

Research Article

A Variable Step-Size Leaky LMS Algorithm

Sihai Zhao , Jiange Xu , and Yuyan Zhang 

School of Mechanical Electrical and Information Engineering, China University of Mining & Technology Beijing Campus, Beijing 100083, China

Correspondence should be addressed to Sihai Zhao; zsh@cumtb.edu.cn

Received 17 September 2021; Accepted 15 October 2021; Published 9 November 2021

Academic Editor: Deepak Kumar Jain

Copyright © 2021 Sihai Zhao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The leaky LMS algorithm has been extensively studied because of its control of parameter drift. This unexpected parameter drift is linked to the inadequacy of excitation in the input sequence. And generally leaky LMS algorithms use fixed step size to force the performance of compromise between the fast convergence rate and small steady-state misalignment. In this paper, variable step-size (VSS) leaky LMS algorithm is proposed. And the variable step-size method combines the time average estimation of the error and the time average estimation of the normalized quantity. Variable step-size method proposed incorporating with leaky LMS algorithm can effectively eliminate noise interference and make the early convergence, and final small misalignments are obtained together. Simulation results demonstrate that the proposed algorithm has better performance than the existing variable step-size algorithms in the unexcited environment. Furthermore, the proposed algorithm is comparable in performance to other variable step-size algorithms under the adequacy of excitation.

1. Introduction

The least-mean-square (LMS) algorithm is widely used in many fields, such as system identification [1, 2], echo cancellation [3], adaptive channel equalization [4], adaptive antenna array [5], and adaptive spectral line enhancement [6], due to its good robustness, low computational complexity, and simple structure. However, it is well known that the convergence rate and steady-state error of the algorithm are directly related to the adaptation step size of a conventional LMS algorithm. The convergence rate increases with the increase of the adaption step size, but the steady-state error changes in the opposite direction [7]. Therefore, the adaptive step size is a contradiction in the iterations, and variable step-size (VSS) LMS algorithm, which is adjusted according to the proximity between the actual value and the optimal value in the iterations, is therefore proposed solving this contradiction. In addition, the evaluation criteria for adjusting step size include squared instantaneous error [8, 9], sign changes of successive samples of the gradient [10–12], cross-correlation of input and error [13], and so on.

Although VSS-LMS can balance the convergence rate and steady-state error of LMS algorithm to some extent, it may generate unbounded estimation of parameters due to the existence of inadequacy of excitation. In order to solve this problem, the leakage factor is introduced into the iterations [14–16]. The leaky LMS algorithm can improve the robustness of channel equalization transmission [17, 18] and noise reduction [19], all of which depend on its handling of filter drift. Furthermore, some studies have shown that step size has a certain impact on the performance of leaky LMS algorithm [20], and this paper tries to enhance the performance of leaky LMS algorithm by variable step size.

In this paper, the convergence condition of the leaky LMS algorithm with variable step size is analyzed, and the quantization range of step size is given. In Section 2, based on the analysis of leakage LMS and variable step-size algorithm, a variable step-size leakage LMS algorithm is proposed. Section 3 analyzes that the performance of the proposed algorithm is conducted, and Section 4 details that the simulation is provided to show the validity of the analytical results. The conclusions are presented in Section 5.

2. VVS-L-LMS Adaptive Algorithm

In the conventional L-LMS algorithm [14], the coefficients of the adaptive filter are updated as:

$$\mathbf{w}(n+1) = \gamma\mathbf{w}(n) + \mu\mathbf{x}(n)e(n), \quad (1)$$

where γ is leakage factor and $0 \ll \gamma < 1$, μ is adaptive step size, n is discrete time series, $\mathbf{w}(n)$ is weight coefficient vector of the adaptive filter at n th time, and $\mathbf{x}(n)$ is input signal vector with time series n . $\mathbf{w}(n)$ and $\mathbf{x}(n)$ can be described as:

$$\begin{aligned} \mathbf{x}(n) &= [\mathbf{x}(n), \mathbf{x}(n-1), \dots, \mathbf{x}(n-L+1)]^T, \\ \mathbf{w}(n) &= [\mathbf{w}_0(n), \mathbf{w}_1(n-1), \dots, \mathbf{w}_{L-1}(n-L+1)]^T, \end{aligned} \quad (2)$$

where L is filter order. The instantaneous estimation error $e(n)$ is

$$e(n) = d(n) - \mathbf{x}^T(n)\mathbf{w}(n), \quad (3)$$

where the desired signal $d(n)$ is

$$d(n) = \mathbf{x}^T(n)\mathbf{w}^*(n) + t(n), \quad (4)$$

where $t(n)$ is independent noise sequence with zero mean and $\mathbf{w}^*(n)$ is time-varying optimal weight coefficient vector.

Substituting Equation (3) into Equation (1) and rearranging the terms yields

$$\mathbf{w}(n+1) = [\gamma\mathbf{I} - \mu\mathbf{x}(n)\mathbf{x}^T(n)]\mathbf{w}(n) + \mu\mathbf{x}(n)e(n). \quad (5)$$

The range of step size can be provided by solving the state equation Equation (5) [20]:

$$\forall n \quad \frac{\gamma-1}{\|\mathbf{x}(n)\|^2} < \mu < \frac{\gamma+1}{\|\mathbf{x}(n)\|^2}, \quad (6)$$

where $\gamma \in (0, 1]$, for $\|\mathbf{x}(n)\|^2 \geq 0$; therefore, μ could be negative [20]. The L-LMS algorithm that is same as the conventional LMS algorithm uses a fixed step size in its coefficient update recursion thus inheriting the limitation in the full-update algorithm of having to compromise between fast convergence speed and low level of steady state error.

This problem above can be tackled by changing fixed step-size μ in Equation (1) to time-varying step-size $\mu(n)$, which is adjusted based on a criterion that tries to measure the proximity of the adaptive filter parameters to the optimal ones. Therefore, Equation (1) can be written as:

$$\mathbf{w}(n+1) = \gamma\mathbf{w}(n) + \mu(n)\mathbf{x}(n)e(n). \quad (7)$$

According to Kwong's VVS-LMS algorithm, recursive expression of step size is

$$\mu(n+1) = \alpha\mu(n) + \nu e^2(n), \quad (8)$$

where the positive constant parameters $\alpha(0 < \alpha < 1)$ and $\lambda(\lambda > 0)$ control the prediction error. But Kwong's VVS-

LMS algorithm cannot exclude the influence of noise on the instantaneous error. And in order to solve this problem, it is proposed to use $q(n)$ instead of $e(n)$; $\mu(n)$ can be expressed as:

$$\mu(n+1) = \alpha\mu(n) + \nu q^2(n), \quad (9)$$

where

$$q(n) = p(n)r(n), \quad (10)$$

where $p(n)$ is smoothing value for the cross-correlation function of error $e(n)$ is used to replace the power of error $e^2(n)$. $r(n)$ is input signal power, and they can be described as:

$$p(n) = \beta p(n-1) + (1-\beta)e(n)e(n-1), \quad (11)$$

$$r(n) = \beta r(n-1) + (1-\beta)\frac{e^2(n)}{\mathbf{x}^T(n)\mathbf{x}(n)}, \quad (12)$$

where $\beta(0 < \beta < 1)$ is the weight parameter that controls the convergence time. And the criteria for the step recursion are as follows:

$$\mu(n+1) = \begin{cases} \mu_{\max} & \text{if } \mu(n+1) > \mu_{\max}, \\ \mu_{\min} & \text{if } \mu(n+1) < \mu_{\min}, \\ \mu(n+1) & \text{otherwise,} \end{cases} \quad (13)$$

where $\mu(0) = \mu_{\max}$ and $p(0) = \mu(0)$. According to Equation (6), μ_{\max} and μ_{\min} should satisfy sufficient conditions which guarantee the stability of leaky LMS algorithm and $\mu_{\max} < (\gamma+1)/\|\mathbf{x}(n)\|^2$ and $\mu_{\min} > (\gamma-1)/\|\mathbf{x}(n)\|^2$.

