PAPER *Special Section on Selected Papers from the 11th Workshop on Circuits and Systems in Karuizawa*

Steady State Analysis of 2-D LMS Adaptive Filters Using the Independence Assumption

Maha SHADAYDEH†, ***Student Member and*** **Masayuki KAWAMATA**†, ***Member***

SUMMARY In this paper, we consider the steady state mean square error (MSE) analysis for 2-D LMS adaptive filtering algorithm in which the filter's weights are updated along both vertical and horizontal directions as a doubly-indexed dynamical system. The MSE analysis is conducted using the well-known independence assumption. First we show that computation of the weight-error covariance matrix for doubly-indexed 2-D LMS algorithm requires an approximation for the weight-error correlation coefficients at large spatial lags. Then we propose a method to solve this problem. Further discussion is carried out for the special case when the input signal is white Gaussian. It is shown that the convergence in the MSE sense occurs for step size range that is significantly smaller than the one necessary for the convergence of the mean. Simulation experiments are presented to support the obtained analytical results.

key words: *2-D LMS, steady state analysis, doubly-indexed system*

1. Introduction

The main advantage of using 2-D adaptive filters for processing nonstationary 2-D signals, such as images, is in their ability to change the filtering characteristics based on the local statistics of the processed data. Least mean square (LMS) type 2-D adaptive filters have, in particular, received considerable research interest, mainly because of its simplicity in computation. Hadhoud and Thomas [1] have proposed a 2-D LMS algorithm, which is called TDLMS, by direct extension of the 1-D LMS algorithm [2]. In the TDLMS, the weights' update process is carried out using either vertical or horizontal 1-D indexing scheme. Accordingly, the authors of [1] have shown that the weights' update equation of the TDLMS can be written in a form which is mathematically equivalent to the well known 1-D LMS [2]; hence, the analysis procedures and results of the 1-D LMS can be directly applied to the TDLMS. The drawback of the algorithm [1], however, is that it can only exploit the correlation information of the image pixels in the direction of the indexing scheme used to process the 2-D data. To overcome such problem, the authors of [3] have proposed a 2-D LMS algorithm in which the filter's weights are updated along both the vertical and horizontal directions as a doubly-indexed dynamical system [4]. Such update mechanism enables efficient use of the 2-D correlation information of the image pixels in both vertical and horizontal directions and hence, provides better performance in nonstationary environments [3].

The convergence of the mean for the 2-D LMS [3] (in what follows, it will be referred to as 2-D LMS) has been investigated in [3] using stability theory of 2-D Fornasini and Marchesini (F-M) state space model [4]. Convergence of the mean does not, however, guarantee finite mean square error (MSE) for the adaptive algorithm.

In this paper, we consider the MSE analysis of the 2-D LMS. The analysis presented in this paper is the first attempt in the literature to investigate the steady state MSE analysis for a doubly-indexed 2-D LMS algorithm. The MSE analysis is carried out using the assumption that the successive input vectors are statistically independent, jointly Gaussian-distributed random variables. This assumption, generally referred to as the *independence assumption* [6], is widely used in the convergence analysis of 1-D LMS for two main reasons. The first is due to the simplification in analysis obtained under such assumption. The second is due to the good agreement between the analytical results obtained using the independence assumption and experimental results [5]–[9].

Though the 2-D MSE analysis will be significantly simplified when invoking the independence assumption, the use of 2-D indexing scheme in the weights' update equation of the 2-D LMS results in a new problem that is not encountered in the 1-D case. For the 1-D LMS, as well as for the TDLMS, the adaptive filter's weight-vector update equation is a 1-D first order difference equation given by

$$\mathbf{H}_{j+1} = \mathbf{H}_j - \mu \mathbf{G}_j \tag{1}$$

where j is the iteration number; $\mathbf{H}_j$ is the adaptive filter's weight-vector; μ is a scalar parameter that controls the convergence rate of the LMS algorithm, and $\mathbf{G}_j$ is the instantaneous gradient of the MSE at iteration j. From Eq. (1), it follows that the weight-error covariance matrix is calculated by a set of 1-D first order difference equations. According to [8], [9], this set of difference equations maintains stability under a general condition imposed on the used step size parameter μ. For the 2-D LMS, however, the adaptive filter's weight-vector update equation is described by the 2-D first

Manuscript received June 27, 1998.

Manuscript revised September 26, 1998.

†The authors are with the Dept. of Electronic Engineering, Graduate School of Engineering, Tohoku University, Sendai-shi, 980–8579 Japan.

order difference equation

$$\mathbf{H}_{m+1,n+1} = f_h\mathbf{H}_{m,n+1} + f_v\mathbf{H}_{m+1,n} - \mu_h\mathbf{G}_{m,n+1} - \mu_v\mathbf{G}_{m+1,n} \quad (2)$$

where m and n are two spatial indices in the vertical and horizontal direction respectively. $\mathbf{H}_{m,n}$ is the adaptive filter weight-vector at spatial indices (m,n); f_h, f_v, μ_h and μ_v are scalar parameters, and $\mathbf{G}_{m,n}$ is the instantaneous gradient of the MSE at spatial indices (m,n). From Eq. (2), and as will be shown in the sequel, the weight-error covariance matrix for the 2-D LMS is calculated by a set of 2-D second order difference equations. Stability analysis for such set of equations is, however, very difficult to handle mathematically.

In this paper, we show that for the steady state, this set of 2-D second order difference equations can be reduced to a set of linear simultaneous equations in the coefficients of weight-error correlation matrices at different spatial lags; however, the number of the unknowns in this set exceeds the number of equations. To solve this problem, we propose a method for the approximation of the coefficients of weight-error correlation matrices at large spatial lags. The approximation method is based on the extension of the direct averaging method [5] to 2-D case. It can also serve as an approximation method for the weight-error covariance matrix without invoking the independence assumption providing that the step size parameters are sufficiently small.

Although the 2-D MSE analysis using the independence assumption may not provide very accurate estimate for the MSE when the input data are correlated, there are three aims from considering the analysis under such assumption. The first is to derive an analytical expression for the MSE that gives some insight to the performance of a truly 2-D LMS even when the input data are correlated. The second is to derive a good approximation of the bounds on the step size parameters that guarantee convergence in the MSE sense. And the third aim is to shed a light on the problem that arises in 2-D MSE analysis for doubly-indexed LMS algorithm. This problem forms a major obstacle in extending the 1-D MSE analysis approaches that do not invoke the independence assumption, such as [10] and [11], to 2-D case.

