On the Convergence Analysis of a Variable Step-Size LMF Algorithm of the Quotient Form

Syed Muhammad Asad, Azzedine Zerguine, and Muhammad Moinuddin
Electrical Engineering Department
King Fahd University of Petroleum & Minerals
Dhahran, 31261, Saudi Arabia.
E-mail: {syedasad, azzedine, moinudin}@kfupm.edu.sa

Abstract—The least-mean fourth (LMF) algorithm is best known for its fast convergence and low steady-state error especially in non-Gaussian noise environments. Recently, there has been a surge of interest in the LMF algorithm with different variants being proposed. The fact that different variable step-size least-mean square algorithms have shown to outperform its fixed step-size counterpart, a variable step-size least-mean fourth algorithm of the quotient form (VSSLMFQ) is proposed here. Therefore in this work, the proposed algorithm is analysed for its performance in the steady-state and it is shown to achieve a lower steady-state error then the traditional LMF algorithm. Finally, a number of computer simulations are carried out to substantiate the theoretical findings.

Index Terms—Adaptive Filters, LMS, LMF, Variable Step-Size LMF, Quotient LMF.

I. INTRODUCTION

The LMF algorithm [1] is the class of algorithms based on the stochastic-gradient descent approach that minimises the mean-fourth error. The power of LMF algorithm lies in its faster convergence rates and lower steady-state error as compared to the least-mean square (LMS) algorithm [2]–[4] in non-Gaussian environments. Variants of the LMS algorithm, specifically that implement the time-varying step-size model, have been extensively analysed and significant performance improvements have been made in terms of rate of convergence and excess mean-square error (EMSE) [5]–[6]. No such improvements have been made in terms of rate of convergence and excess mean-square error (EMSE) [5]–[6]. No such implementations for the LMF algorithm exist that exploit the potential of the adaptive filter, especially in non-Gaussian environments. V ariants of the LMS algorithm, compared to the least-mean square (LMS) algorithm [2]–[4] have been shown to achieve a much better EMSE [7] by implementing the time-varying step-size parameter. The resulting algorithm is said to have a fixed step-size.

Taking into account the fact that the time-varying step-size parameter, in the case of LMS algorithm, has exhibited greater flexibility in the design of adaptive filters, resulting in significant performance improvement at minimal cost of complexity, the model is extended to the case of LMF algorithm. The proposed variable step-size algorithm uses the approach to adjust the step-size based on a quotient of filtered quadratic error, recently proposed in [7]. The weight vector update equation corresponding to the proposed variable step-size model, called variable step-size least-mean fourth algorithm of the quotient form (VSSLMFQ) is therefore defined as

\[ w_{n+1} = w_n + \mu_n x_n e_n^q, \]

where \( \mu_n \) is the time-varying step-size and the mechanism of update of this time-varying step-size is given as

\[ \mu_{n+1} = \alpha \mu_n + \gamma \theta_n, \]

\[ \theta_n = \frac{\sum_{j=0}^{n} a^j e_{n-j}^2}{\sum_{j=0}^{n} b^j e_{n-j}}, \]

where \( \alpha \) and \( \gamma \) are constant parameters, \( a \) and \( b \) are decaying factors used for the exponential windows in numerator and denominator, respectively, and bounded as \( 0 < a < b < 1 \).

II. ALGORITHM FORMULATION

For the general problem of a system identification, the system’s output is given by

\[ d_n = x_n^T w^o + z_n, \]

where \( x_n \) is a vector of length \( N \) (\( N \) being the length of the adaptive filter) that represents the input regressor to the adaptive filter, \( w^o \) corresponds to the system’s impulse response and \( z_n \) represents the measurement noise with a zero mean. If \( w_n \) represents the impulse response of the adaptive filter, then the minimisation in the mean-fourth sense is

\[ J_w = E[e_n^4], \]

where \( e_n \), the error between the output of the system and the output of the adaptive filter, is defined by

\[ e_n = d_n - w_n^T x_n. \]

The weight vector update equation that is based on the minimisation of (2) is the least-mean fourth (LMF) algorithm and defined as

\[ w_{n+1} = w_n + \mu_n x_n e_n^4, \]

where \( \mu \) is the step-size that controls the convergence rate of the algorithm. The step-size in (4) is a constant scalar parameter. The resulting algorithm is said to have a fixed step-size.

