshold, it can be shown that condition (20) is satisfied if γ is appropriately bounded, which is easy to be satisfied in practice. Due to page limitation, the details will be reported elsewhere. Under this circumstance, $E[\lambda^2(n)]$ converges and using (9), (20) and (21), the limiting value of $E[\mu^2(n)]$ is

$$E[\mu^2(\infty)] = \alpha^2 (1 + 2E[\lambda(\infty)] + E[\lambda^2(\infty)]), \quad (22)$$

$$E[\lambda^2(\infty)] \approx \frac{(1-\beta)}{2(2-\beta)} (J_* - b\sigma_{\eta}^2)^2 + \frac{\beta}{4(2-\beta)} (b_1 \sigma_{\eta}^2 J_* + b_2 \sigma_{\eta}^4). \quad (23)$$

where $J_* = \text{Tr}(\mathbf{R}_{xx} \mathbf{E}(\infty)) = \text{Tr}(\Lambda \Phi(\infty))$ is the steady-state EMSE.

Consequently, at the steady state, Eqn. (18) reduces to the following cubic equation by ignoring \mathbf{I}_1 , since from the numerical results \mathbf{I}_1 is much smaller than $\mathbf{I}_{2,3}$

$$\begin{aligned}
& 2(1 + \gamma(J_* - b\sigma_{\eta}^2)/2) J_* \\
& = \alpha(A_0 + A_1 J_* + A_2 J_*^2) (J_* + \sigma_{\eta}^2) \phi_{\text{NLMS}}, \quad (24)
\end{aligned}$$

where $\phi_{\text{NLMS}} = \text{Tr}(\mathbf{D}'_{\Lambda}^{-1} \Lambda \mathbf{D}'(\Lambda))$, $A_0 = (1 - \gamma b \sigma_{\eta}^2 + \frac{(\beta b_2 + 2(1-\beta)b^2)\gamma^2}{4(2-\beta)} \sigma_{\eta}^4)$,

$A_1 = (\gamma + \gamma^2 \frac{\beta b_1 - 4(1-\beta)b}{4(2-\beta)} \sigma_{\eta}^2)$, and $A_2 = \frac{\gamma^2(1-\beta)}{2(2-\beta)}$. From (24), it follows that

$$\begin{aligned}
& (1 - \frac{\gamma b}{2} \sigma_{\eta}^2 - \frac{1}{2} \alpha(A_1 \sigma_{\eta}^2 + A_0) \phi_{\text{NLMS}}) J_* \\
& + (\frac{\gamma}{2} - \frac{1}{2} \alpha(A_2 \sigma_{\eta}^2 + A_1) \phi_{\text{NLMS}}) J_*^2 - \alpha A_2 \phi_{\text{NLMS}} J_*^3 \\
& = \alpha A_0 \sigma_{\eta}^2 \phi_{\text{NLMS}}. \quad (25)
\end{aligned}$$

Assuming J_* is small, which is the usual case for VSS algorithms, the terms involving J_*^2 and J_*^3 can be dropped to obtain a good approximation of J_* as

$$J_* \approx \frac{\frac{1}{2} \alpha A_0 \sigma_{\eta}^2 \phi_{\text{NLMS}}}{1 - \frac{\gamma b}{2} \sigma_{\eta}^2 - \frac{1}{2} \alpha(A_1 \sigma_{\eta}^2 + A_0) \phi_{\text{NLMS}}} \approx \frac{1}{2} \alpha \sigma_{\eta}^2 \phi_{\text{NLMS}} (1 + \delta), \quad (26)$$

where $\delta = (A_0 + \frac{1}{2} \gamma b \sigma_{\eta}^2 - 1) / (1 - \frac{1}{2} \gamma b \sigma_{\eta}^2)$. It can be seen that the steady-state EMSE of the PU algorithms is approximately identical to that of the FU NC-NLMS algorithm [9].