3. Performance Analysis

In this section, the performance of proposed algorithm in stationary and nonstationary environment is described. And in stationary environment, the optimal weight coefficient which is in Equation (4) can be described as a constant:

$$\mathbf{w}^*(n) = \mathbf{w}^*. \quad (14)$$

Furthermore, in stationary environment, $\mathbf{w}^*(n)$ varies randomly as described:

$$\mathbf{w}^*(n+1) = a\mathbf{w}^*(n) + z(n), \quad (15)$$

where $a < 1$, but a is very close to 1. $z(n)$ is a sequence of zero means with covariance $\sigma_z^2\mathbf{I}$, and \mathbf{I} is the identity matrix. According to the above description, $\mathbf{w}^*(n)$ in the stationary environment can be viewed as $\mathbf{w}^*(n)$ in the nonstationary environment where $a = 1$ and the $\sigma_z^2 = 0$.

3.1. Convergence of the Mean Weight Vector. In order to discuss the convergence problem of mean weight vector, translated weight vector $\mathbf{V}(n) = \mathbf{w}(n) - \mathbf{w}^*(n)$ needs to be

introduced, which can reflect the weight error. In addition, the autocorrelation matrix of the input signal is defined as $\mathbf{R} = E\{\mathbf{x}^T(n)\mathbf{x}(n)\}$, which can also be described as $\mathbf{R} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$ where $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$. In the previous formula, $\mathbf{\Lambda}$ is the eigenvalue matrix of \mathbf{R} where $\mathbf{\Lambda} = \text{diag}\{\lambda_N, \lambda_{N-1}, \dots, \lambda_{N-L+2}, \lambda_{N-L+1}\}$, $\lambda_N \leq \lambda_{N-1} \leq \dots \leq \lambda_{N-L+1}$, and \mathbf{Q} is unitary matrix. Then, the rotation matrix, $\hat{\mathbf{V}}(n) = \mathbf{Q}^T\mathbf{v}(n)$, $\hat{\mathbf{X}}(n) = \mathbf{Q}^T\mathbf{x}(n)$, $\hat{\mathbf{V}}(n) = \mathbf{Q}^T\mathbf{v}(n)$, $\hat{\mathbf{W}}^*(n) = \mathbf{Q}^T\mathbf{w}^*(n)$, $\hat{\mathbf{Z}}(n) = \mathbf{Q}^T\mathbf{Z}(n)$, is defined to describe the convergence of the mean weight vector. Substitute Equation (3), Equation (4), and Equation (15) into Equation (7) and the following expression will be found by rewriting:

$$\begin{aligned} \hat{\mathbf{V}}(n+1) &= [\gamma\mathbf{I} - \mu(n)\hat{\mathbf{X}}^T(n)\hat{\mathbf{X}}(n)]\hat{\mathbf{V}}(n) \\ &+ \mu(n)t(n)\hat{\mathbf{X}}(n) + (\gamma - a)\hat{\mathbf{W}}^*(n) - \hat{\mathbf{Z}}(n). \end{aligned} \quad (16)$$

Using assumptions which describes that $\hat{\mathbf{X}}(n)$ and $\hat{\mathbf{V}}(n)$ are independent to find expectations on both sides of Equation (16) yields:

$$\begin{aligned} E\{\hat{\mathbf{V}}(n+1)\} &= [\gamma\mathbf{I} - E\{\mu(n)\}\mathbf{\Lambda}]E\{\hat{\mathbf{V}}(n)\} \\ &+ (\gamma - a)E\{\hat{\mathbf{W}}^*(n)\}. \end{aligned} \quad (17)$$

According to Equation (17), the following conditions can be obtained to ensure the convergence of the mean weight vector:

$$0 < E\{\mu(n)\} < \frac{\gamma + 1}{\lambda_{\max}}, \quad (18)$$

where λ_{\max} is the maximum eigenvalue of the input signal autocorrelation matrix \mathbf{R} . According to the above, λ_{\max} can be rewritten as

$$\lambda_{\max} \leq \text{tr}[\mathbf{R}]. \quad (19)$$

Therefore, $\text{tr}[\mathbf{R}]$ is the trace of the input signal matrix \mathbf{R} , and the condition Equation (18) for guaranteeing the convergence of weight vectors can be rewritten as

$$0 < E\{\mu(n)\} < \frac{\gamma + 1}{\text{tr}[\mathbf{R}]}. \quad (20)$$

But convergence of mean weight vector is not a sufficient condition to guarantee convergence of mean square error.

3.2. Mean Square Error Behavior. In this section, the necessary and sufficient conditions for the mean square error to be stable are discussed. MSE is given by [21]:

$$\epsilon(n) = \epsilon_{\min} + \epsilon_{ex}(n) = \epsilon_{\min} + E\{\hat{\mathbf{V}}^T(n)\mathbf{\Lambda}\hat{\mathbf{V}}(n)\}, \quad (21)$$

where ϵ_{\min} is the minimum of the mean square error where $\epsilon_{\min} = E\{t^2(n)\}$ and $\epsilon_{ex}(n)$ is the excess mean square error (EMSE). In addition, it can be seen from Equation (21) that the convergence of MSE is directly related to diagonal matrix $E\{\hat{\mathbf{V}}^T(n)\mathbf{\Lambda}\hat{\mathbf{V}}(n)\}$. Assuming that $\mu(n)$ is independent of $\hat{\mathbf{V}}(n)$ and $\hat{\mathbf{X}}(n)$, the expression of $E\{\hat{\mathbf{X}}(n)\hat{\mathbf{X}}^T(n)\hat{\mathbf{V}}(n)\hat{\mathbf{V}}^T(n)\hat{\mathbf{X}}(n)\hat{\mathbf{X}}^T(n)\}$ can be simplified into a sum of second-order moments [8] by using Gaussian factoring theorem. So taking the expectation value of $E\{\hat{\mathbf{V}}(n+1)\hat{\mathbf{V}}^T(n+1)\}$ yields:

$$\begin{aligned} E\{\hat{\mathbf{V}}(n+1)\hat{\mathbf{V}}^T(n+1)\} &= \gamma^2 E\{\hat{\mathbf{V}}(n)\hat{\mathbf{V}}^T(n)\} - \gamma\mathbf{\Lambda}E\{\mu(n)\}E\{\hat{\mathbf{V}}(n)\hat{\mathbf{V}}^T(n)\} \\ &- \gamma E\{\mu(n)\}E\{\hat{\mathbf{V}}(n)\hat{\mathbf{V}}^T(n)\mathbf{\Lambda}\} \\ &+ 2E\{\mu^2(n)\}\mathbf{\Lambda}E\{\hat{\mathbf{V}}(n)\hat{\mathbf{V}}^T(n)\mathbf{\Lambda}\} \\ &+ E\{\mu^2(n)\}\mathbf{\Lambda}\text{tr}\left[\mathbf{\Lambda}E\{\hat{\mathbf{V}}(n)\hat{\mathbf{V}}^T(n)\}\right] \\ &+ \gamma(\gamma - a)E\{\hat{\mathbf{V}}(n)\}\mathbf{P}^T\mathbf{\Lambda}^{-1} \\ &- (\gamma - a)\mathbf{\Lambda}E\{\mu(n)\hat{\mathbf{V}}(n)\}\mathbf{P}^T\mathbf{\Lambda}^{-1} \\ &+ \gamma(\gamma - a)\mathbf{\Lambda}^{-1}\hat{\mathbf{P}}E\{\hat{\mathbf{V}}^T(n)\} \\ &- (\gamma - a)\mathbf{\Lambda}^{-1}\hat{\mathbf{P}}E\{\hat{\mathbf{V}}^T(n)\mu(n)\}\mathbf{\Lambda} \\ &+ (\gamma - a)^2\mathbf{\Lambda}^{-1}\hat{\mathbf{P}}(n)\mathbf{P}^T(n)\mathbf{\Lambda}^{-1} + E\{\mu^2(n)\}\epsilon_{mi}\mathbf{\Lambda} + \sigma_z^2\mathbf{I}, \end{aligned} \quad (22)$$