The organization of this paper is as follows. In Sect. 2, a brief review of the 2-D LMS algorithm [3] is given. In Sect. 3, the steady state MSE analysis of the 2-D LMS algorithm using the independence assumption is considered and a method for computing the 2-D weight-error covariance matrix is presented. In Sect. 4, the special case when the input signal is white Gaussian is further discussed, and the condition required to ensure the convergence in the MSE sense is derived. Comparison between experimental and analytical results for the simplified case are presented in Sect. 5. Finally, conclusions are drawn in Sect. 6.

2. Preliminaries

Consider the N by N, causal, 2-D adaptive FIR filter shown in Fig. 1. The filter's input $x(m,n)$ is a 2-D stationary signal of size $M_1 \times M_2$. The filter's output $y(m,n)$ is calculated by

$$y(m,n) = \mathbf{H}^t_{m,n}\mathbf{X}_{m,n} \quad (3)$$

where $\mathbf{H}_{m,n}$ and $\mathbf{X}_{m,n}$ are respectively the adaptive filter's weight-vector and the input data vector given at spatial indices (m,n) by

$$\begin{aligned}\mathbf{X}_{m,n} &= [x(m,n),\cdots,x(m-N+1,n), \\ &\quad \cdots,x(m-N+1,n-N+1)]^t \\ \mathbf{H}_{m,n} &= [h_{m,n}(0,0),\cdots,h_{m,n}(N-1,0), \\ &\quad \cdots,h_{m,n}(N-1,N-1)]^t. \end{aligned} \quad (4)$$

The 2-D LMS updates the filter weight-vector along both the vertical and horizontal directions such that the error between the filter output $y(m,n)$ and the desired signal $d(m,n)$ is minimized in the MSE sense. The MSE is defined as

$$\begin{aligned}\text{MSE} &= \text{E}\{e^2(m,n)\} \\ &= \text{E}\{\big(d(m,n) - \mathbf{H}^t_{m,n}\mathbf{X}_{m,n}\big)^2\}.\end{aligned} \quad (5)$$

The update equation for the 2-D LMS is given by

$$\begin{aligned}&\mathbf{H}_{m+1,n+1} = f_h\mathbf{H}_{m,n+1} + f_v\mathbf{H}_{m+1,n} \\ &\quad + \mu_h e(m,n+1)\mathbf{X}_{m,n+1} + \mu_v e(m+1,n)\mathbf{X}_{m+1,n}; \\ &\quad \mathbf{H}_{m,0} = \mathbf{0}, m = 0\cdots M_1; \mathbf{H}_{0,n} = \mathbf{0}, n = 0\cdots M_2; \\ &\quad f_h + f_v = 1 \end{aligned} \quad (6)$$

where μ_h and μ_v denote the step size parameters in the horizontal and vertical directions respectively.

The optimal solution $\mathbf{H}_{opt}$ that minimizes the MSE is given by the Wiener-Hopf equation

$$\mathbf{H}_{opt} = \mathbf{R}^{-1}\mathbf{P} \quad (7)$$

where

$$\begin{aligned}\mathbf{R} &= \text{E}\{\mathbf{X}_{m,n}\mathbf{X}^t_{m,n}\} \\ \mathbf{P} &= \text{E}\{\mathbf{X}_{m,n}d(m,n)\}.\end{aligned} \quad (8)$$

In [3], it has been shown that the 2-D LMS converges

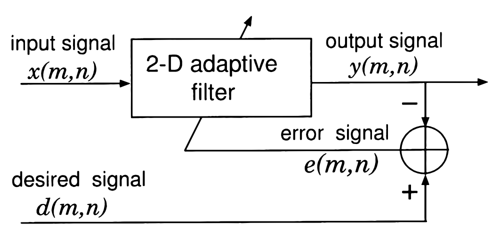

Fig. 1 2-D adaptive filter.

to the optimal solution i.e., $\mathrm{E}\{\mathbf{H}_{m+1,n+1} - \mathbf{H}_{opt}\} \to 0$ as $m+n \to \infty$, if the following condition holds:

$$|f_h - \mu_h \lambda_i| + |f_v - \mu_v \lambda_i| < 1 \tag{9}$$

where $\lambda_i,\ i = 0, \cdots, N^2-1$, are the eigenvalues of the input correlation matrix $\mathbf{R}$.

Condition (9) is, however, not sufficient to guarantee convergence of the 2-D LMS in the MSE sense. Moreover, convergence of the mean does not provide any information about the performance of the adaptive algorithm. In the following section we will present the steady state MSE analysis for the 2-D LMS.

3. Steady State MSE Analysis for the 2-D LMS

3.1 MSE Calculation

The MSE analysis will be carried out using the independence assumption [5], consisting of the following points:

A.1 The input vectors $\mathbf{X}_{0,0}, \mathbf{X}_{1,0}, \cdots, \mathbf{X}_{m,n}$ are zero mean, statistically independent, Gaussian-distributed random variables.

A.2 The error

$$\varepsilon(m,n) = d(m,n) - \mathbf{H}_{opt}^t \mathbf{X}_{m,n} \tag{10}$$

is a zero mean, white Gaussian noise of variance σ_ε^2, and is statistically independent of the input vector $\mathbf{X}_{m,n}$.

Let us define the adaptive filter weight-error vector

$$\mathbf{C}_{m,n} = \mathbf{H}_{m,n} - \mathbf{H}_{opt}. \tag{11}$$

Then, using Eqs. (10) and (11), the error signal $e(m,n)$ can be given by

$$\begin{aligned} e(m,n) &= d(m,n) - \mathbf{H}_{m,n}^t \mathbf{X}_{m,n} \\ &= \varepsilon(m,n) + \mathbf{H}_{opt}^t \mathbf{X}_{m,n} - \mathbf{H}_{m,n}^t \mathbf{X}_{m,n} \\ &= \varepsilon(m,n) - \mathbf{C}_{m,n}^t \mathbf{X}_{m,n}. \end{aligned} \tag{12}$$

Now if we substitute Eq. (12) in (5) and make use of assumptions A.1 and A.2, we can find that the steady state MSE is given by

$$\begin{aligned} \epsilon_\infty &= \lim_{m+n\to\infty} \mathrm{E}\{e(m,n)^2\} \\ &= \sigma_\varepsilon^2 + \lim_{m+n\to\infty} \mathrm{E}\{\mathbf{C}_{m,n}^t \mathbf{X}_{m,n} \mathbf{X}_{m,n}^t \mathbf{C}_{m,n}\} \quad (13) \\ &= \sigma_\varepsilon^2 + \lim_{m+n\to\infty} \mathrm{tr}(\mathbf{R}\ \mathrm{K}_{m,n;m,n}) \quad (14) \end{aligned}$$

where

$$\mathrm{K}_{m,n;m,n} = \mathrm{E}\{\mathbf{C}_{m,n} \mathbf{C}_{m,n}^t\} \tag{15}$$

is the weight-error covariance matrix.