To prevent (26) from being unbound when the denominator is equal to zero, we get the following approximated condition on the maximum possible nominal step-size for mean squares convergence

$$\alpha < \frac{2 - \gamma b \sigma_{\eta}^2}{(A_0 + A_1 \sigma_{\eta}^2) \phi_{\text{NLMS}}} \equiv \alpha_{\text{max}}. \quad (27)$$

C. Switching Threshold and Parameter Selection

1) Selection of T : From [10], the steady-state EMSE of the NLMS is approximately given by $\frac{1}{2} \mu \sigma_{\eta}^2 \phi_{\text{NLMS}}$. It can be shown that for small step-size the S-NLMS has a similar EMSE as the NLMS algorithm [12] and hence its EMSE is also approximately $\frac{1}{2} \mu \sigma_{\eta}^2 \phi_{\text{NLMS}}$. In addition, when the MCM converges, $\text{var}[\bar{\lambda}(\infty)] = E[\bar{\lambda}^2(\infty)] - (E[\bar{\lambda}(\infty)])^2 = \frac{\beta}{2(2-\beta)} ((1 - \frac{1}{8} \mu_{\text{max}} c_f) \mu_{\text{max}} c_f + 1) \sigma_{\eta_{\text{max}}}^4$ with $b=0$ and $c_f = \phi_{\text{NLMS}}$. Assuming $\bar{\lambda}(\infty)$ is Gaussian distributed, T can be chosen as the “ $\kappa \sigma$ ” upper bound of $\bar{\lambda}(\infty)$, i.e.

$T = (\frac{1}{4} \mu_{\text{max}} c_f + \kappa \sqrt{\frac{\beta}{2(2-\beta)} ((1 - \frac{1}{8} \mu_{\text{max}} c_f) \mu_{\text{max}} c_f + 1)}) \sigma_{\eta_{\text{max}}}^2$. κ can be adjusted experimentally and appropriate values are around 4 to 5. In practice, ϕ_{NLMS} is found to vary between 1 and 2.5, which can be employed if the eigenvalues of the input covariance matrix are unavailable.

2) Choice of β and $\bar{\beta}$: Generally, we observe that the parameter $(1-\beta)$ (or $(1-\bar{\beta})$) acts as a forgetting factor and controls the averaging process of the instantaneous MSE. The best value of β (or $\bar{\beta}$) depends mildly on the convergence speed and hence on the correlation of the input signal. For white (colored) input signals, the convergence rate of the algorithm is faster (slower). Consequently, a larger (smaller) value of β can be chosen for white (colored) input signals. For MSM, the convergence speed is usually fast. Thus, a larger $\bar{\beta}$ should be used.

3) Choice of α , δ and γ : According to (26), the product of α and $(1+\delta)$ is fixed for a desired EMSE. Since α contributes more to the convergence speed during the noise-constrained mode, it is advantageous to increase α and decrease δ . A typical value of δ is 0.1. Finally, γ can be computed from the definition of δ in (26). If σ_{η}^2 is not exactly known, we recommend to use the upper bound of $\sigma_{\eta_{\text{max}}}^2$ in (26) to compute a conservative γ .

