STOCHASTIC MODELLING OF THE TRANSFORM DOMAIN εLMS ALGORITHM FOR A TIME-VARYING ENVIRONMENT

Elen M. Lobato, Orlando J. Tobias, and Rui Seara

LINSE – Circuits and Signal Processing Laboratory
Department of Electrical Engineering
Federal University of Santa Catarina
88040-900 – Florianópolis – SC – Brazil
E-mails: {elen, orlando, seara}@linse.ufsc.br

ABSTRACT

This paper presents a stochastic model of the transform domain εLMS (εTDLMS) algorithm for a nonstationary environment. The proposed model is derived for Gaussian inputs, high-order adaptive filter and slow adaptation condition. Through simulations, we can verify a very good agreement between the results obtained by the Monte Carlo method and the predictions from the proposed analytical model.

1. INTRODUCTION

Due to simplicity, robustness and low computational complexity the LMS algorithm is widely used in adaptive filtering applications [1]-[3]. This algorithm however presents a severe limitation when the input signal is correlated; its speed of convergence is very slow. Narayan *et al*. [4] have proposed the LMS algorithm in the transform domain (TDLMS) as an alternative solution to the convergence problem of the ordinary LMS algorithm. TDLMS consists simply of the standard LMS algorithm but having its input signal preprocessed by an orthogonal transform followed by a power normalization operation. In practical applications, the transform domain εLMS (εTDLMS) algorithm is used instead of the ordinary TDLMS one. The former allows overcoming problems of insufficient spectral excitation of the input signal, preventing thus possible instability of the adaptive algorithm.

The statistical analysis of the TDLMS algorithm for stationary environment can be found in [5]-[8]. However, for nonstationary environment only a few analyses are presented in the open literature. In [9] an approach for nonstationary algorithm analysis by using energy relations is discussed. An interesting feature of such an analysis is its independence on both particular data nonlinearity and Gaussian inputs. However, the energy-based approach does not permit to obtain theoretical models to predict the evolution of the mean weight behavior as well as the learning curve.

The aim of this paper is to present stochastic models for the first and second moment of the adaptive filter weight vector of the εTDLMS algorithm. The derived expressions are obtained by considering a system identification problem with a time-varying plant. Section 2 presents the model used for a time-varying plant. In Section 3 the analytical expressions for the first and second moments of the adaptive weight vector are derived. Section 4 shows some simulation results, which ratify the proposed statistical model. Finally, in Section 5 some conclusions of this paper are presented.

2. PROBLEM STATEMENT

Let us consider a system identification problem in which the plant is time-varying with its output given by

$$d(n) = \mathbf{x}^{\mathrm{T}}(n)\mathbf{w}^{\mathrm{o}}(n) + z(n) \; , \qquad (1)$$

where $\mathbf{x}(n) = [x(n) \;\; x(n-1) \cdots x(n-N+1)]^{\mathrm{T}}$ denotes the input signal vector, being $\{x(n)\}$ a Gaussian, zero-mean, and stationary process. Vector $\mathbf{w}^{\mathrm{o}}(n) = [w_0^{\mathrm{o}}(n) \;\; w_1^{\mathrm{o}}(n) \cdots w_{N-1}^{\mathrm{o}}(n)]^{\mathrm{T}}$ represents the time-varying plant. The measurement noise $z(n)$ is i.i.d., zero-mean with variance σ_z^2, and uncorrelated with any other signal in the system. The time-varying plant vector in the transform domain is given by $\mathbf{w}_{\mathrm{T}}^{\mathrm{o}}(n) = \mathbf{T}\mathbf{w}^{\mathrm{o}}(n) = [\overline{w}_0^{\mathrm{o}}(n) \;\; \overline{w}_1^{\mathrm{o}}(n) \cdots \overline{w}_{N-1}^{\mathrm{o}}(n)]^{\mathrm{T}}$ with $\mathbf{T}$ being the orthogonal transform. Similarly, the input signal vector in the transform domain is $\mathbf{x}_{\mathrm{T}}(n) = \mathbf{T}\mathbf{x}(n) = [\overline{x}_0(n) \;\; \overline{x}_1(n) \cdots \overline{x}_{N-1}(n)]^{\mathrm{T}}$. By using the transformed vectors, we can rewrite (1) as follows:

$$d(n) = \mathbf{x}_{\mathrm{T}}^{\mathrm{T}}(n)\mathbf{w}_{\mathrm{T}}^{\mathrm{o}}(n) + z(n) \, . \qquad (2)$$

