STEADY-STATE ANALYSIS OF THE NORMALIZED LEAST MEAN FOURTH ALGORITHM WITHOUT THE INDEPENDENCE AND SMALL STEP SIZE ASSUMPTIONS

Muhammad Moinuddin and Azzedine Zerguine

Electrical Engineering Department
King Fahd University of Petroleum & Minerals
Dhahran, 31261, Saudi Arabia.
E-mail:{moinudin, azzedine}@kfupm.edu.sa

ABSTRACT

In this work, the steady-state analysis of the Normalized Least Mean Fourth (NLMF) algorithm under very weak assumptions is investigated. No restrictions are made on the dependence between input successive regressors, the dependence among input regressor elements, the length of the adaptive filter, the distribution of noise and the filter input. Moreover, in our approach, there is no restriction made on the step size value and therefore the analysis holds for all the values of the step size in the range where the NLMF algorithm is stable. The analysis is based on the effective weight deviation vector performance measure [1]. This vector is the component of weight deviation vector in the direction of the input regressor. The asymptotic time-averaged convergence for the mean square effective weight deviation, the mean absolute excess estimation error, and the mean square excess estimation error for the NLMF algorithm are derived. Finally, a number of simulation results are carried out to corroborate the theoretical findings.

Index Terms — Adaptive filters, NLMF algorithm, Convergence Analysis.

1. INTRODUCTION

The Normalized Least Mean Fourth (NLMF) algorithm [2] is the normalized version of the Least Mean Fourth (LMF) algorithm [3]. The analysis of the NLMF becomes difficult because of the normalization term. Therefore, until now, the analysis of the NLMF algorithm is carried out using some strong assumptions [4, 5], for example, using the independence assumption [6] or the long filter assumption [7]. Recently a new performance measure, the effective weight deviation vector, is introduced for the convergence analysis of the NLMS algorithm [1]. This vector is the component of weight deviation vector in the direction of input regressor vector. It is shown that the effective weight deviation is the only component that contributes to the excess estimation error [1]. Therefore, the analysis based on the study of this component can give more insight on the performance of the adaptive algorithm. In this work, we have used the framework of [1] for the analysis of the NLMF algorithm using the concept of the effective weight deviation vector.

The main contribution of this paper is a rigorous convergence analysis of the NLMF algorithm that has the following advantages: (1) it holds for arbitrary dependence among successive regressor vectors, (2) it holds for arbitrary dependence among the elements of regressor vector, (3) this analysis is not restricted to the class of long filters, (4) it holds for arbitrary distributions of the filter input and the noise, and (5) it holds for all the values of the step size in the range that insures the stability of the NLMF algorithm.

The paper is organized as follows. After introducing the system model in the following subsection, a brief overview the newly introduced performance measure is presented in Section 3. In Section 4, asymptotic time-averaged convergence analysis for the mean square effective weight deviation, the mean absolute excess estimation error, and the mean square excess estimation error of the NLMF algorithm is carried out. Simulation results are presented to validate the theoretical findings in Section 5 and paper is ended with concluding remarks in Section 6.

2. SYSTEM MODEL

Consider the case of adaptive plant identification problem [6, 7]. The output y_k of the plant is given by

$$y_k = \mathbf{c}^T \mathbf{x}_k + \eta_k, \tag{1}$$

where

$$\mathbf{c} = [c_1, c_2, \ldots, c_N]^T \tag{2}$$

is the vector of the unknown system, and

$$\mathbf{x}_k = [x_{1,k}, x_{2,k}, \ldots, x_{N,k}]^T \tag{3}$$

is the input data vector at time k, η_k is the plant noise, N is the number of plant parameters, and $[\cdot]^T$ is the transpose operation. The inputs $x_{1,k}, x_{2,k}, \ldots,$ and $x_{N,k}$ may be successive samples of same signal, such as in the case of adaptive echo cancelation [8] and adaptive line enhancement [9]. They may also be the instantaneous output of N parallel sensors, such as in the case of adaptive beamforming [6]. The identification of the plant is made by an adaptive FIR filter whose weight vector $\mathbf{w}_k$, assumed of dimension N, is adapted on the basis of error e_k given by

$$e_k = y_k - \mathbf{w}_k^T \mathbf{x}_k. \tag{4}$$

The adaptation algorithm considered in this paper is NLMF algorithm [4] described by

