The variable step-size wavelet transform-domain LMS adaptive filter algorithm

M. Shams Esfand Abadi*,a, H. Mesgaranib, and S.M. Khademiyanb

a. *Faculty of Electrical Engineering, Shahid Rajaee Teacher Training University, Tehran, P.O. Box 16785-163, Iran.*
b. *Faculty of Science, Shahid Rajaee Teacher Training University, Tehran, P.O. Box 16785-163, Iran.*

Abstract. In this paper, the Wavelet Transform-Domain Least-Mean-Squares (WT-DLMSs) adaptive algorithm with Variable Step-Size (VSS) is established. The step-size changes according to the largest decrease of mean square deviation. To keep computational complexity low, the Haar Wavelet Transform (HWT) is utilized as a transform. In addition, the mean square performance analysis of the VSS-WTDLMS is studied in the stationary and nonstationary environments, and the theoretical relations for transient and steady-state performances are established. The simulation results show that the proposed VSS-WTDLMS has a faster convergence rate and lower maladjustment, compared to conventional WTDLMS. The theoretical relations are verified by presenting various experimental results.

1. Introduction

Adaptive filters are utilized in many applications such as system identification, channel equalization, and Acoustic Echo Cancellation (AEC) [1–3]. The Least Mean Square (LMS) and Normalized LMS (NLMS) are widely used as the adaptive filter algorithms due to their simplicity and robustness. However, the convergence speed of these algorithms is slow for highly colored input signals. To solve this problem, various adaptive algorithms, such as transform domain adaptive filters, have been proposed [4].

Transform domain adaptive filters exploit de-correlation properties of some well-known signal transforms, such as the discrete Fourier transform (DFT) and the Discrete Cosine Transform (DCT), in order to pre-whiten the input data and speed up filter convergence [5–7]. In the Wavelet Transform Domain Least Mean Square (WTDLMS) adaptive filtering, the projection of the input signal into the orthogonal subspaces is used as inputs to a linear combiner. Weights of the linear combiner can, hence, be updated by the LMS algorithm while normalizing the power at each resolution level to achieve faster and uniform convergence of all weights to the optimum [8,9].

In these algorithms, the fixed step-size can change the convergence rate and the steady-state Mean Square Error (MSE). By selecting the step-size during the adaptation optimally, fast convergence rate and low steady-state MSE are obtained at the same time. In the case of Variable Step-Size (VSS) methods, various approaches have been proposed in the literature [10–15]. One of the most important strategies in this issue was presented in [16]. This approach was successfully extended to different adaptive filter algorithms [17–20]. In [20], the Variable Step-Size Normalized Sub-

*. *Corresponding author. Tel/Fax: +98 21 22970003*
E-mail address: mshams@srttu.edu (M. Shams Esfand Abadi)

band Adaptive Filter (VSS-NSAF) algorithm was introduced. The VSS-NSAF had a faster convergence speed than NSAF. However, the computational complexity of VSS-NSAF was significantly higher than that of NSAF. On the other hand, the theoretical performance analysis of VSS-NSAF was not studied in [20]. It is implied that any theoretical relations are not presented for VSS-NSAF in the stationary and nonstationary environments. Therefore, characteristics of fast convergence speed, low steady-state error, and low computational complexity at the same time are highly desirable in adaptive filter algorithms. Furthermore, the theoretical performance analysis of the derived algorithms is very important.

This paper introduces the VSS-WTDLMS adaptive algorithm. In the proposed VSS-WTDLMS, the step-size changes according to the largest decrease of Mean Square Deviation (MSD). In comparison with conventional WTDLMS, the VSS-WTDLMS has a faster convergence speed and lower steady-state MSE. To keep the computational complexity low, the Haar Wavelet Transform (HWT) is utilized as a transform. Due to the structure of HWT ($+1$ and -1), the number of multiplications in the proposed VSS-WTDlMS is significantly reduced. For performance evaluation of any proposed adaptive algorithm, a theoretical analysis is essential. Therefore, in the following, the mean square performance analysis of VSS-WTDLMS is presented in the stationary and nonstationary environments, and the theoretical relations for transient and steady-state MSEs are established.

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

- The VSS-WTDLMS adaptive algorithm is established. The convergence rate of VSS-WTDLMS is faster than that of conventional WTDLMS. In addition, the steady-state error of VSS-WTDLMS is lower than that of WTDLMS. Furthermore, using the HWT keeps the computational complexity low;

- The mean-square performance analysis of VSS-WTDLMS is studied in the stationary environment, and the theoretical expressions for steady-state and transient performances are introduced;

- The tracking performance analysis of VSS-WTDLMS is studied in the nonstationary environment. In the present study, the unknown system changes according to the random walk model. The transient and steady-state behaviors of VSS-WTDLMS are predicted, and closed-form relations are derived;

- The performance of the proposed algorithm is demonstrated through comprehensive experiments in system identification and AEC applications. The

theoretical results are also justified by several simulations.

The current paper is organized as follows. In Section 2, the WTDLMS is briefly reviewed. Section 3 presents the proposed VSS-WTDLMS adaptive algorithm. Section 4 involves the performance analysis of the VSS-WTDLMS. Section 5 presents the performance analysis in the nonstationary environment. Section 6 presents the computational complexity of the proposed algorithm. Finally, before concluding the paper, the usefulness of this algorithm and the validity of the theoretical relations are demonstrated by presenting several experimental results.

Throughout the paper, $(.)^T$ represents transpose, $||.||^2$ takes the squared Euclidean norm, $||\mathbf{t}||_{\mathbf{\Phi}}^2$ shows the $\mathbf{\Phi}$-weighted Euclidean norm of a column vector $\mathbf{t}$ defined as $\mathbf{t}^T \mathbf{\Phi} \mathbf{t}$, diag$\{\ldots\}$ stands for a diagonal matrix with entries $\{\ldots\}$, vec($\mathbf{T}$) creates an $M^2 \times 1$ column vector $\mathbf{t}$ through stacking the columns of $M \times M$ matrix $\mathbf{T}$, and vec($\mathbf{t}$) creates an $M \times M$ matrix $\mathbf{T}$ from $M^2 \times 1$ column vector $\mathbf{t}$.

2. The WTDLMS adaptive algorithm

A linear data model for $d(n)$ is considered as follows:

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

where $\mathbf{w}_t$ is the unknown M-dimensional vector that we expect to estimate, $v(n)$ is the measurement noise with variance σ_v^2, and $\mathbf{x}(n) = [x(n), x(n-1), \ldots, x(n-M+1)]^T$ denotes an M-dimensional input (regressor) vector. It is assumed that $v(n)$ is zero mean, white, Gaussian, and independent of $\mathbf{x}(n)$. Figure 1 shows the structure of the WTDLMS algorithm [8]. In this figure, $M \times M$ matrix $\mathbf{T}$ is an orthogonal matrix that is derived from a uniform N-band filter bank with filters denoted by $h_0, h_1, \ldots, h_{N-1}$ following the procedure given in [8]. In the matrix form, the orthogonal Wavelet Transform (WT) can be expressed as $\mathbf{z}(n) = \mathbf{T}\mathbf{x}(n)$. This vector can be represented as follows: $\mathbf{z}(n) = [\mathbf{z}_{h_0}^T(n), \mathbf{z}_{h_1}^T(n), \ldots, \mathbf{z}_{h_{N-1}}^T(n)]^T$ where $\mathbf{z}_{h_i}(n)$'s are the output vectors of the N-band filter bank. By splitting the WT domain adaptive filter coefficients $\mathbf{g}(n)$ into N subfilters at time n, each with $\frac{M}{N}$ coefficients, $\mathbf{g}(n) = [\mathbf{g}_{h_0}^T(n), \mathbf{g}_{h_1}^T(n), \ldots, \mathbf{g}_{h_{N-1}}^T(n)]^T$, the output signal can be stated as follows:

$$y(n) = \sum_{i=0}^{N-1} \mathbf{g}_{h_i}^T(n)\mathbf{z}_{h_i}(n), \tag{2}$$

where the error signal is obtained by $e(n) = d(n) - y(n)$. The update equation to each subfilter in WTDLMS is given by:

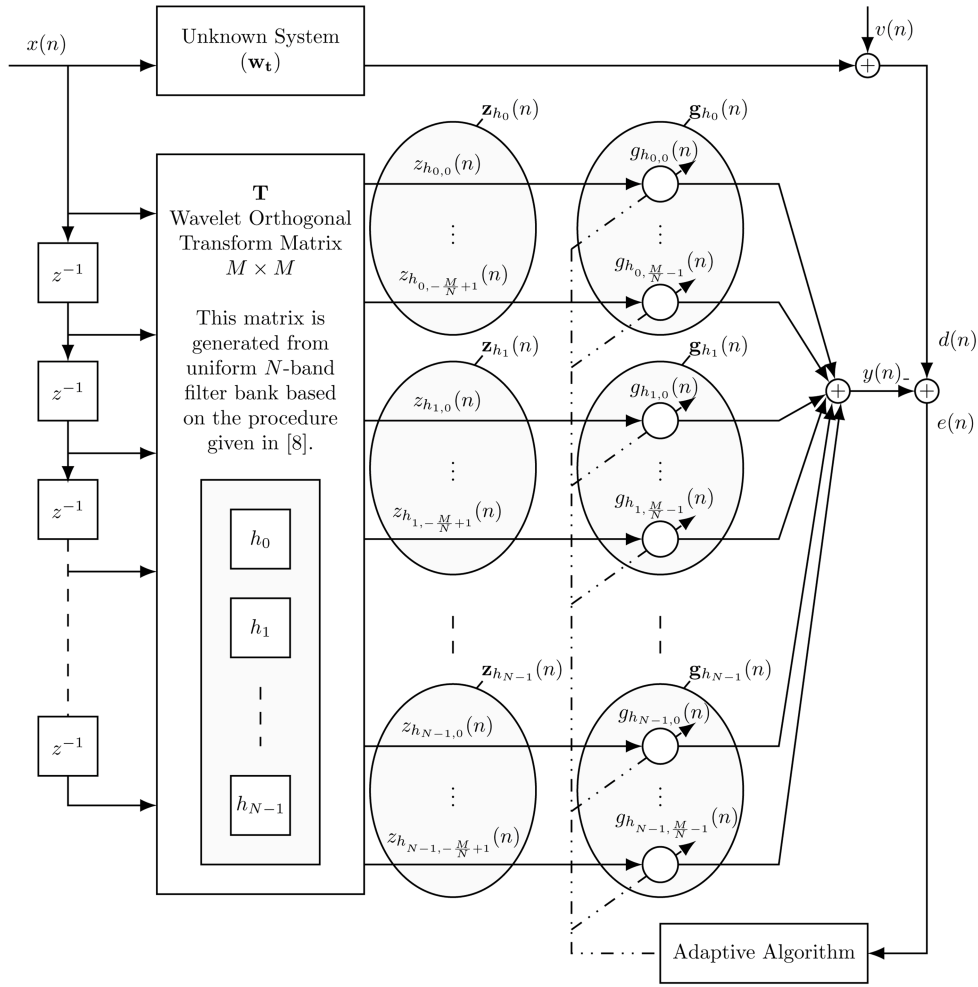$$\mathbf{g}_{h_i}(n+1) = \mathbf{g}_{h_i}(n) + \mu \frac{\mathbf{z}_{h_i}(n)}{\sigma_{h_i}^2(n)} e(n), \tag{3}$$

