STOCHASTIC MODEL FOR THE NLMS ALGORITHM WITH CORRELATED GAUSSIAN DATA

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 proposes a new stochastic model for the normalized LMS (NLMS) algorithm under correlated input data. The proposed model is derived without invoking the simplifying assumption that $\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)$ has a chi-square distribution to determine $E\{1/[\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)/N]\}$. Under correlated input data that assumption is not correct and thus the resulting model becomes inaccurate. Without considering such simplifying assumption, a high-order hyperelliptic integral has to be computed. The proposed model is based on tackling the solution of that integral. Numerical simulations verify the quality of the proposed model.

1. INTRODUCTION

When an adaptive algorithm is modeled some simplifying assumptions must be considered which make the involved mathematics more tractable. The key point to obtain a reasonable modeling (accurate) resides on the degree of truth of the simplifying assumptions used. Accurate models are frequently associated with a complex modeling mathematics; in other words, the assumptions allowed for in such a process should be confirmed for a wide range of operating conditions. For instance, the stochastic modeling of the filtered-X LMS (FXLMS) algorithm under the light of the independence theory (typically used for modeling the LMS-like algorithms) [1] is not consistent, since such an assumption disregards the correlations between input signal vectors at different time lags created by the secondary path [2], [3]. In this way, the resulting model is simple but inaccurate to describe the algorithm behavior. On the other hand, by considering all correlations of the input signal vectors the obtained model is more complex; however, it becomes much more accurate. It should be pointed out that, in general, more than a single simplifying assumption is used for algorithm modeling. However, the model quality depends on the validity of all assumptions considered. Concerning the modeling of the normalized LMS (NLMS) algorithm, it reveals a particular obstacle to calculate the expectation $E\{\mathbf{x}(n)\mathbf{x}^{\mathrm{T}}(n)/[\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)/N]\}$, which is needed to obtain mean weight behavior. In [4], such an expectation is determined by using the multivariate Gaussian density function of the input signal vector. In doing so, there is no approximation in this process; however, such a procedure leads to the computation of a high-order hyperelliptic integral. Since in the open literature there is no available solution for that kind of integral some shortcuts must be adopted to determine such an expectation. Recently in [5], the expectation $E\{\mathbf{x}(n)\mathbf{x}^{\mathrm{T}}(n)/[\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)/N]\}$ has been approximated by $E\{1/[\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)/N]\}E[\mathbf{x}(n)\mathbf{x}^{\mathrm{T}}(n)]$ by invoking the Averaging Principle [6], where $E\{1/[\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)/N]\}$ is obtained by assuming a chi-square distribution. Such assumption is true only for white input signals. For correlated signals it fails, resulting in a model mismatched with respect to the simulations, mainly during the transient phase. In this work, we return to the Bershad's procedure [4], which provides an exact modeling, and we tackle the solution of the hyperelliptic integral problem. We propose here an accurate model for the NLMS algorithm with correlated input data, obtained by solving a high-order indefinite hyperelliptic integral. As a result, the new stochastic model presents a very good agreement for both transient and steady-state phases as compared with numerical simulations considering correlated Gaussian input data.

2. PRELIMINARIES

This section presents the weight update equation of the NLMS algorithm [1]. In addition, the simplifying assumptions permitting to obtain the model expression, which is solved by using a new approach, are stated. The weight updating equation is given by [4]

$$\mathbf{w}(n+1) = \mathbf{w}(n) + 2\mu \frac{e(n)\mathbf{x}(n)}{\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)/N}, \tag{1}$$

where $e(n) = d(n) - \mathbf{w}^{\mathrm{T}}(n)\mathbf{x}(n) + z(n)$ is the error signal and $z(n)$ represents a measurement noise, i.i.d., zero-mean, with variance σ_z^2 uncorrelated with any other signal in the system. The input vector is denoted by $\mathbf{x}(n) = [x(n)\; x(n-1) \cdots x(n-N+1)]^{\mathrm{T}}$ and its variance is σ_x^2. By taking the expected value on both sides of (1), it results in

$$\begin{aligned} E[\mathbf{w}(n+1)] = E[\mathbf{w}(n)] &+ 2\mu E\left[\frac{\mathbf{x}(n)d(n)}{\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)/N}\right] \\ &- 2\mu E\left[\frac{\mathbf{x}(n)\mathbf{x}^{\mathrm{T}}(n)\mathbf{w}(n)}{\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)/N}\right] \end{aligned} \tag{2}$$

To determine the expectations in (2), the following simplifying assumptions are used:

A1) $\mathbf{w}(n)$ and $\mathbf{x}(n)$ are statistically independent [1].