where $\hat{\mathbf{W}}^*(n) = \mathbf{\Lambda}^{-1}\hat{\mathbf{P}}$ and $\hat{\mathbf{P}} = \mathbf{Q}^T\mathbf{P}$ where \mathbf{P} is cross-correlation between the input vector and the expected vector; it can be written as $\mathbf{P} = E\{d(n)\mathbf{x}(n)\}$. A second moment $\mathbf{Z}_1(\mathbf{n})$ including $E\{\hat{\mathbf{V}}(n)\hat{\mathbf{V}}^T(n)\}$ is defined, and $\mathbf{Z}_2(\mathbf{n}) = E\{\hat{\mathbf{V}}(n)\}$. The vector can be described as:

$$\mathbf{Z}(\mathbf{n} + 1) = \begin{bmatrix} \mathbf{Z}_1(\mathbf{n} + 1) \\ \mathbf{Z}_2(\mathbf{n} + 1) \end{bmatrix}. \quad (23)$$

According to Equation (17) and Equation (22), $\mathbf{Z}(\mathbf{n} + 1)$ can be rewritten as:

$$\mathbf{Z}(\mathbf{n} + 1) = \mathbf{K}\mathbf{Z}(\mathbf{n}) + \mathbf{C}, \quad (24)$$

$$\mathbf{K} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{D} \end{bmatrix}, \quad (25)$$

where \mathbf{O} is a 0 matrix. The matrix \mathbf{K} can be described as:

$$\mathbf{A} = \gamma^2\mathbf{I} - 2\gamma\mathbf{\Lambda}E\{\mu(n)\} + 2\mathbf{\Lambda}^2E\{\mu^2(n)\} + E\{\mu^2(n)\}\mathbf{\Gamma}\mathbf{\Gamma}^T, \quad (26)$$

$$\mathbf{B} = [2\gamma(\gamma - a)\mathbf{\Lambda}^{-1} - 2(\gamma - a)E\{\mu(n)\}]\mathbf{Y}, \quad (27)$$

$$\mathbf{D} = \gamma\mathbf{I} - E\{\mu(n)\}\mathbf{\Lambda}, \quad (28)$$

where $\mathbf{Y} = \text{diag}\{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N\}$; the i th part of \mathbf{P} is \mathbf{p}_i . $\mathbf{\Gamma} = \{\lambda_1, \lambda_2, \dots, \lambda_N\}^T$. $E\{\mu^2(\infty)\}$ is steady state of $E\{\mu^2(n)\}$, and $E\{\mu(\infty)\}$ is also steady state of $E\{\mu(n)\}$ and assume that:

$$\begin{aligned} E\{\mu(\infty)\} &= \bar{\mu}, \\ E\{\mu^2(\infty)\} &= \overline{\mu^2}. \end{aligned} \quad (29)$$

In Equation (24), C has no effect on our research, so it can be ignored. And making the following definition is to study further:

$$E\{\dot{\mathbf{V}}(n+1)\dot{\mathbf{V}}^T(n+1)\} = a_{ij}E\{\dot{\mathbf{V}}(n)\dot{\mathbf{V}}^T(n)\}, \quad (30)$$

where a_{ij} can be described as follows:

$$a_{ij} = \gamma^2 - \gamma\bar{\mu}(\lambda_i + \lambda_j) + 2\overline{\mu^2}\lambda_i\lambda_j. \quad (31)$$

But,

$$\begin{aligned} a_{ii}a_{ij} - a_{ij}^2 &= \left[(\gamma - \bar{\mu}\lambda_i)^2 + (2\overline{\mu^2} - \bar{\mu}^2)\lambda_i^2 \right] \\ &\quad \cdot \left[(\gamma - \bar{\mu}\lambda_j)^2 + (2\overline{\mu^2} - \bar{\mu}^2)\lambda_j^2 \right] \\ &\quad - \left[(\gamma - \bar{\mu}\lambda_i)(\gamma - \bar{\mu}\lambda_j) + (2\overline{\mu^2} - \bar{\mu}^2)\lambda_i\lambda_j \right]^2 \\ &= (2\overline{\mu^2} - \bar{\mu}^2) \left[\lambda_i(\gamma - \bar{\mu}\lambda_j) + \lambda_j(\gamma - \bar{\mu}\lambda_i) \right]^2 \geq 0. \end{aligned} \quad (32)$$

Under the condition that can guarantee MSE converges, a_{ij} must satisfy the condition as:

$$a_{ij} < 1. \quad (33)$$

In order to make sure that $E\{\dot{\mathbf{V}}(n+1)\dot{\mathbf{V}}^T(n+1)\}$ converges, we are concerned that we need to make sure that Equation (24) is exponential stable. And a_j is

$$a_j = \gamma^2 - 2\gamma\bar{\mu}\lambda_j + 2\overline{\mu^2}\lambda_j^2. \quad (34)$$

And if the above conditions are satisfied, the following must be true:

$$\det[\mathbf{K} - \rho\mathbf{I}] = \det[\mathbf{D} - \rho\mathbf{I}] \det[\mathbf{A} - \rho\mathbf{I}], \quad (35)$$

where ρ_i is the roots of $\det[\mathbf{K} - \rho\mathbf{I}]$ and the $\det[\mathbf{D} - \rho\mathbf{I}]$ can be got from Equation (28):

$$\det[\mathbf{D} - \rho\mathbf{I}] = \prod_{j=1}^N (\gamma - \bar{\mu}\lambda_j), \quad (36)$$

$$\begin{aligned} \det[\mathbf{A} - \rho\mathbf{I}] &= \det[\text{diag}(a_1 - \rho, \dots, a_n - \rho)] \\ &\quad \cdot \det \left[\mathbf{I} + \text{diag} \left(\frac{1}{a_1 - \rho}, \dots, \frac{1}{a_n - \rho} \right) \cdot \overline{\mu^2} \mathbf{\Lambda}^2 \mathbf{I} \right] \\ &= \left[\prod_{j=1}^N (a_j - \rho) \right] \left[1 + \sum_{j=1}^n \frac{\lambda_j^2}{a_j - \rho} \right]. \end{aligned} \quad (37)$$

According to Equation (37), the conditions that can guarantee the root of Equation (37) in the unit circle can be described as:

$$\begin{aligned} a_j &< 1 \quad j = 1, 2, 3, \dots, N, \\ 1 + \overline{\mu^2} \sum_{i=1}^n \frac{\lambda_j^2}{a_j - 1} &> 0. \end{aligned} \quad (38)$$

Following the same argument in Equation (38). The above conditions can be equivalent to:

$$\frac{1 - \lambda_i^2 (2\overline{\mu^2} - \bar{\mu}^2)}{(\gamma - \bar{\mu}\lambda_i)^2} > 1 \quad j = 1, 2, 3, \dots, N, \quad (39)$$

$$\sum_{i=1}^n \frac{\overline{\mu^2}\lambda_j^2}{(\gamma - \bar{\mu}\lambda_i)^2 + (2\overline{\mu^2} - \bar{\mu}^2)\lambda_i^2 - 1} > -1. \quad (40)$$