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \frac{\mu}{||\mathbf{x}_k||^2} \mathbf{x}_k e_k^3, \tag{5}$$

where $\mu > 0$ is the algorithm step size and the norm of a vector $\mathbf{x}$ is defined as $||\mathbf{x}||^2 \equiv \mathbf{x}^T \mathbf{x}$. The error e_k can be decomposed to two terms which are the plant noise η_k and the excess estimation error ε_k defined by

$$\varepsilon_k = e_k - \eta_k. \tag{6}$$

The parameter ε_k is also termed as adaptation noise [9] since it represents the noise that appears at the filter output due to adaptation. The signal behavior of the adaptive filter is described by the evolution of the moments of ε_k with the time. The weight deviation vector is defined by

$$\mathbf{v}_k = \mathbf{w}_k - \mathbf{c}. \tag{7}$$

due to (1), (4), (6), and (7), it can be shown that:

$$\varepsilon_k = -\mathbf{v}_k^T \mathbf{x}_k. \tag{8}$$

3. THE EFFECTIVE WEIGHT DEVIATION VECTOR PERFORMANCE MEASURE

In this section, a brief overview of the recently proposed performance measure called the effective weight deviation vector [1] is presented. Let $\mathbf{u}_k$ denote a unit vector along the direction of the vector $\mathbf{x}_k$, that is,:

$$\mathbf{u}_k = \begin{cases} \frac{\mathbf{x}_k}{||\mathbf{x}_k||} & \text{if } \mathbf{x}_k \neq 0 \\ \text{an arbitrary unit vector} & \text{if } \mathbf{x}_k = 0 \end{cases}$$

Consequently, the weight deviation vector $\mathbf{v}_k$ can be decomposed to two orthogonal components; the first component $\overline{\mathbf{v}}_k$ is the projection of $\mathbf{v}_k$ along the direction of vector $\mathbf{x}_k$ while the second component $\widetilde{\mathbf{v}}_k$ is orthogonal to $\mathbf{x}_k$. The vectors $\overline{\mathbf{v}}_k$ and $\widetilde{\mathbf{v}}_k$ are given by:

$$\overline{\mathbf{v}}_k = (\mathbf{u}_k^T \mathbf{v}_k)\mathbf{u}_k, \tag{9}$$

$$\widetilde{\mathbf{v}}_k = \mathbf{v}_k - \overline{\mathbf{v}}_k. \tag{10}$$

Due to the unit vector $\mathbf{u}_k$ and (9), the vector $\overline{\mathbf{v}}_k$ satisfies

$$\overline{\mathbf{v}}_k = \frac{\mathbf{v}_k^T \mathbf{x}_k}{||\mathbf{x}_k||^2}\mathbf{x}_k. \tag{11}$$

Equations (10) , (11), and (8) imply that:

$$\overline{\mathbf{v}}_k^T \mathbf{x}_k = \mathbf{v}_k^T \mathbf{x}_k = -\varepsilon_k, \tag{12}$$

Ultimately, it can be shown that:

$$\widetilde{\mathbf{v}}_k^T \mathbf{x}_k = 0. \tag{13}$$

Thus, only the component $\overline{\mathbf{v}}_k$ contributes to the excess estimation error. The reminder, $\widetilde{\mathbf{v}}_k$, of the weight deviation vector $\mathbf{v}_k$ does not contribute to the excess estimation error. For this reason , $\overline{\mathbf{v}}_k$ is called the effective weight deviation vector [1]. From (9) and (12), it can be shown that:

$$|\varepsilon_k| = ||\overline{\mathbf{v}}_k||||\mathbf{x}_k||. \tag{14}$$

Equation (14) shows that what matters in determining the magnitude of excess estimation error is the length of vector $\overline{\mathbf{v}}_k$ rather than the length of $\mathbf{v}_k$. Thus, studying the behavior of $||\overline{\mathbf{v}}_k||$ gives a generally brighter insight on the performance of the algorithm than studying the behavior of $||\mathbf{v}_k||$. The theoretical advantage of $\overline{\mathbf{v}}_k$ in the context of the NLMF algorithm is that it can be analyzed without the need to calculate the mathematical expectations of quantities normalized by $||\mathbf{x}_k||^2$. This is due to the fact that the normalization by $||\mathbf{x}_k||^2$ is already included in the definition of $\overline{\mathbf{v}}_k$, as seen by (11). Therefore, $\overline{\mathbf{v}}_k$ enables a rigorous analysis of the NLMF algorithm under weak assumptions.