Note that from A.1 it follows that the input vector $\mathbf{X}_{m,n}$ and the weight-error vector $\mathbf{C}_{m,n}$ are statistically independent. Accordingly, the expectation term in Eq. (13) can be treated as a product of two expectation terms. Strictly speaking, in adaptive filtering applications, these two vectors are dependent since the successive input vectors are statistically dependent. However, even when this statistical dependency is ignored, the independence assumption still preserve the correlation structure for $\mathrm{E}\{\mathbf{X}_{m,n}\mathbf{X}_{m,n}^t\}$ as well as for $\mathrm{E}\{\mathbf{C}_{m,n}^t \mathbf{C}_{m,n}\}$. Hence, the analysis under such assumption still retains enough information about the behavior of the adaptive process even when the input signal is correlated, (see [5], [6] and references therein).

In the rest of this section we will consider the calculation of the weight-error covariance matrix. In this calculation, we assume that the condition (9), which is necessary for the convergence of the mean, holds.

3.2 Weight-Error Covariance Matrix

To calculate the weight-error covariance matrix we need first to derive the update equation for the weight-error vector. Indeed, if we subtract $\mathbf{H}_{opt}$ from both sides of Eq. (6) and make use of Eq. (12), we get

$$\begin{aligned} \mathbf{C}_{m+1,n+1} &= \mathbf{H}_{m+1,n+1} - \mathbf{H}_{opt} \\ &= (f_h \mathbf{I} - \mu_h \mathbf{X}_{m,n+1} \mathbf{X}_{m,n+1}^t) \mathbf{C}_{m,n+1} \\ &\quad + (f_v \mathbf{I} - \mu_v \mathbf{X}_{m+1,n} \mathbf{X}_{m+1,n}^t) \mathbf{C}_{m+1,n} \\ &\quad + \mu_h \varepsilon(m,n+1) \mathbf{X}_{m,n+1} \\ &\quad + \mu_v \varepsilon(m+1,n) \mathbf{X}_{m+1,n}. \end{aligned} \tag{16}$$

Now before proceeding, we need to define some necessary notations. Since the input correlation matrix $\mathbf{R}$ is symmetric, there exists an orthogonal matrix $\mathbf{Q}$ such that

$$\begin{aligned} &\mathbf{Q}\mathbf{R}\mathbf{Q}^t = \mathbf{\Lambda} = \mathrm{diag}(\lambda_0, \lambda_1, \cdots, \lambda_{N^2-1}) \\ &\mathbf{Q}^t = \mathbf{Q}^{-1}. \end{aligned} \tag{17}$$

Thus, we can define the transformed matrix:

$$\begin{aligned} \mathbf{\Gamma}_{m_1,n_1;m_2,n_2} &= \mathbf{Q}\mathrm{E}\{\mathbf{C}_{m_1,n_1} \mathbf{C}_{m_2,n_2}^t\}\mathbf{Q}^t \\ &= \left[\gamma_{m_1,n_1;m_2,n_2}^{i,j}\right]; \\ &i,j = 0, \cdots, N^2-1. \end{aligned} \tag{18}$$

In Eq. (18), the superscripts (i,j) in the notation $\gamma_{m_1,n_1;m_2,n_2}^{i,j}$ is used to point to the element at the ith row and jth column of the matrix $\mathbf{\Gamma}_{m_1,n_1;m_2,n_2}$.

Now, multiplying each side of Eq. (16) with its transpose, taking the expected values, and making use of the orthogonal transform $\mathbf{Q}$ we arrive at

$$\begin{aligned} &\mathbf{\Gamma}_{m+1,n+1;m+1,n+1} = f_h^2 \mathbf{\Gamma}_{m,n+1;m,n+1} \\ &- \mu_h f_h \mathbf{\Gamma}_{m,n+1;m,n+1} \mathbf{\Lambda} - \mu_h f_h \mathbf{\Lambda} \mathbf{\Gamma}_{m,n+1;m,n+1} \\ &+ 2\mu_h^2 \mathbf{\Lambda} \mathbf{\Gamma}_{m,n+1;m,n+1} \mathbf{\Lambda} + \mu_h^2 \mathrm{tr}(\mathbf{\Gamma}_{m,n+1;m,n+1} \mathbf{\Lambda}) \mathbf{\Lambda} \\ &+ f_h f_v \mathbf{\Gamma}_{m+1,n;m,n+1} - \mu_h f_v \mathbf{\Gamma}_{m+1,n;m,n+1} \mathbf{\Lambda} \\ &- \mu_v f_h \mathbf{\Lambda} \mathbf{\Gamma}_{m+1,n;m,n+1} + \mu_h \mu_v \mathbf{\Lambda} \mathbf{\Gamma}_{m+1,n;m,n+1} \mathbf{\Lambda} \\ &+ f_h f_v \mathbf{\Gamma}_{m,n+1;m+1,n} - \mu_v f_h \mathbf{\Gamma}_{m,n+1;m+1,n} \mathbf{\Lambda} \\ &- \mu_h f_v \mathbf{\Lambda} \mathbf{\Gamma}_{m,n+1;m+1,n} + \mu_h \mu_v \mathbf{\Lambda} \mathbf{\Gamma}_{m,n+1;m+1,n} \mathbf{\Lambda} \end{aligned}$$

$$+ f_v^2 \mathbf{\Gamma}_{m+1,n;m+1,n} - \mu_v f_v \mathbf{\Gamma}_{m+1,n;m+1,n}\mathbf{\Lambda} - \mu_v f_v \mathbf{\Lambda}\mathbf{\Gamma}_{m+1,n;m+1,n} + 2\mu_v^2 \mathbf{\Lambda}\mathbf{\Gamma}_{m+1,n;m+1,n}\mathbf{\Lambda} + \mu_v^2 \mathrm{tr}(\mathbf{\Gamma}_{m+1,n;m+1,n}\mathbf{\Lambda})\mathbf{\Lambda} + (\mu_h^2+\mu_v^2)\sigma_\varepsilon^2\mathbf{\Lambda}. \quad (19)$$

