VARIABLE STEP-SIZE LEAST-MEAN FOURTH ALGORITHM OF THE QUOTIENT FORM

BY

SYED MUHAMMAD ASAD

A Thesis Presented to the

DEANSHIP OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the

Requirements for the Degree of

MASTER OF SCIENCE

In

Electrical Engineering

June 2009

KING FAHD UNIVERSITY OF PETROLEUM AND MINERALS

DHAHRAN 31261, SAUDI ARABIA

DEANSHIP OF GRADUATE STUDIES

This thesis, written by

SYED MUHAMMAD ASAD

under the direction of his thesis advisor and approved by his thesis committee, has been presented to and accepted by the Dean of Graduate Studies, in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE IN ELECTRICAL ENGINEERING

Thesis Committee

Dr. Zerguine, A.(Chairman)

Dr. Deriche, M. (Member)

Dr. Zidouri, A.C. (Member)

29 JUN 2009

Dr. Samir H. Adbul-Jauwad
Department Chairman

Dr. Salam A. Zummo
Dean of Graduate Studies

27/9/09
Date

Dedicated

to the Loving Memories of my

Mother

and

to the perseverance of my

Father and Wife

ACKNOWLEDGEMENTS

In the name of Allah, The Most Gracious, The Most Merciful.

All praise and glory is for Allah (Subhanahu Wa Ta'aala) who gave me the strength and determination to under take this work. My utmost gratitude towards **King Fahd University of Petroleum & Minerals** for the admittance to this prestigious institute and towards the Electrical Engineering Department for providing the learned faculty members for imparting the unmatched quality education and research.

My heartfelt gratitude and appreciation goes to my thesis adviser, **Dr. Azzedine Zerguine**, for his foresight and guidance lead me every step of the way. His attention and devotion to my research work, with valuable suggestions, was the impetus that lead to the completion of this work. It was my great pleasure to have worked with him in this friendly and inspiring environment.

I would like to extend my deepest appreciation to my thesis committee members, **Dr. M. Deriche** and **Dr. Abdelmalek C. Zidouri** for their optimism shown towards my work and their valuable thought-provoking comments that contributed towards my work.

The contribution of **Dr. Mohammad Moinuddin** towards my research cannot be understated. I am deeply humbled by his assistance and devotion at every juncture of my research.

I would like to thank my colleagues and friends at KFUPM for their support and motivation. To name a few, Kashif, Omar, Asif, Babar, Zeeshan, Ahmed, Saqib, Saad, Jamal, Umair, Zeehasham, Salman, Faraz, Usama and the list goes on and on. They made the stay at KFUPM, specially building 903, an absolute pleasure, with the last few months, the most memorable.

I would like to thank my father and my in-laws from the depth of my heart, for they have played a pivotal role in my success. This work could not have been completed without their prayers.

May the principles of Islam, enlightened in The Qur'an and Sunnah, be a guiding light for our salvation in the hereafter. Ameen

Contents

List of Figures

List of Tables

Nomenclature

$\mathbf{R}$	Autocorrelation matrix of the input regressor.
$\mathbf{v}_n$	Weight error vector.
$\mathbf{w}^o$	Unknown system weight vector.
$\mathbf{w}_n$	Weight vector representing the adaptive filter.
$\mathbf{x}_n$	The input sequence vector.
e_{an}	a-priori estimation error.
e_{pn}	a-posteriori estimation error.
J_n	Cost Function.
J_{min}	Minimum mean square error.
y_n	Output of adaptive filter.
z_n	Addtitive Noise.
d_n	Desired response.
e_n	Output estimation error.

$g[e_n]$ General error function.

N Length of the adaptive filter.

∇ Gradient Operator.

μ Step-size used for adapting filter weights.

μ_n Time-varying step-size.

σ_z^2 Noise variance.

σ_x^2 Variance of input regressor.

ψ_n^k k^{th} moment of the noise.

$E[]$ Expectation Operator.

$tr()$ Trace operator.

lim Limit.

LMS Least Mean Square.

LMF Least Mean Fourth.

VSSLMFQ Variable Step-Size Least Mean Fourth of the Quotient Form.

EMSE Excess Mean Square Error.

MSE Mean Square Error.

SNR Signal-to-Noise Ratio.

BPSK Binary Phase Shift Keying.

MSD Mean Square Deviation.

Abstract

Name:	SYED MUHAMMAD ASAD
Title:	VARIABLE STEP-SIZE LEAST-MEAN FOURTH ALGORITHM OF THE QUOTIENT FORM
Degree:	MASTER OF SCIENCE
Major:	ELECTRICAL ENGINEERING
Date of Degree:	JUNE 2009

The aim of this thesis is to exploit the research in adaptive algorithms based on gradient descent approach, namely the least-mean square (LMS) algorithm, particularly in the model of time-varying step-size parameter and extend it to the least-mean fourth (LMF) algorithm. In his work, we have developed a variable step-size least-mean fourth algorithm with the aim in mind to achieve performance enhancement than the traditional least-mean fourth algorithm in terms of the excess mean-square error and retain its inherent dominance over the least-mean square algorithm in non-Gaussian noise environments. The motivation behind it being that the time-varying step-size model of the LMS algorithm have contributed a lot in the performance enhancement of LMS algorithm. In contrast, the LMF algorithm that gives better convergence rate in non-Gaussian noise environment has not been researched in its full spectrum.

In particular, this research has assessed the algorithm's performance through a comprehensive analysis in terms of its steady-state performance, tracking and transient behaviour through analytical means. The concept of energy conservation is employed to carry out all the analysis. Finally, a number of simulation results are carried out to corroborate the theoretical findings which as expected has yielded improved performance.

ملخص

الاسم :سيد محمد اسد

العنوان:الخوارزمية الخطوة المتغيرة القوة الرابعة للمتوسطة من نموذج القسمة

الدرجة:ماجستير علوم

التخصص: الهندسة الكهربائية

تاريخ التخرج:يونيو ٢٠٠٩

الهدف من هذه الرسالة هو استغلال البحوث في خوارزميات التكيف على أساس نهج التدرج النسب ، وهما قوة المربعة للمتوسط ، وخاصة في متباينة التوقيت خطوة متغيرة وتوسيع نطاقه ليشمل الخوارزمية القوة الرابعة للمتوسط. فى هذا العمل ، وقد وضعنا خطوة متغيرة القوة الرابعة للمتوسط بهدف في الاعتبار لتحقيق تحسين الأداء بالمقارنة مع القوة الرابعة للمتوسط التقليدية ، من حيث الخطا المربعة للمتوسط والاحتفاظ بهيمنتها الأصيل على القوة الرابعة للمتوسط في بيئات غير ضوضاء تمويه. الدافع وراء ذلك هو أن متباينة التوقيت خطوة من طراز القوة المربعة للمتوسط قد أسهمت كثيرا في تحسين الأداء من قوة المربعة للمتوسط . في المقابل ،الخوارزمية القوة الرابعة للمتوسط التي تعطي أفضل معدل التقارب في غير غاوسي ضوضاء البيئة لم يتم بحثها في الطيف الكامل

بوجه الخاص ، البحوث وتقييمها الخوارزمية في الأداء من خلال تحليل شامل من حيث الأداء المطرد الدولة ، وتتبع سلوك عابر من خلال الوسائل التحليلية. تم استخدام مبدأ الحفاظ على الطاقة للحصول على التحليل. وأخيرا ، فإن عددا من نتائج المحاكاة تتم لثبيت هذه النتائج النظرية التي كما هو متوقع لم تسفر عن تحسين الأداء.

Chapter 1

Introduction

With the advancement in digital electronics and communication techniques, there has been extensive research in various domains of signal processing that mainly deal in achieving close-to-optimal performance for various systems like in system identification and inverse modelling. Recently digital wireless communication has been a central field of active research due to the promise it holds.

Signal processing plays a key role in all digital communication systems where certain system behaviours are dealt with signal processing techniques. One of these is *adaptive signal processing* where a certain system, often called a filter, tries to learn the behaviour of another without any outside intervention, hence adaptive. This adaptation mechanism of adaptive filters make them an ideal choice for systems that are blind to certain statistics of a system.

Due to their ease of implementation as a digital filter, their field of application can be as diverse as digital communication systems, industrial equipment, medical instrumentation, geophysical sciences and military where they can be utilised for equalisation [1], system identification [2], noise cancellation [3] and linear prediction [4]. Hence there is extensive research in the performance enhancement of adaptive filters.

With this in mind, we will discuss some key principles and applications of adaptive filters and later on carry out a thorough analysis of their performance characteristics which is basically the broader aim of this study.

1.1 Adaptive Filters : Basics

The need for adaptive filters arise in the case where certain statistics of the system are unknown. The argument to use adaptive filters can be motivated through the fact that the Wiener filter gives the optimum steady-state solution in a stationary environment for the case where all the statistics of the system are known [2]. In particular, it means that the statistics of the input signal like its autocorrelation matrix and the cross correlation between the desired and the input signals are known. Hence the goal of an adaptive filter is to "find and track" the optimum filter corresponding to the same signal operating environment with complete knowledge of the required statistics. In this context, optimum filters provide both guidance for the development of adaptive algorithms and a yardstick to evaluate the theoretical performance of adaptive filters.

The process of "find and track" is usually achieved in a recursive manner. The adaptive filter is initialised with some conditions independent of the stochastic environment it is operating in. The mechanism of adaptation tracks some statistical criteria against which the filter learns and adapts. As time proceeds, the recursive mechanism achieves the solution that is close to the optimum in some statistical sense. This adaptability, with no outside intervention, makes adaptive filters the most elegant solution that can be employed in such applications.

The performance of adaptive filters is evaluated using the concepts of stability, speed of adaptation, quality of adaptation, and tracking capabilities. This is a formidable task due to the fact that adaptive filters are inherently nonlinear devices whose performance depend on sequences that are themselves nonlinear and time-varying. Further complications arise due to the fact that these signals are stochastic in nature.

The parameters of performance evaluation of an adaptive filter is its *steady-state performance*, *tracking performance* and *transient performance*. Steady state performance analysis is carried to study the amount of error that remains (i.e. residual error) when the adaptation reaches its final state after sufficient time has passed. Tracking performance analyses the filter's capability in tracking changes in the system. The transient analysis studies the time-evolution of the adaptive filter. The main aim is to study its convergence rate and stability.

1.2 Adaptive Filters : Applications

As described earlier, the usefulness of adaptive filter manifests in the situation where practical applications cannot successfully be implemented by using fixed digital filters. This is due to the fact that either we do not have sufficient information to design a digital filter with fixed coefficients or the design criteria change during the normal operation of the filter. Some applications where adaptive filters are used are briefed here to emphasise their diversity and necessity.

1.2.1 System Identification

In the class of applications dealing with system identification, an adaptive filter is used to provide a linear model that represents the best fit (in some sense) to an unknown plant. The depiction of this scenario is given in Figure 1.1. The plant and the adaptive filter are driven by the same input $\{x_n\}$. The plant output supplies the desired response, d_n, for the adaptive filter. If the plant is dynamic in nature, the model will be time-varying. This application is abundantly used in wireless communication for channel estimation [5].

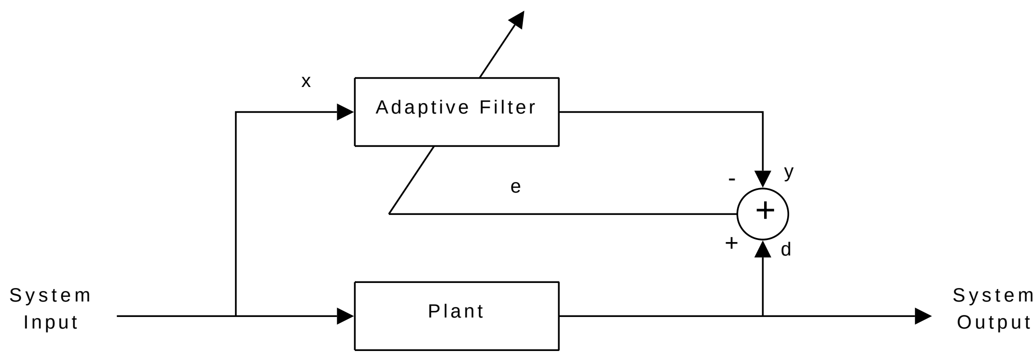

Figure 1.1: Adaptive system identification

1.2.2 Adaptive Equalisation

Every waveform propagating through a channel suffers a certain amount of time dispersion because the frequency response of the channel does not have constant magnitude and linear phase. As a result, the tails of adjacent pulses interfere with the measurement of the current pulse (inter-symbol interference) and can lead to an incorrect decision. Equalisation corresponds to adjusting the relative phases of different frequencies to achieve a constant group delay.

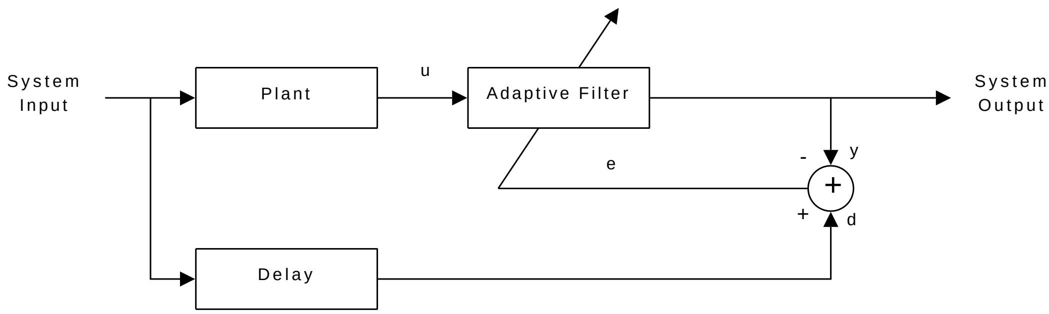

Figure 1.2: Adaptive equaliser

Since the channel can be modelled as a linear system, we can compensate for its distortion by using a linear equaliser as depicted in Figure 1.2. The goal of the equaliser is to restore the received pulse, as closely as possible, to its original shape. The equaliser learns the channel and inverses its effects. Equalisers are therefore categorised as *inverse modelling filters* because they actually behave as inverse of the channel. The combined response of the channel and the adaptive filters should thus be a delta function. Adaptive equalisers are an essential part of any digital communication system and have thus been analysed extensively in literature [1] .

1.2.3 Linear Prediction

The function of the adaptive filter shown in Figure 1.3 provides the best prediction of the present value of a random signal. Thus the current value of the signal serves the purpose of a desired response, d_n, for the adaptive filter. Past values of the signal, $\{u_n\}$, are input to the adaptive filter. Depending on the application of interest, the adaptive filter output or the estimation (prediction) error may serve as the system output. In the first case, the system operates as a predictor; in the latter case, it operates as a prediction-error filter.

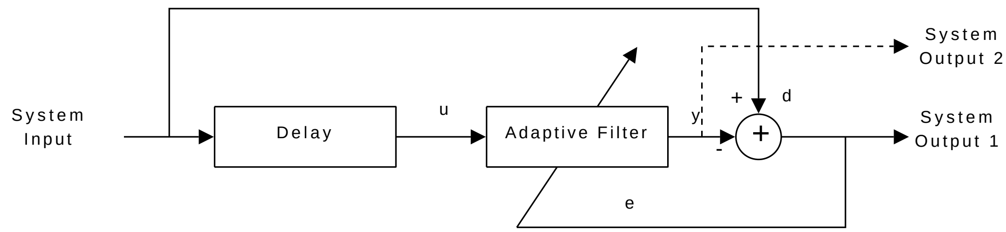

Figure 1.3: Linear prediction

1.2.4 Interference Cancellation

In this application, the adaptive filter is used to cancel unknown interference contained in a primary signal which also contains the information-bearing signal component, with the cancellation being optimised in some sense. Figure 1.4 depicts the filter operation. The primary signal serves as the desired response for the adaptive filter. A reference signal is employed as the input to the adaptive filter. The reference signal is derived in relation to the primary signal in such a way that the information-bearing signal component is weak or essentially undesirable. These type of filters are used in noise cancellation applications [3].

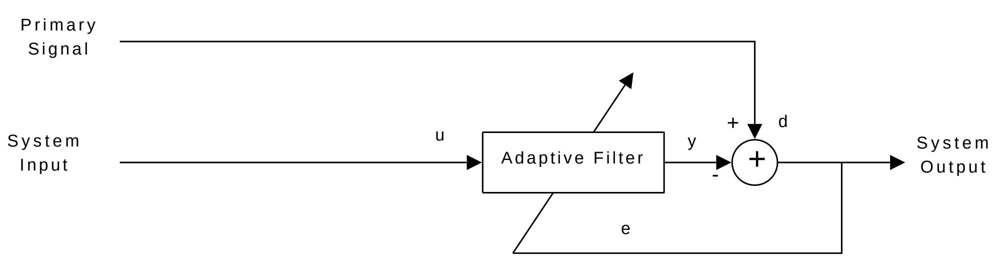

Figure 1.4: Interference cancellation

1.3 Adaptive Filters : Features

Every adaptive filter involves one or more input signals and a desired response signal that may or may not be accessible to the adaptive filter. Every adaptive filter consists of three modules as depicted in Figure 1.5:

- *Filtering structure.* This module is formed by the digital filter design techniques that are available. It can either be a finite impulse response (FIR) digital filter or an infinite impulse response (IIR) digital filter. The filtering structure is linear if the output is obtained as a linear combination of the input measurements; otherwise it is said to be nonlinear. FIR filter is an example of linear structure, where it can be implemented with a direct or lattice structure. The structure is fixed by the designer, and its parameters are adjusted by the adaptive algorithm. Adaptive filters use FIR filters as a favourable design implementation where the filter is commonly termed a transversal filter or a tapped delay line.

- *Criterion of performance.* This module processes the desired response (when available) and output of the adaptive filter by the criterion of performance to assess its quality with respect to the requirements of the particular application. The choice of the criterion is a balanced compromise between what is acceptable to the user of the application and what is mathematically tractable; that is, it can be manipulated to derive an adaptive algorithm.

- *Adaptation algorithm.* The adaptive algorithm uses the value of the criterion of performance, or some function of it, and the measurements of the input and desired response (when available) to decide how to modify the parameters of the filter to improve its performance. The complexity and the characteristics of the adaptive algorithm are functions of the filtering structure and the criterion of performance.

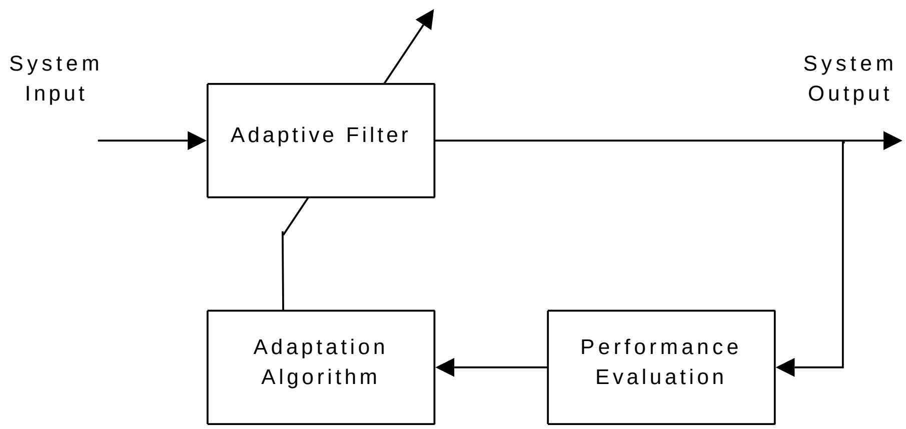

Figure 1.5: Principle of adaptive filters

The design of any adaptive filter is highly dependent on the *a priori* information about the signals and the dynamics of the application. Unreliable *a priori* information and/or incorrect assumptions about the signal can lead to serious performance degradations or even unsuccessful adaptive filter applications. The conversion of the performance assessment to a successful parameter adjustment strategy, that is, the design of an adaptive algorithm, is the most difficult step in the design and application of adaptive filters.