The purpose of the adaptive algorithm is to follow the variations of $\mathbf{w}_{\mathrm{T}}^{\mathrm{o}}(n)$, which are governed by

$$\mathbf{w}_{\mathrm{T}}^{\mathrm{o}}(n+1) = \mathbf{w}_{\mathrm{T}}^{\mathrm{o}}(n) + \mathbf{g}(n) \; , \qquad (3)$$

where vector $\mathbf{g}(n)$ denotes the plant perturbation process, which is white, zero-mean with variance σ_g^2.

3. ANALYSIS

In this section we derive analytical expressions for the first and second moments of the adaptive weight vector considering (3). Let us start by considering the weight update equation in the transform domain, given by [8]

$$\mathbf{w}_{\mathrm{T}}(n+1) = \mathbf{w}_{\mathrm{T}}(n) + 2\mu\mathbf{D}^{-1}(n)e(n)\mathbf{x}_{\mathrm{T}}(n) \; , \qquad (4)$$

where $\mathbf{w}_{\mathrm{T}}(n) = [\overline{w}_0(n) \;\; \overline{w}_1(n) \cdots \overline{w}_{N-1}(n)]^{\mathrm{T}}$ represents the adaptive filter weight vector, and $\mathbf{D}(n) = \mathrm{diag}[\overline{\sigma}_0^2(n) \;\; \overline{\sigma}_1^2(n) \cdots \;\; \overline{\sigma}_{N-1}^2(n)]$ is the step-size normalizing matrix, with the elements recursively obtained by [8]

$$\overline{\sigma}_i^2(n) \cong \overline{\sigma}_i^2(n-1) + \frac{1}{M}\left[\overline{x}_i^2(n) - \overline{\sigma}_i^2(n-1)\right], \;\; i = 0,1,\ldots,N-1, \quad (5)$$

where M is the observation window length.

In the practical algorithm a small positive regularization parameter ε is added to (5). This parameter, considered in the model expressions, prevents division by zero and stabilizes the solution.

The error signal is obtained by

$$e(n) = d(n) - y(n) \, . \qquad (6)$$

Substituting (2) and (6) into (4), we get

$$\mathbf{w}_\mathrm{T}(n+1)=\mathbf{w}_\mathrm{T}(n) + 2\mu\mathbf{D}^{-1}(n)\mathbf{x}_\mathrm{T}(n)[\mathbf{x}_\mathrm{T}^\mathrm{T}(n)\mathbf{w}_\mathrm{T}^\mathrm{o}(n)+z(n)-\mathbf{x}_\mathrm{T}^\mathrm{T}(n)\mathbf{w}_\mathrm{T}(n)]. \quad (7)$$

By defining the weight-error vector in the transform domain as $\mathbf{v}_\mathrm{T}(n+1)=\mathbf{w}_\mathrm{T}(n+1)-\mathbf{w}_\mathrm{T}^\mathrm{o}(n+1)$, (7) is then rewritten as

$$\mathbf{v}_\mathrm{T}(n+1)=\mathbf{v}_\mathrm{T}(n)+\mathbf{w}_\mathrm{T}^\mathrm{o}(n)-\mathbf{w}_\mathrm{T}^\mathrm{o}(n+1)+2\mu\mathbf{D}^{-1}(n)\mathbf{x}_\mathrm{T}(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)\mathbf{w}_\mathrm{T}^\mathrm{o}(n) + 2\mu\mathbf{D}^{-1}(n)\mathbf{x}_\mathrm{T}(n)z(n)-2\mu\mathbf{D}^{-1}(n)\mathbf{x}_\mathrm{T}(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)[\mathbf{v}_\mathrm{T}(n)+\mathbf{w}_\mathrm{T}^\mathrm{o}(n)]. \quad (8)$$

Now, by substituting (3) into (8), the following update expression is obtained

$$\mathbf{v}_\mathrm{T}(n+1)=[\mathbf{I}-2\mu\mathbf{D}^{-1}(n)\mathbf{x}_\mathrm{T}(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)]\mathbf{v}_\mathrm{T}(n) + 2\mu\mathbf{D}^{-1}(n)\mathbf{x}_\mathrm{T}(n)z(n)-\mathbf{g}(n). \quad (9)$$

In the next sections, the first and second moments of (9) are derived.