The derivation of Eq. (19) is very tedious but straightforward. We have mainly made use of assumptions A.1, A.2, and the following property:

- For zero mean, Gaussian random variables, it can be shown that [8]

$$\mathrm{E}\{\mathbf{X}_{m,n}\mathbf{X}_{m,n}^t(\mathbf{X}_{m,n}^t\mathbf{C}_{m,n})^2\} = 2\mathrm{E}\{\mathbf{R}\mathbf{\Gamma}_{m,n;m,n}\mathbf{R}\} + \mathrm{tr}(\mathbf{R}\mathbf{\Gamma}_{m,n;m,n})\mathbf{R}.$$

Analyzing the stability of the set of second-order coupled 2-D difference equations (19) is a very complicated task. Thus, we propose to simplify the analysis by making use of the following two facts.

1. For the transformed weight-error correlation matrix defined in Eq. (18), it follows from Schwartz' inequality [12] that

$$(\gamma_{m,n,m,n}^{i,j})^2 \le \gamma_{m,n,m,n}^{i,i}\cdot\gamma_{m,n,m,n}^{j,j}. \quad (20)$$

   That is to say, the boundedness of the diagonal terms of the weight-error correlation matrices ensures the boundedness of the off-diagonal ones. Hence, it is sufficient to analyze the stability of the diagonal terms of the matrix equation (19). Note that, as for the MSE evaluation (see Eq. (14)), we are only interested in the diagonal terms since

$$\mathrm{tr}(\mathbf{R}\mathrm{K}_{m,n;m,n}) = \mathrm{tr}(\mathbf{\Lambda}\mathbf{\Gamma}_{m,n;m,n}) = \sum_{j=0}^{N^2-1}\gamma_{m,n;m,n}^{j,j}\lambda_j. \quad (21)$$

2. Let, for notational convenience, $\gamma_k^{i,i},\ k=0,1,\cdots,$ denote the steady state values of the weight error correlation coefficients at spatial lag $(k,-k)$. That is

$$\gamma_k^{i,i} = \lim_{m+n\to\infty}\gamma_{m+1,n+1-k;m+1-k,n+1}^{i,i}. \quad (22)$$

   Now, if the adaptive algorithm reaches the steady state, the signal $\mathbf{C}_{m,n}$ becomes stationary random signal. Consequently, if the weight-error covariance coefficient $\gamma_{m+1,n+1;m+1,n+1}^{i,i}$, $i=0,\cdots,N^2-1$, has a steady state value, say $\gamma_0^{i,i}$, then the following equality should hold:

$$\lim_{m+n\to\infty}\gamma_{m+1,n;m+1,n}^{i,i} = \lim_{m+n\to\infty}\gamma_{m,n+1;m,n+1}^{i,i} = \gamma_0^{i,i}. \quad (23)$$

   Similarly, if the weight-error correlation coefficient $\gamma_{m+1,n;m,n+1}^{i,i}$, $i=0,\cdots,N^2-1$, has a steady state value, say $\gamma_1^{i,i}$, then the following equality should hold:

$$\lim_{m+n\to\infty}\gamma_{m+1,n;m,n+1}^{i,i} = \lim_{m+n\to\infty}\gamma_{m,n+1;m+1,n}^{i,i} = \gamma_1^{i,i}. \quad (24)$$

Consequently, for the steady state, the N^2 diagonal coefficients of Eq. (19) should obey the equality

$$\gamma_0^{i,i} = (f_h^2+f_v^2-2(\mu_h f_h+\mu_v f_v)\lambda_i + 2(\mu_h^2+\mu_v^2)\lambda_i^2)\gamma_0^{i,i} + 2(f_h f_v - (\mu_h f_v+\mu_v f_h)\lambda_i + \mu_h\mu_v\lambda_i^2)\gamma_1^{i,i} + (\mu_h^2+\mu_v^2)\lambda_i\sum_{j=0}^{N^2-1}\gamma_0^{j,j}\lambda_j + (\mu_h^2+\mu_v^2)\sigma_\varepsilon^2\lambda_i. \quad (25)$$

There is a need for another set of equations in the unknowns $\gamma_0^{i,i}$ and $\gamma_1^{i,i}$.

If we apply the same way of analysis to evaluate the matrix

$$\lim_{m+n\to\infty}\mathbf{\Gamma}_{m+1,n;m,n+1} = [\gamma_1^{i,j}],\ i,j=0,\cdots,N^2-1$$

we can find that for the steady state, i.e. $m+n\to\infty$, the diagonal terms of the correlation matrix $\mathbf{\Gamma}_{m+1,n;m,n+1}$ should obey the equality

$$\gamma_1^{i,i} = \left(f_h f_v - (\mu_h f_v+\mu_v f_h)\lambda_i + 2\mu_h\mu_v\lambda_i^2\right)\gamma_0^{i,i} + \left(f_h^2+f_v^2-2(\mu_h f_h+\mu_v f_v)\lambda_i + (\mu_h^2+\mu_v^2)\lambda_i^2\right)\gamma_1^{i,i} + \left(f_h f_v - (\mu_h f_v+\mu_v f_h)\lambda_i + \mu_h\mu_v\lambda_i^2\right)\gamma_2^{i,i} + \mu_h\mu_v\lambda_i\sum_{j=0}^{N^2-1}\gamma_0^{j,j}\lambda_j + \mu_h\mu_v\lambda_i\sigma_\varepsilon^2 \quad (26)$$

where

$$\gamma_2^{i,i} = \lim_{m+n\to\infty}\gamma_{m+1,n-1;m-1,n+1}^{i,i}. \quad (27)$$

If we continue in similar way evaluating the weight-error correlation matrices for higher spatial lags, i.e.