1.4 Adaptive Filters : Algorithms

Generally adaptive filter algorithms are implemented in an iterative or recursive manner. With each iteration, they improve the performance of the filter according to the criteria of performance. Specifically if we let $\mathbf{w}_n$ to be a vector of length N whose elements are the time-varying coefficients of an FIR filter at time index n, then the adaptation mechanism of the vector can be given by

$$\mathbf{w}_{n+1} = \mathbf{w}_n - \mu \left[\nabla_{\mathbf{w}} J_{\mathbf{w}} \right], \tag{1.1}$$

where μ is the learning parameter called the step-size that controls the amount of correction applied to the weight vector in each iteration and ensures the convergence of the adaptive algorithm. The term $\nabla_{\mathbf{w}} J_{\mathbf{w}}$ is the gradient of the cost function with respect to the weight vector of the algorithm. Equation (1.1) represents a special approach in adaptive algorithms called the *steepest descent* [2]. In this approach, the solution to the problem, i.e., the filter that best suits the application, is sought in the direction where the cost function is *minimised*. Hence the term $\nabla_{\mathbf{w}} J_{\mathbf{w}}$ is given as

$$\nabla_{\mathbf{w}} J_{\mathbf{w}} = \frac{\partial J}{\partial \mathbf{w}} = \frac{\partial}{\partial \mathbf{w}} E \left[g \left(e_n \right) \right], \tag{1.2}$$

where $g\left(e_n\right)$ is a function of the error and consequently

$$J_{\mathbf{w}} = E \left[g \left(e_n \right) \right] \tag{1.3}$$

is the cost function. According to the steepest descent approach, the gradient of the cost function can be written as

$$\nabla_{\mathbf{w}} J_{\mathbf{w}} = \mathbf{R}_x \mathbf{w}_n - \mathbf{R}_{dx}, \tag{1.4}$$

where $\mathbf{x}_n$ is the tap input regressor taken from the input sequence $\{x_n\}$, $\mathbf{R}_x = E\left[\mathbf{x}_n \mathbf{x}_n^T\right]$ is the autocorrelation of the input regressor and $\mathbf{R}_{dx} = E\left[\mathbf{x}_n d_n\right]$ is the cross-correlation of the input regressor and the desired response. The expectation that appears in (1.2) requires the knowledge of the statistics of the tap input vector that, in practise, is not available. Hence the gradient has to be estimated by dropping the expectation and taking the sample value of the tap input. This introduces a randomness or stochastic behaviour to all such adaptive algorithms and are hence termed *stochastic gradient algorithms*. The approximation made for (1.2) can be further explored where [5]

$$\begin{aligned} -\hat{\nabla}_{\mathbf{w}} J_{\mathbf{w}} &= \hat{\mathbf{R}}_{dx} - \hat{\mathbf{R}}_x \mathbf{w}_n, \\ &= \mathbf{x}_n \left[d_n - \mathbf{x}_n^T \mathbf{w}_n\right], \\ -\hat{\nabla}_{\mathbf{w}} J_{\mathbf{w}} &= \mathbf{x}_n e_n, \end{aligned} \tag{1.5}$$

or in general it can be setup as

$$-\hat{\nabla}_{\mathbf{w}} J_{\mathbf{w}} = \mathbf{x}_n g\left(e_n\right), \tag{1.6}$$

where $g\left(e_n\right)$ is a function of the error,$\hat{\nabla}_{\mathbf{w}} J_{\mathbf{w}}$ can be viewed as the gradient applied to the instantaneous error function, and

$$\hat{\mathbf{R}}_{dx} = d_n \mathbf{x}_n, \qquad \hat{\mathbf{R}}_x = \mathbf{x}_n \mathbf{x}_n^T$$

are the instantaneous cross-correlation between the desired and tap input vector and the autocorrelation of the tap input vector, respectively. Substituting (1.6) in (1.1) will give us the generic adaptation equation for stochastic gradient algorithms as

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu \mathbf{x}_n g\left(e_n\right), \tag{1.7}$$

where μ is the step-size of the algorithm. Selecting the appropriate error function, $g\left(e_n\right)$, will yield different algorithms that behave totally differently and require extensive analysis for their proper behaviour to be characterised.

1.4.1 Least-Mean Square (LMS) Algorithm

By far the most popular stochastic gradient algorithm is the Least-Mean Square (LMS) algorithm developed by Widrow and Hopf. The name signifies its cost function as the minimisation of the mean squared error. For the filters of type (1.7), the error function for LMS is given by $g\left(e_n\right) = e_n$, so the LMS weight update recursion will be

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu \mathbf{x}_n e_n. \tag{1.8}$$

Due to its simple yet elegant mathematics, this algorithm has been extensively used in various applications. The most motivating factor of its usage is its simplicity of implementation. The LMS algorithm achieves good performance characteristics when the condition of operation are rightly suited for it. This performance is seen where the noise environment is Gaussian. Even then there are limitations to its performance that has seen its fair share of research. The motivation behind improving the LMS algorithm is its slow convergence rate and higher steady-state error. The slow convergence of the LMS algorithm is due the fact that it is based on the minimisation of the mean-squared error and is only dependent upon the second order moment of the noise; this results in identical convergence rates in various noise environments. Also the steady-state error of the LMS algorithm is dependent on the second order moment of noise which, as we shall see later, compared to algorithms based on higher order moments of error, results in higher steady-state error. Moreover, the steady-state error and rate of convergence are highly dependent upon the step-size of the algorithm. In fact the steady-state error is inversely proportional to the step-size parameter. This highlights a compromise that has to be made in every design of fixed step-size algorithms.

1.4.2 Variable Step-Size (VSS) LMS Algorithms

A popular approach for the improvement of fixed step-size algorithms is to implement the step-size that is time-varying in the steepest descent manner. The weight update equation only changes in the sense that now the learning parameter μ becomes time-varying. So (1.8) becomes

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu_n \mathbf{x}_n e_n, \tag{1.9}$$

where μ_n is the time-varying step-size. The variable step-size LMS (VSSLMS) provides improved performance while maintaining the inherent simplicity and robustness of the conventional fixed step-size algorithm.

Generally, the approach with VSSLMS is to devise step-size rules that give large steps when the estimated error is large and small steps when the error is small, thereby avoiding the trade-off between convergence rate and steady-state error for fixed step-size LMS. This mechanism is mostly data-dependent and is implemented in an iterative manner (steepest descent). For this reason, there has been a lot of research in the VSSLMS field [6–15] and their stability [16–18]. The two main aims of these researches are to either improve the convergence rate of the algorithm for a given *Excess Mean-Square Error* (EMSE) or to improve the EMSE for a given convergence rate. In terms of the convergence rate, the desired performance enhancement is the speed with which an algorithm attains the steady-state. With high data rates, it is a desirable feature of any algorithm to achieve the steady-state in the minimum number of iterations. Conversely, in terms of the excess mean-square error, the desired performance is to attain lower steady-states at a certain convergence rate. Lower excess mean-square error means the filter attains a steady-state error that is closer to the minimum achievable steady-state error, that is, the *minimum mean-square error*. With the advancement of modern digital communication systems, both performance aspects are of great importance.

Some of the most popular implementations of the VSSLMS algorithms are discussed below.

1. Kwong and Johnston (1992) [12] used the mean-square error value to update the step-size in the sense that higher value resulted in a steep step while smaller error resulted in smaller steps, allowing the adaptive filter to track changes in

the system as well as produce a small steady-state error. The update equation of the step-size is given as

$$\mu_{n+1} = \alpha\mu_n + \gamma e_n^2, \tag{1.10}$$

where $0 < \alpha < 1$ and $\gamma > 0$.

2. Mathews and Xie (1993) [15] used a gradient descent algorithm designed to reduce the squared estimation error during each iteration to update the step-size as

$$\mu_{n+1} = \mu_n + \rho e_n e_{n-1} \mathbf{x}_{n-1}^T \mathbf{x}_n, \tag{1.11}$$

where ρ is a small positive constant that controls the adaptive behaviour of the step-size sequence.

3. Aboulnasr and Mayyas (1997) [6] recognised the sensitivity of the previous VSSLMS implementations to the measurement noise and proposed using an estimate of the autocorrelation between successive estimation errors as follows:

$$\mu_{n+1} = \alpha\mu_n + \gamma p_n^2, \tag{1.12}$$

$$p_n = \beta p_{n-1} + (1-\beta)\, e_n e_{n-1}. \tag{1.13}$$

The use of p_n in the update has two advantages. First, the error autocorrelation is generally a good measure of the proximity to the optimum. It ensures the adaptation in the direction of minimisation. Second, it rejects the effect of the uncorrelated noise sequence on the step-size update as consecutive noise samples would be uncorrelated. In the early stages of adaptation, the error autocorrelation estimate p_n^2 is large, resulting in a large step-size. As we approach the optimum, the error autocorrelation approaches zero, resulting in a smaller step-size. This provides the fast convergence due to large initial μ_n while ensuring low misadjustment near optimum due to the small final μ_n even in the presence of measurement noise.

4. Zhao et al (2008) [13] have recently proposed a new VSSLMS algorithm where they have improved the previous algorithms in terms of their performance against measurement noise [6, 8, 9, 12]. The applications where signal-to-noise ratios are low, are bound to have a critical impact on the performance of adaptive filters. The idea here is to adjust the variable step-size using a *quotient form of filtered versions of the quadratic error*. The filtered estimates of the error are based on exponential windows, applying different decaying factors for the estimations in the numerator and denominator. This scheme is given as

$$\mu_{n+1} = \alpha\mu_n + \gamma\theta_n, \tag{1.14}$$

$$\theta_n = \frac{\sum_{i=0}^{n} a^i e_{n-i}^2}{\sum_{j=0}^{n} b^j e_{n-j}^2}. \tag{1.15}$$

where a and b are decaying factors used for the exponential windows of (1.15) and bounded as $0 < a < b < 1$.

1.4.3 Least-Mean Fourth (LMF) Algorithm

Adaptive filters based on higher order statistics are known to perform better than the mean square estimation employed in the Least-Mean Square algorithm in some scenarios. An algorithm based on the minimisation of the fourth moment of the output estimation error, namely the Least-Mean Fourth (LMF) algorithm [19] is one such example. For the filters of type (1.7), the error function for LMF is given by $g(e_n) = e_n^3$. Then by substituting in (1.7), we get the LMF recursion as

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu \mathbf{x}_n e_n^3. \tag{1.16}$$

The LMF algorithm exhibits lower steady-state error relative to LMS algorithm as shown in non-Gaussian environments [19]. This is due to the fact that the excess mean-square error of the LMS algorithm is dependent only on the second order moment of the noise. The second order moment, or variance of the noise evaluates to be the same for all the noise environments. In contrast, the excess mean-square error of the LMF algorithm depends on higher order moments of the noise that results in lower steady-state error as compared to the LMS algorithm. Recognising this feature of the LMF algorithm, it was also known that the convergence behaviour of the LMS and LMF algorithms are susceptible to the *condition number*, i.e., on the ratio of the maximum to the minimum eigenvalues of the input signal autocorrelation matrix ($\mathbf{R}$) [19].

To overcome this dependency on the condition number, the normalised LMF (NLMF) algorithm was introduced [20–23] where the LMF recursion was modified to include an inverse norm of the input regressor. This resulted in a faster convergence of the algorithm and more stability.

However, this higher-order statistics of the error in the LMF algorithm requires a much smaller step-size to ensure stable adaptation [21]. The cubic error in the LMF recursion can cause severe initial instability. In order to ensure the stable adaptation of the algorithm, a normalisation of the step-size was proposed [24–26]. The step-size was normalised by weighted signal and error powers through a mixing parameter.

It is evident from above that research in LMF algorithm and its variants is largely focused on their steady-state performance. Recently though, quite a number of studies have emerged particularly dealing with convergence and stability analysis [27–32]. Particularly in [31], it has been shown that the LMF algorithm is never mean-square stable for input regressors that are unbounded. There is always a non-zero probability of divergence, even for Gaussian distribution. Nevertheless, results based on standard mean-square stability analysis are useful for practical design purposes. The probability of divergence which was shown to be a function of the step-size, tends to rise abruptly only when it moves past a given threshold. Before that, the probability of divergence tends to be sufficiently small to grant the practical applicability of the LMF algorithm. In case the error becomes unbounded, a re-initialisation scheme can be included in practical scenarios. The estimation of the region of the quick rise of the probability has been tackled in [30]. Moreover, signal amplitudes are necessarily limited in practical applications, which contributes to reduce the probability of divergence for step sizes smaller than the threshold.

1.5 Thesis Objectives and Organisation

This thesis work is carried out to develop a **Variable Step-Size Least-Mean Fourth algorithm with a Quotient form**. The proposed algorithm exploits the fact that time-varying step-size LMS algorithms have outperformed the ones with fixed step-size. There seems to be a lack of the time-varying step-size model and analysis of LMF algorithm in the strict sense of the word. This study aims to tackle this untapped design flexibility of the variable step-size to be utilised with LMF algorithm.

The main objectives of this study can be enumerated as:

1. To come up with the weight update equation of the VSSLMFQ algorithm from the traditional LMF algorithm.

2. To analyse the performance of the algorithm when it has converged, that is, the steady-state performance of the proposed algorithm in terms of its excess mean-square error.

3. To establish the necessary conditions for the convergence of the algorithm in terms of its stability in the mean and mean-square sense.

4. To analyse the performance of the algorithm to time-variations or nonstationarity in the system in terms of the tracking excess mean-square error.

5. To analyse the performance of the algorithm in terms of the time-evolution of the weight vector and derive its mean-square deviation and mean-square error.

6. To carry out experiments to corroborate the analytical analysis.

The thesis has been organised to fulfil all these objectives. In Chapter 2, the proposed algorithm is presented and its recursion is described. Moreover, this Chapter also describes the mathematical layout of the thesis by describing the concept of energy conservation [5], which underlies the whole framework of the analysis carried out. In Chapter 3, the steady-state analysis is carried out and expressions for the excess mean-square error have been derived. Additionally, the advantage of using the proposed algorithm is highlighted in detail. While Chapter 4 discusses the tracking analysis of the proposed algorithm and obtains expression for excess mean-square error for the tracking scenario, Chapter 5 details the transient behaviour through a rigorous analysis and derives expressions for the mean-square error and misadjustment.

Chapter 6 reports the different simulation results carried out in different noise environments to corroborate the theoretical findings for the proposed algorithm. Finally Chapter 7 summarises the thesis contributions and motivations for future research.

Chapter 2

A Variable Step-Size Least-Mean Fourth Algorithm of the Quotient Form

2.1 Introduction

An overview of the various adaptive algorithms and their variants in the previous chapter has given us an insight on how adaptive filters are classified and characterised. The two main performance factors that are highlighted are its convergence rate (speed) and residual error (accuracy). It has also been established that these are reciprocating factors, meaning both cannot be optimised at the same time. So any analysis has to have a compromise between the two.

One main performance degrading factor that all adaptive filter suffer from is the measurement noise statistics. In fact, according to the Wiener filter theory, measurement noise is the minimum achievable error floor [2]. Due to the stochastic nature of adaptive filters, this minimum error floor is not achievable. This becomes a motivating factor for research to find ways in which this difference can be minimised.

Most of the adaptive algorithms discussed in the previous chapter recognised this and proposed ways to minimise the excess error. Particularly the approaches devised in [6, 8, 13] achieved better results. Zhao et al [13] have proposed a time-varying step-size based on the quotient of the filtered quadratic error. This affects the performance of the algorithm profoundly in terms of the excess mean-square error.

2.2 System Model

Before delving into the details of the proposed algorithm, it would be appropriate to layout the system for which the performance of the algorithm would be tested. It has been a common practise to employ a problem of system identification for the performance analysis of adaptive filters. The system is depicted in Figure 1.1.

Based on the system identification problem, two types of channel models are going to be used in the analysis; namely time-invariant and time-varying. The time-invariant channel model is given as

$$d_n = \sum_{i=0}^{N-1} x_i w^o_{n-i} + z_n = \mathbf{x}_n^T \mathbf{w}^o + z_n, \quad n = 0, 1, 2, 3... \tag{2.1}$$

whereas the time-varying channel model is given as

$$d_n = \sum_{i=0}^{N-1} x_i w^o_{n,n-i} + z_n = \mathbf{x}_n^T \mathbf{w}_n^o + z_n, \quad n = 0, 1, 2, 3... \tag{2.2}$$

where in both the equations $\{x_n\}$ is the input process with zero mean and variance σ_x^2, $\{z_n\}$ is the stationary noise process with zero mean and variance σ_z^2 and $[\mathbf{w}^o, \mathbf{w}_n^o]$ are the time-invariant and time-varying channel impulse responses, respectively, with taps N. The adaptive filter response corresponding to this problem of estimation can be given as

$$y_n = \sum_{i=0}^{N-1} x_i w_{n,n-i} = \mathbf{x}_n^T \mathbf{w}_n, \quad n = 0, 1, 2, 3... \tag{2.3}$$

where $\{x_n\}$ is again the input process with zero mean and variance σ_x^2 and $\mathbf{w}_n$ is an $N \times 1$ weight vector of the adaptive filter updated through an algorithm. With this model at hand, we can express the problem of minimisation of the mean fourth error as

$$\min_w E\left[|e_n|^4\right], \tag{2.4}$$

or

$$\nabla_{\mathbf{w}} J_{\mathbf{w}} = \frac{\partial J}{\partial \mathbf{w}} = \frac{\partial}{\partial \mathbf{w}} E\left[\left|d_n - \mathbf{x}_n^T \mathbf{w}_n\right|^4\right], \tag{2.5}$$

where $\frac{\partial}{\partial \mathbf{w}}$ is the differentiation with respect to a vector and the gradient is thus given as

$$\frac{\partial J}{\partial \mathbf{w}} = \frac{1}{2} \begin{bmatrix} \frac{\partial J}{\partial x_0} + j\frac{\partial J}{\partial y_0} \\ \frac{\partial J}{\partial x_1} + j\frac{\partial J}{\partial y_1} \\ \vdots \\ \frac{\partial J}{\partial x_{N-1}} + j\frac{\partial J}{\partial y_{N-1}} \end{bmatrix}, \tag{2.6}$$

where generally each element of the vector $\mathbf{w}$ is given as $w_k = x_k + jy_k$. Evaluating the gradient for (2.4) gives us

$$\nabla_{\mathbf{w}} J_{\mathbf{w}} = -2E\left[\mathbf{x}_n e_n^3\right]. \tag{2.7}$$

Replacing the expectation by instantaneous values, we can express (2.7) as [5]

$$\hat{\nabla}_{\mathbf{w}} J_{\mathbf{w}} \approx -\mathbf{x}_n e_n^3, \tag{2.8}$$

where we have used the convention of incorporating the constant multiplier with the step-size of the algorithm thus eliminating it in the recursion. This leads us to the LMF recursion given in (1.16).

2.3 Proposed Algorithm

The proposed Variable Step-Size Least-Mean Fourth algorithm of the Quotient form (VSSLMFQ) is based on the variable step-size update equation recently proposed by Zhao et al (2008) given in [13]. For the filters employing variable step-size, (1.16) will become

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu_n \mathbf{x}_n e_n^3, \tag{2.9}$$

where now μ_n represents the time-varying step-size parameter governed by some update mechanism.

The variable step-size update mechanism employed in the proposed algorithm is designed to achieve lower steady-state EMSE or misadjustment, faster tracking property in nonstationary environment and enhanced performance for the applications like noise cancellation as compared to the traditional LMF algorithm. It will be shown through analysis that the parameter adjustment of the proposed variable step-size is independent of the measurement noise, making it a more robust design. It also adds to the design flexibility of the algorithm by introducing four design parameters as compared to one in the traditional LMF algorithm. The overall performance enhancement achieved at the expense of implementation complexity makes the algorithm a viable alternative to the traditional LMF algorithm.

The variable step-size algorithm adjusts the time-varying step-size based on a quotient of filtered quadratic error. The proposed scheme is listed as follows:

$$\mu_{n+1} = \alpha \mu_n + \gamma \theta_n, \tag{2.10}$$

$$\theta_n = \frac{\sum_{i=0}^{n} a^i e_{n-i}^2}{\sum_{j=0}^{n} b^j e_{n-j}^2}. \tag{2.11}$$

where a and b are decaying factors used for the exponential windows of (2.11) and bounded as $0 < a < b < 1$. The constant parameters α and γ are adjusted as in [13].