In this work, we derived an upper bound on the long term average of mean square effective weight deviation ($E\left[||\overline{\mathbf{v}}_k||^2\right]$), that is,:

$$\underset{k\to\infty}{Limsup}\,\frac{1}{k}\sum_{j=1}^{k} E\Big(||\overline{\mathbf{v}}_j||^2\Big), \tag{15}$$

where the notation "$Limsup$" is defined by

$$\underset{k\to\infty}{Limsup}\, s_k \equiv \underset{k\to\infty}{Lim}\Big(\underset{i\geq k}{sup}\; s_i\Big), \tag{16}$$

where "sup" means supremum. The smaller is the value of long term average (15), the finer is the steady state performance of algorithm and vice versa. The upper bound of the long term average (15) is used along with (14) to derive boundedness results for mean square excess estimation error ($E\left[\varepsilon_k^2\right]$) and mean absolute excess estimation error ($E\left[|\varepsilon_k|\right]$).

4. CONVERGENCE ANALYSIS OF THE NLMF ALGORITHM

In the ensuing analysis, the following assumptions are used in the convergence analysis of the NLMF Algorithm. These are quite similar to what is usually assumed in the literature [1], [2], [3], [6], [10] and which can also be justified in several practical instances:

A1 The sequences $\{\mathbf{x}_k\}$ and $\{\eta_k\}$ are mutually independent.

A2 The sequence $\{\mathbf{x}_k\}$ is stationary with finite $E\left[1/||\mathbf{x}_k||^2\right]$.

A3 The sequence $\{\eta_k\}$ is a stationary sequence of independent zero mean random variables with finite variance σ_η^2.

Assumptions $\mathbf{A1}$ and $\mathbf{A3}$ are well known independence assumptions while assumption $\mathbf{A2}$ can be well justified as in the case of the NLMS algorithm [1].

4.1. Analysis of the effective weight deviation vector

The update recursion for the weight deviation vector ($\mathbf{v}_k$) is obtained using (1), (4), (5), and (7) and can be shown to be:

$$\mathbf{v}_{k+1} = \mathbf{v}_k + \frac{\mu}{||\mathbf{x}_k||^2}\mathbf{x}_k(\eta_k - \mathbf{v}_k^T\mathbf{x}_k)^3. \tag{17}$$

As we are going to derive the upper bound for steady-state scenario, the higher order terms of $\mathbf{v}_k^T\mathbf{x}_k$ can be ignored (since excess estimation error is very small at steady-state) and thus we can use the following approximation:

$$\mathbf{v}_{k+1} \approx \mathbf{v}_k + \frac{\mu}{||\mathbf{x}_k||^2}\mathbf{x}_k(\eta_k^3 - 3\eta_k^2\mathbf{v}_k^T\mathbf{x}_k). \tag{18}$$

Now, squaring both sides of the above equation and using assumptions $\mathbf{A1}$ and $\mathbf{A3}$, it is found that:

$$\begin{aligned} E[||\mathbf{v}_{k+1}||^2] &= E[||\mathbf{v}_k||^2] - (6\mu\sigma_\eta^2 - 9\mu^2\phi_\eta^4)E\left[\frac{(\mathbf{v}_k^T\mathbf{x}_k)^2}{||\mathbf{x}_k||^2}\right] \\ &+ \mu^2\phi_\eta^6 E\left[\frac{1}{||\mathbf{x}_k||^2}\right], \end{aligned} \tag{19}$$

where ϕ_η^4 and ϕ_η^6 are the fourth and sixth order moments of the noise sequence η_k, respectively. Using the definition of weight deviation vector given in (11), we can rewrite the above equation as follows:

$$\begin{aligned} E[||\mathbf{v}_{k+1}||^2] &= E[||\mathbf{v}_k||^2] - (6\mu\sigma_\eta^2 - 9\mu^2\phi_\eta^4)E\left[||\overline{\mathbf{v}}_k||^2\right] \\ &+ \mu^2\phi_\eta^6 E\left[\frac{1}{||\mathbf{x}_k||^2}\right]. \end{aligned} \tag{20}$$