$\mathbf{Z}_1(\mathbf{n}+1)$ convergence can be guaranteed if the above conditions Equation (39) and Equation (40) are satisfied. And Equation (21) can be rewritten as:

$$\epsilon(n) = \epsilon_{\min} + \epsilon_{ex}(n) = \epsilon_{\min} + \mathbf{\Gamma}^T \mathbf{Z}_1(\mathbf{n}), \quad (41)$$

where $\check{\mathbf{Y}} = [\check{\mathbf{p}}_1, \check{\mathbf{p}}_2, \dots, \check{\mathbf{p}}_N]^T$ and Equation (11) can be rewritten as:

$$p(n) = (1 - \beta) \sum_{i=0}^{n-1} \beta^i e(n-i) e(n-i-1). \quad (42)$$

Taking the expectation of both sides of Equation (42) fields:

$$E\{p(n)\} = \sqrt{\frac{2}{\pi}} E\{p^2(n)\}, \quad (43)$$

where

$$\begin{aligned} E\{p^2(n)\} &= (1 - \beta)^2 \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \beta^{2i} \beta^{2j} \times E \\ &\quad \cdot \{e^2(n-i)e^2(n-i+1)e^2(n-j)e^2(n-j+1)\}, \end{aligned}$$

$$E\{p^2(\infty)\} = \frac{(1 - \beta)[\epsilon_{\min} + \epsilon_{ex}(n)]^2}{(1 + \beta)}. \quad (44)$$

Taking the expectation of both sides of Equation (12) fields:

$$E\{r(n)\} = \beta E\{r(n-1)\} + (1-\beta)E\left\{\frac{e^2(n)}{\mathbf{x}^T(n)\mathbf{x}(n)}\right\}, \quad (45)$$

where

$$\begin{aligned} E\left\{\frac{e^2(n)}{\mathbf{x}^T(n)\mathbf{x}(n)}\right\} &= \frac{\epsilon_{\min} + \epsilon_{ex}(n)}{\sum_N^{N-L+1} E\{\mathbf{x}_{mi}^2(n)\}}, \\ E\{r^2(n)\} &= \beta^2 E\{r^2(n-1)\} \\ &\quad + 2\beta(1-\beta)E\{r(n-1)\}E\left\{\frac{e^2(n)}{\mathbf{x}^T(n)\mathbf{x}(n)}\right\} \\ &\quad + (1-\beta)^2 E\left\{\frac{e^4(n)}{\mathbf{x}^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{x}(n)}\right\}, \end{aligned} \quad (46)$$

where

$$\begin{aligned} E\left\{\frac{e^4(n)}{\mathbf{x}^T(n)\mathbf{x}(n)\mathbf{x}^T(n)\mathbf{x}(n)}\right\} &= \frac{E\{e^4(n)\}}{\sum_N^{N-L+1} \sum_N^{N-L+1} E\{\mathbf{x}_{mi}^2(n)\mathbf{x}_{mi}^2(n)\}}, \\ E\{e^4(n)\} &= 3\epsilon_{\min} + 6\epsilon_{\min} \text{tr}\left[\Lambda E\{\dot{\mathbf{V}}(n)\dot{\mathbf{V}}^T(n)\}\right] + 3\left\{\text{tr}\left[\Lambda E\{\dot{\mathbf{V}}(n)\dot{\mathbf{V}}^T(n)\}\right]\right\}^2 \\ &\quad + 6\text{tr}\left[E\left\{\Lambda \dot{\mathbf{V}}(n)\dot{\mathbf{V}}^T(n)\right\}\right]^2 - 6\text{tr}\left[E\left\{(\mathbf{I} - \mu(n-1)\Lambda)^4\right\}E\{\dot{\mathbf{V}}(n-1)\}\right. \\ &\quad \cdot \left.\left\{\dot{\mathbf{V}}^T(n-1)\Lambda^2\right\}\right], E\{r^2(\infty)\} = \frac{(2\beta+1)E\{e^4(n)\}}{\sum_N^{N-L+1} \sum_N^{N-L+1} E\{\mathbf{x}_{mi}^2(n)\mathbf{x}_{mi}^2(n)\}}, \end{aligned} \quad (47)$$

where

$$E\{e^4(n)\} = 3([\epsilon_{\min} + \Gamma^T \mathbf{Z}_1(n)]) + 6\Gamma^T \mathbf{Z}_1(n) \mathbf{Z}_1^T \Gamma, \quad (48)$$

where $\Gamma^T \mathbf{Z}_1(n) \mathbf{Z}_1^T \Gamma \ll \epsilon_{\min} + \Gamma^T \mathbf{Z}_1(n)$. According to argument in [8], the sufficient conditions to guarantee MSE convergence can be described as:

$$\frac{E\{\mu^2(\infty)\}}{E\{\mu(\infty)\}} = \frac{2\alpha\nu E\{q^2(\infty)\}}{1-\alpha^2} < \frac{6\alpha\nu(1-\beta)(1+2\beta)\epsilon_{\min}^3}{(1-\alpha^2)(1+\beta)E_{xx}^4}, \quad (49)$$

where $E_{xx}^4 = \sum_N^{N-L+1} \sum_N^{N-L+1} E\{\mathbf{x}_{mi}^2(n)\mathbf{x}_{mi}^2(n)\}$. If the input signal is Gaussian white noise signal, then $E_{xx}^4 = 3N\sigma_x^4$ where σ_x^4 is the input signal variance.

3.3. Steady State Misadjustment. In this section, misadjustment of steady state is studied. And the misadjustment is defined as:

$$M = \frac{\epsilon_{ex}(\infty)}{\epsilon_{\min}}, \quad (50)$$

and the $\epsilon_{ex}(n)$ can be represented as:

$$\epsilon_{ex}(n) = \epsilon(n) - \epsilon_{\min} = \Gamma^T \mathbf{Z}_1(n). \quad (51)$$

Accordingly, the solution of $\mathbf{Z}_1(n)$ converges to $\mathbf{Z}_1(\infty)$, and they can be described as:

$$\mathbf{Z}_1(n) = \mathbf{A}\mathbf{Z}_1(n) + \mathbf{B}\mathbf{Z}_2(n) + (\gamma-a)^2(\Lambda^{-1})^2 \check{\mathbf{Y}} + \bar{\mu}\epsilon_{\min}\Gamma + \sigma_z^2 \mathbf{I}. \quad (52)$$

And in the stationary environment ($\sigma_z^2 = 0$), the $\mathbf{Z}_1(\infty)$ can be expressed as:

$$\mathbf{Z}_1(\infty) = [\mathbf{I} - \mathbf{A}]^{-1} [\mathbf{B}\mathbf{Z}_2(\infty) + (\gamma-a)^2(\Lambda^{-1})^2 \check{\mathbf{Y}} + \bar{\mu}\epsilon_{\min}\Gamma], \quad (53)$$

where $\check{\mathbf{Y}} = [\check{\mathbf{p}}_1^2, \check{\mathbf{p}}_2^2, \dots, \check{\mathbf{p}}_N^2]$ and $[\mathbf{I} - \mathbf{A}]^{-1}$ can be computed as:

$$\begin{aligned} [\mathbf{I} - \mathbf{A}]^{-1} &= \left[(1-\gamma^2)\mathbf{I} + 2\gamma\bar{\mu}\Lambda - 2\bar{\mu}^2\Lambda^2 - \bar{\mu}^2\Gamma\Gamma^T \right]^{-1} \\ &= \check{\mathbf{A}} + \frac{\bar{\mu}^2 \check{\mathbf{A}} \Gamma \Gamma^T \check{\mathbf{A}}}{1 - \bar{\mu}^2 \Gamma^T \check{\mathbf{A}} \Gamma} \\ &= \frac{\check{\mathbf{A}}}{1 - \bar{\mu}^2 \sum_{i=1}^n \left(\lambda_i^2 / 2\lambda_i (1 - \lambda_i) (\gamma\bar{\mu} - \bar{\mu}^2) + (1 - \gamma^2) \right)}, \end{aligned} \quad (54)$$

where

$$\check{\mathbf{A}} = \text{diag} \left\{ \frac{1}{2\lambda_1(1-\lambda_1)(\gamma\bar{\mu} - \bar{\mu}^2) + (1-\gamma^2)}, \frac{1}{2\lambda_2(1-\lambda_2)(\gamma\bar{\mu} - \bar{\mu}^2) + (1-\gamma^2)} \dots \right. \\ \left. , \frac{1}{2\lambda_n(1-\lambda_n)(\gamma\bar{\mu} - \bar{\mu}^2) + (1-\gamma^2)} \right\}. \quad (55)$$