The use of a quotient form in (2.10) and (2.11) serves two objectives. First, the quotient is expected for a smoothing decrease of the step-size, where the transient behaviour of the proposed variable step-size in stationary environment may be described by a reformulation of (2.11) as follows:

$$\theta_n = \frac{A_n}{B_n} = \frac{aA_{n-1} + e_n^2}{bB_{n-1} + e_n^2}, \tag{2.12}$$

$$\approx \frac{a}{b}\theta_{n-1} + \frac{e_n^2}{bB_{n-1}}. \tag{2.13}$$

Note that in derivation of (2.13), we have neglected the value of e_n^2 in the denominator, since compared to e_n^2, the error cumulant bB_{n-1} becomes much larger during adaptation because the decaying factor b is very close to one. If we assume that the initial step-size is set to μ_{max} for fast convergence, then from (2.13) we see that the ratio of e_n^2 and bB_{n-1} decreases with the decrease in the error power. This ensures that the step-size also decreases with the decrease in θ_n. Second, in steady-state, the EMSE of the algorithm should be much smaller compared to the power of measurement noise. This implies that the measurement noise dominates the numerator and denominator of (2.11). In statistic sense, the power of measurement noise in the numerator and denominator could be cancelled out, leaving the steady-state mean step-size determined only by the constant parameters. Therefore, the decaying factors a and b could be designed beforehand for a desired steady-state mean step-size level.

In the next section we are going to layout the mathematical basis on which we are going to base our analysis. It is called the fundamental energy conservation method [5]. It neatly unifies the performance analysis of adaptive filters into a single concept of energy conservation.

2.4 Energy Conservation Method

The energy conservation relation described here will hold for any general data $\{d_n, x_n\}$. The generic form of the filter update equation is given as

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu_n \mathbf{x}_n g\left(e_n\right), \tag{2.14}$$

where $\mathbf{x}_n$ is input regressor, μ_n is the time-varying step-size and $g\left(e_n\right)$ is a function of the error. Defining the weight-error vector as

$$\mathbf{v}_n \triangleq \mathbf{w}^o - \mathbf{w}_n, \tag{2.15}$$

subtracting $\mathbf{w}^o$ from both sides of (1.7), we can rewrite the recursion (2.14) as

$$\mathbf{v}_{n+1} = \mathbf{v}_n - \mu_n \mathbf{x}_n g\left(e_n\right). \tag{2.16}$$

We now define two new error quantities namely *a priori estimation error* , e_{an}, and *a posteriori estimation error* , e_{pn}, obtained by pre-multiplying the above equation by $\mathbf{x}_n^T$ to get

$$e_{pn} = e_{an} - \mu_n \left\|\mathbf{x}_n\right\|^2 g\left(e_n\right), \tag{2.17}$$

where

$$e_{an} \triangleq \mathbf{x}_n^T \mathbf{v}_n, \qquad e_{pn} \triangleq \mathbf{x}_n^T \mathbf{v}_{n+1}. \tag{2.18}$$

One important result that is worth noting here is the relationship between the *a priori* estimation error and the EMSE. By definition,

$$\text{EMSE} = \lim_{n\to\infty} E\left[e_n^2\right] - J_{min}, \tag{2.19}$$

where $J_{min} = \sigma_z^2$. Using

$$\begin{aligned} e_n &= d_n - \mathbf{x}_n^T \mathbf{w}_n, \\ &= z_n + \mathbf{x}_n^T \left(\mathbf{w}^o - \mathbf{w}_n\right). \end{aligned}$$

Then by using (2.18) we get

$$e_n = z_n + e_{an}. \tag{2.20}$$

Equation (2.20) will later be used to find an alternate expression for the EMSE.

The equations (2.17) and (2.18) provide an alternative description of adaptive filters in terms of error quantities e_{pn}, e_{an}, $\mathbf{v}_{n+1}$, $\mathbf{v}_n$ and $g\left(e_n\right)$. This description is useful since we are often interested in questions related to the behaviour of these errors, such as:

1. **Steady-state behaviour**, which relates to determining the steady-state values of the weight error variance, $E\left[\|\mathbf{v}_n\|^2\right]$, *a priori* estimation error variance, $E\left[|e_{an}|^2\right]$and mean-square error, $E\left[|e_n|^2\right]$.

2. **Stability**, which relates to determining the range of values of the step-size μ over which the variances $E\left[\|\mathbf{v}_n\|^2\right]$ and $E\left[|e_{an}|^2\right]$ remain bounded.

3. **Transient behaviour**, which relates to studying the time-evolution of the weight error variance, $E\left[\|\mathbf{v}_n\|^2\right]$, mean weight error vector, $E\left[\mathbf{v}_n\right]$ and mean-square error, $E\left[|e_n|^2\right]$.

For the above analysis to be carried out, we will be relying on the energy conservation relation that relates the squared norms of the error quantities.

In order to derive the energy relation, we will first remove the error nonlinearity function $g(e_n)$ from (2.16) by solving (2.17) for $g(e_n)$ and substituting in (2.16), meaning that the resulting energy relation will hold irrespective of the error nonlinearity. Consider the following two case:

1. $\underline{\mathbf{x}_n = 0.}$ This is a degenerate situation. In this case, it is obvious from (2.16) and (2.17) that $\mathbf{v}_{n+1} = \mathbf{v}_n$ and $e_{an} = e_{pn}$ so that $\|\mathbf{v}_{n+1}\|^2 = \|\mathbf{v}_n\|^2$ and $|e_{an}|^2 = |e_{pn}|^2$.

2. $\underline{\mathbf{x}_n \neq 0.}$ In this case, we solve for $g(e_n)$ from (2.17) to get

$$g\left[e_n\right] = \frac{1}{\mu_n \|\mathbf{x}_n\|^2}\left[e_{an} - e_{pn}\right]. \tag{2.21}$$

   Substituting the above equation in (2.16) we get

$$\mathbf{v}_{n+1} = \mathbf{v}_n - \frac{\mathbf{x}_n}{\|\mathbf{x}_n\|^2}\left[e_{an} - e_{pn}\right]. \tag{2.22}$$

The above equation gives the relation in terms of the four error quantities $\mathbf{v}_{n+1}$, $\mathbf{v}_n$, e_{an}

and e_{pn}. It is also worth noting that even the step-size is cancelled out. Rearranging (2.22) we get

$$\mathbf{v}_{n+1} + \frac{\mathbf{x}_n}{\|\mathbf{x}_n\|^2} e_{an} = \mathbf{v}_n + \frac{\mathbf{x}_n}{\|\mathbf{x}_n\|^2} e_{pn}. \tag{2.23}$$

By evaluating the energies on both sides, the following energy equality holds:

$$\|\mathbf{v}_{n+1}\|^2 + \frac{1}{\|\mathbf{x}_n\|^2} |e_{an}|^2 = \|\mathbf{v}_n\|^2 + \frac{1}{\|\mathbf{x}_n\|^2} |e_{pn}|^2 . \tag{2.24}$$

Considering both zero and nonzero regressors by defining $\overline{\mathbf{x}}_n$ such that

$$\overline{\mathbf{x}}_n \triangleq \begin{cases} 1/\|\mathbf{x}_n\|^2 & \text{if } \mathbf{x}_n \neq 0 \\ 0 & \text{otherwise} \end{cases} . \tag{2.25}$$

Using $\overline{\mathbf{x}}_n$ we arrive at the general energy relation as

$$\|\mathbf{v}_{n+1}\|^2 + \overline{\mathbf{x}}_n |e_{an}|^2 = \|\mathbf{v}_n\|^2 + \overline{\mathbf{x}}_n |e_{pn}|^2 . \tag{2.26}$$

This result will the used extensively in the analysis of adaptive filters [5].

Chapter 3

Steady-State Analysis of the VSSLMFQ Algorithm

3.1 Introduction

In this chapter, the steady-state analysis of the proposed algorithm in terms of its Excess Mean Square Error (EMSE) is carried out. For this purpose we are going to make use of the energy conservation relationship derived in the previous chapter. The analysis will be general, meaning it will require the least amount of assumptions. This benefit follows from the use of energy conservation method as will be discussed later. Before we proceed, a few assumptions do need to be made that are going to be used in the analysis of the algorithm [5]:

A1 The input process $\{x_n\}$ is zero-mean with autocorrelation matrix $\mathbf{R} = E\left[\mathbf{x}_n\mathbf{x}_n^T\right]$.

A2 The noise process $\{z_n\}$ is a zero mean and (i.i.d.) with variance σ_z^2 and independent of the input process $\{x_n\}$.

A3 The step-size μ_n is statistically independent of the weight update vector $\mathbf{w}_n$.

A4 The noise process $\{z_n\}$ is independent of the *a priori* estimation error e_{an} and weight-error vector $\{\mathbf{v}_n \ \ for\ j < n\}$.

The assumptions taken here are quite general in nature. Assumption **A1** is a consequence of using the energy conservation method. Hence the analysis will be valid for any general data that meets the condition in **A1**. **A2** is required on the basis that to keep the analysis tractable and is very common in literature. It is known as the *independence assumption.* we require the noise process to have such statistics. **A3** is required because of the stochastic nature of the step-size μ_n. Although **A3** is not true in general because μ_n and $\mathbf{w}_n$ are functions of $\mathbf{x}_n$ and z_n for any given n and therefore are dependent but in the steady-state, since the step-size varies slowly around its mean value, justifies **A3**. **A4** follows from **A2**.

3.2 Mean Square Analysis of the VSSLMFQ

The proposed algorithm's weight update equation given in Chapter 2 can be written as

$$\mathbf{w}_{n+1} \quad = \quad \mathbf{w}_n + \mu_n \mathbf{x}_n e_n^3, \tag{3.1}$$

where μ_n varies as given in (2.10) and (2.11). With reference to the energy conservation method discussed in Chapter 2 where the energy conservation relation was derived in (2.26) for a generic algorithm, the error function used was given as $g(e_n)$. The selection of this error function lead to various algorithms whose performance analysis is totally different for any given $g(e_n)$. In the present case of VSSLMFQ, this error function is given as $g(e_n) = e_n^3$.

Taking into account the definition of steady-state operation of an adaptive filter as

$$E[\mathbf{v}_{n+1}] = E[\mathbf{v}_n] = s \quad as\ n \to \infty, \tag{3.2}$$

$$E\left[\|\mathbf{v}_{n+1}\|^2\right] = E\left[\|\mathbf{v}_n\|^2\right] = c \quad as\ n \to \infty, \tag{3.3}$$

where $\mathbf{v}_n$ is the weight-error vector defined in (2.15), we take the expectation of the energy conservation relation (2.26) to arrive at

$$E\left[\|\mathbf{v}_{n+1}\|^2\right] + E\left[\overline{\mathbf{x}}_n |e_{an}|^2\right] = E\left[\|\mathbf{v}_n\|^2\right] + E\left[\overline{\mathbf{x}}_n |e_{pn}|^2\right]. \tag{3.4}$$

Evaluating (3.4) for $n \to \infty$(steady-state) then using (3.2) and (3.3) we can write

$$E\left[\overline{\mathbf{x}}_n e_{an}^2\right] = E\left[\overline{\mathbf{x}}_n \left(e_{an} - \mu_n \|\mathbf{x}_n\|^2 e_n^3\right)^2\right], \tag{3.5}$$

where we have used (2.17) with $g(e_n) = e_n^3$. Expanding the right hand side of (3.5) and simplifying gives us

$$\overline{\mathbf{x}}_n \left(e_{an} - \mu_n \left\|\mathbf{x}_n\right\|^2 e_n^3\right)^2 = \overline{\mathbf{x}}_n \left(e_{an}^2 + \mu_n^2 \left\|\mathbf{x}_n\right\|^4 e_n^6 - 2e_{an}\mu_n \left\|\mathbf{x}_n\right\|^2 e_n^3\right). \tag{3.6}$$

Substituting in (3.5) we get

$$E\left[\overline{\mathbf{x}}_n e_{an}^2\right] = E\left[\overline{\mathbf{x}}_n e_{an}^2\right] + E\left[\mu_n^2 \left\|\mathbf{x}_n\right\|^2 e_n^6\right] - 2E\left[e_{an}\mu_n e_n^3\right]. \tag{3.7}$$

Using **A4** we can write (3.7) as

$$E\left[\overline{\mathbf{x}}_n e_{an}^2\right] = E\left[\overline{\mathbf{x}}_n e_{an}^2\right] + E\left[\mu_n^2\right] E\left[\left\|\mathbf{x}_n\right\|^2 e_n^6\right] - 2E\left[\mu_n\right] E\left[e_{an} e_n^3\right],$$

$$\overline{\mu_n^2} E\left[\left\|\mathbf{x}_n\right\|^2 e_n^6\right] = 2\overline{\mu_n} E\left[e_{an} e_n^3\right], \tag{3.8}$$

where $\overline{\mu_n} = E[\mu_n]$ and $\overline{\mu_n^2} = E[\mu_n^2]$. The relation in (3.8) is known as the *variance relation*. We see that the relation involves e_{an}. It can be shown that the *a priori* estimation error can be used to evaluate the EMSE of an adaptive filter. Recall the definition of EMSE from (2.19), the result of (2.20) and invoking **A4**, we get

$$E\left[e_n^2\right] = \sigma_z^2 + E\left[e_{an}^2\right], \tag{3.9}$$

and therefore an alternative expression for the EMSE can be given by substituting the above in (2.19) to get

$$\text{EMSE} = \lim_{n\to\infty} E\left[e_{an}^2\right] = J_{ex}. \tag{3.10}$$

Using the identities

$$e_n^6 = e_{an}^6 + 6e_{an}^5 z_n + 6e_{an} z_n^5 + 15e_{an}^4 z_n^2 + 15e_{an}^2 z_n^4 + 20e_{an}^3 z_n^3 + z_n^6, \tag{3.11}$$

and

$$e_n^3 = e_{an}^3 + z_n^3 + 3e_{an}^2 z_n + 3e_{an} z_n^2, \tag{3.12}$$

in (3.8) and ignoring third and higher-order terms in e_{an} we obtain

$$\overline{\mu_n^2} E\left[\|\mathbf{x}_n\|^2 \left(15e_{an}^2 z_n^4 + z_n^6\right)\right] = 6\overline{\mu_n} E\left[e_{an}^2 z_n^2\right]. \tag{3.13}$$

Let $E\left[z_n^m\right] = \psi_z^m$ and using assumptions **A2** and **A4**, the above expression can be written as

$$6\overline{\mu_n}\sigma_z^2 E\left[e_{an}^2\right] \quad = \quad \overline{\mu_n^2}\psi_z^6 tr(\mathbf{R}) + 15\overline{\mu_n^2}\psi_z^4 E\left[\|\mathbf{x}_n\|^2 e_{an}^2\right]. \tag{3.14}$$

The relation given in (3.14) is the variance relation for the proposed VSSLMFQ algorithm and can now be used to evaluate the steady- state EMSE subject to a few assumptions. These assumptions are mainly introduced to solve the expectation $E\left[\|\mathbf{x}_n\|^2 e_{an}^2\right]$ involved in (3.14). We shall examine two scenarios namely *separation principle* and *Gaussian regressor* [5]. But before proceeding to it, we would like to examine the steady-state behaviour of the time-varying step-size of the proposed algorithm. The results will be used in conjunction with (3.14) to arrive at the final expression of the EMSE.

3.2.1 Steady-State Mean and Mean-Square Behaviour of Step-Size

The step-size update equation for the proposed algorithm is given as

$$\mu_{n+1} = \alpha \mu_n + \gamma \theta_n, \tag{3.15}$$

where θ_n, the quotient of the filtered quadratic error is given as

$$\theta_n = \frac{\sum_{i=0}^{n} a^i e_{n-i}^2}{\sum_{j=0}^{n} b^j e_{n-j}^2}, \tag{3.16}$$

which can be written in a recursive manner as

$$\begin{aligned} \theta_n &= \frac{A_n}{B_n} = \frac{aA_{n-1} + e_n^2}{bB_{n-1} + e_n^2}, \\ &\approx \frac{a}{b}\theta_{n-1} + \frac{e_n^2}{bB_{n-1}}\,. \end{aligned} \tag{3.17}$$

For the steady-state mean behaviour of step-size we take the expectation of both sides of (3.15) and assuming that $n \to \infty$ we have

$$\lim_{n\to\infty} E\left[\mu_{n+1}\right] = \alpha \lim_{n\to\infty} E\left[\mu_n\right] + \gamma \lim_{n\to\infty} E\left[\theta_n\right]. \tag{3.18}$$

At the steady-state, we can expect to have $\lim_{n\to\infty} E[\mu_{n+1}] \approx \lim_{n\to\infty} E[\mu_n]$ then (3.18) can be approximated as

$$\overline{\mu_\infty} \approx \frac{\gamma E[\theta_\infty]}{1-\alpha}. \tag{3.19}$$

The term $E[\theta_\infty]$ in (3.19) is given by

$$E[\theta_\infty] = E\left[\frac{A_\infty}{B_\infty}\right]. \tag{3.20}$$

From (3.17) we know that A_n and B_n are of recursive form. Assuming $n \to \infty$, they can be written as

$$\begin{aligned}
\lim_{n\to\infty} A_n &= a \lim_{n\to\infty} A_{n-1} + \lim_{n\to\infty} e_n^2, \\
A_\infty &= aA_\infty + e_\infty^2, \\
E[A_\infty] &= aE[A_\infty] + E\left[e_\infty^2\right].
\end{aligned}$$

Therefore

$$E[A_\infty] \approx \frac{1}{1-a} E\left[e_\infty^2\right]. \tag{3.21}$$

Similarly

$$E[B_\infty] \approx \frac{1}{1-b} E\left[e_\infty^2\right]. \tag{3.22}$$

To proceed with (3.20) we make the assumption that

$$E\left[\frac{A_\infty}{B_\infty}\right] \approx \frac{E\left[A_\infty\right]}{E\left[B_\infty\right]}. \tag{3.23}$$

The assumption is taken to be valid because A_n and B_n both will vary slowly at the steady-state and can be assumed to be independent. A description into the conditions where such simplification can be done is detailed in Appendix 7.3. By using the results in (3.21) and (3.22) we can approximate (3.20) as

$$E\left[\theta_\infty\right] \approx \frac{1-b}{1-a}. \tag{3.24}$$

Substituting (3.24) in (3.19) we get

$$\overline{\mu_\infty} \approx \frac{\gamma\left(1-b\right)}{\left(1-\alpha\right)\left(1-a\right)}. \tag{3.25}$$

Remark One important point worth noting is that (3.25) is free from steady-state MSE, i.e., $J_{min} + J_{ex}$. This may lead us to predict that the proposed algorithm will exhibit less sensitivity to the measurement noise as compared to other LMF algorithms. Also the mean behaviour is proportional to the ratio $\left(\frac{1-b}{1-a}\right)$. This prompts us to predict that steady-state MSE will be dependent of parameters a and b. Particularly, larger a and smaller b will result in poor MSE performance but for smaller a and larger b we will have lower MSE. This behaviour will be seen through simulations as well.

For the mean squared behaviour of step-size, squaring (3.15), taking the expectation and assuming $n \to \infty$ gives us

$$\lim_{n\to\infty} E\left[\mu_{n+1}^2\right] = \lim_{n\to\infty} \left\{\alpha^2 E\left[\mu_n^2\right] + 2\alpha\gamma E\left[\mu_n\right] E\left[\theta_n\right] + \gamma^2 E\left[\theta_n^2\right]\right\}. \tag{3.26}$$

As γ is usually very small therefore the term involving γ^2 can be neglected. The approximation leads to

$$\overline{\mu_\infty^2} \approx \frac{2\alpha\gamma E\left[\mu_\infty\right] E\left[\theta_\infty\right]}{1-\alpha^2}. \tag{3.27}$$

Substituting results from (3.24) and (3.25) in (3.27) we get

$$\overline{\mu_\infty^2} \approx \frac{2\alpha\gamma^2 \left(1-b\right)^2}{\left(1-\alpha^2\right)\left(1-\alpha\right)\left(1-a\right)^2}. \tag{3.28}$$

Equations (3.25) and (3.28) also imply a relationship between them and can thus be written as:

$$\overline{\mu_\infty^2} = \frac{2\alpha}{1+\alpha}\left(\overline{\mu_\infty}\right)^2. \tag{3.29}$$

The relation in (3.29) will be used to gain more insight into the EMSE of the algorithm as we shall see later. Now that we have evaluated both the steady-state mean and mean square step-size, we can replace them in (3.30) and (3.34) to evaluate the steady-state EMSE for the proposed algorithm.