Figure 1. Structure of the Wavelet Transform-Domain Least-Mean-Square (WTDLMS) algorithm.

where μ is the step-size, and $\sigma_{h_i}^2(n)$ is computed iteratively by:

$$\sigma_{h_i}^2(n) = \alpha \sigma_{h_i}^2(n-1) + (1-\alpha)\|\mathbf{z}_{h_i}(n)\|^2, \qquad (4)$$

with a smoothing factor α $(0 \ll \alpha < 1)$.

3. The VSS-WTDLMS

In this section, the VSS-WTDLMS adaptive algorithm is established according to the largest decrease of MSD. In addition, the theoretical performance analysis is presented in the next section. To do so, the following assumptions are made throughout the paper [8,20–26]:

- **Assumption 1.** $\mathbf{z}(n)$ is independent and an identically distributed sequence vector;

- **Assumption 2.** $\mathbf{z}(n)$ is independent of $e(n)$;

- **Assumption 3.** $\mu(n)$ is independent of $\mathbf{z}(n)$, $v(n)$, and $e(n)$;

- **Assumption 4.** In the steady state, error signal $e(n)$ is approximately equal to noise $v(n)$.

By defining the weight error vector as $\tilde{\mathbf{g}}_{h_i}(n) = \mathbf{g}_{h_i}^o - \mathbf{g}_{h_i}(n)$, where $\mathbf{g}_{h_i}^o$ is the true unknown subfilter coefficient, the weight error vector update equation for WTDLMS for each subfilter can be represented as follows:

$$\tilde{\mathbf{g}}_{h_i}(n+1) = \tilde{\mathbf{g}}_{h_i}(n) - \mu_{h_i}(n)\frac{\mathbf{z}_{h_i}(n)}{\sigma_{h_i}^2(n)}e(n). \qquad (5)$$

In Eq. (5), $\mu_{h_i}(n)$ is a VSS in subfilter. By applying the squared Euclidean norm and, then, taking the expectation of both sides of Eq. (5), the following yielded:

$$E[\|\tilde{\mathbf{g}}_{h_i}(n+1)\|^2] = E\left[\|\tilde{\mathbf{g}}_{h_i}(n)\|^2\right] - \Delta, \qquad (6)$$

where:

$$\Delta = 2\mu_{h_i}(n)E\left[\frac{\tilde{\mathbf{g}}_{h_i}^T(n)\mathbf{z}_{h_i}(n)e(n)}{\sigma_{h_i}^2(n)}\right]$$

$$- \mu_{h_i}^2(n)E[\frac{e^2(n)\|\mathbf{z}_{h_i}(n)\|^2}{\sigma_{h_i}^4(n)}]. \qquad (7)$$

Maximizing Δ with respect to $\mu_{h_i}(n)$ leads to the optimum step-size as follows:

$$\mu_{h_i}^o(n) = \frac{E\left[\frac{\tilde{\mathbf{g}}_{h_i}^T(n)\mathbf{z}_{h_i}(n)e(n)}{\sigma_{h_i}^2(n)}\right]}{E\left[\frac{e^2(n)||\mathbf{z}_{h_i}(n)||^2}{\sigma_{h_i}^4(n)}\right]}. \tag{8}$$

Since $e(n) = e_a(n) + v(n)$ and using $e_a(n) \approx \tilde{\mathbf{g}}_{h_i}^T(n)\mathbf{z}_{h_i}(n)$, Eq. (8) can be written as follows [8]:

$$\mu_{h_i}^o(n) \approx \frac{E\left[\frac{e_a^2(n)}{\sigma_{h_i}^2(n)}\right]}{E\left[\frac{e_a^2(n)||\mathbf{z}_{h_i}(n)||^2}{\sigma_{h_i}^4(n)}\right] + \sigma_v^2 E\left[\frac{||\mathbf{z}_{h_i}(n)||^2}{\sigma_{h_i}^4(n)}\right]}. \tag{9}$$

Defining $\mathbf{q}_{h_i}(n) = \frac{\mathbf{z}_{h_i}(n)e_a(n)}{\sigma_{h_i}^2(n)}$, we obtain that $||\mathbf{q}_{h_i}(n)||^2 = \frac{e_a^2(n)||\mathbf{z}_{h_i}(n)||^2}{\sigma_{h_i}^4(n)}$. Then, the optimum step-size can be expressed as follows:

$$\mu_{h_i}^o(n) \approx \frac{E[||\mathbf{q}_{h_i}(n)||^2]}{E[||\mathbf{q}_{h_i}(n)||^2] + \frac{\sigma_v^2}{\sigma_{h_i}^2(n)}}. \tag{10}$$

Taking the expectation of $\mathbf{q}_{h_i}(n)$ yields:

$$E[\mathbf{q}_{h_i}(n)] = E\left[\frac{\mathbf{z}_{h_i}(n)e(n)}{\sigma_{h_i}^2(n)}\right]. \tag{11}$$

Now, we propose to estimate $E[\mathbf{q}_{h_i}(n)]$ by time averaging as follows:

$$\hat{\mathbf{q}}_{h_i}(n) = \beta\hat{\mathbf{q}}_{h_i}(n-1) + (1-\beta)\frac{\mathbf{z}_{h_i}(n)e(n)}{\sigma_{h_i}^2(n)}, \tag{12}$$

where $0 \ll \beta < 1$. Defining $\hat{\mathbf{q}}_{h_i}(0) = \frac{\mathbf{z}_{h_i}(0)e(0)}{\sigma_{h_i}^2(0)}$, we have:

$$\hat{\mathbf{q}}_{h_i}(n) = \beta^n\hat{\mathbf{q}}_{h_i}(0) + (1-\beta)\sum_{j=1}^n\left[\beta^{n-j}\frac{\mathbf{z}_{h_i}(j)e(j)}{\sigma_{h_i}^2(j)}\right]. \tag{13}$$

Taking the squared Euclidean norm and, then, the expectation of both sides of Eq. (13) and using Assumptions 1 and 2, we obtain:

$$E[||\hat{\mathbf{q}}_{h_i}(n)||^2] = \beta^{2n}E[G(0,0)]$$
$$+ (1-\beta)^2\sum_{j=1}^n[\beta^{2n-2j}E[G(j,j)]], \tag{14}$$

where $G(j,k) = \frac{e(j)\mathbf{z}_{h_i}^T(j)\mathbf{z}_{h_i}(k)e(k)}{\sigma_{h_i}^2(j)\sigma_{h_i}^2(k)}$, and:

$$E[G(j,k)] = \begin{cases} 0 & k \neq j \\ \frac{E[e^2(j)]}{\sigma_{h_i}^2(j)} & k = j \end{cases} \tag{15}$$

Eq. (14) can be recursively computed as follows:

$$E[||\hat{\mathbf{q}}_{h_i}(n)||^2] = \beta^2 E[||\hat{\mathbf{q}}_{h_i}(n-1)||^2]$$
$$+ (1-\beta)^2 E[G(n,n)]. \tag{16}$$

Finally, Eq. (16) can be rewritten as follows:

$$E[||\hat{\mathbf{q}}_{h_i}(n)||^2] = \beta^2 E[||\hat{\mathbf{q}}_{h_i}(n-1)||^2]$$
$$+ (1-\beta)^2\frac{E[e^2(n)]}{\sigma_{h_i}^2(n)}, \tag{17}$$

where $E[||\hat{\mathbf{q}}_{h_i}(0)||^2] = \frac{E[e^2(0)]}{\sigma_{h_i}^2(0)}$. Therefore, Eq. (10) is obtained by:

$$\mu_{h_i}^o(n) \approx \frac{E[||\hat{\mathbf{q}}_{h_i}(n)||^2]}{E[||\hat{\mathbf{q}}_{h_i}(n)||^2] + \frac{\sigma_v^2}{\sigma_{h_i}^2(n)}}. \tag{18}$$

By multiplying the numerator and denominator of Eq. (18) by $\sigma_{h_i}^2(n)$, the optimum step-size can be expressed as follows:

$$\mu^o(n) \approx \frac{E[||\hat{\mathbf{p}}(n)||^2]}{E[||\hat{\mathbf{p}}(n)||^2] + C}, \tag{19}$$

where $C = \sigma_v^2$ and:

$$E[||\hat{\mathbf{p}}(n)||^2] = \beta^2 E[||\hat{\mathbf{p}}(n-1)||^2] + (1-\beta)^2 E[e^2(n)]. \tag{20}$$

In Eq. (20), $E[||\hat{\mathbf{p}}(0)||^2] = E[e^2(0)]$. Finally, the update equation for VSS-WTDLMS is established as follows:

$$\mathbf{g}_{h_i}(n+1) = \mathbf{g}_{h_i}(n) + \mu(n)\frac{\mathbf{z}_{h_i}(n)}{\epsilon + \sigma_{h_i}^2(n)}e(n), \tag{21}$$

where ϵ is the regularization parameter and is introduced to avoid being divided by zero. Table 1