And using Equation (17) and Equation (27), the $\mathbf{B}\mathbf{Z}_2(n)$ can be solved:

$$\mathbf{B}\mathbf{Z}_2(\infty) = \frac{2\check{\mathbf{p}}_j^2(\gamma-1)^2(\gamma-\bar{\mu}\lambda_i)}{\lambda_i^2(1-\gamma\bar{\mu}\lambda_i)}. \quad (56)$$

Substituting Equation (53)–Equation (56) into Equation (51) yields:

$$\begin{aligned} \epsilon_{ex}(\infty) &= \sum_{i=1}^n \left(2\check{\mathbf{p}}_j^2(\gamma-1)^2(\gamma-\bar{\mu}\lambda_i) / 2\lambda_i^2(1-\gamma\bar{\mu}\lambda_i)(1-\lambda_i) (\gamma\bar{\mu} - \bar{\mu}^2) \right. \\ &\quad \left. + \lambda_i(1-\gamma\bar{\mu}\lambda_i)(1-\gamma^2) \right) + \sum_{i=1}^n \\ &\quad \cdot \left(\frac{(\gamma-1)^2\check{\mathbf{p}}_i^2 + \bar{\mu}\epsilon_{\min} / 2\lambda_i^2(1-\lambda_i) (\gamma\bar{\mu} - \bar{\mu}^2) + \lambda_i(1-\gamma^2)}{1 - \bar{\mu}^2 \sum_{i=1}^n \left(\lambda_i^2 / 2\lambda_i (1 - \lambda_i) (\gamma\bar{\mu} - \bar{\mu}^2) + (1 - \gamma^2) \right)} \right). \end{aligned} \quad (57)$$

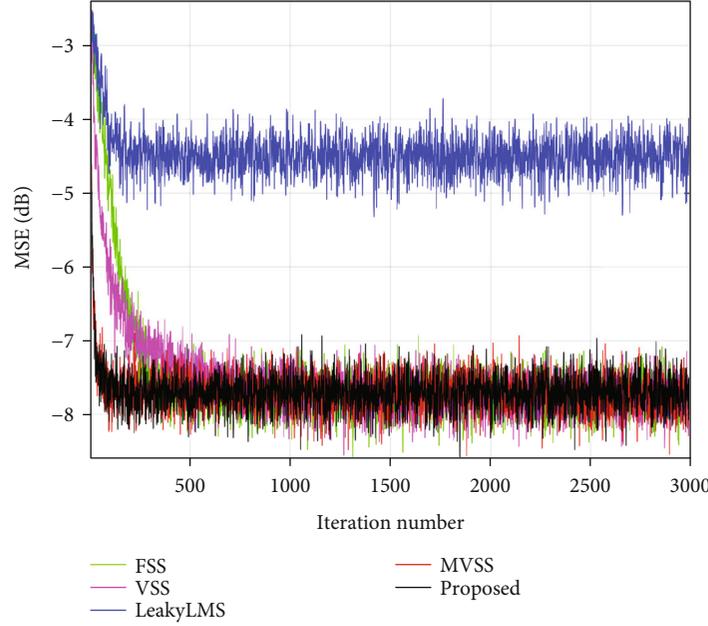


FIGURE 1: Compare the MSE of different algorithms for white input case SNR = 10 dB.

For a leakage factor approaching 1, the above equation can be approximated as:

$$\epsilon_{ex}(\infty) \approx \frac{\sum_{i=1}^n (\bar{\mu} \epsilon_{\min} / 2\lambda_i^2 (1 - \lambda_i) (\gamma \bar{\mu} - \bar{\mu}^2))}{1 - \bar{\mu}^2 \sum_{i=1}^n (\lambda_i^2 / 2\lambda_i (1 - \lambda_i) (\gamma \bar{\mu} - \bar{\mu}^2))}. \quad (58)$$

From Equation (49), the following formula can be got:

$$g = \frac{\bar{\mu}^2}{\bar{\mu}} \approx \frac{6\alpha\nu(1 - \beta)(1 + 2\beta)\epsilon_{\min}^3}{(1 - \alpha^2)(1 + \beta)E_{xx}^4}. \quad (59)$$

Substituting Equation (59) into Equation (58) yields

$$\epsilon_{ex}(\infty) \approx \frac{\sum_{i=1}^n (1/2\lambda_i^2 (1 - \lambda_i) (\gamma - g))}{1 - \sum_{i=1}^n (g\lambda_i^2 / 2\lambda_i (1 - \lambda_i) (\gamma - g))} \epsilon_{\min}. \quad (60)$$

4. Simulation Results

This part analyzes the proposed algorithm in nonstationary environment and focuses on system identification. The performance of the proposed algorithm is compared with MVSS [9], FSS [8], conventional leaky LMS [14] algorithm, and conventional VSS. In order to compare the performance of each algorithm fairly, the algorithms mentioned above are all placed in the same experimental environment, and the parameters of each algorithm are selected to meet the needs of the experiment. Monte Carlo method was used to conduct 200 independent experiments for the following experiments to ensure the reliability of experimental results.

4.1. Example 1: White Input, Low SNR. In the Example 1, the unknown moving average system which is excited by a zero-mean uncorrelated white Gaussian signal of unity variance

ϵ_{\min} has four time-invariant coefficients. And the FIR filter has the same order as the unknown system $L = 4$. The desired signal $d(n)$ is corrupted by zero-mean, uncorrelated Gaussian noise of variance which is also the sequence of input signals. The VSS, proposed algorithm, and MVSS choose $\alpha = 0.97$ which is stable in both stationary and nonstationary environments. The ν for VSS is set to $\nu_{VSS} = 4.8 \times 10^{-5}$ according to the literature, and the ν for proposed algorithm, and MVSS be chosen as 1×10^{-4} and the $\beta = 0.98$. The fixed step-size (FSS) algorithm and conventional leaky LMS algorithm all use the step size $\mu_{FSS} = 3.5 \times 10^{-3}$. Variable step-size iteration algorithm boundary lines are $\mu_{\max} = 0.1$ and $\mu_{\min} = 1 \times 10^{-4}$. The leakage factor $\gamma = 0.999$. Figure 1 shows that both the proposed algorithm and MVSS have good convergence and have basically the same misalignment as FSS algorithm at low SNR. In addition, the leakage algorithm misalignment is basically the same as that of FSS with white noisy input [22] and that is the reason for the big gap between the leaky algorithm and the other algorithms. Figure 2 illustrates that the step-size convergence process of the proposed algorithm is similar to MVSS algorithm and better than VSS algorithm at low SNR.

4.2. Example 2: Correlated Input, Low SNR. In this example, the unknown system to be identified and the filter are set up in the same way as in the Example 1. But the filter input signal is changed from white noise input to correlation input which can be described as:

$$\mathbf{x}(n) = 0.9\mathbf{x}(n-1) + \mathbf{c}(n), \quad (61)$$

where $\mathbf{c}(n)$ is uncorrelated Gaussian white noise with unit variance. And Figure 3 depicts this input signal.