3.2.2 Steady-State EMSE of the VSSLMFQ Algorithm

As described at the end of Section 3.2, the EMSE will be evaluated for two scenarios mainly due to the expectation involved in (3.14). The two scenarios are *separation principle* and *Gaussian regressor* [5].

Separation Principle

The separation principle is used when the step-size does not converge to a small enough value but still guarantees the convergence of the algorithm. In this section we will make the assumption that

A5 At steady-state, $\|\mathbf{x}_n\|^2$ is independent of e_{an}.

The assumption is valid under special cases for example when $\|\mathbf{x}_n\|^2$ have constant Euclidean norms. So (3.14) can now be evaluated assuming $n \to \infty$ as

$$\begin{aligned} 6\overline{\mu_\infty}\sigma_z^2 J_{ex} &= \overline{\mu_\infty^2}\psi_z^6 tr(\mathbf{R}) + 15\overline{\mu_\infty^2}\psi_z^4 tr(\mathbf{R}) J_{ex}, \\ J_{ex} &= \frac{\overline{\mu_\infty^2}\psi_z^6 tr(\mathbf{R})}{6\overline{\mu_\infty}\sigma_z^2 - 15\overline{\mu_\infty^2}\psi_z^4 tr(\mathbf{R})}. \end{aligned} \tag{3.30}$$

Substituting the relations derived for $\overline{\mu}_\infty$ and $\overline{\mu_\infty^2}$ in (3.25) and (3.28) in (3.30) we get

$$\begin{aligned} J_{ex} &= \frac{\frac{2\alpha\gamma^2(1-b)^2}{(1-\alpha^2)(1-\alpha)(1-a)^2}\psi_z^6 tr(\mathbf{R})}{6\frac{\gamma(1-b)}{(1-\alpha)(1-a)}\sigma_z^2 - 15\frac{2\alpha\gamma^2(1-b)^2}{(1-\alpha^2)(1-\alpha)(1-a)^2}\psi_z^4 tr(\mathbf{R})}, \\ &= \frac{2\alpha\gamma^2(1-b)^2\psi_z^6 tr(\mathbf{R})}{6\gamma(1-\alpha^2)(1-a)(1-b)\sigma_z^2 - 30\alpha\gamma^2(1-b)^2\psi_z^4 tr(\mathbf{R})}, \end{aligned}$$

$$J_{ex} = \frac{\alpha\gamma\left(1-b\right)\psi_z^6 tr(\mathbf{R})}{3\left(1-\alpha^2\right)\left(1-a\right)\sigma_z^2 - 15\alpha\gamma\left(1-b\right)\psi_z^4 tr(\mathbf{R})}. \tag{3.31}$$

White Gaussian Regressor

In this case we are going to assume that the regressor input $\mathbf{x}_n$ has a circular Gaussian distribution. Using white Gaussian regressor means that the individual entries of the regressor $\mathbf{x}_n = \{x_{N-1}, x_{N-2}, \ldots, x_2, x_1\}^T$ are uncorrelated with each other. This results in the autocorrelation matrix of the input regressor to be of the form

$$\mathbf{R} = \sigma_x^2 \mathbf{I}. \tag{3.32}$$

Also we shall assume in this section that

A6 At steady-state, $\mathbf{v}_n$, is independent of $\mathbf{x}_n$.

With these assumptions at our disposal, we will be able to find out the expectation $E\left[\|\mathbf{x}_n\|^2 e_{an}^2\right]$ explicitly. It can be shown that the expectation $E\left[\|\mathbf{x}_n\|^2 e_{an}^2\right]$ can be written as a scaled multiple of $E\left[e_{an}^2\right]$, i.e.,

$$E\left[\|\mathbf{x}_n\|^2 e_{an}^2\right] = \left(N+2\right)\sigma_x^2 E\left[e_{an}^2\right]. \tag{3.33}$$

The derivation of (3.33) is highlighted in Appendix 7.3. Using the above relation in (3.14), we get assuming $n \to \infty$

$$\begin{aligned} 6\overline{\mu_\infty}\sigma_z^2 J_{ex} &= \overline{\mu_\infty^2}\psi_z^6\sigma_x^2 N + 15\overline{\mu_\infty^2}\psi_z^4 (N+2)\sigma_x^2 J_{ex}, \\ 6\overline{\mu_\infty}\sigma_z^2 J_{ex} &- 15\overline{\mu_\infty^2}\psi_z^4 (N+2)\sigma_x^2 J_{ex} = \overline{\mu_\infty^2}\psi_z^6\sigma_x^2 N, \\ J_{ex} &= \frac{\overline{\mu_\infty^2} N\psi_z^6\sigma_x^2}{6\overline{\mu_\infty}\sigma_z^2 - 15\overline{\mu_\infty^2}(N+2)\psi_z^4\sigma_x^2}. \end{aligned} \tag{3.34}$$

Substituting the relations derived for $\overline{\mu_\infty}$ and $\overline{\mu_\infty^2}$ in (3.25) and (3.28) in (3.34), we get

$$\begin{aligned} J_{ex} &= \frac{\frac{2\alpha\gamma^2(1-b)^2}{(1-\alpha^2)(1-\alpha)(1-a)^2}\psi_z^6\sigma_x^2 N}{6\frac{\gamma(1-b)}{(1-\alpha)(1-a)}\sigma_z^2 - 15\frac{2\alpha\gamma^2(1-b)^2}{(1-\alpha^2)(1-\alpha)(1-a)^2}\psi_z^4 (N+1)\sigma_x^2}, \\ &= \frac{2\alpha\gamma^2 (1-b)^2 \psi_z^6\sigma_x^2 N}{6\gamma(1-\alpha^2)(1-a)(1-b)\sigma_z^2 - 30\alpha\gamma^2(1-b)^2\psi_z^4(N+1)\sigma_x^2}, \end{aligned}$$

$$J_{ex} = \frac{\alpha\gamma(1-b)\psi_z^6\sigma_x^2 N}{3(1-\alpha^2)(1-a)\sigma_z^2 - 15\alpha\gamma(1-b)\psi_z^4\sigma_x^2 N}. \tag{3.35}$$

Remark Comparing the parameter of design of FSSLMF and VSSLMFQ algorithms, their is only one behaviour controlling parameter in the FSSLMF algorithm, i.e., the step-size μ. The VSSLMFQ algorithm, on the other hand, offers greater flexibility in design by introducing four design parameters namely α, γ, a and b that control the convergence and steady-state behaviour of the step-size. Furthermore referring to (3.29), we can rewrite equations (3.31) and (3.35) as

$$J_{ex} = \frac{\alpha\overline{\mu_\infty}\psi_z^6 tr(\mathbf{R})}{3(1+\alpha)\sigma_z^2 - 15\alpha\overline{\mu_\infty}\psi_z^4 tr(\mathbf{R})}, \tag{3.36}$$

and

$$J_{ex} = \frac{\alpha\overline{\mu_\infty}N\psi_z^6\sigma_x^2}{3(1+\alpha)\sigma_z^2 - 15\alpha\overline{\mu_\infty}(N+2)\psi_z^4\sigma_x^2}, \tag{3.37}$$

respectively. For $\alpha = 1$, the equations reduce to the EMSE of LMF algorithm. Conventionally $\alpha < 1$ that shows that the VSSLMFQ algorithm will exhibit a lower EMSE than the traditional LMF algorithm.

3.3 Computation of LMF and VSSLMFQ Algorithms

Comparing the computational complexities of the FSSLMF and VSSLMFQ algorithms, we find that the VSSLMFQ algorithm additionally requires six multiplications, three additions and a division per iteration by using the recursive forms of A_n and B_n in (3.17). Note that there is a division computation in the VSSLMFQ algorithm, which usually requires some more processing time compared to additions or multiplications. Table (3.1) lists the computational cost of various stochastic-gradient algorithms. Considering the processing speeds that now exist and the success of the variable step-size in adaptive algorithms, it can safely be said that this algorithm can be utilised for run-time applications like channel estimation/tracking and channel equalisation.

Table 3.1: Computational complexity per iteration for different algorithms for real-valued data in terms of the real multiplications, real additions and real divisions.

Algorithm	$\times$	$+$	$/$
LMS	2N+1	2N	
VSSLMS (Zhao et al)	2N+7	2N+3	1
LMF	2N+3	2N	
VSSLMFQ	2N+9	2N+3	1

Chapter 4

Tracking Analysis of the VSSLMFQ Algorithm

4.1 Introduction

Tracking analysis plays a vital role in the performance analysis of adaptive filters. The main aim is to analyse the algorithm's ability to track changes in the channel. Most practical channel are modelled as random processes so their behaviour is time-varying. In this Chapter, the proposed algorithm is analysed for its tracking performance. Two time-varying channel models are considered namely Random Walk and Rayleigh fading models. Again we are going to use the energy conservation method for the analysis as given in [5]. First we are going to describe the Random-Walk model.

4.2 Random-Walk Model

The first order Random Walk model for a time-varying channel having weight $\mathbf{w}_n^o$ can be given as

$$\mathbf{w}_{n+1}^o = \mathbf{w}_n^o + \mathbf{q}_n, \tag{4.1}$$

where

A7 $\mathbf{q}_n$ is assumed to be i.i.d., zero mean, with covariance matrix $E\left[\mathbf{q}_n\mathbf{q}_n^T\right] = \mathbf{Q}$ and independent of $\{\mathbf{x}_n\}$ and $\{z_n\}$ for all n.

It can be seen from (4.1) that $E\left[\mathbf{w}_{n+1}^o\right] = E\left[\mathbf{w}_n^o\right]$, so that $E\left[\mathbf{w}_n^o\right]$ will have a constant mean given by $E\left[\mathbf{w}_n^o\right] = w^o$. Although the model of (4.1) is valid but analysis shows the covariance matrix of $\mathbf{w}_n^o$ grows unbounded with time. This can be seen by

$$\mathbf{w}_{n+1}^o - w^o = \mathbf{w}_n^o - w^o + \mathbf{q}_n, \tag{4.2}$$

then

$$E\left[\left(\mathbf{w}_{n+1}^o - w^o\right)\left(\mathbf{w}_{n+1}^o - w^o\right)^T\right] = E\left[\left(\mathbf{w}_n^o - w^o\right)\left(\mathbf{w}_n^o - w^o\right)^T\right] + \mathbf{Q}. \tag{4.3}$$

This shows that a non-negative definite matrix is added to the covariance matrix of $\mathbf{w}_n^o$ at each iteration. This naturally leads to an unbounded matrix with time. This is more appropriately dealt with a model that slightly changes the original one to give

$$\mathbf{w}_{n+1}^o - w^o = \eta\left(\mathbf{w}_n^o - w^o\right) + \mathbf{q}_n, \tag{4.4}$$

where the scalar factor $\eta < 1$. This leads to a more stable model which would converge to a steady-state given by

$$\lim_{n \to \infty} \left(\mathbf{w}_{n+1}^o - w^o\right) \left(\mathbf{w}_{n+1}^o - w^o\right)^T = \frac{Q}{1 - |\eta|^2}. \tag{4.5}$$

The analysis of this model being quite involved, the model of (4.1) is used in the tracking analysis with the value of η being very close to one [5]. This simplifies the analysis to a great deal.

4.2.1 Tracking Analysis : VSSLMFQ Algorithm for Random-Walk Model

The proposed algorithm's weight update equation given in Chapter 2 can be written as

$$\mathbf{w}_{n+1} \quad = \quad \mathbf{w}_n + \mu_n \mathbf{x}_n e_n^3, \tag{4.6}$$

where again μ_n varies as given in (2.10) and (2.11). With the weight-error vector being defined as $\mathbf{v}_{n+1} = \mathbf{w}_{n+1}^o - \mathbf{w}_{n+1}$, there is a slight modification in the expression of the *a priori* estimation error due to the time-varying nature of the channel and is defined as

$$e_{an} \triangleq \mathbf{x}_n^T \left(\mathbf{w}_{n+1}^o - \mathbf{w}_n\right), \qquad e_{pn} \triangleq \mathbf{x}_n^T \left(\mathbf{w}_{n+1}^o - \mathbf{w}_{n+1}\right). \tag{4.7}$$

We can write (4.6) in terms of the weight-error vector $\mathbf{v}_n$ to get

$$\mathbf{w}^o_{n+1} - \mathbf{w}_{n+1} = \left(\mathbf{w}^o_{n+1} - \mathbf{w}_n\right) - \mu_n \mathbf{x}_n e_n^3. \tag{4.8}$$

Multiplying both sides of this equation by $\mathbf{x}_n^T$ from the left we find that the *a priori* and *a posteriori* estimation errors,$\{e_{pn},\, e_{an}\}$, are related via

$$e_{pn} = e_{an} - \mu_n \left\|\mathbf{x}_n\right\|^2 e_n^3. \tag{4.9}$$

This is an identical equation that we got previously in Chapter 2 for the stationary case. By following the same arguments, we can write

$$\left\|\mathbf{w}^o_{n+1} - \mathbf{w}_{n+1}\right\|^2 + \overline{\mathbf{x}}_n e_{an}^2 = \left\|\mathbf{w}^o_{n+1} - \mathbf{w}_n\right\|^2 + \overline{\mathbf{x}}_n e_{pn}^2, \tag{4.10}$$

where again

$$\overline{\mathbf{x}}_n \triangleq \begin{cases} 1/\left\|\mathbf{x}_n\right\|^2 & \text{if } \mathbf{x}_n \neq 0 \\ 0 & \text{otherwise} \end{cases}. \tag{4.11}$$

One difference worth noting is that the first term on the left hand side is $\left\|\mathbf{v}_{n+1}\right\|^2$ but this is not the case for the first term in right hand side. This difference occurs due to the fact that the expression of the *a priori* estimation error in the nonstationary case is different compared to the stationary case. Substituting (4.1) in the right hand side of (4.10) we get

$$\begin{aligned}\left\|\mathbf{w}^o_{n+1} - \mathbf{w}_{n+1}\right\|^2 + \overline{\mathbf{x}}_n e^2_{an} &= \left\|\mathbf{w}^o_n + \mathbf{q}_n - \mathbf{w}_n\right\|^2 + \overline{\mathbf{x}}_n e^2_{pn}, \\ \left\|\mathbf{v}_{n+1}\right\|^2 + \overline{\mathbf{x}}_n e^2_{an} &= \left\|\mathbf{v}_n\right\|^2 + \overline{\mathbf{x}}_n e^2_{pn} + \left\|\mathbf{q}_n\right\|^2 . \end{aligned} \tag{4.12}$$

Taking the expectation of (4.12) we arrive at the following expression [5]:

$$E\left[\left\|\mathbf{v}_{n+1}\right\|^2\right] + E\left[\overline{\mathbf{x}}_n e^2_{an}\right] = E\left[\left\|\mathbf{v}_n\right\|^2\right] + E\left[\overline{\mathbf{x}}_n e^2_{pn}\right] + tr(\mathbf{Q}). \tag{4.13}$$

Compared to the stationary case equation (3.4), (4.13) only differs in the term $tr(\mathbf{Q})$ on the right hand side of the equation. Therefore the arguments made in the arriving at (3.8) hold here as well and we can write

$$\overline{\mu^2_n} E\left[\left\|\mathbf{x}_n\right\|^2 e^6_n\right] + tr(\mathbf{Q}) = 2\overline{\mu_n} E\left[e_{an} e^3_n\right]. \tag{4.14}$$

Recognising that all the terms have been defined in Chapter 3, following a similar procedure we can arrive at the variance relation of the proposed VSSLMFQ algorithm for tracking case as well. Therefore, expression for finding the tracking steady-state EMSE of the VSSLMFQ algorithm is given as

$$6\overline{\mu_n}\sigma^2_z E\left[e^2_{an}\right] = \overline{\mu^2_n}\psi^6_z tr(\mathbf{R}) + 15\overline{\mu^2_n}\psi^4_z E\left[\left\|\mathbf{x}_n\right\|^2 e^2_{an}\right] + tr(\mathbf{Q}). \tag{4.15}$$

Considering the two scenarios for which the EMSE can be defined, we find the tracking steady-state EMSE as:

Separation Principle

By using the same arguments presented in Section 3.2.2, the EMSE for the separation principle in the nonstationary case is given as

$$J_{ex} = \frac{\overline{\mu_{\infty}^2}\psi_z^6 tr(\mathbf{R}) + tr(\mathbf{Q})}{6\overline{\mu_{\infty}}\sigma_z^2 - 15\overline{\mu_{\infty}^2}\psi_z^4 tr(\mathbf{R})}. \tag{4.16}$$

Substituting (3.25) and (3.28) in (4.16) we get

$$\begin{aligned} J_{ex} &= \frac{\frac{2\alpha\gamma^2(1-b)^2}{(1-\alpha^2)(1-\alpha)(1-a)^2}\psi_z^6 tr(\mathbf{R}) + tr(\mathbf{Q})}{6\frac{\gamma(1-b)}{(1-\alpha)(1-a)}\sigma_z^2 - 15\frac{2\alpha\gamma^2(1-b)^2}{(1-\alpha^2)(1-\alpha)(1-a)^2}\psi_z^4 tr(\mathbf{R})}, \\ &= \frac{2\alpha\gamma^2\left(1-b\right)^2\psi_z^6 tr(\mathbf{R}) + \left(1-\alpha^2\right)\left(1-\alpha\right)\left(1-a\right)^2 tr(\mathbf{Q})}{6\gamma\left(1-\alpha^2\right)\left(1-a\right)\left(1-b\right)\sigma_z^2 - 30\alpha\gamma^2\left(1-b\right)^2\psi_z^4 tr(\mathbf{R})}, \end{aligned}$$

$$\begin{aligned} J_{ex} &= \frac{\alpha\gamma\left(1-b\right)\psi_z^6 tr(\mathbf{R})}{3\left(1-\alpha^2\right)\left(1-a\right)\sigma_z^2 - 15\alpha\gamma\left(1-b\right)\psi_z^4 tr(\mathbf{R})} \\ &\quad + \frac{\left(1-\alpha^2\right)\left(1-\alpha\right)\left(1-a\right)^2 tr(\mathbf{Q})}{6\gamma\left(1-\alpha^2\right)\left(1-a\right)\left(1-b\right)\sigma_z^2 - 30\alpha\gamma^2\left(1-b\right)^2\psi_z^4 tr(\mathbf{R})}. \end{aligned} \tag{4.17}$$

Gaussian Regressor

Using again the same arguments presented in Section 3.2.2, the EMSE for the Gaussian regressor in the nonstationary case is given as

$$J_{ex} = \frac{\overline{\mu_{\infty}^2}N\psi_z^6\sigma_x^2 + tr(\mathbf{Q})}{6\overline{\mu_{\infty}}\sigma_z^2 - 15\overline{\mu_{\infty}^2}\left(N+2\right)\psi_z^4\sigma_x^2}. \tag{4.18}$$

Substituting (3.25) and (3.28) in (4.18) we get

$$
\begin{aligned}
J_{ex} &= \frac{\frac{2\alpha\gamma^2(1-b)^2}{(1-\alpha^2)(1-\alpha)(1-a)^2}\psi_z^6\sigma_x^2 N + tr(\mathbf{Q})}{6\frac{\gamma(1-b)}{(1-\alpha)(1-a)}\sigma_z^2 - 15\frac{2\alpha\gamma^2(1-b)^2}{(1-\alpha^2)(1-\alpha)(1-a)^2}\psi_z^4\left(N+1\right)\sigma_x^2}, \\
&= \frac{2\alpha\gamma^2\left(1-b\right)^2\psi_z^6\sigma_x^2 N + \left(1-\alpha^2\right)\left(1-\alpha\right)\left(1-a\right)^2 tr(\mathbf{Q})}{6\gamma\left(1-\alpha^2\right)\left(1-a\right)\left(1-b\right)\sigma_z^2 - 30\alpha\gamma^2\left(1-b\right)^2\psi_z^4\left(N+1\right)\sigma_x^2},
\end{aligned}
$$

$$
\begin{aligned}
J_{ex} &= \frac{\alpha\gamma\left(1-b\right)\psi_z^6\sigma_x^2 N}{3\left(1-\alpha^2\right)\left(1-a\right)\sigma_z^2 - 15\alpha\gamma\left(1-b\right)\psi_z^4\sigma_x^2 N} \\
&\quad + \frac{\left(1-\alpha^2\right)\left(1-\alpha\right)\left(1-a\right)^2 tr(\mathbf{Q})}{6\gamma\left(1-\alpha^2\right)\left(1-a\right)\left(1-b\right)\sigma_z^2 - 30\alpha\gamma^2\left(1-b\right)^2\psi_z^4\sigma_x^2 N}. \qquad (4.19)
\end{aligned}
$$