For the general problem of a system identification, the system’s output is given by

\[ d_n = x_n^T w^o + z_n, \]

where \( x_n \) is a vector of length \( N \) (\( N \) being the length of the adaptive filter) that represents the input regressor to the adaptive filter, \( w^o \) corresponds to the system’s impulse response and \( z_n \) represents the measurement noise with a zero mean. If \( w_n \) represents the impulse response of the adaptive filter, then the minimisation in the mean-fourth sense is

\[ J_w = E[e_n^4], \]

where \( e_n \), the error between the output of the system and the output of the adaptive filter, is defined by

\[ e_n = d_n - w_n^T x_n. \]

The weight vector update equation that is based on the minimisation of (2) is the least-mean fourth (LMF) algorithm and defined as

\[ w_{n+1} = w_n + \mu_n x_n e_n^4, \]

where \( \mu \) is the step-size that controls the convergence rate of the algorithm. The step-size in (4) is a constant scalar parameter. The resulting algorithm is said to have a fixed step-size.
The use of a quotient form in (6) and (7) has two main properties worth noting. First, the quotient ensures that the step-size decays smoothly, where the transient behaviour of the proposed VSS in stationary environment may be described by a reformulation of (7) as follows [7]:

$$\theta_n = \frac{A_n}{B_n} = \frac{a_n a_{n-1} + e_n^2}{b_n b_{n-1} + e_n^2}$$

(8)

Second, as the quotient form tends to cancel out the power of the measurement noise ($e_n^2$ being in both numerator and denominator), in the steady-state, the EMSE of the algorithm should be much smaller compared to the power of measurement noise. As a result, the steady-state behaviour of the variable step-size, being free of the steady-state error, can be completely described by the design parameters of the algorithm. Therefore, the decaying factors $a$ and $b$ could be designed beforehand for a desired steady-state mean step-size level.

III. CONVERGENCE ANALYSIS OF THE VSSLMFQ ALGORITHM

If $w^o$ is the unknown impulse response to be estimated, then we can define the weight error vector as follows:

$$v_n = w^o - w_n.$$  

(9)

Considering the well established assumptions in the literature [9], one can proceed for the convergence analysis of the proposed algorithm. Using (9), it can be shown that in the system identification scenario, $e_n$ can be set up into the following relation:

$$e_n = x_n^T v_n + z_n.$$  

(10)

Now subtracting $w^o$ from both sides of (5), the recursion of the weight error vector is given by

$$v_{n+1} = v_n - \mu_n x_n e_n^3.$$  

(11)

which can be setup as

$$v_{n+1} = v_n - \mu_n \sum_{i=0}^{3} \binom{3}{i} x_n \left[ x_n^T v_n \right] v_n x_n^{3-i}.$$  

(12)

Taking the expectation of both sides of (12), results in

$$E[v_{n+1}] = [I - 3\rho \rho E[z_n] R] E[v_n],$$  

(13)

where $\rho = E[\mu_n]$. Ultimately, a necessary condition for the convergence of (13) is

$$0 < \rho < \frac{2}{3\sigma_e^2 \lambda_{max}},$$  

(14)

where $\lambda_{max}$ is the largest eigenvalue of the correlation matrix ($R = E[x_n x_n^T]$) of the input regressor $x_n$.

IV. STEADY-STATE ANALYSIS OF THE VSSLMFQ ALGORITHM

The problem of system identification considered in this study has a time varying system response modelled as a random walk process [8]. So the desired system’s response becomes

$$d_n = x_n^T w^o_n + z_n.$$  

(15)

where $w^o_n$ is the time-varying unknown system’s response, modelled as a random walk is given as

$$w^o_{n+1} = w^o_n + q_n,$$  

(16)

where $q_n$ is some random perturbation vector independent of $x_n$ and $z_n$. The sequence $q_n$ is assumed to be i.i.d., zero mean, with covariance matrix

$$Q = E[q_n q_n^T],$$  

(17)

and $\sigma_q^2$ is the variance of $q_n$.

In this work we are going to introduce two estimation errors called $a$ priori estimation error ($e_{an}$) and $a$ posteriori estimation error ($e_{pm}$) defined, respectively, as

$$e_{an} \triangleq x_n^T w^o_{n+1} - x_n^T w_n, \; \text{and} \; e_{pm} \triangleq x_n^T w^o_{n+1} - x_n^T w_{n+1}.$$  

Now we can rewrite (5) by subtracting it from $w^o_{n+1}$ from its both sides to get

$$v_{n+1} = (w^o_{n+1} - w_n) - \mu_n x_n e_n^3,$$  

(18)

and then multiplying (18) by $x_n^T$ from the left to get

$$e_{pm} = e_{an} - \mu_n \|x_n\|^2 e_n^3.$$  

(19)

Solving (19) for $e_n^3$ and replacing it in (18) and introducing $x_n \triangleq \frac{1}{\|x_n\|^2}$ would result in

$$v_{n+1} = (w^o_{n+1} - w_n) - x_n x_n [e_{an} - e_{pm}].$$  

(20)