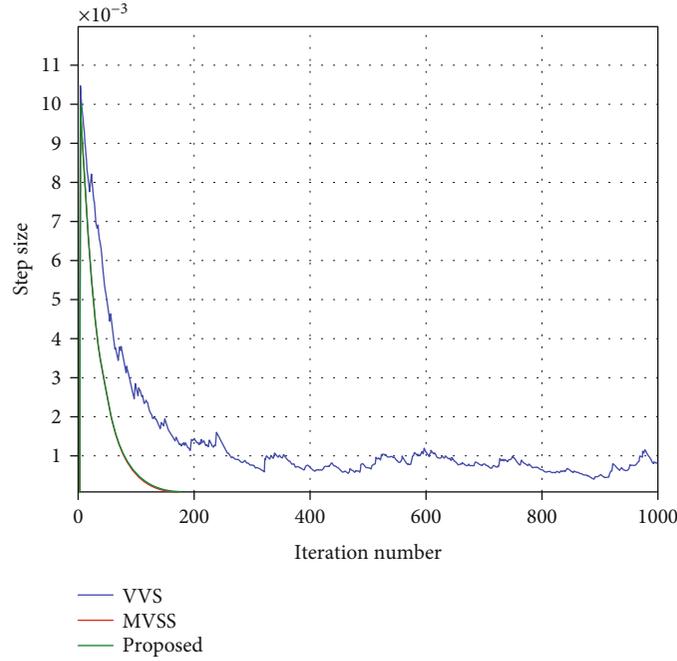


FIGURE 2: Compare the convergence process of variable step-size algorithms for white input case SNR = 10 dB.

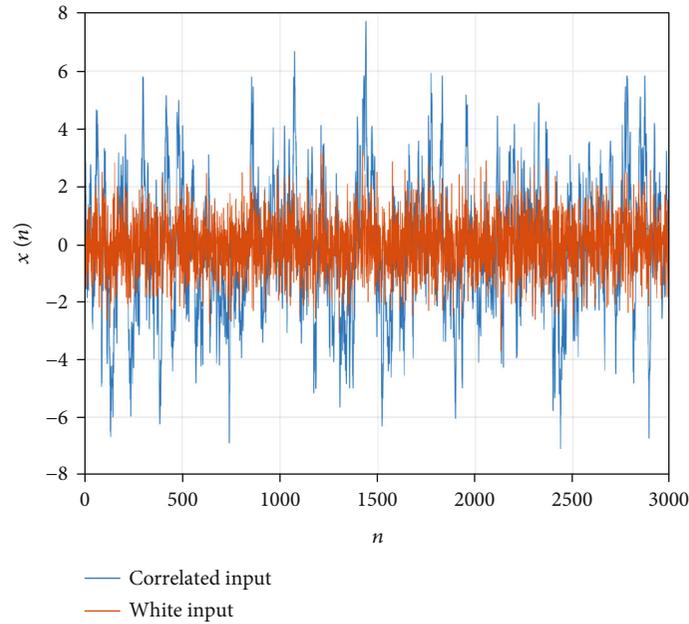


FIGURE 3: Compare the white input and correlated input with $n = 3000$.

The step size of FSS algorithm is $\mu_{\text{FSS}} = 3 \times 10^{-4}$. To ensure that the convergence value of conventional leaky LMS is close to that of the FSS algorithm, the step size can be set to $\mu_{\text{Leaky}} = 3 \times 10^{-3}$, and the leakage factor is $\gamma = 0.999$. The parameters of VSS algorithm with $\mu_{\text{max}} = 1 \times 10^{-3}$ and $\mu_{\text{min}} = 1 \times 10^{-5}$ are set to $\alpha_{\text{VSS}} = 0.97$ and $\nu_{\text{VSS}} = 8 \times 10^{-5}$, respectively. The proposed algorithm and MVSS algorithm have the same α and ν as VSS algorithm, and $\beta = 0.98$. The step-size range of the variable step-size algo-

rithm above is $\mu_{\text{max}} = 1 \times 10^{-2}$ and $\mu_{\text{min}} = 1 \times 10^{-5}$. Figure 4 shows that the convergence speed of the proposed algorithm is close to that of MVSS under the condition of correlated input, and they are both slightly faster than that of FSS algorithm. Under the action of leakage factor, the same MSE convergence value can be obtained only when the step length of leakage algorithm is larger than that of FSS algorithm. As can be seen from Figure 5, the convergence of MVSS step size is slightly faster than that of the proposed algorithm.

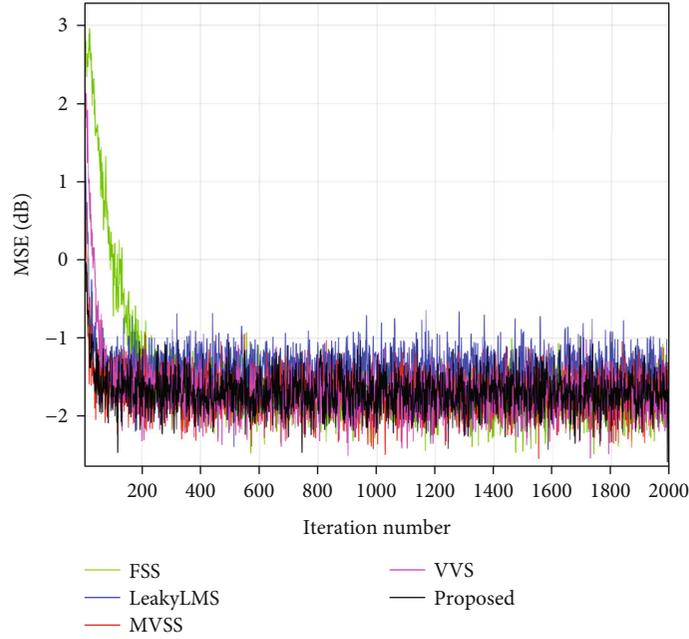


FIGURE 4: Compare the MSE of different algorithms for correlated input case SNR = 10 dB.

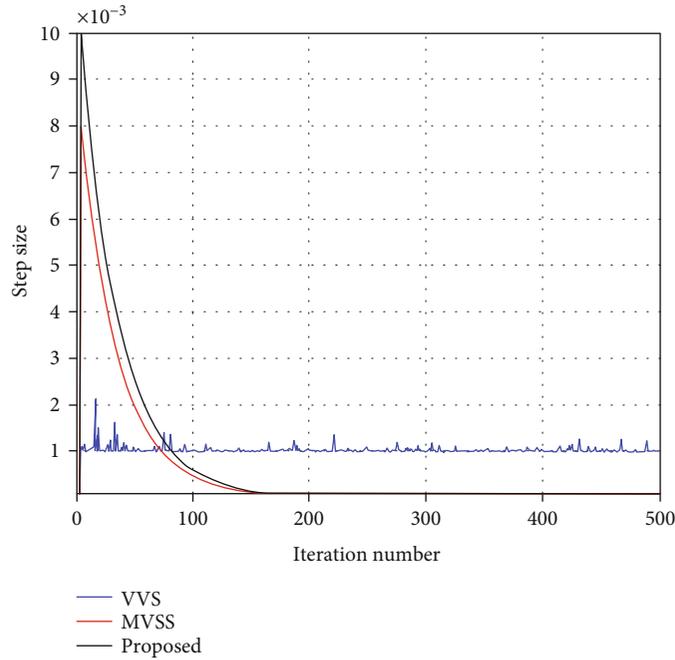


FIGURE 5: Compare the convergence process of variable step-size algorithms for correlated input case SNR = 10 dB.