Remark It is evident that due to the nonstationarity of the channel, the additional term in (4.17) and (4.19), will increase the EMSE as compared to the stationary case. Furthermore referring to (3.29), we can rewrite equations (4.16) and (4.18) as

$$
J_{ex} = \frac{2\alpha\overline{\mu_\infty}\psi_z^6 tr(\mathbf{R}) + \left(1+\alpha\right)\overline{\mu_\infty}^{-1} tr(\mathbf{Q})}{6\left(1+\alpha\right)\sigma_z^2 - 30\alpha\overline{\mu_\infty}\psi_z^4 tr(\mathbf{R})}, \qquad (4.20)
$$

and

$$
J_{ex} = \frac{2\alpha\overline{\mu_\infty}N\psi_z^6\sigma_x^2 + \left(1+\alpha\right)\overline{\mu_\infty}^{-1} tr(\mathbf{Q})}{6\left(1+\alpha\right)\sigma_z^2 - 30\alpha\overline{\mu_\infty}\left(N+2\right)\psi_z^4\sigma_x^2}, \qquad (4.21)
$$

respectively. Observe again that for $\alpha = 1$, the equations reduce to the tracking EMSE of LMF. Also noticeable is the inverse variable step-size that is multiplied with the trace of the perturbation covariance matrix, $tr(\mathbf{Q})$. This means, the larger the step-size, the lower the effects of nonstationarity [5]. This follows from the fact that larger step-size signifies faster adaptation hence better tracking while a smaller one may result in degradation in tracking performance. This motivates an introduction of the optimum choice of step-size. This can be obtained by minimising (4.20) and (4.21) with respect to μ. Differentiating (4.20) and (4.21) and equating to zero, we get

$$\overline{\mu}_{opt} = \sqrt{\frac{(1+\alpha)\,tr\,(\mathbf{Q})}{2\alpha\psi_z^6 tr\,(\mathbf{R})} + \frac{25\,(tr\,(\mathbf{Q}))^2}{4\,(\sigma_z^4)^2}} - \frac{5tr\,(\mathbf{Q})}{2\sigma_z^4}, \tag{4.22}$$

and

$$\overline{\mu}_{opt} = \sqrt{\frac{(1+\alpha)\,tr\,(\mathbf{Q})}{2\alpha N\psi_z^6 tr\,(\mathbf{R})} + \frac{25\,(N+2)^2\,(tr\,(\mathbf{Q}))^2}{4\,(\sigma_z^4)^2}} - \frac{5\,(N+2)\,tr\,(\mathbf{Q})}{2N\sigma_z^4}. \tag{4.23}$$

These can be substituted in (4.20) and (4.21) to get the minimum EMSE.

4.3 Rayleigh Fading Channel Model

It is known that in a wireless communications environment, signals suffer from multiple reflections while travelling from the transmitter to the receiver so that the receiver ends up getting several (almost simultaneous) replicas of the transmitted signal. The reflections are received with different amplitude and phase distortions, and the overall received signal is the combined sum of the reflections. Based on the relative phases of the reflections, the signals may add up constructively or destructively at the receiver. Furthermore, if the transmitter is moving with respect to the receiver, these destructive and constructive interferences will vary with time. This phenomenon is known as channel fading.

4.3.1 Single Path Channel

The impulse response of a single-path (i.e., single-tap) fading channel can be described as

$$h_n = \zeta s_n \delta_{n-n_o} \tag{4.24}$$

where $\{s_n\}$ is a time-variant complex sequence that models the time-variations in the channel, and n_o is the channel delay. Without loss of generality, the sequence $\{s_n\}$ is assumed to have unit variance, and the scalar ζ is used to model the actual path loss that is introduced by the channel. That is, ζ^2 is equal to the power attenuation that a signal will undergo when it travels through the channel.

There are several mathematical models can be used to characterise the fading properties of $\{s_n\}$,but one that is widely used is known as Rayleigh fading. The amplitude $|s_n|$ in this scenario is assumed to have a Rayleigh distribution for all n given by

$$f_{|s_n|}\left(|s_n|\right) = |s_n|\, e^{-|s_n|^2/2}, \qquad |s_n| \geq 0 \tag{4.25}$$

while the phase $\angle\,|s_n|$ is assumed to be uniformly distributed within $[-\pi, \pi]$

$$f\left(\angle\,|s_n|\right) = \frac{1}{2\pi}, \qquad -\pi \leq \angle\,|s_n| \leq \pi. \tag{4.26}$$

In addition, it is further assumed that all scatterers are uniformly distributed on a circle around the receiver. Then the model widely used in literature to fit the autocorrelation function of $\{s_n\}$ is the zeroth-order Bessel function of the first kind given by [5]

$$r_n \triangleq s_n s_{n-k} = \mathcal{J}_o(2\pi f_D T_s k), \qquad k = \ldots, -1, 0, 1, \ldots \tag{4.27}$$

where T_s, is the sampling period of the sequence $\{s_n\}$, f_D is called the maximum Doppler frequency of the Rayleigh fading channel, and the function $\mathcal{J}_o(.)$ is defined by

$$\mathcal{J}_o(y) = \frac{1}{\pi}\int_0^{\pi} cos\left(ysin\left(t\right)\right) dt. \tag{4.28}$$

The Doppler frequency f_D is related to the speed of the mobile user, v, and to the carrier frequency, f_c, as follows:

$$f_D = \frac{v f_c}{c}, \tag{4.29}$$

where c denotes the speed of light, $c = 3 \times 10^{-8} m/s$. The power spectral density of the channel fading gain $\{s_n\}$, in continuous-time, would have the following well-known U-shaped spectrum:

$$S\left(f\right) = \frac{1}{\pi f_D \sqrt{1 - \left(\frac{f}{f_D}\right)^2}}, \qquad |f| \leq f_D. \tag{4.30}$$

4.3.2 Multipath Channel

It is commonly seen in wireless communication that other reflections might be originated from a far away object such as a mountain or a tall building. These reflections arrive at the receiver with longer delay than the first group of reflections. In such situations, a single-path Rayleigh fading model is not adequate to represent the wireless channel. To model this *multipath* phenomenon, a finite-impulse response model for the channel can be used, say one of the form

$$h_n = \sum_{k=1}^{L} \zeta_k s_{n_k} \delta_{n-n_k}, \tag{4.31}$$

where $\{\zeta_k\}$ and $\{s_{n_k}\}$ are, respectively, the path loss and fading sequence of the $k-th$ cluster of reflectors, and the $\{n_k\}$ are the cluster delays. The sequence $\{s_{n_k}\}$ are modelled as independent Rayleigh fading sequences and the channel is referred to as a multipath Rayleigh fading channel.

In this analysis, we consider a wireless channel with two Rayleigh fading rays with both rays assumed to fade at the same Doppler frequency. The channel impulse response sequence consists of two zero initial samples (i.e., delay of two samples), followed by a Rayleigh fading ray, followed by another zero sample, and by a second Rayleigh fading ray. In other words, we are assuming a channel length of $N = 5$ taps with only two active Rayleigh fading rays, so that the weight vector that we wish to estimate has the form

$$\begin{bmatrix} 0 & 0 & s_{1n} & 0 & s_{2n} \end{bmatrix}. \tag{4.32}$$

According to [5], a first-order approximation for the variation of a Rayleigh fading coefficient s_n is to assume that s_n varies according to the auto-regressive model

$$s_n = r\left(1\right) s_{n-1} + \sqrt{1 - \left|r\left(1\right)\right|^2}\nu_n, \tag{4.33}$$

where $r\left(1\right) = J_o(2\pi f_D T_s)$ and ν_n denotes a white noise process with unit variance.

Now since the multipath rays of the channel are assumed to fade at the same rate, the above approximation indicates that the variations in the channel weight vector could be approximated as

$$\mathbf{w}_{n+1}^{o} = \eta \mathbf{w}_{n}^{o} + \mathbf{q}_{n}, \tag{4.34}$$

where the covariance matrix of sequence $\{q_n\}$ is $\mathbf{Q} = (1 - \eta^2)\,\mathbf{I}$ with $r\,(1) = \eta$. The value of η depends upon the Doppler frequency of the channel. Unless this value is other than unity, the analysis for the Random-Walk model is equally applicable to this scenario as well.

Chapter 5

Transient Analysis of the VSSLMFQ Algorithm

5.1 Introduction

Transient analysis is one of the most important aspects of performance analysis of adaptive filters. It involves the analysis of the time-evolution of the adaptive algorithms under variations to the signal statistics; the main aim being to study the learning mechanism of the adaptive algorithm. Consequently, the algorithm's rate of convergence and stability performance becomes a central part of transient analysis. The methodology used for the transient analysis was given in [33] and described in detail in [5] where again the concept of energy conservation is used to carry out the analysis. The main benefit of using this approach is that it does not require the input regressor to be Gaussian. So as we described in Chapter 3, the analysis presented here would be general for all data. Moreover, there is no need for the recursions of the weight-error covariance matrix to be evaluated.

Before proceeding to the analysis, lets go through the underlying assumptions made for the analysis.

A8 The input process $\{x_n\}$ is an i.i.d. sequence of Gaussian random variable with zero-mean and autocorrelation matrix $\mathbf{R} = E\left[\mathbf{x}_n\mathbf{x}_n^T\right]$.

A9 The noise process $\{z_n\}$ is a zero mean and identically distributed (i.i.d.) with variance σ_z^2 and independent of the input process $\{x_n\}$.

A10 The step-size μ_n is statistically independent of the weight update vector $\mathbf{w}_n$.

Withe the exception of **A8**, these assumptions have been described in detail in Chapter 3 and are relevant for the transient analysis as well. Assumption **A8** is not true in practise but it is very common in literature and many works have shown that the analytical results obtained under this assumption agree closely with simulation results under general conditions.

5.2 Transient Analysis of the VSSLMFQ Algorithm

The proposed algorithm's weight update equation given in Chapter 2 can be written as

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu_n \mathbf{x}_n e_n^3, \tag{5.1}$$

where μ_n varies as given in (2.10) and (2.11). With the weight-error vector defined as $\mathbf{v}_n = \mathbf{w}^o - \mathbf{w}_n$, we can write (5.1) as

$$\mathbf{v}_{n+1} = \mathbf{v}_n - \mu_n \mathbf{x}_n e_n^3. \tag{5.2}$$

With reference to the energy conservation method discussed in Chapter 2, in this section we have used a weighted version of the energy relation. Moreover, as opposed to the steady-state, where the weight-error vector vanished under the limit $n \to \infty$, the transient analysis is based on its time-evolution.

Let us define the weighted squared Euclidean norm of a vector $\mathbf{x}_n$ as

$$\|\mathbf{x}_n\|_\Sigma^2 = \mathbf{x}_n^T \Sigma \mathbf{x}_n, \tag{5.3}$$

where Σ is any positive-definite weighting matrix. Choosing $\Sigma = \mathbf{I}$ results in the standard Euclidean norm on $\mathbf{x}$

$$\|\mathbf{x}_n\|_\mathbf{I}^2 = \mathbf{x}_n^T \mathbf{I} \mathbf{x}_n = \|\mathbf{x}_n\|^2. \tag{5.4}$$

We now define the two weighted error quantities, *weighted a priori estimation error*, e_{an}^Σ and the *weighted a posteriori estimation error*, e_{pn}^Σ as

$$e_{an}^\Sigma \triangleq \mathbf{x}_n^T \Sigma \mathbf{v}_n, \tag{5.5}$$

and

$$e_{pn}^{\Sigma} \triangleq \mathbf{x}_n^T \Sigma \mathbf{v}_{n+1}. \tag{5.6}$$

If we multiply (5.2) by $\mathbf{x}_n^T \Sigma$, then the expression relating e_{an}^{Σ} and e_{pn}^{Σ} is given as

$$e_{pn}^{\Sigma} = e_{an}^{\Sigma} - \mu_n \left\| \mathbf{x}_n \right\|_{\Sigma}^2 e_n^3. \tag{5.7}$$

The expression gives an alternate description of adaptive filter in terms of its error quantities, e_{an}^{Σ}, e_{pn}^{Σ}, $\mathbf{v}_n$, $\mathbf{v}_{n+1}$ and e_n^3. Solving it for e_n^3 and substituting in (5.2) results in

$$\mathbf{v}_{n+1} + \frac{\mathbf{x}_n e_{an}^{\Sigma}}{\left\| \mathbf{x}_n \right\|_{\Sigma}^2} = \mathbf{v}_n + \frac{\mathbf{x}_n e_{pn}^{\Sigma}}{\left\| \mathbf{x}_n \right\|_{\Sigma}^2}. \tag{5.8}$$

Taking the weighted Euclidean norm on both sides and equating gives us

$$\left\| \mathbf{v}_{n+1} + \frac{\mathbf{x}_n e_{an}^{\Sigma}}{\left\| \mathbf{x}_n \right\|_{\Sigma}^2} \right\|_{\Sigma}^2 = \left\| \mathbf{v}_n + \frac{\mathbf{x}_n e_{pn}^{\Sigma}}{\left\| \mathbf{x}_n \right\|_{\Sigma}^2} \right\|_{\Sigma}^2. \tag{5.9}$$

By evaluating the energies on both sides and knowing the fact that $P\left[\left\| \mathbf{x}_n \right\|_{\Sigma}^2 = 0\right] = 0$, the following energy equality holds:

$$\left\| \mathbf{v}_{n+1} \right\|_{\Sigma}^2 + \overline{\mathbf{x}}_n^{\Sigma} \left| e_{an}^{\Sigma} \right|^2 = \left\| \mathbf{v}_n \right\|_{\Sigma}^2 + \overline{\mathbf{x}}_n^{\Sigma} \left| e_{pn}^{\Sigma} \right|^2, \tag{5.10}$$

where

$$\overline{\mathbf{x}}_n^{\Sigma} \triangleq \begin{cases} 1/\left\| \mathbf{x}_n \right\|_{\Sigma}^2 & \text{if } \left\| \mathbf{x}_n \right\|_{\Sigma}^2 \neq 0 \\ 0 & \text{otherwise} \end{cases}. \tag{5.11}$$

The weighted energy conservation relation in (5.10) shows the time-evolution of the error quantities. The free parameter Σ can be appropriately chosen to have different values allowing us to analyse the algorithm for different performance measures. These performance measures will be discussed later.

Now by replacing e^{Σ}_{pn} from (5.6) in (5.10) to get the expression in terms of e^{Σ}_{an}, we get

$$\|\mathbf{v}_{n+1}\|^2_{\Sigma} + \overline{\mathbf{x}}^{\Sigma}_n \left|e^{\Sigma}_{an}\right|^2 = \|\mathbf{v}_n\|^2_{\Sigma} + \overline{\mathbf{x}}^{\Sigma}_n \left|e^{\Sigma}_{an} - \mu_n \|\mathbf{x}_n\|^2_{\Sigma} e^3_n\right|^2 . \tag{5.12}$$

Expanding the right hand side and taking the expectation on both sides gives us

$$E\left[\|\mathbf{v}_{n+1}\|^2_{\Sigma}\right] = E\left[\|\mathbf{v}_n\|^2_{\Sigma}\right] + \overline{\mu^2_n} E\left[\|\mathbf{x}_n\|^2_{\Sigma} e^6_n\right] - 2\overline{\mu_n} E\left[e^{\Sigma}_{an} e^3_n\right], \tag{5.13}$$

where$\overline{\mu^2_n} = E\left[\mu^2_n\right]$ and $\overline{\mu_n} = E\left[\mu_n\right]$. In order to evaluate the expectation in the above equations, we make the following assumptions [5], [33]:

A11 The *a priori* estimation errors $\left\{e_{an}, e^{\Sigma}_{an}\right\}$ are jointly Gaussian.

A12 The *a priori* estimation error e_{an} and the noise process $\{z_n\}$ are independent.

Both of these assumptions are valid enough for long adaptive filters. As $e_{an} = \mathbf{x}^T_n \mathbf{v}_n$, it can be thought of as the sum of $(N-1)$ random variables. According to the Central Limit Theorem [34], [35], as the length of the filter increases, the distribution of the sum can be approximated as Gaussian. Similar arguments hold for $\left\{e^{\Sigma}_{an}\right\}$. This test of Gaussianity was also carried out in [33] and the result was quite consistent with **A11**. Coming back to the equation, we can simplify the expectation $E\left[e^{\Sigma}_{an} e^3_n\right]$ using

Price's Theorem [36], **A11** and the fact that $e_n = e_{an} + z_n$, as

$$E\left[e_{an}^{\Sigma} e_n^3\right] = E\left[e_{an}^{\Sigma} e_{an}\right] \cdot \left(\frac{E\left[e_{an} e_n^3\right]}{E\left[e_{an}^2\right]}\right). \tag{5.14}$$

It is known that from literature that the expectation of a function of Gaussian random variable will depend only on the variance of this variable and not on the higher order moments of it [35]. Therefore, according to **A11**, the expectation $E\left[e_{an}^{\Sigma} e_n^3\right]$ depends on e_{an} only through it second moment, $E\left[e_{an}^2\right]$. Using this result, we can introduce a new term as defined in [5], [37]

$$h_g = \frac{E\left[e_{an} e_n^3\right]}{E\left[e_{an}^2\right]}. \tag{5.15}$$

This simplifies in the case of our proposed VSSLMFQ algorithm as

$$h_g = 3\left(E\left[e_{an}^2\right] + \sigma_z^2\right). \tag{5.16}$$

Hence the expectation of $E\left[e_{an}^{\Sigma} e_n^3\right]$ becomes

$$E\left[e_{an}^{\Sigma} e_n^3\right] = h_g E\left[e_{an}^{\Sigma} e_{an}\right]. \tag{5.17}$$

The other expectation in (5.13), $E\left[\|\mathbf{x}_n\|_{\Sigma}^2 e_n^6\right]$, also needs to be simplified in order to make the analysis tractable. This is done by assuming a long enough filter and that

A13 The weighted error norm of the input regressor $\|\mathbf{x}_n\|_{\Sigma}^2$ is independent of e_n.