3.1. Analysis assumptions

To carry out the stochastic analysis, the following simplifying assumptions are considered.

i) $\mathbf{g}(n)$ is Gaussian, zero-mean, stationary with autocorrelation matrix given by $\mathbf{G}=E[\mathbf{g}(n)\mathbf{g}^\mathrm{T}(n)]=\sigma_q^2\mathbf{I}$.

ii) $\mathbf{g}(n)$ and $\mathbf{x}_\mathrm{T}(n)$, and $\mathbf{v}_\mathrm{T}(n)$ and $\mathbf{x}_\mathrm{T}(n)$ are statistically independent.

iii) $\mathbf{D}^{-1}(n)$ and $\mathbf{x}_\mathrm{T}(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)$ are jointly stationary, such that $\mathbf{D}^{-1}(n)$ is slowly varying with respect to $\mathbf{x}_\mathrm{T}(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)$. This assumption permits to invoke the Averaging Principle [10].

In this way, we can now proceed with the model derivations.

3.2. First moment of $\mathbf{v}_\mathrm{T}(n)$

By taking the expectation on both sides of (9), we obtain

$$E[\mathbf{v}_\mathrm{T}(n+1)]=E[\mathbf{v}_\mathrm{T}(n)]-2\mu E[\mathbf{D}^{-1}(n)\mathbf{x}_\mathrm{T}(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)\mathbf{v}_\mathrm{T}(n)] + 2\mu E[\mathbf{D}^{-1}(n)\mathbf{x}_\mathrm{T}(n)z(n)]-E[\mathbf{g}(n)]. \quad (10)$$

By using (*i*), (*ii*), and (*iii*), we get

$$E[\mathbf{v}_\mathrm{T}(n+1)]=\{\mathbf{I}-2\mu E[\mathbf{D}^{-1}(n)]\mathbf{R}_\mathrm{T}\}E[\mathbf{v}_\mathrm{T}(n)]. \quad (11)$$

The expectation of $\mathbf{D}^{-1}(n)$ in the r.h.s. of (11) is determined by invoking the Averaging Principle, resulting in

$$E[\mathbf{D}^{-1}(n)]=\frac{M}{(M-2)}[\mathrm{diag}(\mathbf{R}_\mathrm{T})]^{-1}-\varepsilon\frac{M^2}{(M-2)(M-4)}[\mathrm{diag}(\mathbf{R}_\mathrm{T}^2)]^{-1}. \quad (12)$$

Concerning the derivation of (12), the reader is referred to [8] for more details. Now, by substituting (12) into (11), we obtain the model expression for the first moment of $\mathbf{v}_\mathrm{T}(n)$, which is given by

$$E[\mathbf{v}_\mathrm{T}(n+1)]=E[\mathbf{v}_\mathrm{T}(n)]-2\mu\frac{M}{(M-2)}[\mathrm{diag}(\mathbf{R}_\mathrm{T})]^{-1}\mathbf{R}_\mathrm{T}E[\mathbf{v}_\mathrm{T}(n)] + 2\mu\varepsilon\frac{M^2}{(M-2)(M-4)}[\mathrm{diag}(\mathbf{R}_\mathrm{T}^2)]^{-1}\mathbf{R}_\mathrm{T}E[\mathbf{v}_\mathrm{T}(n)]. \quad (13)$$