Under the effect of leakage factor, the peak step size of the proposed algorithm has a big gap with that of MVSS algorithm.

4.3. Example 3: High SNR. Example 3 has the same parameters of filter and unknown system as Examples 1 and 2, except that it has a higher SNR which is SNR = 40 dB.

The FSS algorithm adopts the same step size which is $\mu_{\text{FSS}} = \mu_{\text{Leaky}} = 5 \times 10^{-2}$ as the leaky LMS algorithm, and the

leage factor of the leaky LMS algorithm is $\gamma = 0.99$. All variable step-size algorithms, including MVSS, FSS, and proposed algorithm, use the same step-size range which is $\mu_{\text{max}} = 0.1$ and $\mu_{\text{min}} = 1 \times 10^{-5}$. And the ν of MVSS, FSS, and proposed algorithm is $\nu_{\text{FSS}} = 1 \times 10^{-2}$, $\nu_{\text{MVSS}} = 1$, and $\nu_{\text{pro}} = 1 \times 10^{-5}$, respectively. The β of MVSS and proposed algorithm are all set to $\beta = 0.98$. Figure 6 shows that in the case of large SNR, the gap between the convergence the MSE values of FSS and of leaky LMS algorithm becomes

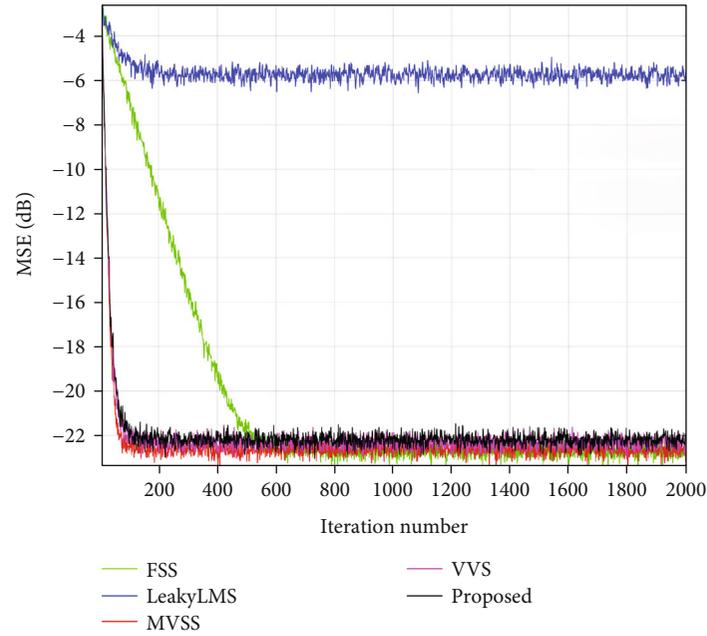


FIGURE 6: Compare the MSE of different algorithms for white input case SNR = 40 dB.

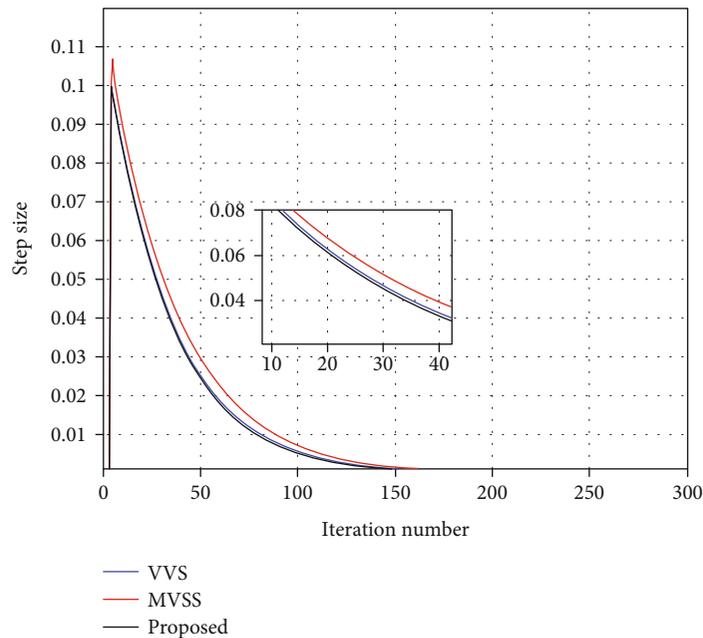


FIGURE 7: Compare the convergence process of variable step-size algorithms for white input case SNR = 40 dB.

larger. And the proposed algorithm converges slightly faster than other variable step-size algorithms. Figure 7 mainly illustrates that the proposed algorithm has advantages in step-size convergence under the condition of high SNR.

4.4. Example 4: Unexcited Input High Order. The parameters of the system to be identified and the filter in Example 4 are the same as those in the above three examples except the order which higher order numbers $L = 14$ are used in this example. The input signal used in Example 4 is a weak white

signal with zero mean and $\sigma_x^2 = 1 \times 10^{-4}$, and the noisy signal which destroys the desired signal $d(n)$ is zero-mean, uncorrelated Gaussian noise of variance.

The FSS algorithm and the leaky LMS algorithm use the same step size $\mu_{\text{FSS}} = \mu_{\text{Leaky}} = 0.01$. And the leakage factor of leaky LMS algorithm and proposed algorithm is $\gamma = 0.999$. The step size of all variable step-size algorithms including FSS, MVSS, and proposed algorithm are set to $\mu_{\text{max}} = 0.05$ and $\mu_{\text{min}} = 1 \times 10^{-5}$, and the α of the all variable step-size algorithms is $\alpha = 0.97$. In addition, the ν of all variable

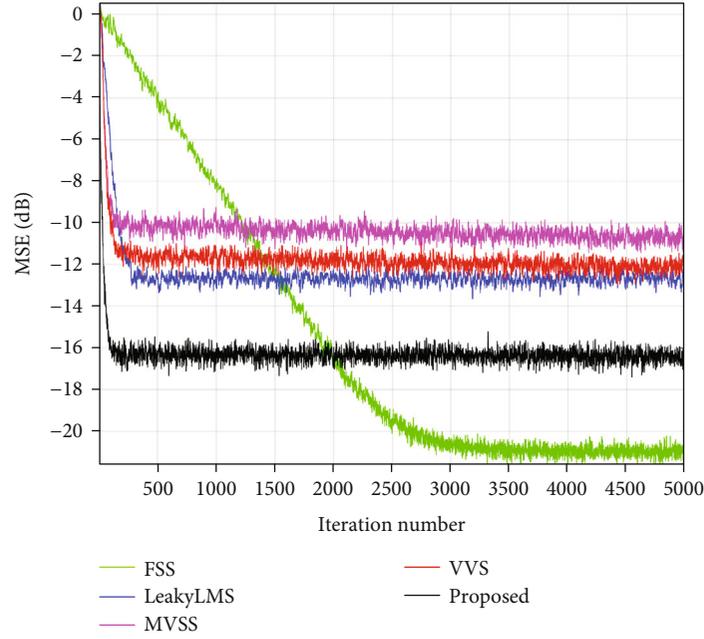


FIGURE 8: Compare the MSE of different algorithms for unexcited input case SNR = 60 dB.