Table 1. The VSS-WTDLMS adaptive algorithm.

for $n = 0, 1, \ldots$
$\quad \mathbf{x}(n) = [x(n), \ldots, x(n-M+1)]^T$
$\quad \mathbf{z}(n) = \mathbf{T}\mathbf{x}(n)$
$\quad \mathbf{g}(n) = [\mathbf{g}_{h_0}^T(n), \mathbf{g}_{h_1}^T(n), \ldots, \mathbf{g}_{h_{N-1}}^T(n)]^T$
$\quad \mathbf{z}(n) = [\mathbf{z}_{h_0}^T(n), \mathbf{z}_{h_1}^T(n), \ldots, \mathbf{z}_{h_{N-1}}^T(n)]^T$
$\quad e(n) = d(n) - \sum_{i=0}^{N-1}\mathbf{g}_{h_i}^T(n)\mathbf{z}_{h_i}(n)$
$\quad ||\hat{\mathbf{p}}(n)||^2 = \beta^2||\hat{\mathbf{p}}(n-1)||^2 + (1-\beta)^2 e^2(n)$
$\quad \mu(n) = \min[\mu_{max}, \mu(n)], \mu(n) = \frac{||\hat{\mathbf{p}}(n)||^2}{||\hat{\mathbf{p}}(n)||^2 + \sigma_v^2}$
$\quad$ for $i = 0, 1, \ldots, N-1$
$\quad\quad \sigma_{h_i}^2(n) = \alpha\sigma_{h_i}^2(n-1) + (1-\alpha)||\mathbf{z}_{h_i}(n)||^2$
$\quad\quad \mathbf{g}_{h_i}(n+1) = \mathbf{g}_{h_i}(n) + \mu(n)\frac{\mathbf{z}_{h_i}(n)}{\epsilon + \sigma_{h_i}^2(n)}e(n)$
$\quad$ end
end

summarizes the procedure of VSS-WTDLMS adaptive algorithm. The step-size in Eq. (21) is bounded by the following relation:

$$\mu(n) = \min[\mu_{\max}, \mu(n)]. \tag{22}$$

In Appendix A, the values of $\mu_{\max}$ are presented. The fully update equation of VSS-WTDLMS can be represented as follows:

$$\mathbf{g}(n+1) = \mathbf{g}(n) + \mu(n)\mathbf{C}(n)\mathbf{z}(n)e(n), \tag{23}$$

where:

$$\mathbf{C}(n) = \begin{pmatrix} \mathbf{C}_{h_0}(n) & 0 & \ldots & 0 \\ 0 & \mathbf{C}_{h_1}(n) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & \mathbf{C}_{h_{N-1}}(n) \end{pmatrix} \tag{24}$$

and $\mathbf{C}_{h_i}(n) = \frac{1}{\epsilon + \sigma_{h_i}^2(n)}\mathbf{I}_{\frac{M}{N} \times \frac{M}{N}}$.

4. Performance analysis of VSS-WTDLMS in stationary environment

The transient behavior of an adaptive filter algorithm is determined by the evolution of the expected squared *a priori* error at time n, i.e., $E\{e_a^2(n)\}$, which is:

$$E\{e_a^2(n)\} = E\{\tilde{\mathbf{g}}^T(n)\mathbf{z}(n)\mathbf{z}^T(n)\tilde{\mathbf{g}}(n)\}, \tag{25}$$

where $\tilde{\mathbf{g}}(n) = \mathbf{g}_t - \mathbf{g}(n)$ is the weight-error vector. Employing the common *independence assumption* [27], we have:

$$E\{e_a^2(n)\} = E\{\tilde{\mathbf{g}}^T(n)\mathbf{R}_{WT}\tilde{\mathbf{g}}(n)\} = E\{\|\tilde{\mathbf{g}}(n)\|_{\mathbf{R}_{WT}}^2\}, \tag{26}$$

where $\mathbf{R}_{WT}$ is the autocorrelation matrix and $\mathbf{R}_{WT} = E\{\mathbf{z}(n)\mathbf{z}^T(n)\}$. Therefore, to obtain the learning curve, we need to find $E\{\|\tilde{\mathbf{g}}(n)\|_{\mathbf{R}_{WT}}^2\}$ as a function of n. We can recursively obtain $E\{\|\tilde{\mathbf{g}}(n)\|_{\mathbf{\Phi}}^2\}$, where $\mathbf{\Phi}$ is a positive definite symmetric matrix whose dimension is commensurate with that of $\tilde{\mathbf{g}}(n)$.

As is known, the relation among output error, a priori error, and the noise signals is as follows;

$$e(n) = e_a(n) + v(n), \tag{27}$$

where $e_a(n) = \mathbf{z}^T(n)\tilde{\mathbf{g}}(n)$ is the a priori error signal. By substituting Eq. (27) into Eq. (23), the weight error vector update equation can be stated as follows:

$$\tilde{\mathbf{g}}(n+1) = \tilde{\mathbf{g}}(n) - \mu(n)\mathbf{C}(n)\mathbf{z}(n)[\mathbf{z}^T(n)\tilde{\mathbf{g}}(n) + v(n)]. \tag{28}$$

By defining $\mathbf{D}(n) = \mathbf{z}^T(n)\mathbf{C}^T(n)$, the $\mathbf{\Phi}$-weighted Euclidean norm from both sides of Eq. (28) yields:

$$\|\tilde{\mathbf{g}}(n+1)\|_{\mathbf{\Phi}}^2 = \|\tilde{\mathbf{g}}(n)\|_{\mathbf{\Psi}(n)}^2 + \mu^2(n)v^2(n)\mathbf{z}^{\mathbf{\Phi}}(n)$$

$$+\{\text{Cross terms involving one instance of } v(n)\}, \tag{29}$$

where:

$$\mathbf{\Psi}(n) = \mathbf{\Phi} - \mu(n)\mathbf{\Phi}\mathbf{D}^T(n)\mathbf{z}^T(n) - \mu(n)\mathbf{z}(n)\mathbf{D}(n)\mathbf{\Phi}$$

$$+\mu^2(n)\mathbf{z}(n)\mathbf{z}^{\mathbf{\Phi}}(n)\mathbf{z}^T(n), \tag{30}$$

and $\mathbf{z}^{\mathbf{\Phi}}(n) = \mathbf{D}(n)\mathbf{\Phi}\mathbf{D}^T(n)$. Taking the expectation of both sides of Eq. (29) and using Assumptions 1, 2, and 3 yield:

$$E\{\|\tilde{\mathbf{g}}(n+1)\|_{\mathbf{\Phi}}^2\} = E\{\|\tilde{\mathbf{g}}(n)\|_{\mathbf{\Psi}(n)}^2\}$$

$$+E\{\mu^2(n)v^2(n)\mathbf{z}^{\mathbf{\Phi}}(n)\}, \tag{31}$$

where:

$$\mathbf{\Psi}(n) = \mathbf{\Phi} - \mathbf{\Phi}E\{\mu(n)\mathbf{D}^T(n)\mathbf{z}^T(n)\}$$

$$-E\{\mu(n)\mathbf{z}(n)\mathbf{D}(n)\}\mathbf{\Phi}$$

$$+E\{\mu^2(n)\mathbf{z}(n)\mathbf{z}^{\mathbf{\Phi}}(n)\mathbf{z}^T(n)\}. \tag{32}$$

Defining $m(n) = E\{\mu(n)\}$ and $m^2(n) = E\{\mu^2(n)\}$ and using Assumption 3, we obtain:

$$\mathbf{\Psi}(n) = \mathbf{\Phi} - m(n)\mathbf{\Phi}E\{\mathbf{D}^T(n)\mathbf{z}^T(n)\}$$

$$-m(n)E\{\mathbf{z}(n)\mathbf{D}(n)\}\mathbf{\Phi}$$

$$+m^2(n)E\{\mathbf{z}(n)\mathbf{z}^{\mathbf{\Phi}}(n)\mathbf{z}^T(n)\}. \tag{33}$$

Considering only the second term of the right-hand side of Eq. (31), we write:

$$E\{\mu^2(n)v^2(n)\mathbf{z}^{\mathbf{\Phi}}(n)\} = m^2(n)\sigma_v^2 E\{\mathbf{z}^{\mathbf{\Phi}}(n)\}. \tag{34}$$

Therefore, Eq. (31) can be rewritten as follows:

$$E\{\|\tilde{\mathbf{g}}(n+1)\|_{\mathbf{\Phi}}^2\} = E\{\|\tilde{\mathbf{g}}(n)\|_{\mathbf{\Psi}(n)}^2\}$$

$$+m^2(n)\sigma_v^2 E\{\mathbf{z}^{\mathbf{\Phi}}(n)\}. \tag{35}$$

Applying vec(.) operator to both sides of Eq. (33) and using $\text{vec}(\mathbf{P}\mathbf{\Sigma}\mathbf{Q}) = (\mathbf{Q}^T \otimes \mathbf{P})\text{vec}(\mathbf{\Sigma})$ [28] yield:

$$\varphi(n) = \phi - m(n)(E\{\mathbf{z}(n)\mathbf{D}(n)\} \otimes \mathbf{I})$$

$$.\phi - m(n)(\mathbf{I} \otimes E\{\mathbf{z}(n)\mathbf{D}(n)\}).\phi$$

$$+m^2(n)(E\{(\mathbf{z}(n)\mathbf{D}(n)) \otimes (\mathbf{z}(n)\mathbf{D}(n))\}).\phi, \tag{36}$$

where $\varphi(n) = \text{vec}(\mathbf{\Psi}(n))$ and $\phi = \text{vec}(\mathbf{\Phi})$. By defining $M^2 \times M^2$ matrix $\mathbf{P}(n)$ as:

$$\mathbf{P}(n) = \mathbf{I} - m(n)E\{\mathbf{z}(n)\mathbf{D}(n)\} \otimes \mathbf{I}$$

$$-m(n)\mathbf{I} \otimes E\{\mathbf{z}(n)\mathbf{D}(n)\}$$

$$+m^2(n)E\{(\mathbf{z}(n)\mathbf{D}(n)) \otimes (\mathbf{z}(n)\mathbf{D}(n))\}, \tag{37}$$