Iterating the above equation backward $(k-1)$ iterations and using the assumption $\mathbf{A2}$, equation (20) can be set up as follows:

$$E[||\mathbf{v}_{k+1}||^2] = E[||\mathbf{v}_1||^2] - (6\mu\sigma_\eta^2 - 9\mu^2\phi_\eta^4)\sum_{j=1}^{k} E\left[||\overline{\mathbf{v}}_j||^2\right] + k\mu^2\phi_\eta^6 E\left[\frac{1}{||\mathbf{x}_k||^2}\right]. \quad (21)$$

Since $E[||\mathbf{v}_{k+1}||^2]$ is a positive quantity and it converges provided that μ is in the range that insures the stability of the NLMF algorithm [2], on dividing the above equation by k, one obtains:

$$0 \leq \frac{1}{k}E[||\mathbf{v}_1||^2] - (6\mu\sigma_\eta^2 - 9\mu^2\phi_\eta^4)\frac{1}{k}\sum_{j=1}^{k} E\left[||\overline{\mathbf{v}}_j||^2\right] + \mu^2\phi_\eta^6 E\left[\frac{1}{||\mathbf{x}_k||^2}\right]. \quad (22)$$

Finally, by taking the limit of the above equation as $k \to \infty$ and using (16), it can be shown that for $0 < \mu < \frac{2\sigma_\eta^2}{3\phi_\eta^4}$ the following bound exists:

$$\underset{k\to\infty}{Limsup}\,\frac{1}{k}\sum_{j=1}^{k} E\Big(||\overline{\mathbf{v}}_j||^2\Big) \leq \frac{\mu\phi_\eta^6}{(6\sigma_\eta^2 - 9\mu\phi_\eta^4)}E\left[\frac{1}{||\mathbf{x}_k||^2}\right]. \quad (23)$$

This relation gives us an upper bound on the long term average of the mean-squared norm of $\overline{\mathbf{v}}_k$. The above result is obtained under very weak assumptions and it has all the points of strength mentioned in Section 1.

4.2. Analysis of the excess estimation error

In analyzing the convergence of excess estimation error, we have considered the following two scenarios of bounded input plant and unbounded input plant. For the analysis with bounded plant input, we need the following assumption [1]:

A4: There exists a positive number D such that $||\mathbf{x}_k|| \leq D \;\forall\; k$.

The above assumption is valid in many practical cases as naturally input data is bounded. Now, using relation (14) and assumption $\mathbf{A4}$, the bound given in (23) is modified to the following:

$$\underset{k\to\infty}{Limsup}\,\frac{1}{k}\sum_{j=1}^{k} E\left[\varepsilon_j^2\right] \leq \frac{\mu\phi_\eta^6 D^2}{(6\sigma_\eta^2 - 9\mu\phi_\eta^4)}E\left[\frac{1}{||\mathbf{x}_1||^2}\right]. \quad (24)$$

This bound shows that the long term average of the mean-squared excess estimation error can be reduced to an arbitrary small value by using very small value of the step-size provided that $0 < \mu < \frac{2}{3\sigma_\eta^2}$. Moreover, in achieving the above bound we have used very weak assumptions and it has the same advantages as in the case of bound (23). Furthermore, this bound emphasizes the fact mentioned in Section 3 that a good behavior of the effective weight deviation vector implies a good behavior of the excess estimation error.

In the case of unbounded input plant, here too, we need the following assumption to simplify the analysis [1]:

A5: The sequence $\{\mathbf{x}_k\}$ is stationary with finite $E[||\mathbf{x}_k||^2]$.

This is a weak assumption as the second order moment of input regressor generally exist. Now, using relation (14) and assumption $\mathbf{A5}$, the bound given in (23) can be set up as follows:

$$\underset{k\to\infty}{Limsup}\,\frac{1}{k}\sum_{j=1}^{k} E\left[|\varepsilon_j|\right] \leq \sqrt{\frac{\mu\phi_\eta^6 E\left[||\mathbf{x}_1||^2\right]}{(6\sigma_\eta^2 - 9\mu\phi_\eta^4)}E\left[\frac{1}{||\mathbf{x}_1||^2}\right]}. \quad (25)$$

This bound implies that the long term average of the absolute excess estimation error can be reduced to an arbitrary small value by using very small value of the step-size provided in the range that insures the stability of the NLMF algorithm. Moreover, it can be noticed that the upper bound on the right hand side of (25) will remain unchanged even if the sequence $\mathbf{x}_k$ is multiplied by a constant. This indicates that the average behavior of the steady-state excess estimation error is not sensitive to the input power of the adaptive filter.