This allows us to split the expectation as

$$E\left[\|\mathbf{x}_n\|_\Sigma^2 e_n^6\right] = E\left[\|\mathbf{x}_n\|_\Sigma^2\right] E\left[e_n^6\right]. \tag{5.18}$$

The same logic used in defining h_g can be used here motivated by the fact that e_{an} is Gaussian and independent of the noise, therefore $E\left[e_n^6\right]$ depends on e_{an} only through its second moment. This allows us to introduce another term as defined in [5], [37]

$$h_U = E\left[e_n^6\right]. \tag{5.19}$$

After some straight forward manipulations, this simplifies in the case of our proposed VSSLMFQ algorithm as

$$h_U = 15\left(E\left[e_{an}^2\right]\right)^3 + 45\sigma_z^2\left(E\left[e_{an}^2\right]\right)^2 + 15\psi_z^4 E\left[e_{an}^2\right] + \psi_z^6. \tag{5.20}$$

Hence the expectation of $E\left[\|\mathbf{x}_n\|_\Sigma^2 e_n^6\right]$ can be written as

$$E\left[\|\mathbf{x}_n\|_\Sigma^2 e_n^6\right] = h_U tr\left(\mathbf{R\Sigma}\right). \tag{5.21}$$

With the expectations simplified, we can substitute (5.17) and (5.21) in (5.13) to get the following version of the *weighted-variance relation*:

$$E\left[\|\mathbf{v}_{n+1}\|_\Sigma^2\right] = E\left[\|\mathbf{v}_n\|_\Sigma^2\right] + \overline{\mu_n^2} h_U tr\left(\mathbf{R\Sigma}\right) - 2\overline{\mu_n} h_g E\left[e_{an}^\Sigma e_{an}\right]. \tag{5.22}$$

This is the weighted variance relation that will be used for the time-evolution analysis of the algorithm. The expression still holds an expectation, i.e., $E\left[e_{an}^{\Sigma} e_{an}\right]$ that needs to be solved before we can proceed. The evaluation of this expectation is difficult due to the dependencies among the regressors $\{\mathbf{x}_n\}$. With assumption **A8**, we can solve this expectation as

$$\begin{aligned} E\left[e_{an}^{\Sigma} e_{an}\right] &= E\left[\mathbf{x}_n^T \Sigma \mathbf{v}_n \mathbf{x}_n^T \mathbf{I} \mathbf{v}_n\right] \\ &= E\left[\|\mathbf{v}_n\|^2_{\Sigma E[\mathbf{x}_n \mathbf{x}_n^T]\mathbf{I}}\right] \\ &= E\left[\|\mathbf{v}_n\|^2_{\Sigma \mathbf{R}}\right]. \end{aligned} \tag{5.23}$$

and therefore (5.22) reduces to

$$E\left[\|\mathbf{v}_{n+1}\|^2_{\Sigma}\right] = E\left[\|\mathbf{v}_n\|^2_{\Sigma}\right] + \overline{\mu_n^2} h_U tr\left(\mathbf{R}\mathbf{\Sigma}\right) - 2\overline{\mu_n} h_g E\left[\|\mathbf{v}_n\|^2_{\mathbf{\Sigma}\mathbf{R}}\right]. \tag{5.24}$$

The most important result underlying the above equation is that the study of transient analysis of the proposed VSSLMFQ algorithm reduces to evaluating h_g and h_U. Recalling the fact discussed earlier that the choice of the free parameter Σ influences the performance analysis of the algorithm, the following analysis is divided into two parts.

5.2.1 Transient Analysis of the VSSLMFQ : White Input Data

In the case when the input data is white, the individual entries of $\{\mathbf{x}_n\}$ are i.i.d., i.e., $\mathbf{R}$ is a diagonal matrix with entries $\mathbf{R} = \sigma_x^2\mathbf{I}$ and $E\left[e_{an}^2\right] = \sigma_x^2 E\left[\|\mathbf{v}_n\|^2\right]$, then for $\Sigma = \mathbf{I}$, the variance relation of (5.24) becomes

$$E\left[\|\mathbf{v}_{n+1}\|^2\right] = E\left[\|\mathbf{v}_n\|^2\right] + \overline{\mu_n^2} h_U \sigma_x^2 N - 2\overline{\mu_n}\sigma_x^2 h_g E\left[\|\mathbf{v}_n\|^2\right]. \tag{5.25}$$

With h_g and h_U now being functions of the $E\left[\|\mathbf{v}_n\|^2\right]$, we can replace these, derived in (5.17) and (5.21), respectively, for the proposed VSSLMFQ algorithm, into the above equation to arrive at

$$\begin{aligned} E\left[\|\mathbf{v}_{n+1}\|^2\right] &= E\left[\|\mathbf{v}_n\|^2\right] + \overline{\mu_n^2}\sigma_x^2 N\left[15\left(E\left[e_{an}^2\right]\right)^3 + 45\sigma_z^2\left(E\left[e_{an}^2\right]\right)^2 + 15\psi_z^4 E\left[e_{an}^2\right] + \psi_z^6\right] \\ &\quad -6\overline{\mu_n}\sigma_x^2 E\left[\|\mathbf{v}_n\|^2\right]\left[\left(E\left[e_{an}^2\right] + \sigma_z^2\right)\right]. \end{aligned} \tag{5.26}$$

This can be compactly expressed as

$$E\left[\|\mathbf{v}_{n+1}\|^2\right] = \mathbf{f} E\left[\|\mathbf{v}_n\|^2\right] + \overline{\mu_n^2}\sigma_x^2\psi_z^6 N, \tag{5.27}$$

where

$$\begin{aligned} \mathbf{f} &= 1 + \sigma_x^2\left(15\overline{\mu_n^2}\sigma_x^2\psi_z^4 N - 6\overline{\mu_n}\sigma_z^2\right) + \sigma_x^4\left(45\overline{\mu_n^2}\sigma_x^2\sigma_z^2 N - 6\overline{\mu_n}\right) E\left[\|\mathbf{v}_n\|^2\right] \\ &\quad +15\overline{\mu_n^2}\sigma_x^8 N\left(E\left[\|\mathbf{v}_n\|^2\right]\right)^2. \end{aligned} \tag{5.28}$$

The expression arrived in (5.27) describes the transient behaviour of the proposed VSSLMFQ algorithm for white input data.

5.2.2 Transient Analysis of the VSSLMFQ : Correlated Input Data

When the data is correlated, it can be seen that different weighting matrices will appear on both sides of (5.24). This was not the case in the uncorrelated data where unweighted norms of $\mathbf{v}_n$ appeared on both sides. For this analysis, we have taken advantage of the free parameter $\mathbf{\Sigma}$ by choosing $\mathbf{\Sigma} = \mathbf{I}$, $\mathbf{R}$, $\mathbf{R}^2, \ldots, \mathbf{R}^{N-1}$. Writing (5.24) for $\mathbf{\Sigma} = \mathbf{I}$ and also keeping in mind the fact that $E\left[e_{an}^2\right] = E\left[\|\mathbf{v}_n\|_{\mathbf{R}}^2\right]$we get

$$E\left[\|\mathbf{v}_{n+1}\|_{\mathbf{I}}^2\right] = E\left[\|\mathbf{v}_n\|_{\mathbf{I}}^2\right] + \overline{\mu_n^2} h_U tr\left(\mathbf{RI}\right) - 2\overline{\mu_n} h_g E\left[\|\mathbf{v}_n\|_{\mathbf{RI}}^2\right]. \tag{5.29}$$

It is seen that a weighting matrix $\mathbf{R}$ appears on the right-hand side of the equation that weights the norm $E\left[\|\mathbf{v}_n\|^2\right]$. This term can be inferred by writing (5.24) for $\mathbf{\Sigma} = \mathbf{R}$, which gives us

$$E\left[\|\mathbf{v}_{n+1}\|_{\mathbf{R}}^2\right] = E\left[\|\mathbf{v}_n\|_{\mathbf{R}}^2\right] + \overline{\mu_n^2} h_U tr\left(\mathbf{R}^2\right) - 2\overline{\mu_n} h_g E\left[\|\mathbf{v}_n\|_{\mathbf{R}^2}^2\right]. \tag{5.30}$$

Again, the the norm $E\left[\|\mathbf{v}_n\|^2\right]$ is weighted by the weighting matrix $\mathbf{R}^2$, appearing on the right-hand side. Again in the same fashion, we can infer the term by writing (5.24) for $\mathbf{\Sigma} = \mathbf{R}^2$ and continue in the same fashion until we write (5.24) for $\mathbf{\Sigma} = \mathbf{R}^{N-1}$ as

$$E\left[\|\mathbf{v}_{n+1}\|^2_{\mathbf{R}^{N-1}}\right] = E\left[\|\mathbf{v}_n\|^2_{\mathbf{R}^{N-1}}\right] + \overline{\mu_n^2} h_U tr\left(\mathbf{R}^N\right) - 2\overline{\mu_n} h_g E\left[\|\mathbf{v}_n\|^2_{\mathbf{R}^N}\right]. \qquad (5.31)$$

In the above equation, the weighted norm $E\left[\|\mathbf{v}_n\|^2\right]$ is now weighted by the weighting matrix $\mathbf{R}^N$ that appears on the right-hand side. In all the above variance relations, the left-hand side is always one variable short of the number of variables on the right-hand side. Fortunately, we do not have to continue in this manner indefinitely since the additional variable $E\left[\|\mathbf{v}_n\|^2_{\mathbf{R}^N}\right]$ can be inferred from the prior weighting factors,

$$\left\{E\left[\|\mathbf{v}_n\|^2\right], E\left[\|\mathbf{v}_n\|^2_{\mathbf{R}}\right], E\left[\|\mathbf{v}_n\|^2_{\mathbf{R}^2}\right], \cdots, E\left[\|\mathbf{v}_n\|^2_{\mathbf{R}^{N-1}}\right]\right\},$$

by expressing $\mathbf{R}^N$ as a linear combination of the "lower order" variables using the Cayley–Hamilton theorem [33]. Therefore, we can write

$$\mathbf{R}^N = -p_0\mathbf{I} + p_1\mathbf{R} - \ldots \cdots - p_{N-1}\mathbf{R}^{N-1}, \qquad (5.32)$$

where

$$\begin{aligned} p(x) &\triangleq \det(x\mathbf{I} - \mathbf{R}), \\ &= p_0 + p_1 x + \cdots + p_{N-1}x^{N-1} + x^N, \end{aligned} \qquad (5.33)$$

is the characteristic polynomial of $\mathbf{R}$. This gives us the desired result and the variance relation (5.24) can be written as

$$\begin{aligned} E\left[\|\mathbf{v}_{n+1}\|^2_{\mathbf{R}^{N-1}}\right] &= E\left[\|\mathbf{v}_n\|^2_{\mathbf{R}^{N-1}}\right] + \overline{\mu_n^2} h_U tr\left(\mathbf{R}^N\right) + 2\overline{\mu_n} h_g \left(p_0 E\left[\|\mathbf{v}_n\|^2_{\mathbf{I}}\right]\right. \\ &\left. + p_1 E\left[\|\mathbf{v}_n\|^2_{\mathbf{R}}\right] + \cdots + p_{N-1} E\left[\|\mathbf{v}_n\|^2_{\mathbf{R}^{N-1}}\right]\right). \end{aligned} \tag{5.34}$$

The variance relations for $\mathbf{\Sigma} = \mathbf{I}$, $\mathbf{R}$, $\mathbf{R}^2, \ldots,$ $\mathbf{R}^{N-1}$ describe the transient behaviour of the proposed VSSLMFQ algorithm and can therefore be written as

$$\begin{cases} E\left[\|\mathbf{v}_{n+1}\|^2_{\mathbf{I}}\right] = E\left[\|\mathbf{v}_n\|^2_{\mathbf{I}}\right] + \overline{\mu_n^2} h_U tr\left(\mathbf{RI}\right) - 2\overline{\mu_n} h_g E\left[\|\mathbf{v}_n\|^2_{\mathbf{RI}}\right], \\ E\left[\|\mathbf{v}_{n+1}\|^2_{\mathbf{R}}\right] = E\left[\|\mathbf{v}_n\|^2_{\mathbf{R}}\right] + \overline{\mu_n^2} h_U tr\left(\mathbf{R}^2\right) - 2\overline{\mu_n} h_g E\left[\|\mathbf{v}_n\|^2_{\mathbf{R}^2}\right], \\ \quad\vdots \\ E\left[\|\mathbf{v}_{n+1}\|^2_{\mathbf{R}^{N-1}}\right] = E\left[\|\mathbf{v}_n\|^2_{\mathbf{R}^{N-1}}\right] + \overline{\mu_n^2} h_U tr\left(\mathbf{R}^N\right) - 2\overline{\mu_n} h_g E\left[\|\mathbf{v}_n\|^2_{\mathbf{R}^N}\right]. \end{cases} \tag{5.35}$$

The equation warrants a compact form and therefore (5.35) can be written in a recursive manner as

$$\mathcal{V}_{n+1} = \mathcal{F}_n \mathcal{V}_n + \overline{\mu_n^2} h_U \mathcal{Y}, \tag{5.36}$$

where

$$\mathcal{V}_n = \left[E\left[\|\mathbf{v}_n\|^2\right] \; E\left[\|\mathbf{v}_n\|^2_{\mathbf{R}}\right] \cdots \; E\left[\|\mathbf{v}_n\|^2_{\mathbf{R}^{N-1}}\right]\right]^T, \tag{5.37}$$

$$\mathcal{Y} = \left[tr\left(\mathbf{R}\right) \; tr\left(\mathbf{R}^2\right) \cdots tr\left(\mathbf{R}^N\right)\right]^T, \tag{5.38}$$

and $\mathcal{F}_n$ is given by

$$\mathcal{F}_n = \begin{bmatrix} 1 & -2\overline{\mu_n}h_g & 0 & \cdots & 0 & 0 \\ 0 & 1 & -2\overline{\mu_n}h_g & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -2\overline{\mu_n}h_g \\ 2\overline{\mu_n}p_0h_g & 2\overline{\mu_n}p_1h_g & 2\overline{\mu_n}p_2h_g & \cdots & 2\overline{\mu_n}p_{N-2}h_g & 1+2\overline{\mu_n}p_{N-1}h_g \end{bmatrix}, \quad (5.39)$$

where (5.36) is a state-space equation,$\mathcal{V}_n$ is the state vector that represents the time-evolution of the filter and $\mathcal{F}_n$ is the coefficient matrix.

The contrast between the transient behaviour for white input data (5.27) and correlated input data (5.36) can be appreciated by elaborating the fact that in the later case, the transient behaviour is described by an N-dimensional state-space equation while the former was a one-dimensional recursion.

Another important consequence of (5.36) being a state vector is that the mean-square deviation (MSD) and the mean-square error (MSE) can be obtained from the first and second entries of this vector, respectively. As $E\left[e_{an}^2\right] = E\left[\|\mathbf{v}_n\|_{\mathbf{R}}^2\right]$, we have

$$\text{MSE} = \lim_{n\to\infty} E\left[\|\mathbf{v}_n\|_{\mathbf{R}}^2\right] + \sigma_z^2. \quad (5.40)$$

Also the MSD would be given as

$$\text{MSD} = \lim_{n\to\infty} E\left[\|\mathbf{v}_n\|^2\right]. \quad (5.41)$$

In all of the discussions done so far, we have not taken into account the transient behaviour of the time-varying step-size. The transient analysis will remain incomplete without a detailed analysis of the time-varying step-size. We now proceed to characterise the transient behaviour of our time-varying step-size to arrive at the expressions for $\overline{\mu_n}$ and $\overline{\mu_n^2}$.

5.2.3 Mean and Mean-Square Behaviour of the Step-Size

In this section, we have discussed the transient behaviour of the time-varying step-size parameter of the proposed VSSLMFQ algorithm. This is essential because the algorithms main controlling parameter is the step-size. It being time-varying, hence a random variable, needs to be properly studied and analysed. For the analysis in this section, we have supposed the input regressor to be white Gaussian.

The step-size update equation is given by the following equation [13]:

$$\mu_{n+1} = \alpha\mu_n + \gamma\theta_n, \tag{5.42}$$

where θ_n, the quotient of the filtered quadratic error is given as

$$\theta_n = \frac{\sum_{i=0}^{n} a^i e_{n-i}^2}{\sum_{j=0}^{n} b^j e_{n-j}^2}, \tag{5.43}$$

which can be written in a recursive manner as

$$\theta_n \quad = \quad \frac{A_n}{B_n} = \frac{aA_{n-1} + e_n^2}{bB_{n-1} + e_n^2}. \tag{5.44}$$

The parameter θ_n is of a special form called the *ratio of quadratic forms of random variables.* In our case, this quadratic form consists of a linear combination of correlated Gaussian variables. Finding the expectation of this term is challenging under normal circumstances. But under certain conditions and assumptions, we can either estimate or find the exact moments of this ratio. Therefore, we have divided the analysis into three cases under various assumptions. Simulations will show that all the three cases discussed agree very closely with the simulation results.

Case 1:

In the first case, we make the assumption that

> **A14** The decaying parameters a and b are very close to unity.

With the assumption **A14** at our disposal, taking the expectation of both sides of (5.42) gives us

$$E\left[\mu_{n+1}\right] = \alpha E\left[\mu_n\right] + \gamma E\left[\theta_n\right]. \tag{5.45}$$

As a consequence of using **A14**, we can separate the expectation in the numerator and the denominator of (5.44), i.e.,

$$E\left[\theta_n\right] = E\left[\frac{A_n}{B_n}\right] \approx \frac{E\left[A_n\right]}{E\left[B_n\right]}. \tag{5.46}$$

Although this assumptions is very weak under normal conditions, we still employ it to simplify the expectation of (5.44). Therefore, both the expectations are dealt with separately and we can write

$$\begin{aligned} E\left[A_n\right] &= aE\left[A_{n-1}\right] + E\left[e_n^2\right], \\ &= aE\left[A_{n-1}\right] + E\left[\left(e_{an} + z_n\right)^2\right], \\ E\left[A_n\right] &= aE\left[A_{n-1}\right] + \sigma_x^2 E\left[\|\mathbf{v}_n\|^2\right] + \sigma_z^2. \end{aligned} \tag{5.47}$$

Similarly

$$E\left[B_n\right] = bE\left[B_{n-1}\right] + \sigma_x^2 E\left[\|\mathbf{v}_n\|^2\right] + \sigma_z^2. \tag{5.48}$$

Substituting in (5.46) we get

$$E\left[\theta_n\right] \approx \frac{aE\left[A_{n-1}\right] + \sigma_x^2 E\left[\|\mathbf{v}_n\|^2\right] + \sigma_z^2}{bE\left[B_{n-1}\right] + \sigma_x^2 E\left[\|\mathbf{v}_n\|^2\right] + \sigma_z^2}. \tag{5.49}$$

For the mean-square behaviour, taking the expectation of the squared of (5.42) we get

$$E\left[\mu_{n+1}^2\right] \approx \alpha^2 E\left[\mu_n^2\right] + \gamma^2 E\left[\theta_n^2\right] + 2\alpha\gamma E\left[\mu_n\right] E\left[\theta_n\right]. \tag{5.50}$$

Note that we have used **A3** to write $E\left[\mu_n\theta_n\right] \approx E\left[\mu_n\right] E\left[\theta_n\right]$. If we neglect the term involving γ^2 ($\gamma \ll 1$) we can approximate (5.50) as

$$E\left[\mu_{n+1}^2\right] \approx \alpha^2 E\left[\mu_n^2\right] + 2\alpha\gamma E\left[\mu_n\right] E\left[\theta_n\right]. \tag{5.51}$$

Now the equation (5.45) along with (5.49) and (5.51) completely describe the mean and mean-square behaviour of the time-varying step-size of the proposed algorithm.

Case 2:

In this case, we have used an approximation of the recursion (5.44) to simplify the expectation of the quadratic ratio. We use the idea that the value of the error ,e_n^2, can be neglected as compared to the error cumulant bB_{n-1} because the decaying factor b is very close to unity. Following this argument, we can rewrite (5.44) as

$$\theta_n \approx \frac{a}{b}\theta_{n-1} + \frac{e_n^2}{bB_{n-1}}. \tag{5.52}$$

Taking the expectation of both sides, we get

$$E\left[\theta_n\right] \approx \frac{a}{b}E\left[\theta_{n-1}\right] + \frac{1}{b}E\left[\frac{e_n^2}{B_{n-1}}\right]. \tag{5.53}$$

Noting the fact the e_n^2 and B_{n-1} are uncorrelated, the expectation of the numerator and the denominator of the second term on the right-hand side of the equation can be taken separately . This leads to

$$E\left[\theta_n\right] \approx \frac{a}{b}E\left[\theta_{n-1}\right] + \frac{1}{b}\frac{E\left[e_n^2\right]}{E\left[B_{n-1}\right]}. \tag{5.54}$$

$$E\left[\theta_n\right] \approx \frac{a}{b}E\left[\theta_{n-1}\right] + \frac{1}{b}\frac{\sigma_x^2 E\left[\|\mathbf{v}_n\|^2\right] + \sigma_z^2}{E\left[B_{n-1}\right]}. \tag{5.55}$$

Considering the same arguments for the mean-squared case given for (5.51), the transient behaviour of the step-size is characterised by (5.45), (5.51) and (5.55).