Eq. (37) becomes:

$$\varphi(n) = \mathbf{P}(n).\phi. \tag{38}$$

The second term on the right-hand side of Eq. (35) can be expressed as follows:

$$E\{\mathbf{z}^{\mathbf{\Phi}}(n)\} = \mathrm{Tr}(E\{\mathbf{D}^T(n)\mathbf{D}(n)\}.\mathbf{\Phi}). \tag{39}$$

Defining γ through:

$$\gamma = \mathrm{vec}(E\{\mathbf{D}^T(n)\mathbf{D}(n)\}), \tag{40}$$

we have:

$$\mathrm{Tr}(E\{\mathbf{D}^T(n)\mathbf{D}(n)\}.\mathbf{\Phi}) = \gamma^T\phi. \tag{41}$$

According to the above relation, the recursion of Eq. (35) can be formulated as follows:

$$E\{||\tilde{\mathbf{g}}(n+1)||^2_\phi\} = E\{||\tilde{\mathbf{g}}(n)||^2_{\mathbf{P}(n)\phi}\} + m^2(n)\sigma_v^2\gamma^T\phi. \tag{42}$$

Focusing again on the learning curve, we substitute $\mathbf{R}_{WT}$ for $\mathbf{\Phi}$, define $\mathbf{r} = \mathrm{vec}(\mathbf{R}_{WT})$, and, write:

$$E\{||\tilde{\mathbf{g}}(n)||^2_{\mathbf{r}}\} = E\{||\tilde{\mathbf{g}}(0)||^2_{\mathbf{P}(0)\mathbf{P}(1)\cdots\mathbf{P}(n-1)\mathbf{r}}\}$$
$$+\sigma_v^2\gamma^T\mathbf{A}\mathbf{r}, \tag{43}$$

where:

$$\mathbf{A} = m^2(0)\mathbf{P}(1)\cdots\mathbf{P}(n-1) + m^2(1)\mathbf{P}(2)\cdots\mathbf{P}(n-1)$$
$$+\cdots+ m^2(n-2)\mathbf{P}(n-1) + m^2(n-1)\mathbf{I}. \tag{44}$$

Based on this recursion, the Excess Mean Square Error (EMSE) can be obtained, when n goes to infinity. The EMSE in the steady state is:

$$\mathrm{EMSE} = \lim_{n\to\infty} E\{e_a^2(n)\}. \tag{45}$$

Based on Eq. (27), the MSE and EMSE are related as follows:

$$\mathrm{MSE} = \mathrm{EMSE} + \sigma_v^2. \tag{46}$$

Therefore, the steady-state EMSE can be stated as follows:

$$\mathrm{EMSE} = m^2(\infty)\sigma_v^2\gamma^T(\mathbf{I} - \mathbf{P}(\infty))^{-1}\mathbf{r}. \tag{47}$$

In addition, the steady-state Mean Square Deviation (MSD) is given by:

$$\mathrm{MSD} = m^2(\infty)\sigma_v^2\gamma^T(\mathbf{I} - \mathbf{P}(\infty))^{-1}\mathrm{vec}(\mathbf{I}). \tag{48}$$

For the theoretical transient and steady-state performances of the step-size, please refer to Appendix B.

5. Performance analysis of VSS-WTDLMS in the nonstationary environment

In the nonstationary environment, the unknown system $(\mathbf{w}_t)$ is assumed time-variant, which changes based on the following random walk model [26,29]:

$$\mathbf{w}_t(n+1) = \mathbf{w}_t(n) + \mathbf{q}(n), \tag{49}$$

where the random sequence of $\mathbf{q}(n)$ is a zero mean and an independent and identically distributed sequence with autocorrelation matrix $\mathbf{Q} = E\{\mathbf{q}(n)\mathbf{q}^T(n)\}$ and independent of $\mathbf{z}(n)$, $\mathbf{v}(n)$, and $\mu(n)$ [29]. Multiplying both sides of Eq. (49) by $\mathbf{T}$, we obtain:

$$\mathbf{g}_t(n+1) = \mathbf{g}_t(n) + \theta(n), \tag{50}$$

where $\mathbf{g}_t(n) = \mathbf{T}\mathbf{w}_t(n)$ and $\theta(n) = \mathbf{T}\mathbf{q}(n)$. Now, the weight error vector update equation of $(\tilde{\mathbf{g}}(n) = \mathbf{g}_t(n) - \mathbf{g}(n))$ in the nonstationary environment can be expressed as follows:

$$\tilde{\mathbf{g}}(n+1) = \tilde{\mathbf{g}}(n) + \theta(n)$$
$$-\mu(n)\mathbf{C}(n)\mathbf{z}(n)[\mathbf{z}^T(n)\tilde{\mathbf{g}}(n) + v(n)]. \tag{51}$$

Taking the $\mathbf{\Phi}$-weighted Euclidean norm and, then, expectation of both sides of Eq. (51) leads to:

$$E\{||\tilde{\mathbf{g}}(n+1)||^2_{\mathbf{\Phi}}\} = E\{||\tilde{\mathbf{g}}(n)||^2_{\mathbf{\Psi}(n)}\} + E\{||\theta(n)||^2_{\mathbf{\Phi}}\}$$
$$+m^2(n)\sigma_v^2 E\{\mathbf{z}^{\mathbf{\Phi}}(n)\}. \tag{52}$$

By following the same strategy in Section 4, Eq. (52) can be expressed as follows:

$$E\{||\tilde{\mathbf{g}}(n+1)||^2_\phi\} = E\{||\tilde{\mathbf{g}}(n)||^2_{\mathbf{P}(n)\phi}\} + E\{||\theta(n)||^2_\phi\}$$
$$+m^2(n)\sigma_v^2\gamma^T\phi. \tag{53}$$

The recently equation is related to $\tilde{\mathbf{g}}(0)$ as in the following:

$$E\{||\tilde{\mathbf{g}}(n)||^2_{\mathbf{r}}\} = E\{||\tilde{\mathbf{g}}(0)||^2_{\mathbf{P}(0)\mathbf{P}(1)\cdots\mathbf{P}(n-1)\mathbf{r}}\}$$
$$+\mathrm{Tr}(\mathbf{Q}_\theta\mathbf{B}\mathbf{r}) + \sigma_v^2\gamma^T\mathbf{A}\mathbf{r}, \tag{54}$$

where $\mathbf{Q}_\theta = E\{\theta(n)\theta^T(n)\}$ and:

$$\mathbf{B} = \mathbf{P}(1)\mathbf{P}(2)\ldots\mathbf{P}(n-1) + \mathbf{P}(2)\mathbf{P}(3)\ldots\mathbf{P}(n-1)$$
$$+\ldots+ \mathbf{P}(n-2)\mathbf{P}(n-1) + \mathbf{P}(n-1) + \mathbf{I}. \tag{55}$$

The steady-state EMSE and MSD in the nonstationary environment are obtained by:

$$\mathrm{EMSE} = m^2(\infty)\sigma_v^2\gamma^T(\mathbf{I} - \mathbf{P}(\infty))^{-1}\mathbf{r}$$
$$+\mathrm{Tr}(\mathbf{Q}_\theta\mathrm{vec}((\mathbf{I} - \mathbf{P}(\infty))^{-1}\mathbf{r})), \tag{56}$$

and:

$$\mathrm{MSD} = m^2(\infty)\sigma_v^2\gamma^T(\mathbf{I} - \mathbf{P}(\infty))^{-1}\mathrm{vec}(\mathbf{I})$$
$$+\mathrm{Tr}(\mathbf{Q}_\theta\mathrm{vec}((\mathbf{I} - \mathbf{P}(\infty))^{-1}\mathrm{vec}(\mathbf{I}))). \tag{57}$$

6. Computational complexity

Table 2 describes the computational complexity of VSS-WTDLMS algorithms. The number of multiplications and divisions have been calculated for each term. In the following, Table 3 compares the computational complexity of various VSS-TDLMS algorithms. These algorithms are derived from [6,10,11,17]. In this table, M is the number of filter coefficients, N is the number of subbands, M_t is the number of past values of the ith transform coefficient, and L is the number of past squared values of the error. In comparison with WTDLMS, the VSS-WTDLMS needs only three more multiplications and one division. The VSS-WTDLMS has also lower computational complexity than other VSS-TDLMS algorithms. It is interesting to note that using the HWT leads to the only $3M+3N+3$ multiplications, significantly useful in some applications such as AEC.

7. Simulation results

We demonstrated the performance of the proposed algorithm by several computer simulations in system identification and AEC scenarios. For the system identification, the unknown impulse response is randomly selected with 16 taps ($M = 16$), and the input signal is an AR(1) signal generated by passing a zero-mean white Gaussian noise through a first-order system of $H(z) = \frac{1}{1-0.9z^{-1}}$. In AEC, the input signal is the real speech, and the unknown system is the car echo path with $M = 256$. An additive white Gaussian noise, with variance $\sigma_v^2 = 10^{-3}$, is added to the system output, setting the signal-to-noise ratio (SNR) to 30 dB. The HWT is used in all simulations, leading to the reduction of computational complexity due to the elements (+1 and -1) in HWT. Parameter C was set to 10^{-3}, and the values of α and β were set to 0.994 and 0.9, respectively. Based on the mean square stability

**Table 2.** The computational complexity of VSS-WTDLMS algorithm.

Equation	In subband		Total	
	$\times$	$\div$	$\times$	$\div$
$\mathbf{z}(n) = \mathbf{T}\mathbf{x}(n)$	-	-	M^2	-
$e(n) = d(n) - \mathbf{g}^T(n)\mathbf{z}(n)$	-	-	M	-
$\|\hat{\mathbf{p}}(n)\|^2 = \beta^2\|\hat{\mathbf{p}}(n-1)\|^2 + (1-\beta)^2 e^2(n)$	-	-	3	-
$\mu(n) = \frac{\|\hat{\mathbf{p}}(n)\|^2}{\|\hat{\mathbf{p}}(n)\|^2 + \sigma_v^2}$	-	-	-	1
$\sigma_{h_i}^2(n) = \alpha\sigma_{h_i}^2(n-1) + (1-\alpha)\|\mathbf{z}_{h_i}(n)\|^2$	$\frac{M}{N}+2$	-	$M+2N$	-
$\mathbf{g}_{h_i}(n+1) = \mathbf{g}_{h_i}(n) + \mu(n)\frac{\mathbf{z}_{h_i}(n)}{\epsilon+\sigma_{h_i}^2(n)}e(n)$	$\frac{M}{N}+1$	1	$M+N$	N

Total Multiplications : $M^2 + 3M + 3N + 3$

Table 3. The computational complexity of various VSS-WTDLMS.