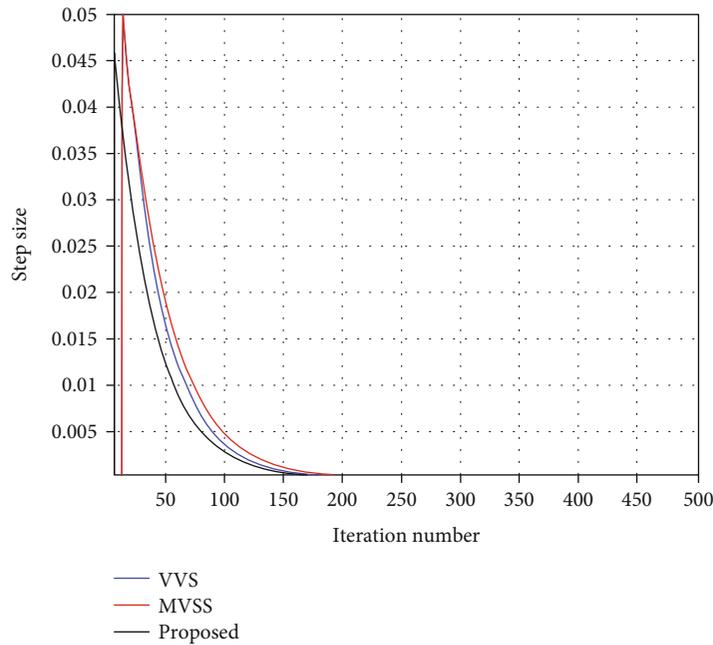


FIGURE 9: Compare the convergence process of variable step-size algorithms for unexcited input case SNR = 60 dB.

step-size algorithms is selected as $\nu_{\text{FSS}} = 1 \times 10^{-5}$, $\nu_{\text{MVSS}} = 1 \times 10^{-3}$, and $\nu_{\text{pro}} = 1 \times 10^{-2}$, respectively. The β of MVSS and proposed algorithm is chosen as $\beta = 0.98$, as can be seen from Figure 8, under the condition of unexcited input, the MSE which leaky algorithms with leakage factor converges to closer to that of FSS. And the proposed algorithm converges faster and can obtain better MSE compared with other algorithms. The existence of leakage factor can better deal with the parameter drift in the case of underexcitation, and the use of variable step size can make the algorithm

convergence faster. Figure 9 shows that the step-size convergence of the proposed algorithm is faster than that of other variable step-size algorithms under the condition of underexcitation.

5. Conclusion

Based on the studies in this paper, the following conclusions can be drawn: (1) The proposed algorithm combines the time average estimation of error autocorrelation with the

time average estimation of normalized quantity. By adjusting the parameters, the proposed algorithm can eliminate uncorrelated noise well. (2) For the same MSE, the convergence speed of the proposed algorithm is close to that of other variable-step algorithms under the condition of low SNR and high SNR. (3) The leakage factor can reduce the parameter drift caused by unexcitation. And in the case of underexcitation, the convergence rate of the proposed algorithm is faster than the other variable step-size algorithms, and the results are closer to wiener solution. In conclusion, the proposed algorithm has better performance in the case of unexcitation.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] N. D. Gaubitch, M. K. Hasan, and P. A. Naylor, "Generalized optimal step-size for blind multichannel LMS system identification," *IEEE Signal Processing Letters*, vol. 13, no. 10, pp. 624–627, 2006.
- [2] Yuantao Gu, Jian Jin, and Shunliang Mei, " L_0 norm constraint LMS algorithm for sparse system identification," *IEEE Signal Processing Letters*, vol. 16, no. 9, pp. 774–777, 2009.
- [3] A. Mader, H. Puder, and G. U. Schmidt, "Step-size control for acoustic echo cancellation filters - an overview," *Signal Processing*, vol. 80, no. 9, pp. 1697–1719, 2000.
- [4] S. M. Shah, R. Samar, N. M. Khan, and M. Raja, "Design of fractional-order variants of complex LMS and NLMS algorithms for adaptive channel equalization," *Nonlinear Dynamics*, vol. 88, no. 2, pp. 839–858, 2017.
- [5] P. Fletcher and M. Dean, "Low complexity implementation of LMS algorithm," *Electronics Letters*, vol. 38, no. 15, pp. 836–837, 2002.
- [6] T. I. Haweel and P. M. Clarkson, "Median adaptive filtering for spectral line enhancement," *The Journal of the Acoustical Society of America*, vol. 85, Supplement 1, pp. S22–S22, 1989.
- [7] A. Khong, W. S. Gan, P. A. Naylor, and M. Brookes, "A low complexity fast converging partial update adaptive algorithm employing variable step-size for acoustic echo cancellation," in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 237–240, Las Vegas, NV, USA, 2008.
- [8] R. H. Kwong and E. W. Johnston, "A variable step size LMS algorithm," *IEEE Transactions on Signal Processing*, vol. 40, no. 7, pp. 1633–1642, 1992.
- [9] T. Aboulnasr and K. Mayyas, "A robust variable step-size LMS-type algorithm: analysis and simulations," *IEEE Transactions on Signal Processing*, vol. 45, no. 3, pp. 631–639, 1997.
- [10] V. J. Mathews and Z. Xie, "A stochastic gradient adaptive filter with gradient adaptive step size," *IEEE Transactions on Signal Processing*, vol. 41, no. 6, pp. 2075–2087, 1993.
- [11] Wee-Peng Ang and B. Farhang-Boroujeny, "A new class of gradient adaptive step-size LMS algorithms," *IEEE Transactions on Signal Processing*, vol. 49, no. 4, pp. 805–810, 2001.
- [12] Jeng-Kuang Hwang and Yuan-Ping Li, "Variable step-size LMS algorithm with a gradient-based weighted average," *IEEE Signal Processing Letters*, vol. 16, no. 12, pp. 1043–1046, 2009.
- [13] T. J. Shan and T. Kailath, "Adaptive algorithms with an automatic gain control feature," *IEEE Transactions on Applied Superconductivity*, vol. 35, no. 1, pp. 122–127, 1988.
- [14] W. A. Sethares, D. A. Lawrence, C. Johnson, and R. R. Bitmead, "Parameter drift in LMS adaptive filters," *IEEE Transactions on Acoustics Speech and Signal Processing*, vol. 34, no. 4, pp. 868–879, 1986.
- [15] K. Mayyas and T. Aboulnasr, "Leaky LMS algorithm: MSE analysis for Gaussian data," *IEEE Trans Signal Process*, vol. 45, no. 4, pp. 927–934, 1997.
- [16] F. Laichi, T. Aboulnasr, and W. Steenaart, "Effect of delay on the performance of the leaky LMS adaptive algorithm," *IEEE Transactions on Signal Processing*, vol. 45, no. 3, pp. 811–813, 1997.
- [17] R. D. Gitlin, H. C. Meadors, and S. B. Weinstein, "The tap-leakage algorithm: an algorithm for the stable operation of a digitally implemented, fractionally spaced adaptive equalizer," *Bell System Technical Journal*, vol. 61, no. 8, pp. 1817–1839, 1982.
- [18] S. Jo, J. Choi, and Y. H. Lee, "Modified leaky LMS algorithm for channel estimation in DS-CDMA systems," *IEEE Communications Letters*, vol. 6, no. 5, pp. 202–204, 2002.
- [19] D. A. Cartes, L. R. Ray, and R. D. Collier, "Experimental evaluation of leaky LMS algorithms for active noise reduction in communication headsets," *The Journal of the Acoustical Society of America*, vol. 108, no. 5, pp. 2483–2484, 2000.
- [20] D. Bismor and M. Pawelczyk, "Stability conditions for the leaky LMS algorithm based on control theory analysis," *Archives of Acoustics*, vol. 41, no. 4, pp. 731–739, 2016.
- [21] B. Widrow, J. Mccool, M. G. Larimore, and C. R. Johnson, "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proceedings of the IEEE*, vol. 64, no. 8, pp. 1151–1162, 1977.
- [22] S. C. Douglas, "Performance comparison of two implementations of the leaky LMS adaptive filter," *IEEE Transactions on Signal Processing*, vol. 45, no. 8, pp. 2125–2129, 1997.