3.3. Second moment of $\mathbf{v}_\mathrm{T}(n)$

The second moment for the weight-error vector in the transform domain is obtained by making $\mathbf{K}(n)=E[\mathbf{v}_\mathrm{T}(n)\mathbf{v}_\mathrm{T}^\mathrm{T}(n)]$. Then, transposing both sides of (9), performing the product $\mathbf{v}_\mathrm{T}(n)\mathbf{v}_\mathrm{T}^\mathrm{T}(n)$, and taking the expectation on both sides of the resulting expression, we obtain

$$\begin{aligned}
E[\mathbf{v}_\mathrm{T}(n+1)\mathbf{v}_\mathrm{T}^\mathrm{T}(n+1)] &= E[\mathbf{v}_\mathrm{T}(n)\mathbf{v}_\mathrm{T}^\mathrm{T}(n)]+2\mu E[z(n)\mathbf{v}_\mathrm{T}(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)\mathbf{D}^{-1}(n)] \\
&\overbrace{-2\mu E[\mathbf{v}_\mathrm{T}(n)\mathbf{v}_\mathrm{T}^\mathrm{T}(n)\mathbf{x}_\mathrm{T}(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)\mathbf{D}^{-1}(n)]}^{\mathrm{A}}-E[\mathbf{v}_\mathrm{T}(n)\mathbf{g}^\mathrm{T}(n)] \\
&+2\mu E[\mathbf{D}^{-1}(n)\mathbf{x}_\mathrm{T}(n)z(n)\mathbf{v}_\mathrm{T}^\mathrm{T}(n)] \\
&\overbrace{+4\mu^2 E[\mathbf{D}^{-1}(n)\mathbf{x}_\mathrm{T}(n)z^2(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)\mathbf{D}^{-1}(n)]}^{\mathrm{B}} \\
&-4\mu^2 E[\mathbf{D}^{-1}(n)\mathbf{x}_\mathrm{T}(n)z(n)\mathbf{v}_\mathrm{T}^\mathrm{T}(n)\mathbf{x}_\mathrm{T}(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)\mathbf{D}^{-1}(n)] \\
&-2\mu E[\mathbf{D}^{-1}(n)\mathbf{x}_\mathrm{T}(n)z(n)\mathbf{g}^\mathrm{T}(n)] \\
&\overbrace{-2\mu E[\mathbf{D}^{-1}(n)\mathbf{x}_\mathrm{T}(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)\mathbf{v}_\mathrm{T}(n)\mathbf{v}_\mathrm{T}^\mathrm{T}(n)]}^{\mathrm{C}} \\
&-4\mu^2 E[\mathbf{D}^{-1}(n)\mathbf{x}_\mathrm{T}(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)\mathbf{v}_\mathrm{T}(n)z(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)\mathbf{D}^{-1}(n)] \\
&\overbrace{+4\mu^2 E[\mathbf{D}^{-1}(n)\mathbf{x}_\mathrm{T}(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)\mathbf{v}_\mathrm{T}(n)\mathbf{v}_\mathrm{T}^\mathrm{T}(n)\mathbf{x}_\mathrm{T}(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)\mathbf{D}^{-1}(n)]}^{\mathrm{D}} \\
&+2\mu E[\mathbf{D}^{-1}(n)\mathbf{x}_\mathrm{T}(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)\mathbf{v}_\mathrm{T}(n)\mathbf{g}^\mathrm{T}(n)]-E[\mathbf{g}(n)\mathbf{v}_\mathrm{T}^\mathrm{T}(n)] \\
&-2\mu E[z(n)\mathbf{g}(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)\mathbf{D}^{-1}(n)] \\
&+2\mu E[\mathbf{g}(n)\mathbf{v}_\mathrm{T}^\mathrm{T}(n)\mathbf{x}_\mathrm{T}(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)\mathbf{D}^{-1}(n)]+E[\mathbf{g}(n)\mathbf{g}^\mathrm{T}(n)].
\end{aligned} \quad (14)$$

To obtain the final form of (14), the expectations in the r.h.s. must be determined. Let us first consider the terms containing $z(n)$, except that with $z^2(n)$. Due to the characteristics of $z(n)$, the concerning terms are equal to zero.