Algorithm	Multiplications	Divisions
DCT-LMS [6]	$M^2 + (M_t+4)M + 1$	$2M$
VSS-TDLMS [10]	$M^2 + 5M + L + 2$	$M+1$
VSS-TDLMS [11]	$M^2 + 8M + 8$	$M+1$
VSS-TDLMS [13]	$M^2 + 5M + 8$	$M+1$
WTDLMS	$M^2 + 3M + 3N$	N
VSS-WTDLMS	$M^2 + 3M + 3N + 3$	$N+1$
VSS-WTDLMS (HWT)	$3M + 3N + 3$	$N+1$

Table 4. Stability bounds of the WTDLMS adaptive algorithm for different values of N.

Algorithm	$\frac{1}{\lambda_{\max}(\mathbf{M}^{-1}\mathbf{N})}$	$\frac{1}{\max(\lambda(\mathbf{H})\in\Re^+)}$	$\boldsymbol{\mu}_{\max}$
WTDLMS ($N=2$)	0.8233	2.4208	0.8233
WTDLMS ($N=4$)	0.3925	1.0996	0.3925
WTDLMS ($N=8$)	0.1977	0.5531	0.1977
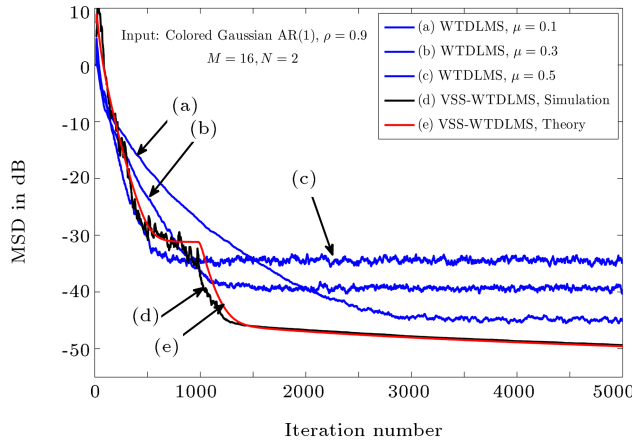
Figure 2. The MSD learning curves of WTDLMS and VSS-WTDLMS algorithms with $N=2$ (input signal: AR(1)).
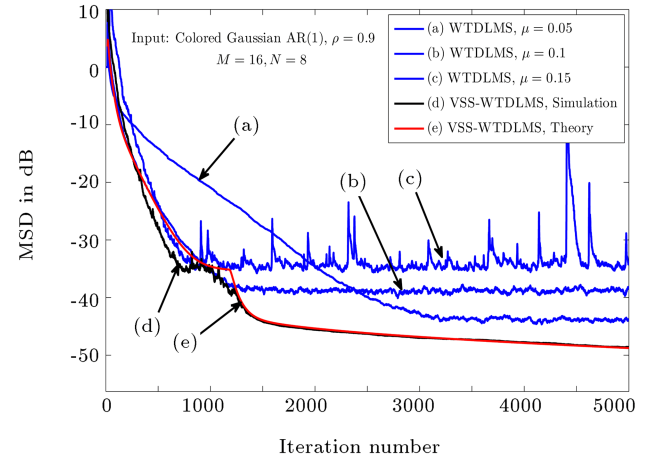
Figure 4. The MSD learning curves of WTDLMS and VSS-WTDLMS algorithms with $N=8$ (input signal: AR(1)).
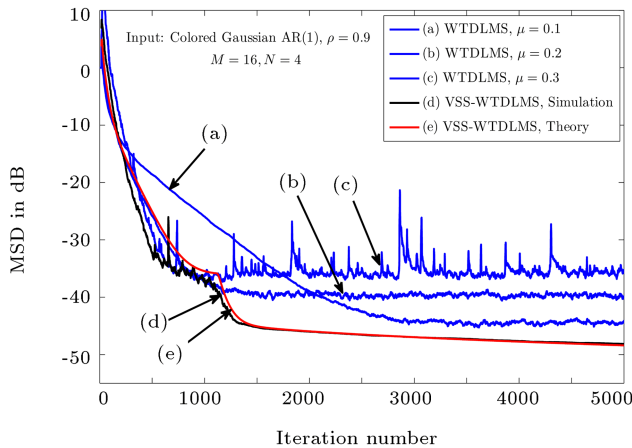
Figure 3. The MSD learning curves of WTDLMS and VSS-WTDLMS algorithms with $N=4$ (input signal: AR(1)).

analysis in Appendix A, the values of $\mu_{\max}$ are selected in VSS-WTDLMS. Table 4 shows $\mu_{\max}$ for $N=2$, 4, and 8. These values have been obtained from Eq. (A.3).

Figures 2–4 show the MSD learning curves of the proposed VSS-WTDLMS and conventional WTDLMS algorithm for different values of N. In WTDLMS, different values of the step-size are selected. Compared to the conventional WTDLMS algorithm, it is observed that VSS-WTDLMS has a faster convergence speed

and lower steady-state error for all values of N. The theoretical learning curves for VSS-WTDLMS are also presented in these figures. The theoretical learning curves are obtained from Eq. (43). Good agreement between simulated and theoretical learning curves can be seen.
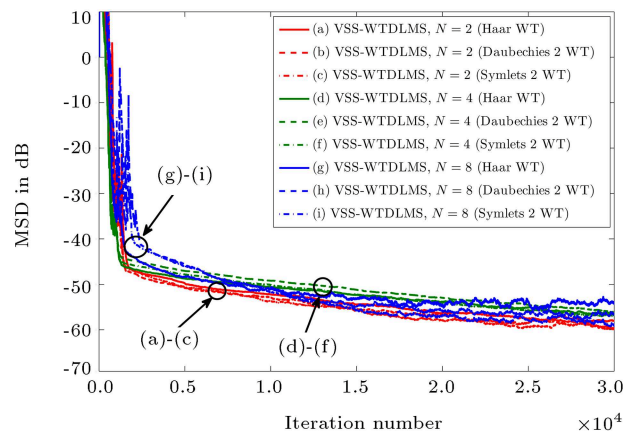
Figure 5 presents the theoretical and simulated values of the step-size during adaptation. The theoretical transient values for the step-size are obtained from Eqs. (19) and (B.1). In addition, the theoretical steady-state step-size is calculated through Eq. (B.4). To clear the steady-state values, we have shown the last values from iteration 1500 to 5000 in Figure 5. Again, good agreement can be seen for different values of N. A comparison between VSS-WTDLMS and recently and famous VSS-TDLMS algorithms is made, as shown in Figure 6 [6,10,11,17]. The parameters in these algorithms are set according to Table 5. This figure indicates that the VSS-WTDLMS has better performance than other algorithms for all values of N. In addition, the computational complexity of the proposed algorithm is lower than that of other TDLMS algorithms due to HWT. In Figure 7, we have justified the performance of VSS-WTDLMS for various WTs. The Haar, daubechies 2 (db2), symlets 2 (sym2) WTs were applied to the transformation matrix. The simulation results show that the performances of VSS-WTDLMS with different WTs are approximately close

Table 5. The parameters in VSS-TDLMS and VSS-WTDLMS algorithms.

DCT-LMS [6]	VSS-TDLMS [10]		
$\beta = 0.9985, \quad \gamma = 8 \times 10^{-3},$	$\alpha = 0.99, \quad \beta = 0.9, \quad \gamma = 10^{-3},$		
$M_t = 10, \quad \epsilon = 2.5 \times 10^{-2}.$	$L = 10, \quad \epsilon = 2.5, \times 10^{-2},$		
	$\mu_{\max} = 5 \times 10^{-2}, \quad \mu_{\min} = 4.7 \times 10^{-3}$		
VSS-TDLMS [11]	**VSS-TDLMS [13]**		
$\beta = 0.98, \quad \gamma = 0.98,$	$\alpha = 0.995, \quad \beta = 0.9, \quad \gamma = 0.9, \quad \mu_{\max} = 0.5,$		
$\epsilon = 2.5 \times 10^{-2}$	$C = M \times \sigma_\nu^2, \quad \epsilon = 2.5 \times 10^{-2}$		
VSS-WTDLMS			
$\alpha = 0.994, \quad \beta = 0.9, \quad \epsilon = 2.5, \times 10^{-2}$			
$C = \sigma_\nu^2, \quad \mu_{\max} = 0.7, 0.3, 0.15, \quad (N = 2, 4, 8).$			

Figure 5. The simulated and theoretical transient and steady-state values of the step-size for different values of N.
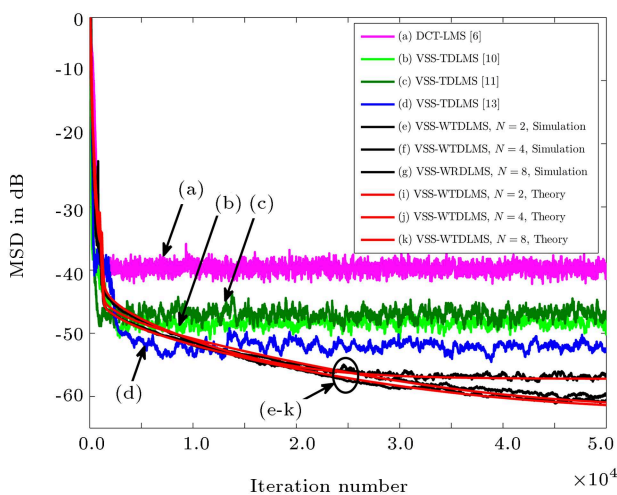
Figure 6. The MSD learning curves of various VSS-TDLMS and VSS-WTDLMS algorithms (input signal: AR(1)).