A2) $1/[\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)/N]$ and $\mathbf{x}(n)\mathbf{x}^{\mathrm{T}}(n)$ are jointly stationary processes, where $1/[\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)/N]$ is slowly varying with respect to $\mathbf{x}(n)\mathbf{x}^{\mathrm{T}}(n)$. In this way, the Averaging Principle can be invoked [6], resulting in

$$E\{\mathbf{x}(n)\mathbf{x}^{\mathrm{T}}(n)/[\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)/N]\} \approx E\{1/[\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)/N]\}E[\mathbf{x}(n)\mathbf{x}^{\mathrm{T}}(n)]. \tag{3}$$

After using assumptions A1 and A2 in (2), we obtain the following expression:

$$\begin{aligned} E[\mathbf{w}(n+1)] = E[\mathbf{w}(n)] &+ 2\mu \underbrace{E\left[\frac{1}{\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)/N}\right]}_{\mathrm{A}}\mathbf{p} \\ &- 2\mu \underbrace{E\left[\frac{1}{\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)/N}\right]}_{\mathrm{A}}\mathbf{R}, \end{aligned} \tag{4}$$

where $\mathbf{R} = E[\mathbf{x}(n)\mathbf{x}^{\mathrm{T}}(n)]$ and $\mathbf{p} = E[d(n)\mathbf{x}(n)]$ are the input autocorrelation matrix and the cross-correlation vector between the desired response and the input vector [1]. Then, the main point now is to determine the expected values A in (4).

3. DETERMINATION OF $E\{1/[\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)/N]\}$

3.1 Problem statement

Regarding the modeling of the NLMS algorithm, a major mathematical complexity arises when the expectation

$$E\left[\frac{1}{\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)/N}\right], \tag{5}$$

must be determined. In [4] there is no approximation for computing the expected value of the term containing $\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)/N$ in the denominator. Such a procedure is notably complex, since it requires at the end that a high-order indefinite hyperelliptic integral is solved. However, the solution of this integral does not yet have a closed formulation as pointed out by the area literature. Recently in [5], expression (5) has been determined considering that $\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)$ has a chi-square distribution with N degrees of freedom. Therefore,

$$E\left[\frac{1}{\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)/N}\right] = \frac{N}{(N-2)\,\sigma_x^2}. \tag{6}$$

Such an assumption is true if and only if $\{x(n)\}$ is an independent Gaussian zero mean random variable [7], thus failing for correlated ones.

3.2 Proposed approach

Since the modeling of the NLMS algorithm also requires the determination of expected values such as

$$E\left\{\frac{1}{[\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)/N]^2}\right\}, \tag{7}$$

we propose a method to solve the general problem of the expected values in question. Thus,

$$E\left\{\frac{N^k}{[\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)]^k}\right\} = \underbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}_{N\ \text{fold}} \frac{N^k}{[\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)]^k} f(\mathbf{x})d\mathbf{x}. \tag{8}$$

In (8) $f(\mathbf{x})$ represents the multivariate Gaussian density function of the input vector $\mathbf{x}(n)$. It is given by

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{N/2}[\det(\mathbf{R})]^{1/2}} \mathrm{e}^{-\frac{\mathbf{x}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{x}}{2}}, \tag{9}$$

where $\det(\cdot)$ denotes the determinant of a matrix.

Now, substituting (9) into (8) and defining a function $F(\omega)$ so that

$$F(\omega) = \frac{N^k}{(2\pi)^{N/2}[\det(\mathbf{R})]^{1/2}} \underbrace{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}}_{N\ \text{fold}} \frac{1}{(\mathbf{x}^{\mathrm{T}}\mathbf{x})^k} \mathrm{e}^{-\omega(\mathbf{x}^{\mathrm{T}}\mathbf{x})}\, \mathrm{e}^{-\frac{\mathbf{x}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{x}}{2}} d\mathbf{x}, \tag{10}$$

we return the expected value in question, making $\omega = 0$. Thus,

$$E\left\{\frac{N^k}{[\mathbf{x}^{\mathrm{T}}(n)\mathbf{x}(n)]^k}\right\} = F(\omega)\big|_{\omega=0} = F(0). \tag{11}$$

By differentiating (10) k times and using the properties of the Gaussian density function, we obtain

$$\frac{d^k F(\omega)}{d\omega^k} = \frac{(-1)^k N^k}{\sqrt{\det(\mathbf{B}^{-1}\mathbf{R})}}, \tag{12}$$

where, from the simple algebraic manipulation, we get

$$\mathbf{B}^{-1} = 2\omega\mathbf{I} + \mathbf{R}^{-1}, \tag{13}$$

where $\mathbf{I}$ represent the identity matrix.