Case 3:

In this case, we have used the exact moments of ratio of quadratic form of random variables to arrive at the expectation of (5.43). Since exact moments are obtained, this method by far is the most accurate in describing the transient behaviour of the time-varying step-size. This method is based on the exact moments of the ratio of quadratic form in normal variables as given in [38]. The general solution and description is provided in Appendix 7.3. Thus, the first and second moment of θ_n can be written as

$$E\left[\theta_n\right] = e^{\left(-(1/2)\mathbf{m}'\mathbf{K}^{-1}\mathbf{m}\right)} \times \int_0^1 |\triangle|\, e^{1/2\zeta'\zeta}\left[Tr\left(\mathbf{G}\right) + \zeta'\mathbf{G}\zeta\right] dt, \tag{5.56}$$

and

$$E\left[\theta_n^2\right] = e^{\left(-(1/2)\mathbf{m}'\mathbf{K}^{-1}\mathbf{m}\right)} \sum_{\mathbf{l}} \delta_2\left(\mathbf{l}\right) \times \int_0^1 |\triangle|\, e^{1/2\zeta'\zeta} \prod_{j=1}^{2}\left[Tr\left(\mathbf{G}^j\right) + \zeta'\mathbf{G}^j\zeta\right]^{r_j} dt. \tag{5.57}$$

There is no close form solution of the integrals in both of the above equations. Therefore, they have to be numerically integrated to arrive at the final result. So the equations (5.45), (5.50), (5.56) and (5.57) completely describe the transient behaviour of the time-varying step-size.

In summary, the variance equations derived for the proposed VSSLMFQ algorithm in (5.27) and (5.36) in conjunction with any of the 3 cases of the transient behaviour of the step-size constitute the transient behaviour of the proposed algorithm.

5.3 Stability Analysis of the VSSLMFQ Algorithm

The stability analysis of an adaptive algorithm provides necessary and sufficient conditions for the algorithm to converge. Hence this is an important design parameter that needs to be analysed properly.

5.3.1 Stability in the Mean Sense

For the proposed VSSLMFQ algorithm, the weight-error vector update equation is given by

$$\mathbf{v}_{n+1} = \mathbf{v}_n - \mu_n \mathbf{x}_n e_n^3. \tag{5.58}$$

where $e_n = \mathbf{x}_n^T \mathbf{v}_n + z_n$. It is shown in [19] and [28] that we can write the (5.58) as

$$\mathbf{v}_{n+1} = \mathbf{v}_n - \mu_n \sum_{i=0}^{3} \begin{pmatrix} 3 \\ i \end{pmatrix} \mathbf{x}_n \left\{ \left[\mathbf{x}_n^T \mathbf{v}_n\right]^i \mathbf{v}_n z_n^{3-i} \right\}. \tag{5.59}$$

Using assumptions **A8**, **A9** and **A10** and taking the expectation, we can write the above equation as

$$E\left[\mathbf{v}_{n+1}\right] = \left[\mathbf{I} - 3\overline{\mu_n} E\left[z_n\right] \mathbf{R}\right] E\left[\mathbf{v}_n\right], \tag{5.60}$$

where $\overline{\mu_n} = E\left[\mu_n\right]$. Changing the coordinates by applying unitary transformation gives us the kth natural mode as

$$v_{n+1}^k = \left[1 - 3\overline{\mu_n}\sigma_z^2 \lambda^k\right] v_n^k, \qquad k = 1, 2, \ldots, N \tag{5.61}$$

In order for the modes to decay to zero, the necessary condition on the step-size is

$$-1 \quad < \quad 1 - 3\overline{\mu_n}\sigma_z^2\lambda_{max} < 1,$$

or

$$0 < \overline{\mu_n} < \frac{2}{3\sigma_z^2\lambda_{max}}, \tag{5.62}$$

where λ_{max} is the largest eigenvalue of the autocorrelation matrix of the input regressor $\mathbf{x}_n$.

5.3.2 Stability in the Mean-Square Sense

For stability in the mean-square sense, we have relied on (5.24) and **A11**. With $\Sigma = \mathbf{I}$, we get

$$E\left[\|\mathbf{v}_{n+1}\|^2\right] = E\left[\|\mathbf{v}_n\|^2\right] + \overline{\mu_n^2} h_U tr\left(\mathbf{R}\right) - 2\overline{\mu_n} h_g E\left[e_{an}^2\right], \tag{5.63}$$

where we have used the fact that $E\left[e_{an}^2\right] = E\left[\|\mathbf{v}_n\|_{\mathbf{R}}^2\right]$. The proposed VSSLMFQ algorithm will be mean-square stable only if

$$\overline{\mu_n^2} h_U tr\left(\mathbf{R}\right) - 2\overline{\mu_n} h_g E\left[e_{an}^2\right] \leq 0. \tag{5.64}$$

The above inequality after substituting the values of h_g and h_U from (5.16) and (5.19) respectively, gives the following condition for the mean-square stability:

$$\frac{\overline{\mu_n^2}}{\overline{\mu_n}} \leq \begin{cases} \frac{2\mathcal{C}_n}{5(\mathcal{C}_n+\sigma_z^2)^3 tr(\mathbf{R})} & \text{for Gaussian Noise,} \\ \frac{6\mathcal{C}_n\left(\mathcal{C}_n+\sigma_z^2\right)}{(15\mathcal{C}_n^3+45\sigma_z^2\mathcal{C}_n^2+15\psi_z^4\mathcal{C}_n+\psi_z^6)tr(\mathbf{R})} & \text{otherwise,} \end{cases} \tag{5.65}$$

where $\mathcal{C}_n \leq E\left[e_{an}^2\right]$ is the Cramer–Rao Bound associated with the problem of estimating the random quantity $\mathbf{x}^T\mathbf{w}^o$ by using $\mathbf{x}^T\mathbf{w}_n$ [39]. The rationale behind using Cramer-Rao Bound for the mean-square stability analysis is thoroughly highlighted in [32].

We can also find the bound on the parameter α, γ, a and b which can guarantee the convergence of the proposed VSSLMFQ algorithm to the steady-state value. Considering for $n \rightarrow \infty$, then we can substitute the values of $\overline{\mu_\infty}$ and $\overline{\mu_\infty^2}$ from (3.25) and (3.28) in (5.65) to get

$$0 \leq \Omega \leq \begin{cases} \frac{2\mathcal{C}_\infty}{5(\mathcal{C}_\infty+\sigma_z^2)^3 tr(\mathbf{R})} & \text{for Gaussian Noise,} \\ \frac{6\mathcal{C}_\infty\left(\mathcal{C}_\infty+\sigma_z^2\right)}{(15\mathcal{C}_\infty^3+45\sigma_z^2\mathcal{C}_\infty^2+15\psi_z^4\mathcal{C}_\infty+\psi_z^6)tr(\mathbf{R})} & \text{otherwise,} \end{cases} \tag{5.66}$$

where we have represented the ratio $\frac{\overline{\mu_\infty^2}}{\overline{\mu_\infty}}$ by $\Omega = \frac{2\alpha\gamma(1-b)}{(1+\alpha)(1-\alpha)(1-a)}$.

5.4 Convergence Time of the VSSLMFQ Algorithm

In this section, using the results of the transient analysis, we have estimated the convergence time of our proposed VSSLMFQ algorithm. For this purpose, we have only taken the case when the input regressor is white as it simplifies the solution. The convergence time of an adaptive filter is defined as *the number of iterations, τ, needed for the mean-square error to reach* $(1+\epsilon)$ *times the steady-state error for a given* $\epsilon > 0$, i.e.,

$$E\left[e_\tau^2\right] = (1+\epsilon)\, E\left[e_\infty^2\right]. \tag{5.67}$$

Then we first need to establish the learning curve recursion for our proposed algorithm. Recalling the variance relation in Section 5.2.1 derived as

$$E\left[\|\mathbf{v}_{n+1}\|^2\right] = \mathbf{f} E\left[\|\mathbf{v}_n\|^2\right] + \overline{\mu_n^2}\sigma_x^2\psi_z^6 N, \tag{5.68}$$

where

$$\begin{aligned}\mathbf{f} \;&=\; 1+\sigma_x^2\left(15\overline{\mu_n^2}\sigma_x^2\psi_z^4 N - 6\overline{\mu_n}\sigma_z^2\right) + \sigma_x^4\left(45\overline{\mu_n^2}\sigma_x^2\sigma_z^2 N - 6\overline{\mu_n}\right) E\left[\|\mathbf{v}_n\|^2\right] \\ &\quad +15\overline{\mu_n^2}\sigma_x^8 N\left(E\left[\|\mathbf{v}_n\|^2\right]\right)^2. \end{aligned}\tag{5.69}$$

Now since $E\left[e_{an}^2\right] = \sigma_x^2 E\left[\|\mathbf{v}_n\|^2\right]$ and $E\left[e_n^2\right] = \sigma_z^2 + E\left[e_{an}^2\right]$ and ignoring the higher powered terms of $E\left[\|\mathbf{v}_n\|^2\right]$ in (5.69), after some mathematical manipulation, we can arrive at the following recursion for the learning curve:

$$\begin{aligned}E\left[e_n^2\right] \;&=\; \left[1+\sigma_x^2\left(15\overline{\mu_n^2}\sigma_x^2\psi_z^4 N - 6\overline{\mu_n}\sigma_z^2\right)\right] E\left[e_{n-1}^2\right] + \overline{\mu_n^2}\sigma_x^2\psi_z^6 N \\ &\quad -\sigma_x^2\sigma_z^2\left(15\overline{\mu_n^2}\sigma_x^2\psi_z^4 N - 6\overline{\mu_n}\sigma_z^2\right). \end{aligned}\tag{5.70}$$

which be expressed compactly as

$$E\left[e_n^2\right] = \rho E\left[e_{n-1}^2\right] + \phi, \tag{5.71}$$

where

$$\begin{aligned} \rho &= \left[1 + \sigma_x^2\left(15\overline{\mu_n^2}\sigma_x^2\psi_z^4 N - 6\overline{\mu_n}\sigma_z^2\right)\right], & (5.72)\\ \phi &= \overline{\mu_n^2}\sigma_x^2\psi_z^6 N - \sigma_x^2\sigma_z^2\left(15\overline{\mu_n^2}\sigma_x^2\psi_z^4 N - 6\overline{\mu_n}\sigma_z^2\right). & (5.73) \end{aligned}$$

Solving (5.71) for steady-state, i.e., as $n \to \infty$, we get

$$\text{MSE} = E\left[e_\infty^2\right] = \frac{\phi_\infty}{1-\rho_\infty}. \tag{5.74}$$

We centre the mean-square error by subtracting the MSE from both sides to get

$$E\left[e_n^2\right] - \frac{\phi_\infty}{1-\rho_\infty} = \rho\left(E\left[e_{n-1}^2\right] - \frac{\phi_\infty}{1-\rho_\infty}\right). \tag{5.75}$$

The solution to this difference equation can be written as

$$E\left[e_n^2\right] - \text{MSE} = \rho^n\left(E\left[e_0^2\right] - \text{MSE}\right). \tag{5.76}$$

This equation can be manipulated to be expressed as the output signal-to-noise ratio (SNR) of the desired filter. Knowing the fact that $E\left[e_{a0}\right] = \sigma_z^2 + \sigma_x^2\left\|\mathbf{w}^o\right\|^2$, we can write the above equation as

$$E\left[e_n^2\right] - \mathrm{MSE} = \rho^n\left(\sigma_z^2\left(1+\mathrm{SNR}\right) - \mathrm{MSE}\right), \tag{5.77}$$

where

$$\mathrm{SNR} = \frac{\sigma_x^2\left\|\mathbf{w}^o\right\|^2}{\sigma_z^2}. \tag{5.78}$$

For $n=\tau$, $E\left[e_\tau^2\right] = (1+\epsilon)\,\mathrm{MSE}$, and solving for τ, we get the expression

$$\tau \ln\rho = \ln\left[\frac{\epsilon\mathrm{MSE}}{\sigma_z^2\left(1+\mathrm{SNR}\right) - \mathrm{MSE}}\right]. \tag{5.79}$$

This result can be rearranged in terms of the misadjustment of the proposed algorithm. Defining misadjustment as $\mathcal{M} = \mathrm{EMSE}/\sigma_z^2$, we can write (5.79) as

$$\tau = \frac{\ln\left[\frac{\epsilon(1+\mathcal{M})}{\mathrm{SNR}-\mathcal{M}}\right]}{\ln\rho}. \tag{5.80}$$

Equation (5.80) shows that the convergence time of the proposed VSSLMFQ depends upon the misadjustment, SNR and the time-varying step-size of the algorithm.

Remark As a corollary to the results achieved in this section, we can arrive at the EMSE and the misadjustment of the proposed algorithm, based solely on the arguments given for the transient analysis. From the time-evolution equation of (5.24) and unity weight matrix, $\Sigma = \mathbf{I}$, we get

$$E\left[\left\|\mathbf{v}_{n+1}\right\|^2\right] = E\left[\left\|\mathbf{v}_n\right\|^2\right] + \overline{\mu_n^2} h_U tr\left(\mathbf{R}\right) - 2\overline{\mu_n} h_g E\left[e_{an}^2\right], \tag{5.81}$$

where we have used the fact that $E\left[e_{an}^2\right] = E\left[\|\mathbf{v}_n\|_{\mathbf{R}}^2\right]$. Replacing (5.19) and (5.16) and ignoring higher order powers of $E\left[e_{an}^2\right]$ in the above equation and furthermore assuming the weight-error vector reaches the steady-state, that is $n \to \infty$, we get

$$\left(6\overline{\mu_\infty}\sigma_z^2 - 15\overline{\mu_\infty^2}\psi_z^4 tr(\mathbf{R})\right) J_{ex} = \overline{\mu_\infty^2}\psi_z^6 tr(\mathbf{R}). \tag{5.82}$$

Evaluating for J_{ex} we get

$$J_{ex} = \frac{\overline{\mu_\infty^2}\psi_z^6 tr(\mathbf{R})}{6\overline{\mu_\infty}\sigma_z^2 - 15\overline{\mu_\infty^2}\psi_z^4 tr(\mathbf{R})}, \tag{5.83}$$

where if we substitute (3.29) in the above equation results in

$$J_{ex} = \frac{\alpha\overline{\mu_\infty}\psi_z^6 tr(\mathbf{R})}{3\left(1+\alpha\right)\sigma_z^2 - 15\alpha\overline{\mu_\infty}\psi_z^4 tr(\mathbf{R})}. \tag{5.84}$$

Comparing equations (3.36) and (5.84) indicates that they are identical. This shows a close interconnect between the steady-state analysis and the transient analysis. In fact the transient analysis is a more general investigation into the performance and behaviour of an adaptive filter. As a consequence, steady-state and tracking analysis become a special case of the transient analysis. The misadjustment defined as $\mathcal{M} = EMSE/\sigma_z^2$ can thus be expressed as

$$\mathcal{M} = \frac{\alpha\overline{\mu_\infty}\psi_z^6 tr(\mathbf{R})}{3\left(1+\alpha\right)\sigma_z^4 - 15\alpha\overline{\mu_\infty}\sigma_z^2\psi_z^4 tr(\mathbf{R})}. \tag{5.85}$$

5.5 Steady-State MSE and MSD

The expressions for the steady-state mean-square error and mean-square deviation can also be derived from the (5.36) as stated earlier. Following the definition given for MSE in (5.40) , we have:

$$\text{MSE} = \lim_{n\to\infty} E\left[\|\mathbf{v}_n\|^2_{\mathbf{R}}\right] + \sigma_z^2,$$

or

$$\text{MSE} \quad = \quad J_{ex} + J_{min}, \tag{5.86}$$

where by using (5.84), it evaluates to

$$\text{MSE} = \frac{\alpha\overline{\mu_\infty}\psi_z^6 tr(\mathbf{R}) + 3\,(1+\alpha)\,\sigma_z^4 - 15\alpha\overline{\mu_\infty}\sigma_z^2\psi_z^4 tr(\mathbf{R})}{3\,(1+\alpha)\,\sigma_z^2 - 15\alpha\overline{\mu_\infty}\psi_z^4 tr(\mathbf{R})}. \tag{5.87}$$

For the mean-square deviation, we choose the weight matrix that is equal to the inverse of the input regressor correlation matrix, that is $\Sigma = \mathbf{R}^{-1}$. Hence (5.24) can be written as

$$E\left[\|\mathbf{v}_{n+1}\|^2_{\mathbf{R}^{-1}}\right] = E\left[\|\mathbf{v}_n\|^2_{\mathbf{R}^{-1}}\right] + \overline{\mu_n^2} h_U tr\left(\mathbf{I}\right) - 2\overline{\mu_n} h_g E\left[\|\mathbf{v}_n\|^2\right]. \tag{5.88}$$

Following the definition of MSD as in (5.41), assuming $n \to \infty$, we get

$$\text{MSD} = \frac{\overline{\mu_\infty^2} h_U N}{2\overline{\mu_\infty} h_g}. \tag{5.89}$$

Substituting the values of h_U and h_g (ignoring third and higher order powers of $E\left[e_{an}^2\right]$) and making use of (3.29), we get

$$\text{MSD} \approx \frac{\alpha\overline{\mu_\infty}\left(15 J_{ex}\psi_z^4 + \psi_z^6\right) N}{3\left(1+\alpha\right)\text{MSE}}. \tag{5.90}$$

Chapter 6

Performance Analysis of the VSSLMFQ Algorithm

In this chapter, we present the result of the computer simulations carried out to investigate the performance characteristics of the proposed VSSLMFQ algorithm. The proposed algorithm is compared with the traditional LMF algorithm under a system identification problem. The simulations; carried out to corroborate the theoretical analysis; have been found to exhibit improved performance for the proposed algorithm over the traditional one.

The major performance criteria chosen for performance comparison is the residual error or the mean-square error that remains in the steady-state. For this, the performance analysis is categorised in three sections. The first pertains to the steady-state mean-square analysis of the proposed algorithm in stationary environments. The second pertains to the tracking performance of the proposed algorithm for non-stationary environments. The final sections pertains to the transient performance of the proposed algorithm.

In order to have a fair comparison between both the algorithms, the parameters governing their behaviour are set such that they have either the same misadjustment or the same convergence rate. Same misadjustment results in the comparison of their convergence rate, that is, how fast the algorithm attain the prescribed misadjustment. Having the same convergence rate results in the comparison of their EMSE (excess mean-square error). In our case, we have used the later and have selected the parameters so as to achieve the same convergence rate for both the algorithms.

6.1 Steady-State Performance Analysis : VSSLMFQ Algorithm

The general layout of the computer simulations runs around the FIR system identification problem that can also be seen as a problem of channel estimation. For the convergence rates to be the same would mean to have the same step-sizes for both the algorithms. As the proposed algorithm has a time-varying step-size, its initial condition μ_0 has been set to be the same as that of the traditional LMF, i.e.,

$$\mu_0^{VSSLMFQ} = \mu^{LMF}.$$

The constant parameters α and γ are set according to the same manner as described in [13]. Furthermore, the FIR channel response to be estimated is a normalised Hanning window (i.e., $\mathbf{w}^{oT}\mathbf{w}^o = 1$) where the length of the adaptive filter is equal to the unknown system. The input sequence $\{x_n\}$ is taken to be BPSK $\{\pm 1\}$ signal. The performance is analysed in three noise environments namely Gaussian, Laplacian and Uniform. The experiments are conducted for SNR ranges of 0 dB, 10 dB and 20 dB. The results obtained are averaged over 500 independent runs.

In order to motivate the idea of using LMF algorithm for non-Gaussian noise environments, we have presented a result where it is seen that the LMF algorithm performance better than the LMS algorithm in terms of the steady-state EMSE. Figure (6.1) gives the EMSE performance of both the algorithms in Uniform noise environment. This result also corroborates the findings in [19] where a similar result was shown.