Figure 7. The MSD learning curves of VSS-TDLMS algorithm with different WTs (input signal: AR(1)).

to each other. In Figure 8, the performance of VSS-WTDLMS is studied for different values of C. Parameter N is set to 2, and the values of C change from 0.01 to 0.0005. The results show that, for large values

of C, the performance is deviated. The learning curves are compared with other algorithms. As observed, the VSS-WTDLMS has performance better than other VSS-TDLMS algorithms. Figure 9 presents the simulated and theoretical learning curves for different values of SNR. It is observed that, for all values of SNR, the proposed VSS-WTDLMS works well.
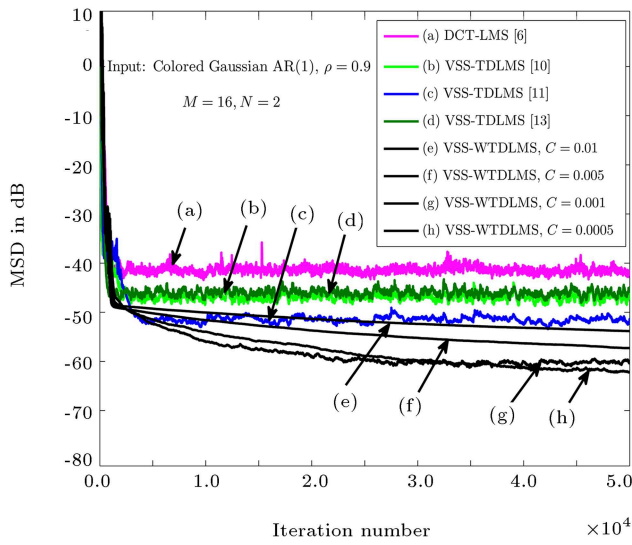
Figure 8. The MSD learning curves of various VSS-TDLMS and VSS-WTDLMS algorithms for different values of C (input signal: AR(1)).
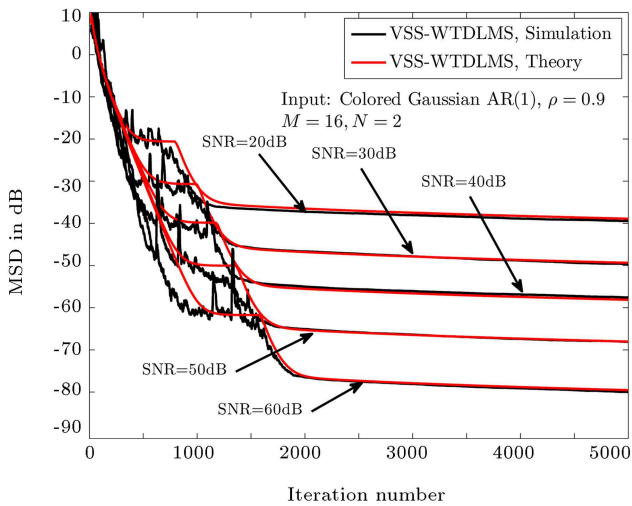

Figure 9. The MSD learning curves of VSS-WTDLMS algorithm for different values of SNR (input signal: AR(1)).

Figure 10. The MSD learning curves of various VSS-TDLMS and VSS-WTDLMS with $N = 8$ for tracking performance analysis (input signal: AR(1)).

Figure 11. The simulated and theoretical MSD learning curves of VSS-TDLMS with $N = 2$ for different values of σ_q^2.

Figure 12. The simulated and theoretical MSD values versus different values of σ_q^2.

In Figure 10, the tracking performance of VSS-WTDLMS is studied. During the adaptation, the unknown system changes in the middle of the iteration. The results show that the VSS-WTDLMS has a tracking ability better than other algorithms. It is clear that the Recursive Least Squares (RLS) algorithm has a faster convergence speed than VSS-WTDLMS. However, when the system changes, the performance of RLS is deviated. The performance of WTDLMS in the nonstationary environment is demonstrated in Figures 11 and 12. The unknown system changes according to the random walk model. We assume an independent and identically distributed sequence for $\mathbf{q}(n)$ with autocorrelation matrix $\mathbf{Q} = \sigma_q^2\mathbf{I}$ [26]. To obtain various degrees of nonstationary, different
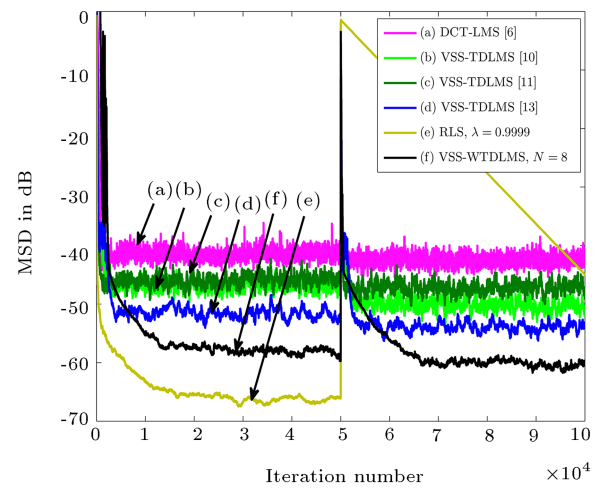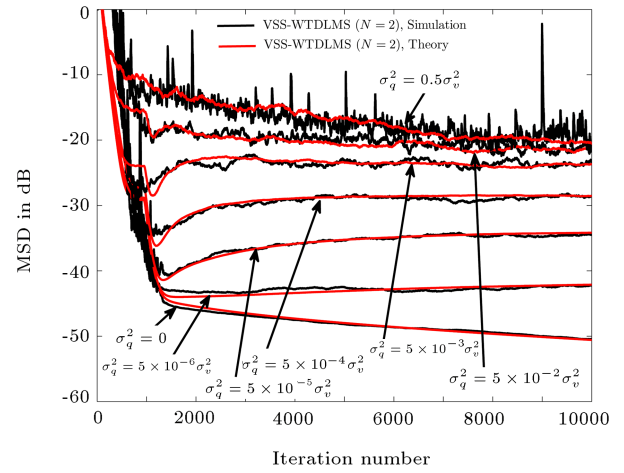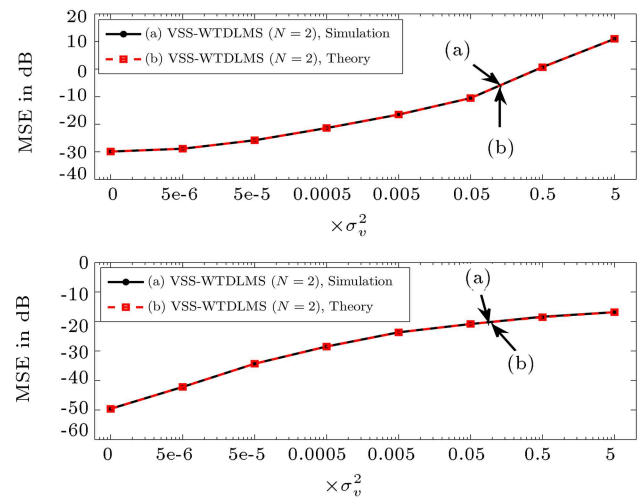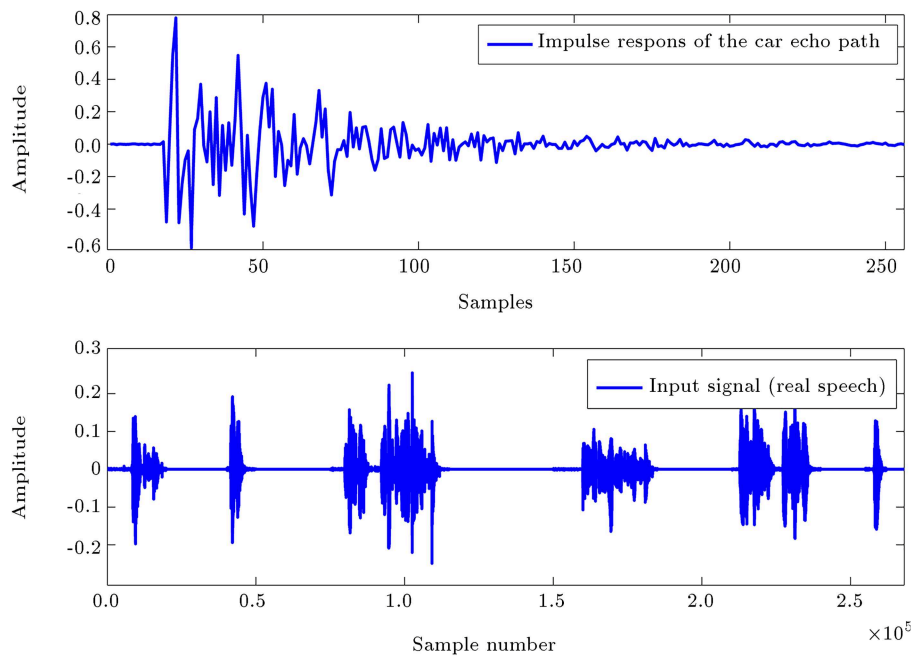
Figure 13. The impulse response of the car echo path and real speech input signal.

values for σ_q^2 have been chosen. Figure 11 presents the simulated and theoretical MSD learning curves of VSS-WTDLMS with $N = 2$. The theoretical learning curves are obtained from Eq. (54). Accordingly, when the variance of $\mathbf{q}(n)$ increases, the performance of VSS-WTDLMS is deviated. In addition, the simulated and theoretical learning curves have good agreement. It means that the derived theoretical relation in Eq. (54) has good ability to predict the performance of the VSS-WTDlMS algorithm. In Figure 12, the simulated and theoretical MSD and MSE versus different values of σ_q^2 are presented. As observed, the theoretical and simulated values have good agreement.