Now, by using the analysis assumptions (*i*)-(*iii*) (Section 3.1), expectations A, B, C, and D in (14) are given by

A) $-2\mu E[\mathbf{v}_\mathrm{T}(n)\mathbf{v}_\mathrm{T}^\mathrm{T}(n)]E[\mathbf{x}_\mathrm{T}(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)]E[\mathbf{D}^{-1}(n)] = -2\mu\mathbf{K}(n)\mathbf{R}_\mathrm{T}E[\mathbf{D}^{-1}(n)]$

B) $4\mu^2 E[\mathbf{D}^{-1}(n)]E[\mathbf{x}_\mathrm{T}(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)]E[\mathbf{D}^{-1}(n)]E[z^2(n)] = 4\mu^2 E[\mathbf{D}^{-1}(n)]\mathbf{R}_\mathrm{T}E[\mathbf{D}^{-1}(n)]\sigma_z^2$

C) $-2\mu E[\mathbf{D}^{-1}(n)]E[\mathbf{x}_\mathrm{T}(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)]E[\mathbf{v}_\mathrm{T}(n)\mathbf{v}_\mathrm{T}^\mathrm{T}(n)] = -2\mu E[\mathbf{D}^{-1}(n)]\mathbf{R}_\mathrm{T}\mathbf{K}(n)$

D) $4\mu^2 E[\mathbf{D}^{-1}(n)]\{2E[\mathbf{x}_\mathrm{T}(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)]E[\mathbf{v}_\mathrm{T}(n)\mathbf{v}_\mathrm{T}^\mathrm{T}(n)]E[\mathbf{x}_\mathrm{T}(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)] + E[\mathbf{x}_\mathrm{T}(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)]\mathrm{tr}\left[E[\mathbf{v}_\mathrm{T}(n)\mathbf{v}_\mathrm{T}^\mathrm{T}(n)]E[\mathbf{x}_\mathrm{T}(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)]\right]\}E[\mathbf{D}^{-1}(n)] = 4\mu^2 E[\mathbf{D}^{-1}(n)]\{2\mathbf{R}_\mathrm{T}\mathbf{K}(n)\mathbf{R}_\mathrm{T}+\mathbf{R}_\mathrm{T}\mathrm{tr}[\mathbf{K}(n)\mathbf{R}_\mathrm{T}]\}E[\mathbf{D}^{-1}(n)]$

Now, we can express (14) as follows:

$$\mathbf{K}(n+1)=\mathbf{K}(n)-2\mu\mathbf{K}(n)\mathbf{R}_\mathrm{T}E[\mathbf{D}^{-1}(n)]-2\mu E[\mathbf{D}^{-1}(n)]\mathbf{R}_\mathrm{T}\mathbf{K}(n) + 4\mu^2 E[\mathbf{D}^{-1}(n)]\{2\mathbf{R}_\mathrm{T}\mathbf{K}(n)\mathbf{R}_\mathrm{T}+\mathbf{R}_\mathrm{T}\mathrm{tr}[\mathbf{R}_\mathrm{T}\mathbf{K}(n)]\}E[\mathbf{D}^{-1}(n)] + 4\mu^2\sigma_z^2 E[\mathbf{D}^{-1}(n)]\mathbf{R}_\mathrm{T}E[\mathbf{D}^{-1}(n)]+\mathbf{G} \quad (15)$$

3.4. Learning curve

From (2) and (6), the error signal can be written as

$$e(n)=\mathbf{x}_\mathrm{T}^\mathrm{T}(n)\mathbf{w}_\mathrm{T}^\mathrm{o}(n)+z(n)-\mathbf{x}_\mathrm{T}^\mathrm{T}(n)[\mathbf{v}_\mathrm{T}(n)+\mathbf{w}_\mathrm{T}^\mathrm{o}(n)] = z(n)-\mathbf{x}_\mathrm{T}^\mathrm{T}(n)\mathbf{v}_\mathrm{T}(n). \quad (16)$$

By squaring (16), taking the expectation on both sides of the resulting expression, and from the definition of $z(n)$, one obtains

$$E[e^2(n)]=E[z^2(n)]+E[\mathbf{v}_\mathrm{T}^\mathrm{T}(n)\mathbf{x}_\mathrm{T}(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)\mathbf{v}_\mathrm{T}(n)]. \quad (17)$$

The second term in the r.h.s. in (17) is rewritten as

$$E[\mathbf{v}_\mathrm{T}^\mathrm{T}(n)\mathbf{x}_\mathrm{T}(n)\mathbf{x}_\mathrm{T}^\mathrm{T}(n)\mathbf{v}_\mathrm{T}(n)]=\{\mathbf{R}_\mathrm{T}E[\mathbf{v}_\mathrm{T}(n)\mathbf{v}_\mathrm{T}^\mathrm{T}(n)]\}. \quad (18)$$