$$\lim_{m+n\to\infty}\mathbf{\Gamma}_{m+1,n+1-k;m+1-k,n+1} = [\gamma_k^{i,j}];\ k=2,3,\cdots,$$

at each stage $k \in 0,1,\cdots,$ we will have a set of $(k+1)\times N^2$ equations in $(k+2)\times N^2$ unknowns, namely, $\gamma_j^{i,i};\ 0\le j\le k+1,\ 0\le i\le N^2-1$. To solve this problem, we propose two methods. The first method, presented in the Appendix, makes use of the direct averaging method [5]. It approximates the stochastic difference Eq. (16) of the weight-error vector with a simpler time-invariant averaged system. The proposed direct averaging-based analysis can be used to derive an approximation of the weight-error correlation matrix $\mathbf{\Gamma}_{m,n-k;m-k,n}$ for an arbitrary integer k without invoking the independence assumption given by A.1 and A.2. For example, we can use this method to obtain an approximation for the weight error correlation coefficients, $\gamma_2^{i,i} = \lim_{m+n\to\infty}\gamma_{m+1,n-1;m-1,n+1}^{i,i}$, $i=1,\cdots,N^2$;

using this approximation in Eq. (26), the two set of Eqs. (25) and (26) can then be solved for $\gamma_0^{i,i}$ and $\gamma_1^{i,i}$.

The second alternative method is to state that, under the white Gaussian assumption for the input vector $\mathbf{X}_{m,n}$ and the error signal $\varepsilon(m,n)$, the weight-error correlation coefficients $\gamma_{k+1}^{i,i}$ for $k \gg 1$ can be approximated with zero. Thus, the available $(k+1)\times N^2$ equations can be solved for the $(k+1)\times N^2$ unknowns to obtain the weight-error covariance coefficients $\gamma_0^{i,i}$, $i = 0,\cdots,N^2-1$. The solution of these $(k+1)\times N^2$ simultaneous equations can be obtained using mathematical tool box for the general case. In the Appendix we show that the error that results from approximating the weight-error correlation coefficients $\gamma_{k+1}^{i,i}$, $k \gg 1$ with zeros, decreases as the spatial lag k increases.

In the following section we will discuss in more details the steady state analysis for the simple case when the input signal is white Gaussian noise.

4. Steady State MSE Analysis with White Gaussian Input Data

In this section we deal with the steady state analysis for the case when the input signal is white Gaussian noise with variance σ_x^2; the correlation coefficients $\gamma_2^{i,i}$, $i = 0,\cdots,N^2-1$ are set to zero; $f_h = f_v$, and $\mu_h = \mu_v = \mu$. We choose to work with this case merely to make the solution of the equations traceable. Similar kind of analysis can be applied to any other case within which A.1 and A.2 hold.

For the white Gaussian input case, $\lambda_0 = \lambda_1,\cdots,\lambda_{N^2-1} = \sigma_x^2$. Accordingly, $\gamma_0^{0,0} = \gamma_0^{1,1} = \gamma_0^{i,i} = \gamma_0$, $i = 0,\cdots,N^2-1$. Hence, solving Eqs. (25) and (26) for γ_0 we get

$$\gamma_0 = \frac{\sigma_\varepsilon^2}{\sigma_x^2}\frac{\zeta^2}{(0.25-\zeta+(2+p)\zeta^2)}\times \frac{0.375-\zeta+(0.5+1.5p)\zeta^2+(6+2p)\zeta^3-(4+2p)\zeta^4}{0.125+3\zeta-(2.5+1.5p)\zeta^2-(6+2p)\zeta^3+(4+2p)\zeta^4} \tag{28}$$

where, for notational convenience, we have defined $p = N^2$, and

$$\zeta = \mu\sigma_x^2. \tag{29}$$

Now, since the weight-error covariance coefficient $\gamma_0^{i,i}$, $i = 0,1,\cdots,N^2-1$ should be positive and finite, the range of the step size μ that ensures the convergence of the 2-D LMS in the MSE sense can be determined by the following condition

$$0 \le \gamma_0^{i,i} < \infty, \quad i = 0,1,\cdots,N^2-1. \tag{30}$$

For this simplified case, analysis of Eq. (28) reveals that in this equation, the first term and the numerator of the second term are always positive for $0 \le \zeta < 1$,

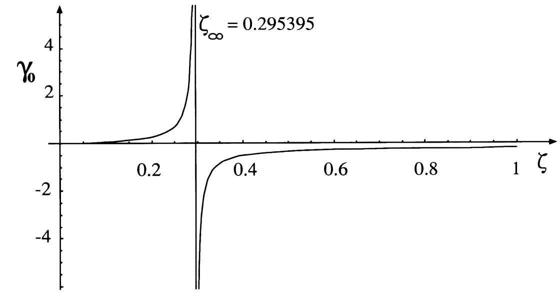

Fig. 2 Weight-error covariance coefficient γ_0 as a function of $\zeta = \mu\sigma_x^2$ for 2 by 2 adaptive FIR filter.

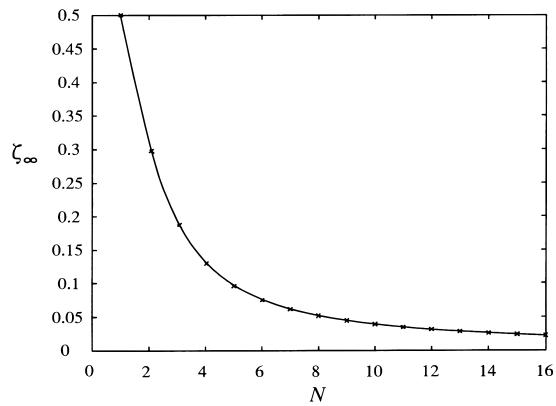

Fig. 3 The root value ζ_∞ versus N.

and that for any value of $N^2 \ge 1$, the polynomial in the denominator of the second term has only one real positive root, say ζ_∞, in the range $0 \le \zeta < 1$ where the sign of this polynomial changes from positive to negative. Thus, we can deduce that the upper bound of the step size value that ensures finite variance is given by

$$0 \le \mu < \frac{\zeta_\infty}{\sigma_x^2}. \tag{31}$$

Figure 2 shows γ_0 as a function of ζ with $\sigma_\varepsilon^2/\sigma_x^2 = 1$ for a 2 by 2, 2-D adaptive FIR filter. Figure 3 shows the values of the root ζ_∞ for different values of N. From Fig. 3, it is clear that for any filter order, $\zeta_\infty < 1$. Accordingly, the condition required for the convergence in the MSE sense, as given in Eq. (31), decreases significantly the convergence region of the 2-D LMS algorithm when comparing to the condition necessary for the convergence of the mean

$$0 \le \mu < \frac{1}{\sigma_x^2} \tag{32}$$

given by Eq. (9).