Figure 13 shows the impulse response of the car echo path and real speech input signal in AEC setup. The variance of the additive noise is set to 10^{-9}. The MSD learning curves of various VSS-TDLMS and proposed VSS-WTDLMS are presented in Figure 14. This figure indicates that some VSS-TDLMS algorithms for the real speech input signal diverge. For all values of N, the VSS-WTDLMS shows better performance than other VSS-TDLMS algorithms. Figure 15 presents the error signals in VSS-WTDLMS. The variation of the step-size for different values of N during the adaptation has been presented in Figure 16. In Figures 17 and 18, the variances of the additive noise increase to 10^{-6} and 10^{-3}. In comparison with other VSS-TDLMS algorithms, the VSS-WTDLMS has a faster convergence speed and lower steady-state error.

Finally, Figure 19 shows the number of filter coefficients versus the filter length for VSS-TDLMS, WTDLMS, and VSS-WTDLMS algorithms. This fig-
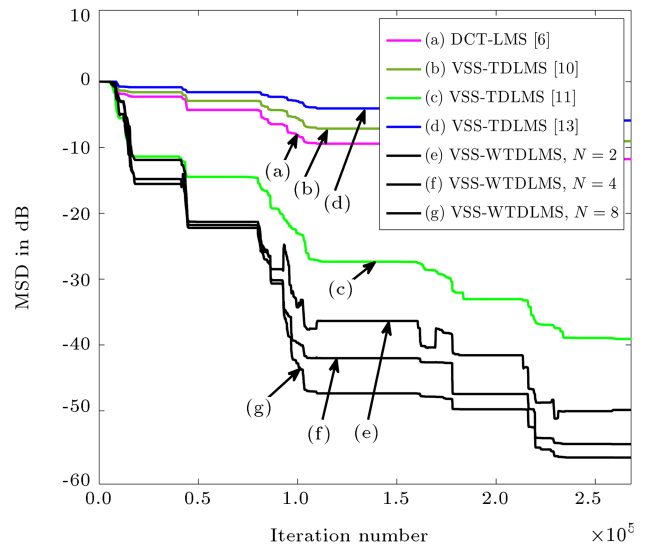

Figure 14. The MSD learning curves of various VSS-TDLMS and VSS-WTDLMS algorithms, $\sigma_v = 10^{-9}$ (Input signal: real speech).

ure shows that the computational complexity of VSS-WTDLMS is lower than that of other algorithms. As is observed, this difference is significantly large for HWT.

8. Conclusion

In this paper, the Wavelet Transform Domain Least Mean Square (WTDLMS) combined with Variable Step-Size (VSS) was established. The step-size changed according to the largest decrease of mean square deviation. In addition, the mean square performance
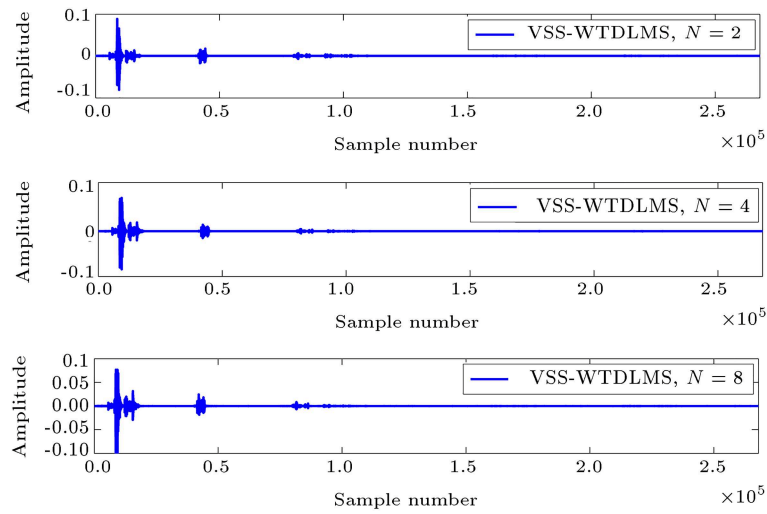
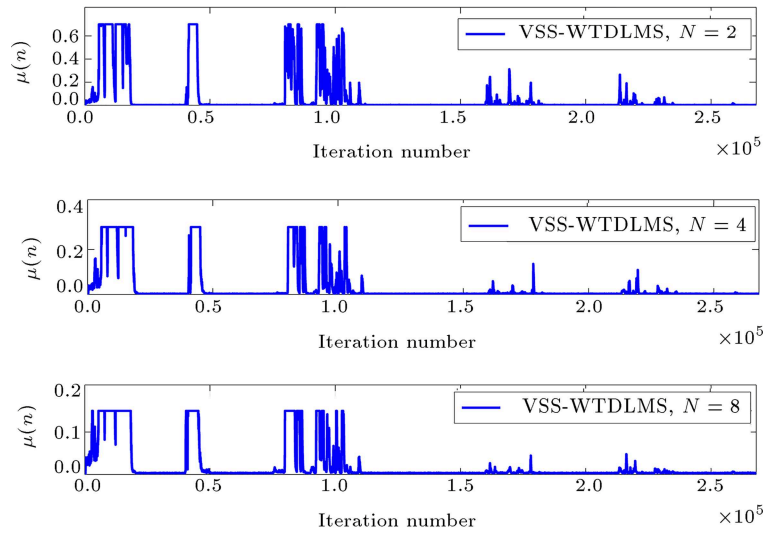**Figure 15.** Error signals in VSS-WTDLMS with $N = 2$, 4, and 8.

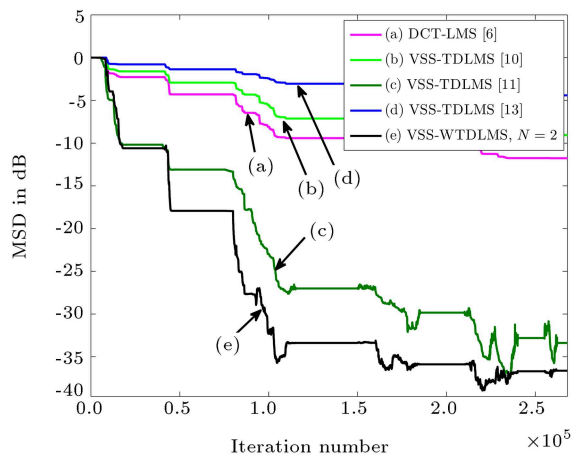**Figure 16.** Variation of the step-size during the adaptation for VSS-WTDLMS in AEC.

Figure 17. The MSD learning curves of various VSS-TDLMS and VSS-WTDLMS algorithms, $\sigma_v^2 = 10^{-6}$ (input signal: real speech).
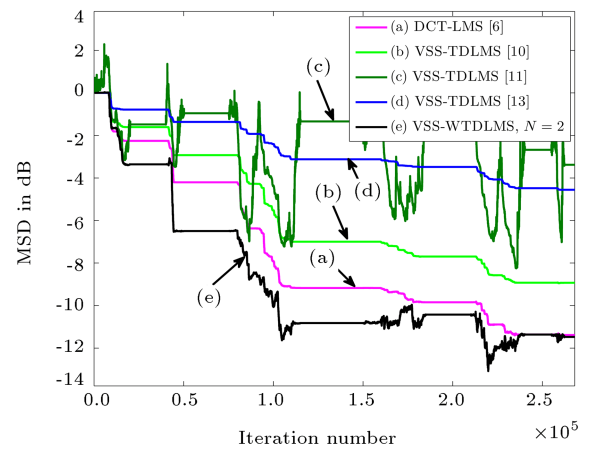
Figure 18. The MSD learning curves of various VSS-TDLMS and VSS-WTDLMS algorithms, $\sigma_v^2 = 10^{-3}$ (input signal: real speech).
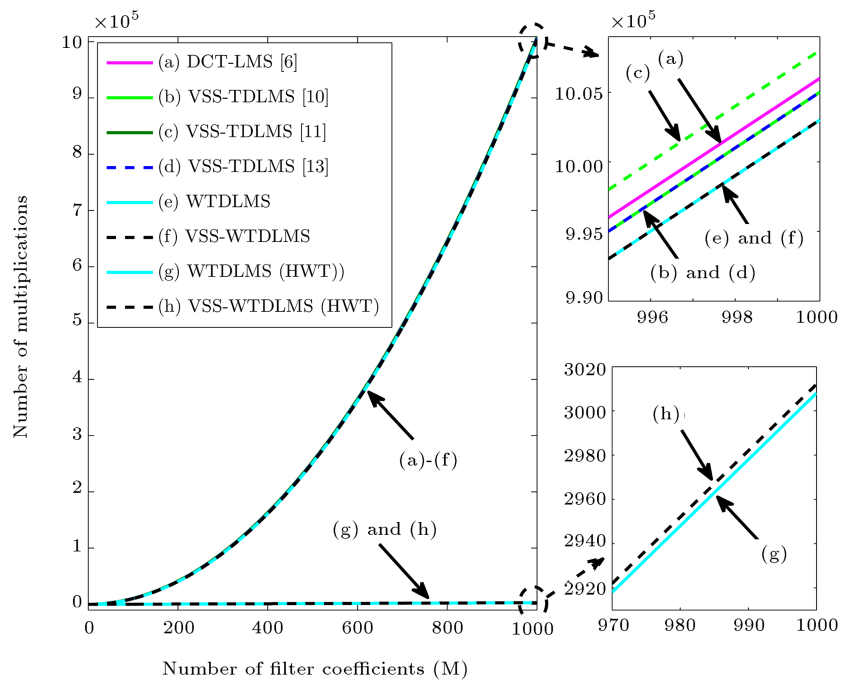
Figure 19. The number of filter coefficients versus the filter length (M) in various VSS-TDLMS, WTDLMS, and VSS-WTDLMS algorithms.

of the VSS-WTDLMS was studied, and theoretical relations for transient and steady-state performances of this algorithm were established in stationary and nonstationary environments. The good performance of the VSS-WTDLMS and the validity of the theoretical relations were confirmed by several simulation results.