Now, substituting (13) into (12), and integrating k times the resulting expression, we obtain

$$F(\omega) = (-1)^k N^k \underbrace{\int\cdots\int}_{k\ \text{fold}} \frac{1}{[\det(\mathbf{I}+2\omega\mathbf{R})]^{1/2}} \underbrace{d\omega\ldots d\omega}_{k\ \text{fold}} + C. \tag{14}$$

The constant C is obtained by considering that $\lim_{\omega\to\infty} F(\omega) = 0$. In the literature there does not exist a closed form for a solution of a high-order indefinite hyperelliptic integral (also known as Abelian integral) (14). However, by considering the particular form of the autocorrelation matrix $\mathbf{R}$, we propose the following approach to solve such an integral:

i) Decompose $\mathbf{R}$ according to $\mathbf{R} = \mathbf{Q\Lambda Q}^{\mathrm{T}}$, where $\mathbf{Q}$ is the eigenvector matrix and $\mathbf{\Lambda}$ is a diagonal matrix containing the eigenvalues λ_i of $\mathbf{R}$, respectively. It results in the polynomial of N^{th}-degree. Thus,

$$\det(\mathbf{I}+2\omega\mathbf{R}) = \prod_{i=1}^{N}(1+2\omega\lambda_i). \tag{15}$$

ii) From (15), the coefficient a_N of ω^N is $a_N = 2^N\prod_{i=1}^{N}\lambda_i$ and the roots of (15) are $\omega_i = -1/(2\lambda_i)$.

iii) Substitute adjacent pairs of roots of (15) by its geometric mean

$$\omega_k' = -\sqrt{\omega_i\omega_j}. \tag{16}$$

In this way, now (15) will have roots with multiplicity two. Thus, we can rewrite (14) as follows:

$$\int\frac{d\omega}{[\det(\mathbf{I}+\beta\mathbf{R})]^{1/2}} = \frac{1}{\sqrt{a_N}}\int\frac{d\omega}{\left|(\omega-\omega_1')\cdots(\omega-\omega_{N/2}')\right|} + C \tag{17}$$

Since we are interested in the result for $\omega = 0$, the absolute value operator can be disregarded.

Then, considering partial fraction expansion of (17) the result of the integral is

$$\int \frac{d\omega}{[\det(\mathbf{I}+\omega\mathbf{R})]^{1/2}} = \frac{1}{\sqrt{a_N}} \sum_{i=1}^{N/2} A_i \ln(\omega-\omega_i') + C \ , \tag{18}$$

where

$$A_i = \frac{1}{\prod\limits_{\substack{j=1 \\ j\neq i}}^{N/2} (\omega_i' - \omega_j')} . \tag{19}$$

Thus, by using the proposed approach, the expected values needed for modeling the NLMS algorithm are now given by

$$E\left[\frac{N}{\mathbf{x}^T(n)\mathbf{x}(n)}\right] = F(\omega)\big|_{\omega=0} = \frac{-N}{\sqrt{a_N}} \sum_{i=1}^{N/2} A_i \ln(-\omega_i), \tag{20}$$

and

$$E\left\{\frac{N^2}{[\mathbf{x}^T(n)\mathbf{x}(n)]^2}\right\} = \frac{N^2}{\sqrt{a_N}} \sum_{i=1}^{N/2} A_i\left[-\omega_i \ln(-\omega_i) + \omega_i\right]. \tag{21}$$

3.3. Learning curve and second moment of the weight-error matrix

By defining the weight-error vector as $\mathbf{v}(n) = \mathbf{w}(n) - \mathbf{w}_{\mathrm{opt}}$, the error signal is given by

$$\begin{aligned} e(n) &= d(n) - \mathbf{x}^{\mathrm{T}}(n)\mathbf{v}(n) - \mathbf{x}^{\mathrm{T}}(n)\mathbf{w}_{\mathrm{opt}} + z(n) \\ &= e_{\mathrm{o}}(n) - \mathbf{x}^{\mathrm{T}}(n)\mathbf{v}(n), \end{aligned} \tag{22}$$

where $e_{\mathrm{o}}(n)$ is defined as

$$e_{\mathrm{o}}(n) = d(n) - \mathbf{x}^{\mathrm{T}}(n)\mathbf{w}_{\mathrm{opt}} + z(n) \ . \tag{23}$$