5. Simulation Results

5.1 Example 1

In this example we aim to test the accuracy of the obtained analytical results for the simplified setting ($\lambda_0 = \lambda_1, \cdots, \lambda_{N^2-1} = \sigma_x^2$, $f_h = f_v = 0.5$, and $\mu_h = \mu_v = \mu$). We performed system identification experiment for the following 2-D FIR filter:

$$d(m,n) = x(m,n) + 0.5x(m-1,n) + 0.5x(m,n-1) + 0.125\,x(m-1,n-1) + \varepsilon(m,n). \quad (33)$$

We used two independent, 2-D white Gaussian sequences with variances $\sigma_x^2 = 1$, and $\sigma_\varepsilon^2 = 1$ for the input signal $x(m,n)$ and the additive noise $\varepsilon(m,n)$ respectively.

As a measure for the performance of the 2-D LMS we used the misadjustment M which is defined as

$$\begin{aligned} \mathrm{M} &= \frac{\epsilon_\infty - \sigma_\varepsilon^2}{\sigma_\varepsilon^2} \\ &= \frac{1}{\sigma_\varepsilon^2} \lim_{m+n\to\infty} \mathrm{tr}(\mathbf{R}\mathbf{K}_{m,n;m,n}) \\ &= \frac{1}{\sigma_\varepsilon^2} \sum_{j=0}^{N^2-1} \gamma_0^{j,j} \lambda_j. \end{aligned} \quad (34)$$

Figure 4 shows a comparison between experimental results and the misadjustment obtained using two different methods. In the first method (referred to as the independent assumption method in Fig. 4), the coefficients of the WECM in Eq. (34) were calculated using Eq. (28). And in the second (referred to as the the direct averaging method in Fig. 4), the WECM in Eq. (34) were calculated using the direct averaging method presented in the Appendix with k set to zero in Eq. (A· 8). The experimental misadjustment is calculated by averaging the results of 30 independent runs. For each run the misadjustment is calculated by averaging 40000 iterations in the steady state.

From Fig. 4, we can observe that the MSE analysis using both the independent assumption and the direct averaging method gives satisfactory results for small step size values. However, as the step size μ increases, the error in the estimated MSE increases. On the other hand, we can notice that the performance of the 2-D LMS is well preserved using the independence assumption based analysis, whereas, the direct averaging method fails completely for large step size values.

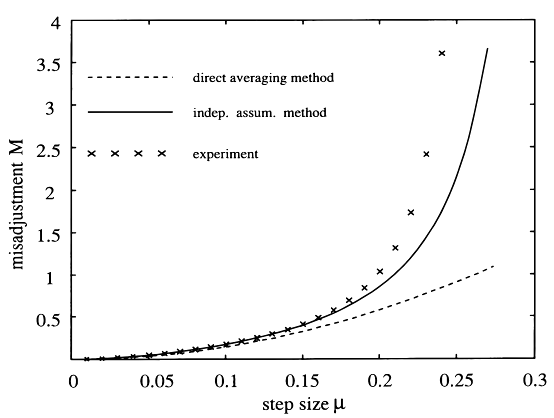

Fig. 4 Comparison of the experimental results with the theoretical values for the misadjustment of the 2-D LMS in the simplified setting ($\lambda_0 = \lambda_1 = \cdots = \lambda_{N^2-1} = \sigma_x^2$, $f_h = f_v = 0.5$, and $\mu_h = \mu_v = \mu$).

5.2 Example 2

In this example we performed the system identification experiment described in Example 1, however, with correlated input signal. The correlated input $x(m,n)$ was generated by filtering a 2-D white Gaussian noise $u(m,n)$ of zero mean and unit variance with the following 2-D, 2×2 filter:

$$x(m,n) = u(m,n) + \alpha u(m-1,n) + \alpha u(m,n-1). \quad (35)$$

Accordingly, the 4×4 input correlation matrix $\mathbf{R}$ is given by:

$$\mathbf{R} = \begin{bmatrix} 1+\alpha^2 & \alpha & \alpha & 0 \\ \alpha & 1+\alpha^2 & \alpha^2 & \alpha \\ \alpha & \alpha^2 & 1+\alpha^2 & \alpha \\ 0 & \alpha & \alpha & 1+\alpha^2 \end{bmatrix}. \quad (36)$$

We repeated the same experiment for different values of α, ($\alpha = 0.2, 0.3, 0.4$), to test the accuracy of the obtained analytical results for different levels of input correlation. In this example, an approximation of the weight-error correlation coefficients $\gamma_2^{i,i}$ was calculated using the proposed direct averaging method (Eq. (A· 8) with $k = 2$). The weight-error correlation coefficients $\gamma_0^{i,i}$, $i = 0, \cdots, N^2 - 1$ were then obtained by solving Eqs. (25) and (26).

Table 1 shows the values of the misadjustment calculated both experimentally and using the proposed independence assumption based analysis. From the table entries, it is seen that the independence assumption based analysis provides accurate results for small step size values. However, for large step size values, the error in estimating the misadjustment of the adaptive filter increases as the level of the input correlation increases.

The upper bound on the step size parameter μ that ensures the convergence in the MSE, say μ_{max}, were calculated for each particular value of α from condition (30). Table 2 shows the obtained numerical results in comparison with the upper bound on the step size parameter μ that ensures convergence of the mean as given by Eq. (9) [3]. It is seen that the maximum step size values that ensure convergence of the MSE are significantly smaller than those that ensure the convergence of the mean.