5. SIMULATION RESULTS

In this section, the steady-state performance of the NLMF algorithm is investigated in an unknown system identification scenario with $\mathbf{c} = [1, 1, \cdots, 1]^T$. The system noise η_k is a zero mean i.i.d. sequence with variance 0.01. The plant input regressor vector $\mathbf{x}_k$ as defined in (3) with x_k being stationary zero mean unity variance correlated sequence obtained as follows:

$$x_k = \beta x_{k-1} + \sqrt{1-\beta^2}q_k, \quad (26)$$

where β is a correlation factor and q_k is a zero mean unity variance i.i.d. sequence. In our simulations, we have used $\beta = 0.95$ showing a highly correlated input sequence. The objective of our simulations is to validate the derived analytical results without restrictions on the dependence between successive regressors, the dependence between the components of regressor, the value of step-size in the range that insures the stability of the NLMF algorithm, the length of adaptive filter, and the distribution of the filter input and the noise.

Figure 1 compares the long term average of the mean-squared effective weight deviation obtained by simulation and the upper bound given in (23) for Gaussian η_k and q_k with filter length equal to 4 showing a good match between theory and simulation. It can be seen that the simulation is carried out over a wide range of step-size (0.01 to 1). The same experiment is repeated for a filter length of 32 and is shown in Fig. 2. Therefore, as depicted from these two figures that the analytical result is valid for both long and short adaptive filter. Moreover, the results show that the analytical bound is applicable for both small and large value of step-size.

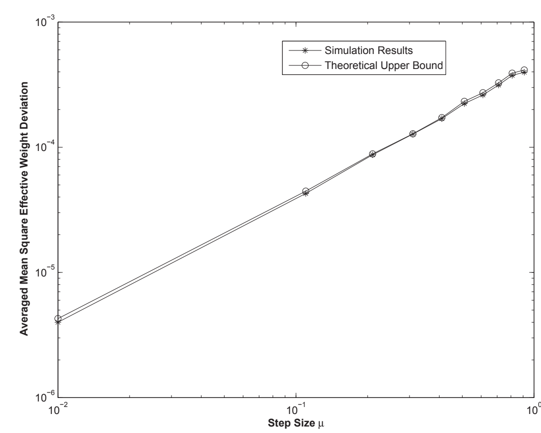

Figure 1: Analytical and simulation results with $N = 4$, Gaussian η_k and q_k.

Next, the same experiment is performed with uniform input q_k and noise η_k and the results are shown in Fig. 3 when the

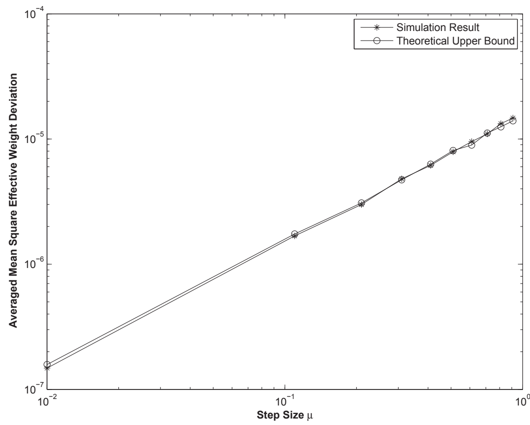

Figure 2: Analytical and simulation results with $N = 32$, Gaussian η_k and q_k.