Rearranging the terms of both sides of the above equation and evaluating the energies would lead us to

$$\|v_{n+1}\|^2 + x_n \|e_{an}\|^2 = \|w^o_{n+1} - w_n\|^2 + x_n \|e_{pm}\|^2.$$  

(21)

This relation will be the basis of the steady-state analysis of the VSSLMFQ algorithm. Taking the expectation of both sides of (21) will result in

$$E[\|v_{n+1}\|^2] + E[x_n \|e_{an}\|^2] = E[\|w^o_{n+1} - w_n\|^2] + E[x_n \|e_{pm}\|^2].$$  

(22)

It can shown that

$$E[\|w^o_{n+1} - w_n\|^2] = E[\|v_n\|^2] + E[\|q_n\|^2]$$

$$+ E[v_n^T q_n] + E[q_n^T v_n],$$

$$= E[\|v_n\|^2] + Tr(Q),$$  

(23)

which can be set up as

$$E[\|v_{n+1}\|^2] + E[x_n \|e_{an}\|^2] = E[\|v_n\|^2]$$

$$+ E[x_n \|e_{pm}\|^2] + Tr(Q).$$  

(24)

At steady-state

$$\lim_{n \to \infty} E[\|v_{n+1}\|^2] = \lim_{n \to \infty} E[\|v_n\|^2],$$  

(25)

therefore (24) looks like the following:

$$E[x_n \|e_{an}\|^2] = E[x_n \|e_{an} - \mu_n \|x_n\|^2 e_n^3 + Tr(Q).$$  

(26)

We assume that the $a$ priori estimation error $e_{an}$ and the measurement noise $z_n$ are independent and related through:
\( e_n = z_n + e_{an} \). With this result, we develop a relation for the excess mean-square error (\( J_{ex} \))

\[
J_{ex} = \lim_{n \to \infty} E \left[ e_{an}^2 \right].
\]

(27)

After some algebraic manipulations, (26) will look like the following:

\[
\mu_\infty^2 E \left[ e_n^2 \right] + Tr(Q) = 2\mu_\infty E \left[ e_{an} e_n \right],
\]

where \( \mu_\infty = E \left[ \mu_a \right] \) and \( \mu_\infty^2 = E \left[ \mu_{an}^2 \right] \). Replacing \( e_n \) in terms of \( e_{an} \) in (28) will lead to the following relation:

\[
\mu_\infty^4 Tr(R) \left( 15J_{ex} \xi_n^3 + \xi_n^4 \right) + Tr(Q) = 0\mu_\infty \sigma_n^2 J_{ex},
\]

(29)

where \( E \left[ z_n^m \right] = \xi_n^m \) for \( m = 4, 6 \). Solving for \( J_{ex} \) we get

\[
J_{ex} = \frac{\mu_\infty^4 \xi_n^4 Tr(R) + Tr(Q)}{0\mu_\infty \sigma_n^2 - 15\mu_\infty \xi_n^2 Tr(R)}.
\]

(30)

Equation (30) provides the EMSE of the VSSLMFQ algorithm in terms of the steady-state mean and mean-square VSS which can be shown, respectively, to be given by

\[
\mu_\infty \approx \frac{\gamma \left( 1 - b \right)}{(1 - \alpha)(1 - a)},
\]

(31)

and

\[
\mu_\infty^2 \approx \frac{2\alpha \gamma^2 \left( 1 - b \right)^2}{(1 - \alpha^2)(1 - a)}.
\]

(32)

Equations (31) and (32) also imply a relationship between them and can thus be written as

\[
\mu_\infty^2 = \frac{2\alpha}{1 + \alpha} \mu_\infty. \quad (33)
\]

Substituting (31) and (32) in (30) to evaluate the steady-state EMSE for the proposed algorithm we get

\[
J_{ex} = \frac{\alpha \gamma \left( 1 - b \right) \xi_n^4 Tr(R)}{3(1 - \alpha^2)(1 - a) \sigma_n^2 - 15\alpha \gamma \left( 1 - b \right) \xi_n^2 Tr(R)}
\]

\[
+ \frac{6 \gamma \left( 1 - \alpha^2 \right)(1 - b) \sigma_n^2 - 30\alpha \gamma^2 \left( 1 - b \right)^2 \xi_n^4 Tr(R)}{4(1 - \alpha^2)(1 - a)(1 - b) \sigma_n^2 - 15\alpha \gamma \left( 1 - b \right) \xi_n^2 Tr(R)}.
\]

(34)

For a stationary environment \( \sigma_n^2 = 0 \) would lead us to the EMSE expression given as

\[
J_{ex} = \frac{\alpha \gamma \left( 1 - b \right) \xi_n^4 Tr(R)}{3(1 - \alpha^2)(1 - a) \sigma_n^2 - 15\alpha \gamma \left( 1 - b \right) \xi_n^2 Tr(R)}.
\]