5.3 Example 3

In this example we aim to test the obtained analytical

Table 1 Simulation results for Example 2, correlated Gaussian input.

step size μ	$\alpha = 0.2$ $\beta = 0.19$		$\alpha = 0.3$ $\beta = 0.16$		$\alpha = 0.4$ $\beta = 0.11$	
	Indep.	Exper.	Indep.	Exper.	Indep.	Exper.
$0.3 \times \beta$	0.0929	0.0929	0.0821	0.0821	0.0544	0.0545
$0.4 \times \beta$	0.1522	0.1522	0.1339	0.1341	0.0877	0.0901
$0.5 \times \beta$	0.2288	0.2289	0.2001	0.2010	0.1283	0.1287
$0.6 \times \beta$	0.3297	0.3316	0.2856	0.2871	0.1776	0.1864
$0.7 \times \beta$	0.4661	0.5011	0.3989	0.4783	0.2381	0.3180
$0.8 \times \beta$	0.6600	0.8611	0.5550	1.1034	0.3130	0.6101

Table 2 Simulation results for Example 2, maximum step size value μ_{max}, Gaussian input, $N = 2$.

α	0	0.1	0.2	0.3	0.4	0.5
Ref. [3]	1	0.81	0.67	0.55	0.46	0.38
Indep.	0.29	0.26	0.22	0.19	0.16	0.14

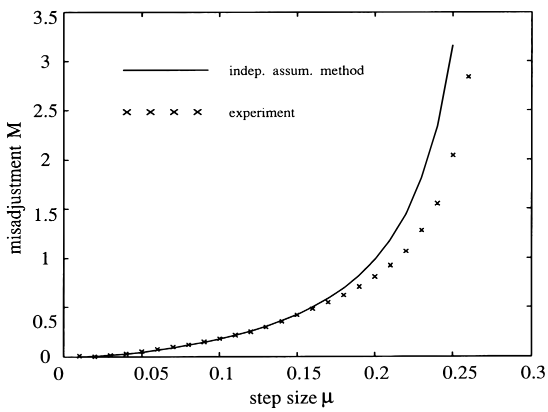

Fig. 5 Comparison of the experimental results with the theoretical values for the misadjustment of the 2-D LMS in the simplified setting ($\lambda_0 = \lambda_1 = \cdots = \lambda_{N^2-1} = \sigma_x^2$, $f_h = f_v = 0.5$, and $\mu_h = \mu_v = \mu$), i.i.d. binary input.

results for non Gaussian input. We performed a system identification experiment similar to that presented in Example 1, however with uncorrelated binary input signal of unit variance. Figure 5 shows the values of the misadjustment obtained both experimentally and using the independence assumption analysis. It can be seen that the independence assumption-based analysis can serve to give good insight to the behavior of the adaptive process even when the Gaussian assumption does not hold.

6. Conclusions

We have considered the steady state MSE analysis for 2-D LMS algorithm using the independence assumption. We have shown that the evaluation of the weight-error covariance matrix for doubly-indexed 2-D LMS algorithm requires approximation of the weight error correlation coefficients at large spatial lags. Then, we have proposed a method to solve this problem. We have shown that the convergence in the MSE sense occurs for step size range that is significantly smaller than the one necessary for the convergence of the mean. Simulation examples were presented to support the analytical results and to show that the analysis using the independence assumption does provide good insight to the performance of the 2-D LMS algorithm.

References

[1] M.M. Hadhoud and D.W. Thomas, "The two-dimensional adaptive LMS (TDLMS) algorithm," IEEE Trans. Circuits & Syst., pp.485–494, May 1988.
[2] B. Widrow and S.D. Stearn, "Adaptive Signal Processing," Prentice-Hall, NJ, 1985.
[3] M. Ohki and S. Hashiguchi, "Two-dimensional LMS adaptive FIR filters," IEEE Trans. Consumer Electronics, vol.37, no.1, pp.66–73, Feb. 1991.
[4] E. Fornasini and G. Marchesini, "Stability analysis of 2-D systems," IEEE Trans. Circuits & Syst., vol.CAS-27, pp.2085–2088, Dec. 1980.
[5] V. Solo and X. Kong, "Adaptive Signal Processing Algorithms," Prentice-Hall, NJ, 1995.
[6] S. Haykin, "Adaptive Filter Theory," Third edition, Prentice-Hall, NJ, 1996.
[7] B. Widrow, J.M. McCool, M.G. Larimore, and C.R. Johnson, "Stationary and nonstationary learning characteristic of the LMS adaptive filter," Proc. IEEE, vol.64, no.8, pp.1151–1162, Aug. 1976.
[8] L.L. Horowitz and K.D. Senne, "Performance advantage of complex LMS for controlling narrow-band adaptive array," IEEE Trans. Circuits & Syst., vol.CAS-28, pp.562–576, June 1981.
[9] M.A. Feuer and E. Weinstein, "Convergence analysis of LMS filters with uncorrelated Gaussian data," IEEE Trans. Acoust., Speech & Signal Process., vol.ASSP-33, pp.222–230, Feb. 1985.
[10] S. Florian and A. Feuer, "Performance analysis of the LMS algorithm with a tapped delay line (two dimensional case)," IEEE Trans. Acoust., Speech & Signal Process., vol.ASSP-34, pp.1542–1549, Dec. 1986.
[11] S.D. Douglas and W. Pan, "Exact expectation analysis of the LMS adaptive filter," IEEE Trans. Signal Processing, vol.43, no.12, pp.2863–2871, Dec. 1995.
[12] A. Papoulis, "Probability, Random Variables, and Stochastic Processes," McGraw-Hill, 1984.

Appendix: Direct Averaging Method for the Approximation of the Weight-Error Correlation Matrix

Providing that the step sizes μ_h and μ_v are small, and based on the direct averaging method [5], the solution of the stochastic difference Eq. (16) can be approximated with that of the following averaged system:

$$\mathbf{C}_{m+1,n+1} = \mathbf{A}_h\mathbf{C}_{m,n+1} + \mathbf{A}_v\mathbf{C}_{m+1,n} + \mu_h\varepsilon_{m,n+1}\mathbf{X}_{m,n+1} + \mu_v\varepsilon_{m+1,n}\mathbf{X}_{m+1,n} \quad \text{(A·1)}$$

where

$$\mathbf{A}_h = f_h\mathbf{I} - \mu_h\mathbf{R}$$
$$\mathbf{A}_v = f_v\mathbf{I} - \mu_v\mathbf{R}.$$

Equation (A·1) is a 2-D F-M state space model with local state space vector $\mathbf{C}_{m,n}$ and input vector $\varepsilon_{m,n}\mathbf{X}_{m,n}$. This 2-D F-M model is exponentially stable if and only if [4]

$$\det(\mathbf{I} - z_1^{-1}\mathbf{A}_h - z_2^{-1}\mathbf{A}_v) \neq 0 \quad \text{(A·2)}$$

in the region

$$U_2 = \{(z_1, z_2)| \quad |z_1| \geq 1, |z_2| \geq 1\}.$$

Note that the condition (A·2) is the same condition required for the convergence of the mean which was reduced in [3] to the condition (9).