Thus, recalling that $\mathbf{K}(n)=E[\mathbf{v}_\mathrm{T}(n)\mathbf{v}_\mathrm{T}^\mathrm{T}(n)]$ and by substituting (18) into (17), the learning curve is given by

$$E[e^2(n)] = \sigma_z^2 + \mathrm{tr}\{\mathbf{R}_\mathrm{T}\mathbf{K}(n)\}\,. \qquad (19)$$

3.5. Excess error

The excess error is given by

$$\xi_{\mathrm{exc}} = \mathrm{tr}[\mathbf{R}_\mathrm{T}\mathbf{K}(\infty)]\,. \qquad (20)$$

Following [1] and after some simple algebra, we can express (15) as follows:

$$\xi_{\mathrm{exc}} = \frac{1}{1-\mu\mathrm{tr}(E[\mathbf{D}^{-1}(n)]\mathbf{R}_\mathrm{T})} \times \left\{ \frac{1}{4\mu}\mathrm{tr}\left[(E[\mathbf{D}^{-1}(n)])^{-1}\mathbf{G}\right] + \mu\sigma_z^2\mathrm{tr}(E[\mathbf{D}^{-1}(n)]\mathbf{R}_\mathrm{T}) \right\}. \qquad (21)$$

The difference between (21) and the excess error for a stationary environment is the term involving matrix **G**. Such a difference leads to a larger excess error in a nonstationary environment as compared with a stationary one [2].

3.6. Misadjustment

Misadjustment $\mathcal{M}$ is obtained from (21) and it is given by

$$\mathcal{M} = \frac{\xi_{\mathrm{exc}}}{\xi_{\min}} = \frac{1}{1-\mu\mathrm{tr}\left[E[\mathbf{D}^{-1}(n)]\mathbf{R}_\mathrm{T}\right]} \times \left\{ \frac{1}{4\mu}\sigma_z^{-2}\mathrm{tr}\left[(E[\mathbf{D}^{-1}(n)])^{-1}\mathbf{G}\right] + \mu\mathrm{tr}(E[\mathbf{D}^{-1}(n)]\mathbf{R}_\mathrm{T}) \right\} \qquad (22)$$

The accuracy of (22) is verified in Table 1, which presents the misadjustment for the cases shown in the section of Simulation Results.

Table 1. Verification of algorithm misadjustment.

Condition		$\mathcal{M}$ (simulation)	$\mathcal{M}$ from (22)	
$\alpha = 2$	$\mu_{\mathrm{opt}}/2$	40.57	43.68	(Fig. 1)
	$\mu_{\mathrm{opt}}/10$	158.90	172.93	
$\alpha = 1$	$\mu_{\mathrm{opt}}/2$	11.06	11.72	(Fig. 2)
	$\mu_{\mathrm{opt}}/10$	42.56	46.36	

3.7. Degree of nonstationarity

The degree of nonstationarity denoted by α is given in [1, p. 640]. Then, for our case, we can write

$$\alpha = \frac{1}{\sigma_z}[\mathrm{tr}(\mathbf{G}\mathbf{R}_\mathrm{T})]^{1/2} = \frac{1}{\sigma_z}[(\sigma_0^2 + \cdots + \sigma_{N-1}^2)\sigma_g^2]^{1/2} = \frac{\sigma_g}{\sigma_z}[\mathrm{tr}(\mathbf{R}_\mathrm{T})]^{1/2} \qquad (23)$$

4. SIMULATION RESULTS

The proposed model is applied to a system identification problem in which the accuracy of the proposed model expressions is assessed for colored Gaussian inputs. The correlated signal is obtained from an AR(2) process, defined by

$$x(n) = \alpha_1 x(n-1) + \alpha_2 x(n-2) + v(n)\,, \qquad (24)$$

where $v(n)$ is white noise with variance σ_v^2 such that the variance of $x(n)$ is equal to 1, α_1 and α_2 are the autoregressive coefficients, with $\alpha_1 = -0.1833$ and $\alpha_2 = 0.85$. The variance of the measurement noise $z(n)$ is 0.0001 (SNR = 40 dB). All the Monte Carlo simulations are obtained from an average of 500 independent runs. The time-varying coefficients of the plant are given by (3), with the starting coefficient vector obtained from $\mathbf{w}_\mathrm{T}^\mathrm{o}(0) = \mathbf{T}[\mathrm{sinc}(0) \quad \mathrm{sinc}(1/N)\cdots\mathrm{sinc}(N-1/N)]^\mathrm{T}$ and the elements of $\mathbf{g}(n)$ from a white noise process with variance given by (23), resulting in

$$\sigma_g^2 = \left(\frac{\alpha\sigma_z}{[\mathrm{tr}(\mathbf{R}_\mathrm{T})]^{1/2}}\right)^2 \qquad (25)$$