References

1. Widrow, B. and Stearns D., *Adaptive Signal Process-ing*, Englewood Cliffs, NJ Prentice Hall, Inc (1985).

2. Haykin, S.S., *Adaptive Filter Theory*, Pearson Educa-tion India, 5th Edn. (2013).

3. Sayed, A.H., *Adaptive Filters*, John Wiley & Sons, Inc. (2008).

4. Farhang-Boroujeny, B., *Adaptive Filters: Theory and Applications*, John Wiley & Sons, Ltd (2013).

5. Narayan, S.S., Peterson, A.M., and Narashima, M.J. "Transform domain LMS algorithm" *IEEE Trans. Acoust., Speech, Signal Processing*, pp. 609–615 (1983).

6. Kim, D.I. and Wilde, P.D. "Performance analysis of the DCT-LMS adaptive filtering algorithm", *Signal Processing*, **80**, pp. 1629–1645 (2000).

7. Zhao, S., Man, Z., Khoo, S., and Wu, H. "Stability and convergence analysis of transform-domain LMS adaptive filters with second-order autoregressive pro-cess" *IEEE Trans. Signal Processing*, **57**, pp. 119–130 (2009).

8. Attallah, S. "The wavelet transform-domain LMS algorithm: a more practical approach," *IEEE Trans.*

Circuits, Syst. II: Analog and Digital Signal Process-ing, **47**(3), pp. 209–213 (2000).

9. Attallah, S. "The wavelet transform-domain LMS adaptive filter with partial subband-coefficient updat-ing", *IEEE Trans. Circuits Syst. II Express Briefs*, **53**(1), pp. 8–12 (2006).

10. Bilcu, R.C., Kuosmanen, P., and Egiazarian, K. "A transform domain LMS adaptive filter with variable step-size", *IEEE Signal Process. Lett.*, **9**(2), pp. 51–53 (2002).

11. Mayyas, K. "A transform domain LMS algorithm with an adaptive step size equation', *Proceedings of the Fourth IEEE International Symposium on Signal Processing and Information Technology*, pp. 229–232 (2004).

12. Huang, F., Zhang, J., and Zhang, S. "Combined-step-size affine projection sign algorithm for robust adaptive filtering in impulsive interference environments" *IEEE Trans. Circuits, Syst. II: Express Briefs*, **63**(5), pp. 493–497 (2016).

13. Zhang, S., Zheng, W.X., and Zhang, J. "A new combined-step-size normalized least mean square algo-rithm for cyclostationary inputs", *Signal Processing*, **141**(12), pp. 261–272 (2017).

14. Huang, F., Zhang, J., and Pang, Y. "A novel combina-tion scheme of proportionate filter", *Signal Processing*, **143**(2), pp. 222–231 (2018).

15. Shams Esfand Abadi, M., Mehrdad, V., Gholipour, A., and Noroozi, M. "A family of variable step-size affine projection adaptive filtering algorithms", *International Journal of Computer and Electrical Engineering*, **2**(3), pp. 447–459 (2010).

16. Shin, H.C., Sayed, A.H., and Song, W.J. "Variable step-size NLMS and affine projection algorithms", *IEEE Signal Processing Letters*, **11**, pp. 132–135 (2004).

17. Zhao, S., Jones, D.L., Khoo, S., and Man, Z. "New variable step-sizes minimizing mean-square deviation for the LMS-type algorithms", *Circuits, Systems, and Signal Processing*, **33**, pp. 2251–2265 (2014).

18. Shams Esfand Abadi, M., Far, A.M., and Moussavi, S.Z. "Variable step-size block normalized least mean square adaptive filter: A unified framework", *International Journal of Science and Technology, Scientia Iranica*, **15**, pp. 195–202 (2008).

19. Shams Esfand Abadi, M., Mehrdad, V., and Gholipour, A. "Family of variable step-size affine projection adaptive filtering algorithms with selective regressors and selective partial update", *International Journal of Science and Technology, Scientia Iranica*, **17**, pp. 81–98 (2010).

20. Shams Esfand Abadi, M. and Shafiee, M.S. "The new normalized subband adaptive filter algorithm with variable step-size" *Automatika, Journal for Control, Measurement, Electronics, Computing and Communications*, **55**(2), pp. 188–198 (2014).

21. Lee, H.S., Kim, S.E., Lee, J.W., and Song, W.J. "A variable step-size diffusion LMS algorithm for distributed estimation", *IEEE Trans. Signal Processing*, **63**(7), pp. 1808–1820 (2015).

22. Huang, H.C. and Lee, J. "A new variable step-size NLMS algorithm and its performance analysis", *IEEE Trans. Signal Processing*, **60**(4), pp. 2055–2060 (2012).

23. Kwong, R.H. and Johnston, E.W. "A variable step size LMS algorithm", *IEEE Trans. Signal Processing*, **40**(7), pp. 1633–1642 (1992).

24. Koike, S. "A class of adaptive step-size control algorithms for adaptive filters", *IEEE Trans. Signal Processing*, **50**(6), pp. 1315–1326 (2002).

25. Mayyas, K. and Momani, F. "An LMS adaptive algorithm with a new step-size control equation", *Journal of the Franklin Institute*, **348**(4), pp. 589–605 (2011).

26. Shin, H.C. and Sayed, A.H. "Mean-square performance of a family of affine projection algorithms", *IEEE Trans. Signal Processing*, **52**(1), pp. 90–102 (2004).

27. Haykin, S., *Adaptive Filter Theory*, NJ: Prentice-Hall, 4th edition (2002).

28. Moon, T.K. and Sterling, W.C., *Mathematical Methods and Algorithms for Signal Processing*, Upper Saddle River: Prentice Hall (2000).

29. Sayed, A.H., *Fundamentals of Adaptive Filtering*, Wiley (2003).

Appendix A

Mean square stability analysis of WTDLMS

To study the theoretical performance of WTDLMS, the fixed step-size is replaced in the VSS-WTDLMS relations. Following the analysis, we obtain:

$$E\{||\tilde{\mathbf{g}}(n+1)||_\phi^2\} = E\{||\tilde{\mathbf{g}}(n)||_{\mathbf{P}\phi}^2\} + \mu^2\sigma_v^2\varphi^T\phi, \quad \text{(A.1)}$$

where:

$$\mathbf{P} = \mathbf{I} - \mu E\{\mathbf{z}(n)\mathbf{D}(n)\} \otimes \mathbf{I} - \mu\mathbf{I} \otimes E\{\mathbf{z}(n)\mathbf{D}(n)\}$$

$$+\mu^2 E\{(\mathbf{z}(n)\mathbf{D}(n)) \otimes (\mathbf{z}(n)\mathbf{D}(n))\}. \quad \text{(A.2)}$$

Eq. (A.1) is stable if matrix $\mathbf{P}$ is stable [26]. Based on Eq. (A.2), it is known that $\mathbf{P} = \mathbf{I} - \mu\mathbf{M} + \mu^2\mathbf{N}$, where $\mathbf{M} = E\{\mathbf{z}(n)\mathbf{D}(n)\} \otimes \mathbf{I} + \mathbf{I} \otimes E\{\mathbf{z}(n)\mathbf{D}(n)\}$, and $\mathbf{N} = E\{(\mathbf{z}(n)\mathbf{D}(n)) \otimes (\mathbf{z}(n)\mathbf{D}(n))\}$. The condition on μ to guarantee the convergence in the mean-square sense of the WTDLMS algorithms is as follows:

$$0 < \mu < \min\left\{\frac{1}{\lambda_{\max}(\mathbf{M}^{-1}\mathbf{N})}, \frac{1}{\max(\lambda(\mathbf{H}) \in \Re^+)}\right\}, \quad \text{(A.3)}$$

where $\mathbf{H} = \begin{bmatrix} \frac{1}{2}\mathbf{M} & -\frac{1}{2}\mathbf{N} \\ \mathbf{I} & \mathbf{0} \end{bmatrix}$.

Appendix B

The theoretical transient and steady-state step-size

The transient study of step-size is obtained from Eq. (19), where:

$$E[||\hat{\mathbf{p}}(n)||^2] = \beta^2 E[||\hat{\mathbf{p}}(n-1)||^2]$$

$$+(1-\beta)^2(E[||\tilde{\mathbf{g}}(n)||_{\mathbf{R}_{WT}}^2] + \sigma_v^2). \quad \text{(B.1)}$$

Under Assumption 4, when n goes to infinity, we have:

$$E[||\hat{\mathbf{p}}(\infty)||^2] = \beta^2 E[||\hat{\mathbf{p}}(\infty)||^2] + (1-\beta)^2\sigma_v^2. \quad \text{(B.2)}$$

Therefore:

$$E[||\hat{\mathbf{p}}(\infty)||^2] = \frac{1-\beta}{1+\beta}\sigma_v^2. \quad \text{(B.3)}$$

According to Eq. (19), the steady-state value for the step-size is given by:

$$m(\infty) = \frac{E[||\hat{\mathbf{p}}(\infty)||^2]}{E[||\hat{\mathbf{p}}(\infty)||^2] + \sigma_v^2} = \frac{1-\beta}{2}. \quad \text{(B.4)}$$

Biographies

Mohammad Shams Esfand Abadi received the BS degree in Electrical Engineering from Mazandaran University, Mazandaran, Iran and the MS degree in Electrical Engineering from Tarbiat Modares University, Tehran, Iran in 2000 and 2002, respectively. He obtained the PhD degree in Biomedical Engineering

from Tarbiat Modares University, Tehran, Iran in 2007. Since 2004, he has been with the Faculty of Electrical Engineering, Shahid Rajaee Teacher Training University, Tehran, Iran, where he is currently an Associate Professor. His research interests include digital filter theory, adaptive distributed networks, and adaptive signal processing algorithms.

Hamid Mesgarani received the MSc and PhD degrees in Applied Mathematics from Iran University of Science and Technology, Tehran, Iran in 1996 and 2002, respectively. He is an Associate Professor at the Department of Mathematics at the Shahid Rajaee Teacher Training University, Tehran, Iran. His research interests include integral equations and wavelet functions.

Seyed Mahmoud Khademiyan received the MSc degree in Applied Mathematics from Iran University of Science and Technology, Tehran, Iran in 2012. Currently, he is a PhD candidate at the Department of Mathematics, Faculty of Science, Shahid Rajaee Teacher Training University, Tehran, Iran. His research interests include digital filter theory and adaptive signal processing algorithms.