IV. SIMULATION RESULTS

In this section, computer simulations are conducted to

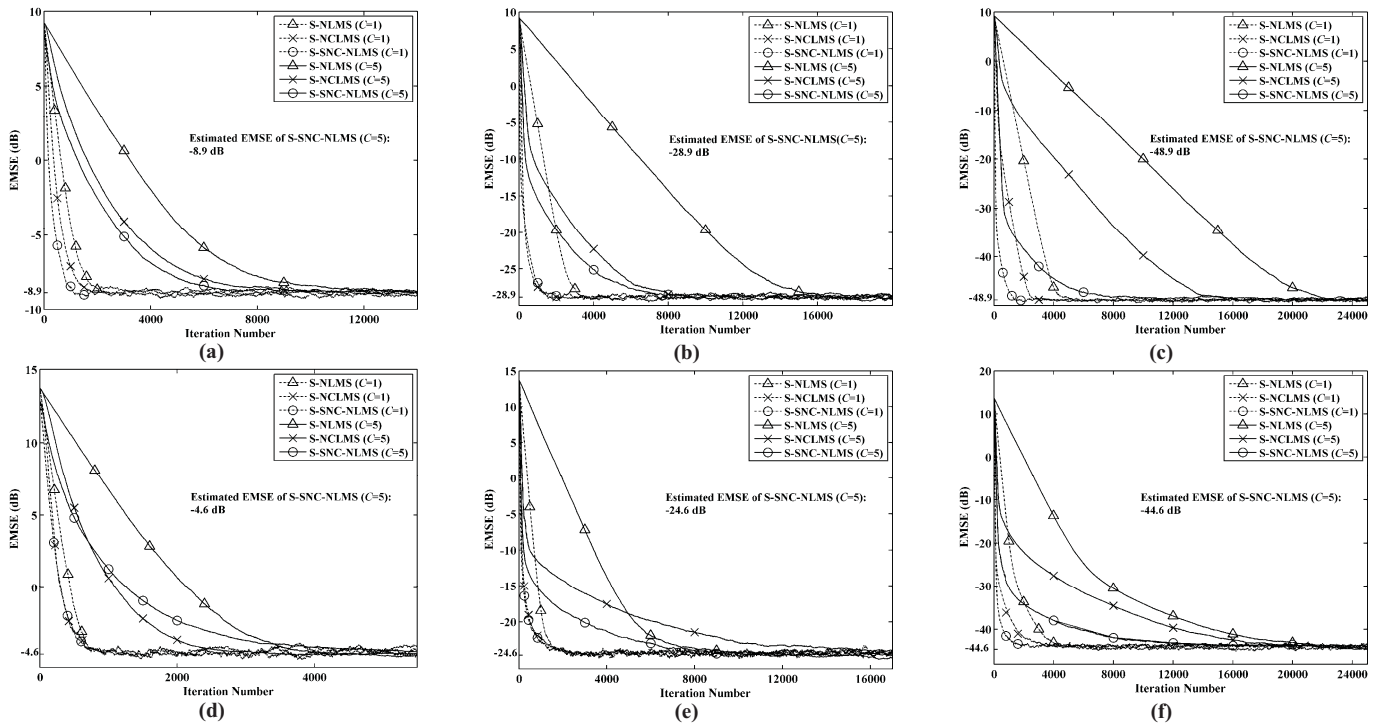


Fig. 1 Learning curves of EMSE for the time-invariant channel identification problem with white Gaussian input in experiment 1 at SNR = (a) 0dB (b) 20dB (c) 40dB, and with first-order AR input in experiment 2 at SNR = (d) 0dB (e) 20dB (f) 40dB.

evaluate the convergence behavior of the proposed algorithm and verify the analytical results obtained in section III. As a comparison, we also consider the NLMS and NCLMS [7] algorithms and their PU counterparts. For simplicity, we only focus on the sequential PU algorithm [11], where the i -th diagonal element of the selection matrix is shifted as follows: $s_i(n) = 1$ if $(n+i) \bmod C = 0$, or $s_i(n) = 0$ otherwise. To investigate the effect of the decimation factor, C is chosen to be 1 and 5. Note when $C=1$, the algorithms actually correspond to their respective conventional FU algorithms.

All simulations below are performed using the system identification model and the unknown system to be estimated is an L -order ($L=15$) FIR filter. Different signal-to-noise ratios (SNRs) at the system output (SNR = 0 dB, 20 dB and 40 dB) are used to examine the performance of the parameter selection scheme proposed in IIIC. In all the simulations, the parameters for deciding the switching threshold of the S-SNC-NLMS are chosen as $\kappa = 5$ and $c_f = \phi_{\text{NLMS}}$. The maximum step-size used is $\mu_{\text{max}} = 1$. All results are obtained by averaging over 300 independent runs.

A. Experiment 1: White Gaussian (WG) Input

The input signal $\{x(n)\}$ is a zero-mean white Gaussian-distributed process with $R_{xx} = I$. In order to achieve a similar steady-state EMSE for all the algorithms to be compared, the input statistics is assumed to be known. The parameters of NCLMS algorithm are chosen in a trial and error manner to obtain its best performance, while for the S-NLMS and S-SNC-NLMS algorithms, their parameters are respectively

chosen according to the theoretical analyses in [10] and Section IIIC. The corresponding parameters are summarized in Table I. The learning curves of EMSE for various algorithms are shown in Figs. 1(a), (b), and (c). It can be seen that the S-NLMS and S-NCLMS algorithms with $C=5$ converge at a relatively lower rates, as compared with $C=1$. The S-SNC-NLMS algorithm with $C=5$, however, speed up the convergence speed more significantly over other algorithms as the SNR increases. Also, it can be seen from Fig. 1 that the estimated steady-state EMSE of the S-SNC-NLMS algorithm with $C=5$ agrees well with simulation results.

B. Experiment 2: Colored Gaussian Input

In this experiment, the following first order autoregressive (AR) process is employed as the input: $x(n) = 0.5x(n-1) + g(n)$, where $g(n)$ is a zero-mean white Gaussian noise with variance $\sigma_g^2 = 0.2$. The algorithm parameters are selected in a similar manner as in experiment 1 and they are summarized in Table I. The learning curves of EMSE are shown in Figs. 1(d), (e), and (f). It can be seen that the estimated steady-state EMSE of the S-SNC-NLMS algorithm ($C=5$) agrees well with simulations. Moreover, it can be seen that the convergence speed of the S-SNC-NLMS algorithm with $C=5$ is generally much faster than other PU algorithms except in Fig. 1(d), where the S-NCLMS algorithm with $C=5$ has a slightly faster convergence rate. It should be noted that the optimal selection of the parameters in the S-NCLMS algorithm requires considerable human intervention so that the improved performance for AR inputs at low SNR may not be realized in practice, especially when the input statistic is unknown.