Squaring both sides of (23) and calculating the expected value of the resulting expression, according to A1 and A2, and using the Orthogonality Principle [1], we obtain

$$\begin{aligned} E[e^2(n)] &= e_{\min} + E[\mathbf{x}^{\mathrm{T}}(n)\mathbf{v}(n)\mathbf{x}^{\mathrm{T}}(n)\mathbf{v}(n)] \\ &= e_{\min} + \mathrm{tr}\{\mathbf{R}E[\mathbf{v}(n)\mathbf{v}^{\mathrm{T}}(n)]\}, \end{aligned} \tag{24}$$

where $e_{\min} = E[e_{\mathrm{o}}^2(n)]$ is the minimum error in steady-state.

Note that (24) is completely determined if the weight-error covariance matrix $\mathbf{K}(n) = E[\mathbf{v}(n)\mathbf{v}^{\mathrm{T}}(n)]$ is known. Then, by subtracting $\mathbf{w}_{\mathrm{opt}}$ from both sides of (1), determining the outer product $\mathbf{v}(n)\mathbf{v}^{\mathrm{T}}(n)$, and taking the expectation on both sides of the resulting expression according to the simplifying assumptions A1 and A2, we obtain

$$\begin{aligned} \mathbf{K}(n+1) = \mathbf{K}(n) &- 2\mu\mathbf{K}(n)\mathbf{R}E_1 - 2\mu E_1\mathbf{R}\mathbf{K}(n) \\ &+ 4\mu^2 E_1\mathbf{R}\mathbf{K}(n)\mathbf{R}E_1 + 4\mu^2 E_2\mathbf{R}\mathbf{K}(n)\mathbf{R} \\ &+ 4\mu^2 E_1\mathbf{R}\,\mathrm{tr}\left\{E_1\mathbf{R}\mathbf{K}(n)\right\} + 4\mu^2 E_1\mathbf{R}E_1\, e_{\min}, \end{aligned} \tag{25}$$