Now, the transfer function between the input $\varepsilon_{m,n}\mathbf{X}_{m,n}$ and the state space vector $\mathbf{C}_{m,n}$ is given by

$$\begin{aligned} H(z_1, z_2) &= (I - \mathbf{A}_h z_1^{-1} - \mathbf{A}_v z_2^{-1})^{-1}(\mu_h z_1^{-1} + \mu_v z_2^{-1}) \\ &= (\mu_h z_1^{-1} + \mu_v z_2^{-1}) \sum_{k=0}^{\infty} (\mathbf{A}_h z_1^{-1} + \mathbf{A}_v z_2^{-1})^k \\ &= (\mu_h z_1^{-1} + \mu_v z_2^{-1}) \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \mathbf{A}^{i,j} z_1^{-i} z_2^{-j} \end{aligned} \quad \text{(A·3)}$$

where the series expansion is absolutely convergent in the region U_2 [4], and

$$\mathbf{A}^{0,0} = \mathbf{I}$$
$$\mathbf{A}^{i,j} = \mathbf{A}_h\mathbf{A}^{i-1,j} + \mathbf{A}_v\mathbf{A}^{i,j-1}, \text{ for } i+j > 0$$
$$\mathbf{A}^{i,j} = \mathbf{0}, \text{ for } i < 0 \text{ or } j < 0. \quad \text{(A·4)}$$

Hence, from Eq. (A·3), the weight-error vector $\mathbf{C}_{m,n}$ can be calculated by

$$\mathbf{C}_{m,n} = \sum_{i=0}^{m}\sum_{j=0}^{n} H(i,j)\varepsilon_{m-i,n-j}\mathbf{X}_{m-i,n-j} \quad \text{(A·5)}$$

with

$$H(i,j) = \mu_h\mathbf{A}^{i-1,j} + \mu_v\mathbf{A}^{i,j-1}.$$

From Eq. (A·5), the weight-error correlation matrix $\mathbf{K}_{m,n-k;m-k,n}$ can be calculated for any spatial lag k as follows:

$$\begin{aligned} &\mathbf{K}_{m,n-k;m-k,n} \\ &= \mathrm{E}\{\mathbf{C}_{m,n-k}\mathbf{C}^t_{m-k,n}\} \\ &= \sum_{i=0}^{m}\sum_{j=0}^{n}\sum_{p=0}^{m}\sum_{q=0}^{n} H(i,j-k) \\ &\quad \mathrm{E}\{\mathbf{V}_{m-i,n-j}\mathbf{V}^t_{m-p,n-q}\}H(p-k,q) \end{aligned} \quad \text{(A·6)}$$

where, for notational convenience, we have defined:

$$\mathbf{V}_{m,n} = \varepsilon_{m,n}\mathbf{X}_{m,n}. \quad \text{(A·7)}$$

If the probability distribution of the input signal $x(m,n)$ and the measurement noise $\varepsilon_{m,n}$ are available, Eq. (A·6) can be used to obtain the weight-error correlation matrix $\mathbf{K}_{m,n-k;m-k,n}$.

For the special case when the measurement noise $\varepsilon(m,n)$ is white Gaussian noise and independent of $x(m,n)$, Eq. (A·6) is reduced to:

$$\begin{aligned} &\mathbf{K}_{m,n-k;m-k,n} \\ &= \sigma_\varepsilon^2 \sum_{i=k}^{m}\sum_{j=k}^{n} H(i,j-k)\ \mathbf{R}\ H(i-k,j). \end{aligned} \quad \text{(A·8)}$$

Stability condition (9) guarantees that the spectral norm of each of the matrices $\mathbf{A}_h$, $\mathbf{A}_v$, and $\mathbf{A}^{i,j}$ are less than unity. And since these matrices are symmetric, it is straightforward to show that, $\lim_{i,j\to\infty} H(i,j) = 0$. Thus, we can deduce that the error that results from using the approximation (A·8) decreases as the spatial lag k increases. For sufficiently large k, the correlation matrix $\mathbf{K}_{m,n;m,n}$ can be approximated with zero as it has been suggested in Sect. 3.2. For $k = 0$, Eq. (A·8) can be used as an approximation of the weight error covariance matrix $\mathbf{K}_{m,n;m,n}$.

Maha Shadaydeh received the B.E. in electronic engineering from Damascus University (Syria) and M.E. degree in information and computer sciences from Toyohashi University of Technology (Japan) in 1990 and 1996, respectively. From 1990 to 1991 she was with the department of automatic control in the High Institute for Applied Sciences and Technology (Syria). From 1991 to 1993 she worked as assistant lecturer in the electronic engineering department, Damascus University. She is currently working towards the doctoral degree of electrical communication engineering at Tohoku University, Sendai, Japan. Her research interests include 1-D and 2-D adaptive signal processing and higher order statistics for signal processing applications.

Masayuki Kawamata received B.E., M.E., and D.E. degrees in electronic engineering from Tohoku University, Sendai, Japan, in 1977, 1979, and 1982, respectively. He was an Associate Professor in the Graduate School of Information Sciences at Tohoku University and is currently a Professor in the Graduate School of Engineering at Tohoku University. His research interests include 1-D and multi-D digital signal processing, intelligent signal processing, fractals and chaos in signal processing, control theory, and linear system theory. Dr. Kawamata received the Outstanding Transactions Award from the Society of Instrument and Control Engineers of Japan in 1984 (with T. Higuchi), the Outstanding Literary Work Award from the Society of Instrument and Control Engineering of Japan in 1996 (with T. Higuchi) and the 11^{th} IBM-Japan Scientific Award in Electronics in 1997. He is a member of the IEEE, the Society of Instrument and Control Engineers of Japan, the Information Processing Society of Japan, and Robotics Society of Japan. He is an IEEE Senior Member.