Nevertheless, according to extensive computer simulations, the proposed S-SNC-NLMS algorithm together with parameter selection rules described in Section III C always outperforms other algorithms for medium and high SNRs. Simulation results with longer filter lengths also show the effectiveness of the parameter selection rules. However, they are not presented here for simplicity.

V. CONCLUSION

A new family of S-SNC-NLMS algorithm is proposed. It accelerates the convergence speed significantly at the expense of moderately increased complexity. The mean and mean square convergence analysis is developed. The effectiveness of the proposed algorithm and its theoretical analysis is verified by computer simulations as presented in Section IV.

REFERENCES

[1] B. Widrow, J. M. McCool, M. G. Larimore, and C. R. Johnson, Jr., "Stationary and nonstationary learning characteristics of the LMS adaptive filter," in *Proc. IEEE*, vol. 64, pp. 1151-1162, Aug. 1976.

[2] J. I. Nagumo and A. Noda, "A learning method for system identification," *IEEE Trans. Automat. Contr.*, vol. AC-12, pp. 282-287, June 1967.

[3] S. Werner, M. L. R. Campos, and P. S. R. Diniz, "Partial-update NLMS algorithms with data-selective updating," *IEEE Trans. Signal Process.*, vol. 52, no. 4, pp. 938-949, Apr. 2004.

[4] M. Godavarti and A. O. Hero, III, "Partial update LMS algorithms," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2382-2399, July 2005.

[5] R. H. Kwong and E. W. Johnston, "A variable step size LMS algorithm," *IEEE Trans. Signal Process.*, vol. 40, no. 7, pp. 1633-1642, July 1992.

[6] T. Aboulnasr and K. Mayyas, "A robust variable step-size LMS-type algorithm: analysis and simulations," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 631-639, Mar. 1997.

[7] Y. Wei, S. B. Gelfand, and J. V. Krogmeier, "Noise-constrained least mean squares algorithm," *IEEE Trans. Signal Process.*, vol. 49, no. 9, pp. 1961-1970, Sep. 2001.

[8] S. C. Chan, Z. G. Zhang, Y. Zhou, and Y. Hu, "A new noise-constrained normalized least mean squares adaptive filtering algorithm," in *Proc. IEEE APCCAS 2008*, Macao, China, Dec. 2008.

[9] S. C. Chan, Y. J. Chu, Z. G. Zhang and Y. Zhou, "On the convergence behavior of the noise-constrained NLMS algorithm," in *Proc. IEEE ISCAS 2009*, Taipei, May 24-27, 2009.

[10] S. C. Chan and Y. Zhou, "Convergence behavior of NLMS algorithm for Gaussian inputs: solutions using generalized Abelian integral functions and step size selection," *Journal of Signal Process. Syst.*, vol. 59, no. 3, pp. 255-265, June 2010.

[11] S. C. Chan, Y. Zhou, K. L. Ho, "A new sequential block partial update normalized least mean M-estimate algorithm and its convergence performance analysis," in *Proc. ISSPIT 2007*, Cairo, Egypt Dec. 2007.

[12] Y. Zhou, S. C. Chan and K. L. Ho, "New sequential partial update least mean M-estimate algorithms for robust adaptive system identification in impulsive noise," submitted to *IEEE Trans. Industrial Electronics*, revised June 2010.

TABLE I
PARAMETERS OF THE FU AND PU ADAPTIVE ALGORITHMS
IN EXPERIMENTS 1 AND 2

Algorithms		Parameters			
		Set	SNR=0	SNR=20	SNR=40
S-NLMS	WG	μ :	0.026	0.026	0.026
	AR		0.025	0.025	0.025
S-NCLMS	WG	α : γ : ($\beta = 0.01$)	0.0017	0.002	0.002
	AR		1	20	25
S-SNC-NLMS	WG	α : β : γ : ($\beta = 0.1$)	0.0017	0.002	0.002
			0.4	10	15
	AR		0.026	0.026	0.026
			0.03	0.03	0.03
AR	0.14	14	1400		
	0.025	0.025	0.025		
		0.01	0.01	0.01	
		0.27	27	2700	

WG: White Gaussian; AR: 1st Order Autoregressive