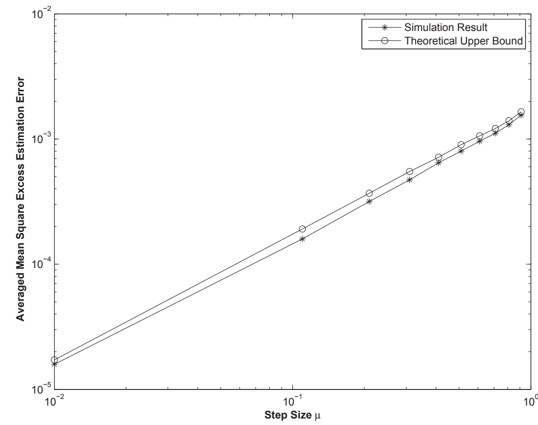

Figure 4: Analytical and simulation results with $N = 4$, uniform η_k and q_k.

filter length 32. It can be depicted from this figure that the derived analytical result is not limited to a particular distribution of input and noise sequences. The analytical upper bound on the long term average of mean-squared excess estimation error given by (24) is investigated in Fig. 4 with uniform input q_k and noise η_k and filter length equal to 4. Here too, the analytical result can well model the simulation results. Finally, similar behavior is obtained for other input and noise distributions (e.g., Laplacian) and different filter lengths, but due to space limitations these are not reported here.

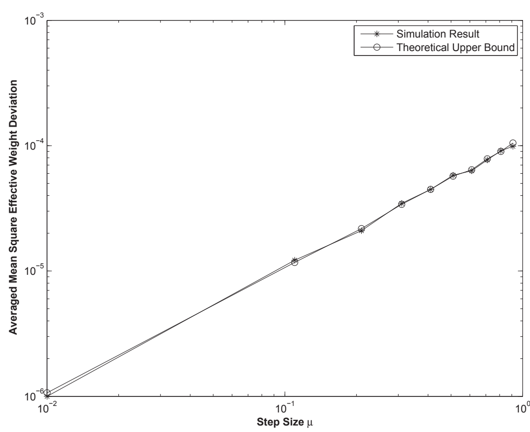

Figure 3: Analytical and simulation results with $N = 32$, uniform η_k and q_k.

6. CONCLUSION

In this work, a rigorous steady-state analysis of the NLMF algorithm is carried out using a newly proposed performance measure called the effective weight deviation vector. Asymptotic time-averaged convergence for the mean square effective weight deviation, the mean absolute excess estimation error, and the mean square excess estimation error for the NLMF algorithm is performed and consequently new explicit upper bounds for the long term average of mean-squared effective weight deviation, the mean-squared excess estimation error, and the mean absolute excess estimation error. Simulation results verified our theoretical findings.

Acknowledgments: This research work is funded by King Fahd University of Petroleum & Minerals (KFUPM) under Research Grant (FT080005).

7. REFERENCES

[1] E. Eweda, "A new performance measure of the normalized LMS adaptive filter," IEEE International Conference on Signal Processing and Communications, Dubai, UAE, November 2007.

[2] A. Zerguine, "Convergence and Steady-State Analysis of the Normalized Least Mean Fourth Algorithm," *Digital Signal Processing*, Vol. 17 (1), pp. 17-31, January 2007.

[3] E. Walach and B. Widrow, "The least mean fourth (LMF) adaptive algorithm and its family," *IEEE Transactions on Information Theory*, vol. 30, no. 2, pp. 275-283, February 1984.

[4] A. Zerguine, "Convergence Behaviour of the Normalized Least Mean Fourth Algorithm," Proceedings of 34th Annual, Asilomar Conference on Signals, Systems, and Computers, pp. 275-278, Oct. 29-Nov. 1, 2000.

[5] M.K. Chan and C.F.N. Cowan, "Using A Normalised LMF Algorithm For Channel Equalisation With Co-Channel Interference," XI Euro. Sig. Process. Conf. Eusipco2002, vol. 2, pp. 48-51, Sept. 2002.

[6] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, 3rd Edition, Upper-Saddle River, NJ, 1996.

[7] A. H. Sayed, *Fundamentals of Adaptive Filtering*, New York: Wiley-Interscience, 2003.

[8] M. M. Sondhi, "An Adaptive Echo Canceller," *Bell Syst. Tech. J., vol.*, 46, No. 3, pp 497-511, Mar. 1967.

[9] B. Widrow, and S. D. Stearns, *Adaptive Signal Processing*, Englewood Cliffs, NJ: Prentice- Hall, 1985.

[10] J. E. Mazo, "On the independence theory of equalizer convergence," *Bell Syst. Tech. J.,* No. 58, pp. 963-993, May-June 1979.