The orthogonal transform considered is DCT. The step size μ used (denoted by μ_{opt}) is obtained from (21) to attain a minimum excess error. The results are obtained by using two nonstationarity degree values α and two step-size values equal to $0.5\mu_{\mathrm{opt}}$ and $0.1\mu_{\mathrm{opt}}$ for each α value. The regularization parameter used for all cases is $\varepsilon = 0.001$.

The curves of Figs. 1 and 2 are obtained for $N = 8$, $M = 32$, and $\mu_{\mathrm{opt}} = 0.0585$. The eigenvalue dispersion of the correlated input signal is 81 and the degree of nonstationarity is 2 and 1, respectively, obtained from (23). Figs. 1(a) and (b) illustrate, respectively, the first and second moments of the adaptive filter weights obtained from Monte Carlo (MC) simulations and from the proposed models (13), (15), and (19). For the curves of Fig. 2, the same parameters previously considered are used, except for $\alpha = 1$. From these results, we can verify a very good matching between the curves obtained from simulations and the proposed analytical models. In addition, from Table 1 we can observe that the predictions obtained from (22) present a satisfactory accuracy, as compared with those obtained by simulation.

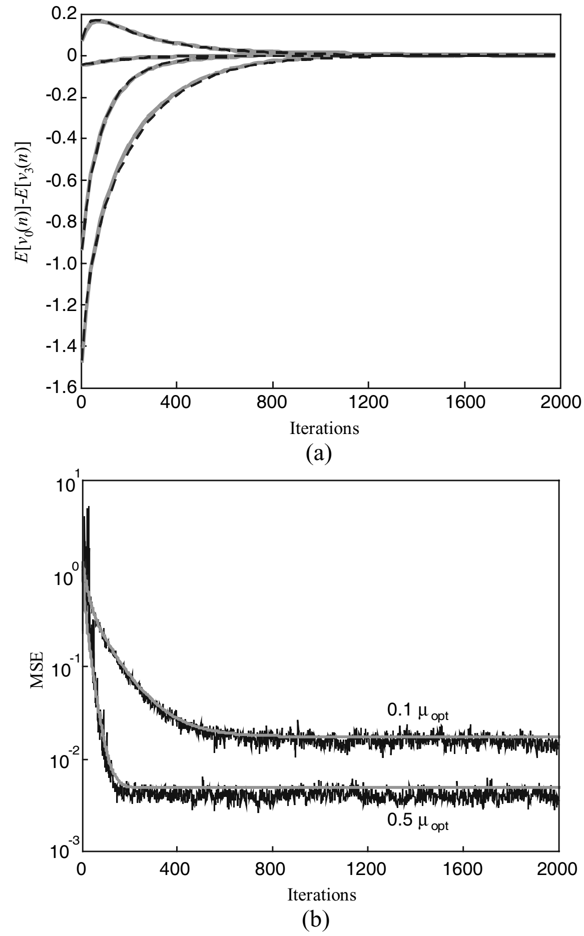
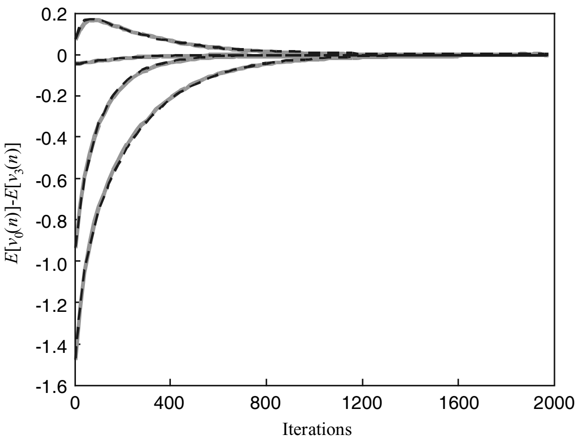
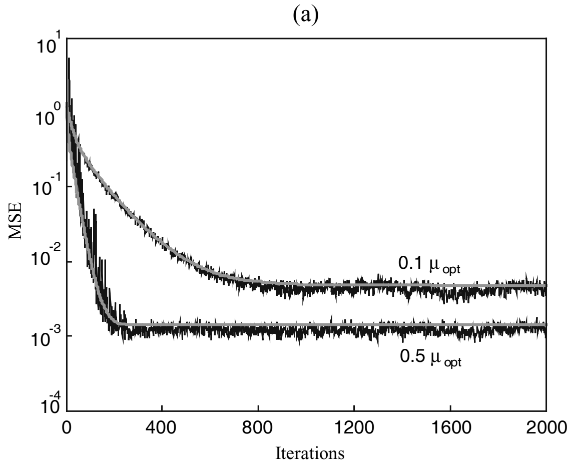