(35)

Also note that by replacing (33) in (30) and considering stationary environment would lead to

\[
J_{ex} = \frac{\alpha \mu_\infty \xi_n^4 Tr(R)}{3(1 + \alpha) \sigma_n^2 - 15\alpha \mu_\infty \gamma \xi_n^2 Tr(R)}.
\]

(36)

For \( \alpha = 1 \), (36) reduces to the EMSE of the LMF algorithm. Conventionally \( \alpha < 1 \) which shows that the VSSLMFQ will exhibit a lower EMSE than the traditional LMF.

We see from (35) that the EMSE has four controlling parameters namely \( a, b, \alpha \) and \( \gamma \) where the steady-state variable step-size is totally governed by parameters \( a \) and \( b \). So these parameters can be designed beforehand to give a particular steady-state EMSE. It is already shown that the time-varying step-size is insensitive to measurement noise, then the EMSE of the proposed algorithm is predicted to be lower than the EMSE of the conventional fixed step-size LMF algorithm. The simulations will demonstrate this clearly.

V. SIMULATION RESULTS

The simulation results carried out here will substantiate the analytical findings. Particularly we are going to investigate the performance analysis of the LMF and VSSLMFQ algorithms for an unknown system identification scenario under different noise environments, e.g., Gaussian, Laplacian and Uniform.

The input signal \( z_n \) is a unit power and white Gaussian with zero mean unless stated otherwise. The unknown plant and the adaptive filter have the same number of taps, i.e., \( N = 10 \). The impulse response of the unknown plant is a Hanning window with unit norm (\( w^T w = 1 \)). Here \( \alpha = 0.9997, \gamma = 2 \times 10^{-6}, a = 0.95 \) and \( b = 0.995 \). For all the simulations \( \mu_{max} = 0.005 \) and \( \mu_{min} = 0 \). All the results produced are obtained by averaging 500 trials in Monte Carlo simulations.

Figure 1 shows the mean behaviour of the VSS for different values of the decaying parameters \( a \) and \( b \) and a signal-to-noise ratio (SNR) of 10 dB. The results demonstrate that a lower value of \( a \) and a higher value of \( b \) achieves the lowest EMSE. This also confirms the analytical findings as shown in Table I.

![Figure 1](image-url)

**Figure 1.** Mean step size behaviour of VSSLMFQ algorithm in white Gaussian noise environment for various decaying factor \( a \) and \( b \).

**Table I**

<table>
<thead>
<tr>
<th>( J_{ex} )</th>
<th>(Theoretical)</th>
<th>(Eq. (35))</th>
<th>Experimental</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 dB</td>
<td>-17.54</td>
<td>-17.62</td>
<td></td>
</tr>
<tr>
<td>10 dB</td>
<td>-31.77</td>
<td>-31.84</td>
<td></td>
</tr>
<tr>
<td>20 dB</td>
<td>-57.78</td>
<td>-57.01</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2 shows the mean behaviour of the VSS for different values of the SNR. Due to the quotient form of the VSS, the steady-state value of the VSS is independent of the noise level and corroborates the motivation in using the quotient form of VSS. Thus the algorithm exhibits extremely good resilience to the measurement noise.

Figure 3 shows the EMSE learning curves of the LMF algorithm and the proposed VSSLMFQ algorithm with Gaussian noise and \( \text{SNR}=10 \) dB. As noted in the remarks for the relation of the EMSE of the proposed algorithm, the VSSLMFQ algorithm improves the EMSE by a factor of almost 10 dB as compared to the LMF algorithm. The time-varying step-size mechanism enables the VSSLMFQ algorithm to achieve a lower EMSE.
Also the plant input signal is a correlated first order Markov process defined as $x_n = a_1 x_{n-1} + u_n$, with $a_1 = 0.7$. Figure 5 depicts this behaviour of the VSSLMFQ and LMF algorithms for this non-stationary environment. It is evident, from Fig. 5, that the proposed algorithm has a better tracking capability than does the LMF algorithm.

VI. CONCLUSION

In this study, a variable step-size LMF algorithm was proposed that achieves a better performance than the LMF algorithm in different noise environments. The algorithm is much less sensitive to measurement noise and achieves lower EMSE even in lower SNR. These enhancements have been analytically proposed and substantiated through extensive simulations.

Finally, comparing the computational complexities of the traditional LMF and VSSLMFQ algorithms, we find that the VSSLMFQ algorithm additionally requires six more multiplications, three more additions and one more division per iteration than the LMF algorithm.

REFERENCES