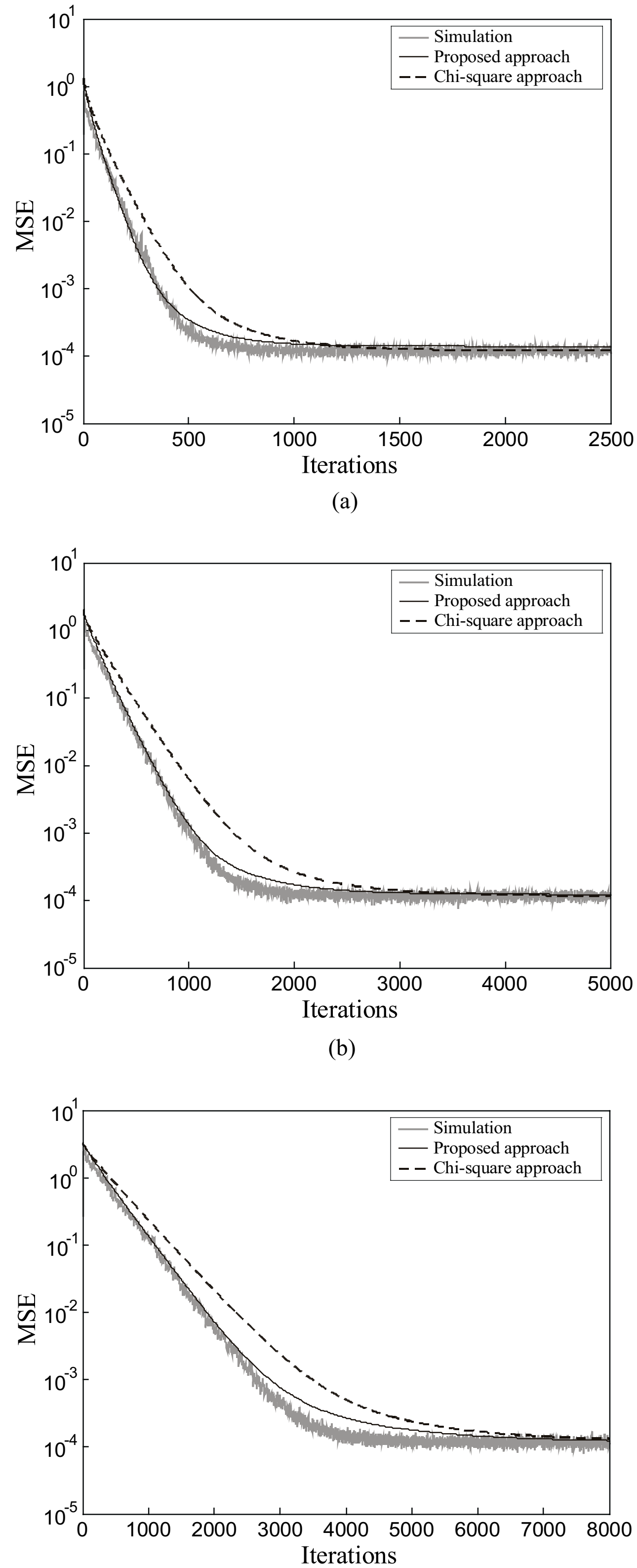
where

$$E_1 = E\left[\frac{N}{\mathbf{x}^T(n)\mathbf{x}(n)}\right], \tag{26}$$

and

$$E_2 = E\left\{\frac{N^2}{[\mathbf{x}^T(n)\mathbf{x}(n)]^2}\right\} \tag{27}$$

are determined from (20) and (21), respectively.



(c)

Fig. 1. Learning curves for correlated input signal. (a) $N = 8$ and eigenvalue dispersion of 46. (b) $N = 16$ and eigenvalue dispersion of 147. (c) $N = 32$ and eigenvalue dispersion of 220.

4. SIMULATION RESULTS

To assess the accuracy of the proposed model some examples are presented considering a system identification problem. The input signal is correlated, obtained from an AR(2) process given by $x(n) = a_1 x(n-1) + a_2 x(n-2) + v(n)$, where $v(n)$ is a white noise signal with variance σ_v^2. The AR coefficients are $a_1 = 1.3214$ and $a_2 = -0.8500$. The measurement noise $z(n)$ has a variance $\sigma_z^2 = 10^{-4}$ (SNR $= 40$dB). In the examples, the step-size values are selected to be $0.1\mu_{\max}$, where $\mu_{\max}$ is the maximum step-size for algorithm stability, determined experimentally.

The learning curves of Fig. 1 are obtained by using adaptive filters with length-8, 16, and 32 and with eigenvalue dispersions of 46, 147, and 220, respectively. In all cases the results obtained with the chi-square approach, proposed model, and Monte Carlo method (average of 200 independent runs) are shown. From these curves, the improved accuracy of the proposed model obtained by now considering the proper correlations of the input data can be observed.

5. CONCLUSIONS

This paper has presented a new stochastic model for the NLMS algorithm under a slow adaptation condition. Due to the analysis procedure used, the proposed model exhibits a satisfactory accuracy for correlated input data. Numerical simulations confirm the quality of the new obtained model.

6. REFERENCES

[1] S. Haykin, *Adaptive Filter Theory*, 4th ed., Upper Saddler River, NJ: Prentice-Hall, 2002.

[2] E. Bjarnason, "Analysis of the filtered-X LMS algorithm," *IEEE Trans. Speech and Audio Processing*, vol. 3, no. 6, pp. 504-514, Nov. 1995.

[3] O. J. Tobias, J. C. M. Bermudez, and N. J. Bershad, "Mean weight behavior of the filtered-X LMS algorithm," *IEEE Trans. Signal Processing*, vol. 48, no. 4, pp. 1061-1075, Apr. 2000.

[4] N. J. Bershad, "Analysis of the normalized LMS algorithm with Gaussian inputs," *IEEE Trans. Acoust., Speech, and Signal Processing*, vol. 34, no. 4, pp. 793-806, Aug. 1986.

[5] M. H. Costa and J. C. M. Bermudez, "An improved model for the normalized LMS algorithm with Gaussian inputs and large number of coefficients," *IEEE Int. Conf. Acoust., Speech, and Signal Processing*, Orlando, USA, vol. II, May 2002, pp. 1385-1388.

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

[7] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, 3ed., McGraw-Hill, 1991.