Fig. 1. Case for $\alpha = 2$. Model performance for colored Gaussian input signal. (a) Mean weight-error vector behavior for the coefficients $E[v_0(n)]$ - $E[v_3(n)]$ for $0.1\mu_{\mathrm{opt}}$. (Black dashed line) MC simulation (average of 500 runs); (Gray solid line) proposed model for $0.1\mu_{\mathrm{opt}}$. (b) MSE curves for $0.1\mu_{\mathrm{opt}}$ and $0.5\mu_{\mathrm{opt}}$. (Black ragged line) MC simulation; (Gray solid line) proposed model.



(a)



(b)

Fig. 2. Case for $\alpha = 1$. Model performance for colored Gaussian input signal. (a) Mean weight-error vector behavior for the coefficients $E[v_0(n)] - E[v_3(n)]$ for $0.1\mu_{opt}$. (Black dashed line) MC simulation (average of 500 runs); (Gray solid line) proposed model for $0.1\mu_{opt}$. (b) MSE curves for $0.1\mu_{opt}$ and $0.5\mu_{opt}$. (Black ragged line) MC simulation; (Gray solid line) proposed model.

5. CONCLUSIONS

A stochastic model for the εTDLMS algorithm in a nonstationary environment is derived. The theoretical models for the first and second moment of the weight vector have been derived for Gaussian inputs, a high-order adaptive filter, and slow adaptation condition. Through numerical simulations, we have verified a very good matching between the results obtained by the Monte Carlo method and the proposed model.

REFERENCES

[1] S. Haykin, *Adaptive Filter Theory*, 4ed., Prentice-Hall, 2002.

[2] B. Farhang-Boroujeny, *Adaptive Filters: Theory and Applications*, John Wiley & Sons, 1998.

[3] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*, Prentice-Hall, 1985.

[4] S. S. Narayan, A. M. Peterson, and M. J. Narasimha, “Transform domain LMS algorithm,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 31, no. 3, pp. 609-615, June 1983.

[5] S. Hosur and A. H. Tewfik, “Wavelet transform domain adaptive FIR filtering,” *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 617-630, Mar. 1997.

[6] S. J. Chern, J. C. Horng, and K. M. Wong, “The performance of the hybrid LMS adaptive algorithm,” *Elsevier Signal Processing*, vol. 44, pp. 67-88, 1995.

[7] D. I. Kim and P. De Wilde, “Performance analysis of the DCT-LMS adaptive filtering algorithm,” *Elsevier Signal Processing*, vol. 80, pp. 1629-1654, 2000.

[8] E. M. Lobato, O. J. Tobias, and R. Seara, “A stochastic model for the transform domain LMS algorithm”, in *Proc. XII European Signal Processing Conf.*, Vienna, Austria, Sep. 2004, pp. 1833-1836.

[9] A. H. Sayed and T. Y. Al-Naffouri, “Mean-square analysis of normalized leaky adaptive filters,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Salt Lake City, USA, vol. 6, May 2001, pp. 3873-3876.

[10] C. G. Samson and U. Reddy, “Fixed point error analysis of the normalized ladder algorithm,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 31, no. 5, pp. 1177-1191, Oct. 1983.