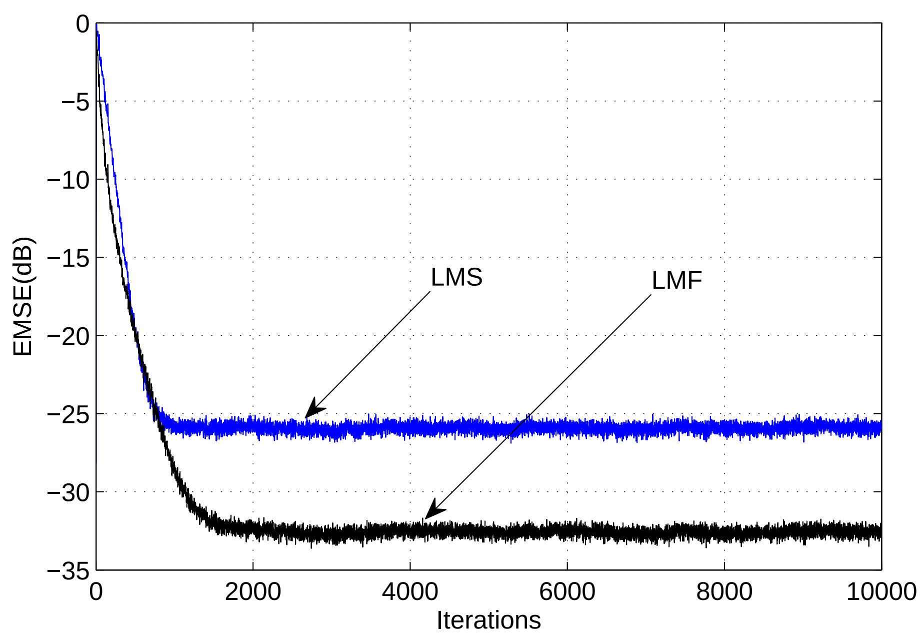

Figure 6.1: Comparison of the EMSE of LMF algorithm and LMS algorithm in Uniform Environment with SNR=10dB.

The main results presented in this section can be divided into 2 parts. The first one deals with the mean behaviour of the time-varying step-size and the proposed algorithm's performance dependence on the design parameters a and b in terms of the mean-square error. The second part deals with the EMSE performance of the proposed VSSLMFQ algorithm against the traditional LMF algorithm. The main results presented in this section are summarised as follows:

- Part 1 - Performance of the time-varying step-size of the proposed VSSLMFQ algorithm in terms of the:

    1. Mean step-size behaviour in the presence of Gaussian noise for different values of parameter a and b.
    2. Mean step-size behaviour in the presence of Laplacian noise for different values of parameter a and b.
    3. Mean step-size behaviour in the presence of Uniform noise for different values of parameter a and b.
    4. Mean step-size behaviour in the presence of Gaussian noise for SNR -10 dB, 0 dB, 10 dB and 20 dB.
    5. Mean step-size behaviour in the presence of Laplacian noise for SNR -10 dB, 0 dB, 10 dB and 20 dB.
    6. Mean step-size behaviour in the presence of Uniform noise for SNR -10 dB, 0 dB, 10 dB and 20 dB.
    7. Analytical and experimental mean-square error for different values of parameters a and b.

8. EMSE for different values of parameters a and b in the presence of Gaussian, Laplacian and Uniform noise environments, respectively.

– Part 2 - Comparison of the EMSE of LMF algorithm with the proposed VSSLMFQ algorithm.

1. In the presence of Gaussian noise for SNR 0 dB, 10 dB and 20 dB.
2. In the presence of Laplacian noise for SNR 0 dB, 10 dB and 20 dB.
3. In the presence of Uniform noise for SNR 0 dB, 10 dB and 20 dB.
4. In the presence of all the three noise environments namely Gaussian, Laplacian and Uniform for SNR 0 dB, 10 dB and 20 dB.

6.1.1 Performance of the time-varying step-size of the proposed VSSLMFQ algorithm

In this section, we have investigated the performance of the time-varying step-size of the proposed VSSLMFQ algorithm, its properties and impact on the MSE (mean-square error). The experiments conducted mainly test the impact of the design parameters a and b on the performance of the algorithm. It is observed that the performance of the proposed algorithm is highly dependent on these parameters. Particularly Figures 6.2, 6.3 and 6.4 illustrate the mean behaviour of the step-size in all the noise environments for different values of parameters a and b with SNR 10 dB. First thing to be observed is the smooth transition of the variable step-size as predicted in the analytical results. Second, it is seen that the convergence rate of the variable step-size is independent of the choices of these parameters. But the steady-state MSE is determined by the parameters a and b. A larger value of a and smaller value of b will result in poor MSE while a smaller a and larger b will result in lower MSE. This is consistent with the analytical results of the algorithm. This observation is further illustrated in Figures 6.8 and 6.9 where both the experimental and theoretical MSE are evaluated for different values of a and b. Both the experimental and analytical results are in close agreement. It would be appropriate to mention here that the theoretical values of the MSE are evaluated using the second entry in the state-space vector given by (5.36) and subsequently (5.40).

The EMSE for the different values of a and b are provided in Figures 6.10, 6.11 and 6.12 for all the three noise environments to provide a clear picture of the level of EMSE attained for these parameter choices.

The results shown in Figures 6.5, 6.6 and 6.7 demonstrate that the steady-state performance of the algorithm is insensitive to the choices of noise levels and environments. This explains the proposed algorithm's resilience to noise and hence lower EMSE as compared to the traditional LMF algorithm.

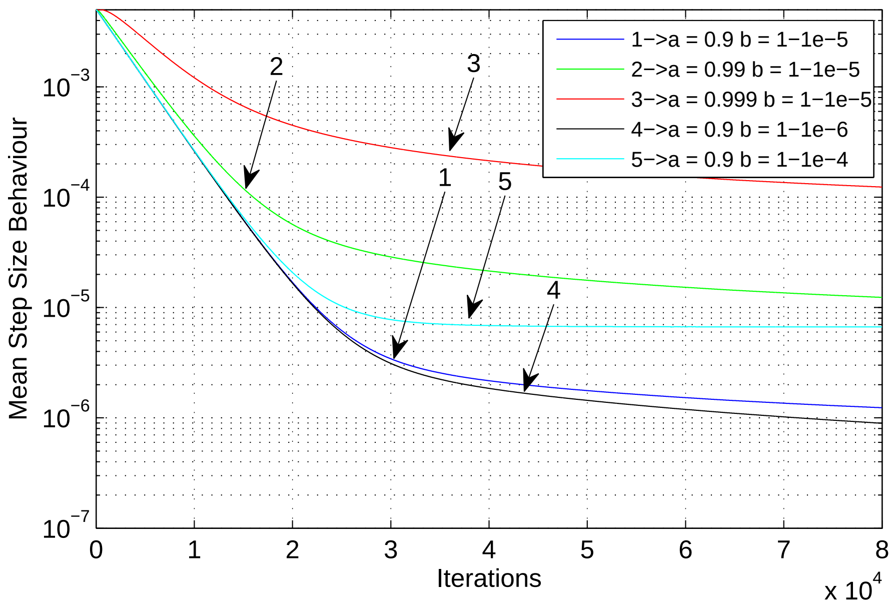

Figure 6.2: Mean behaviour of the variable step-size of VSSLMFQ for various values of parameter a and b in the presence of Gaussian noise with SNR $=$ 10dB.

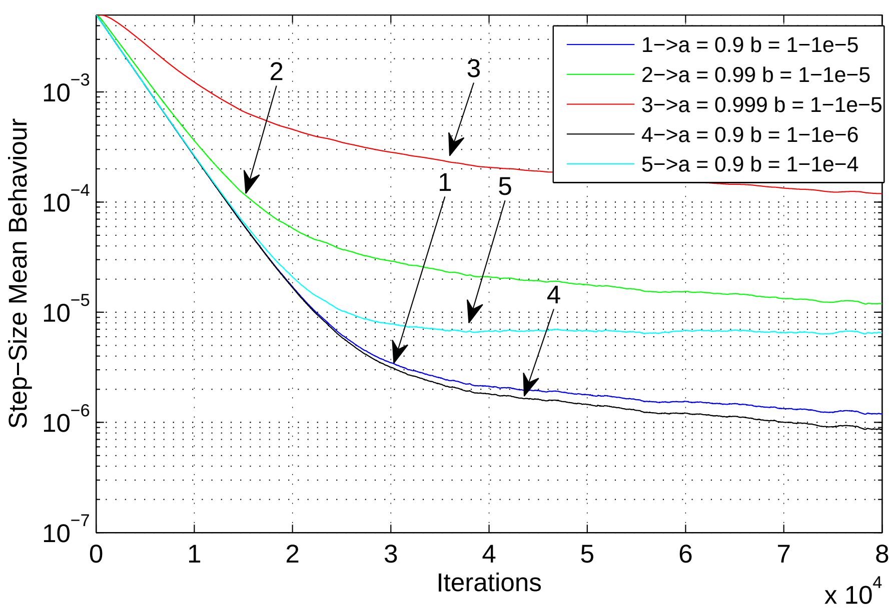

Figure 6.3: Mean behaviour of the variable step-size of VSSLMFQ for various values of parameter a and b in the presence of Laplacian noise with SNR $=$ 10dB.

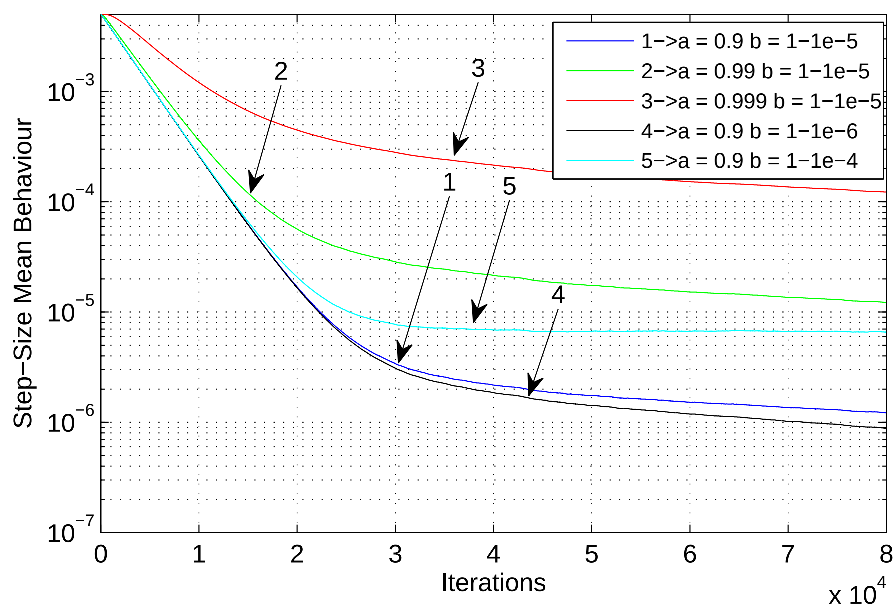

Figure 6.4: Mean behaviour of the variable step-size of VSSLMFQ for various values of parameter a and b in the presence of Uniform noise with SNR $=$ 10dB.

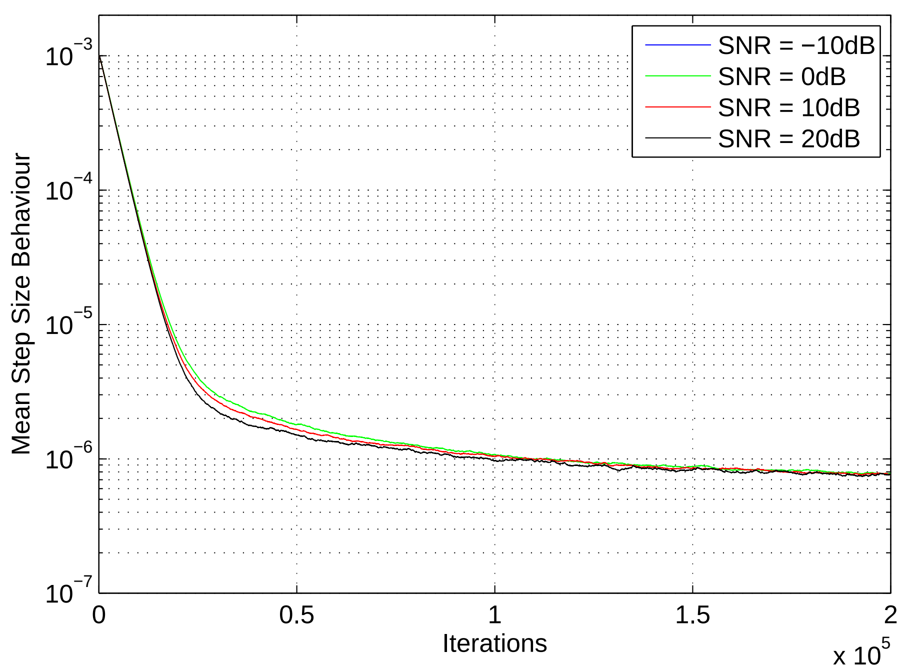

Figure 6.5: Mean behaviour of the variable step-size of VSSLMFQ in the presence of Gaussian environment with various values of SNR.

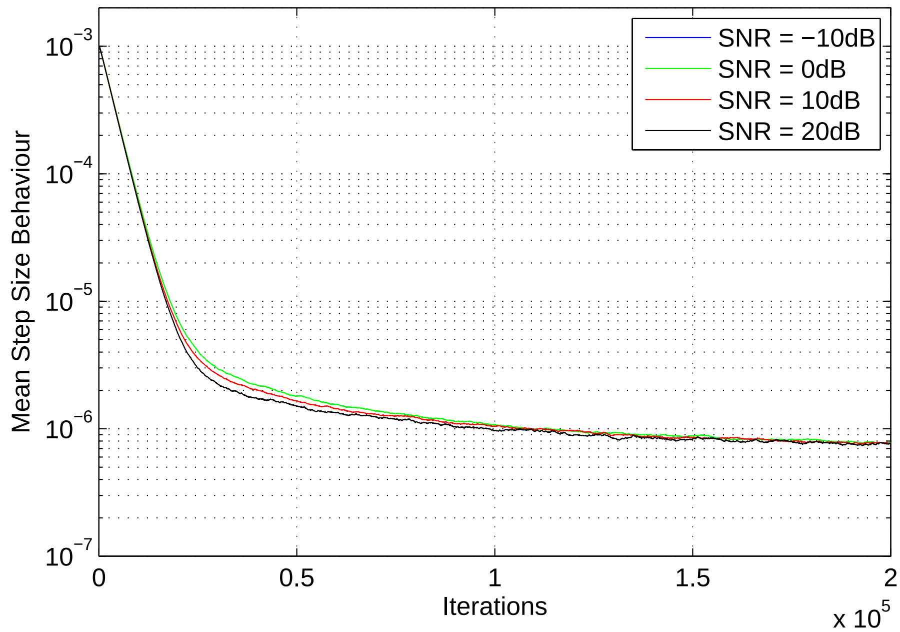

Figure 6.6: Mean behaviour of the variable step-size of VSSLMFQ in the presence of Laplacian environment with various values of SNR.

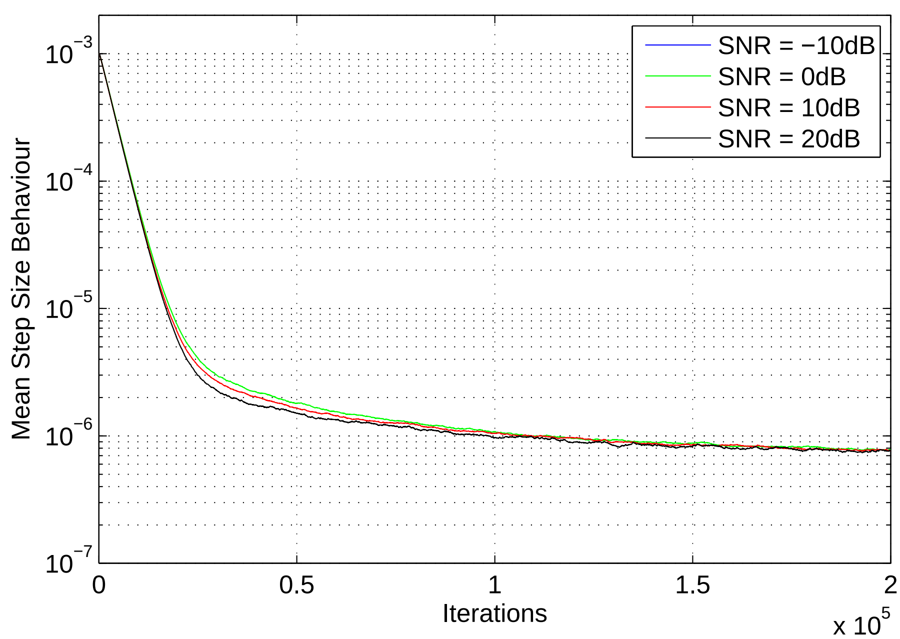

Figure 6.7: Mean behaviour of the variable step-size of VSSLMFQ in the presence of Uniform environment with various values of SNR.

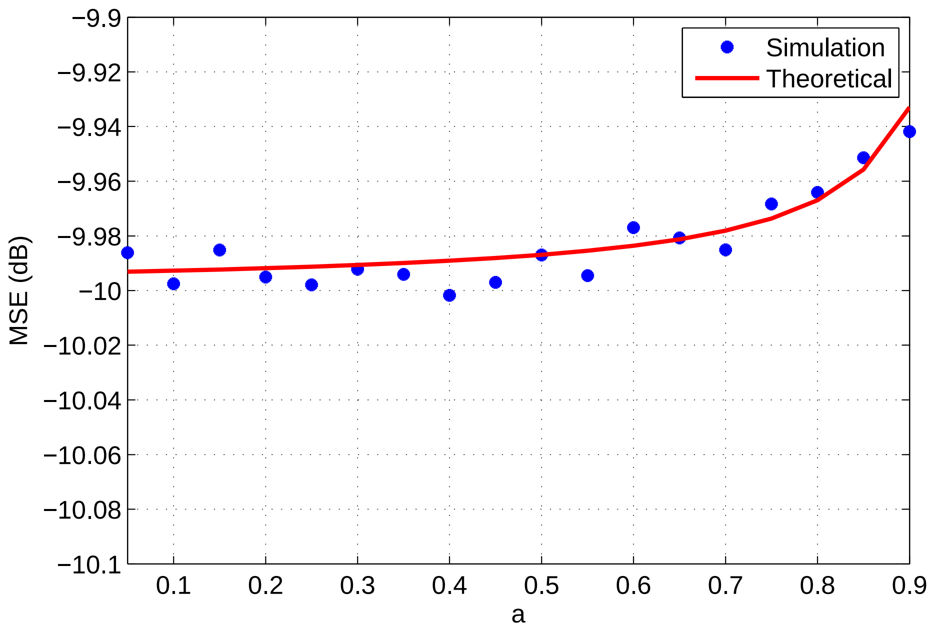

Figure 6.8: Comparison of Analytical and Experimental MSE of the proposed VSSLMFQ algorithm for different values of parameter a.

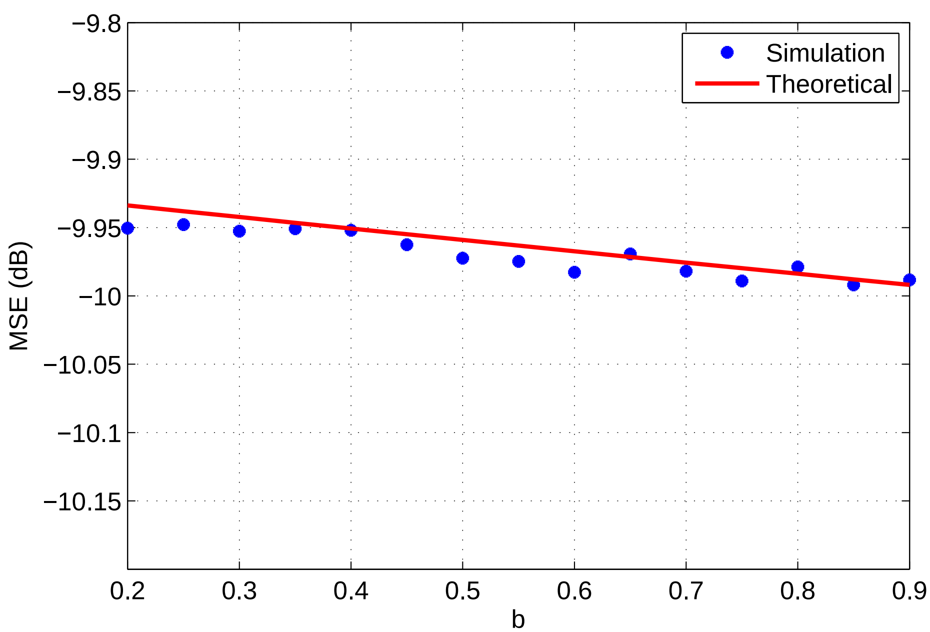

Figure 6.9: Comparison of Analytical and Experimental MSE of the proposed VSSLMFQ algorithm for different values of parameter b.

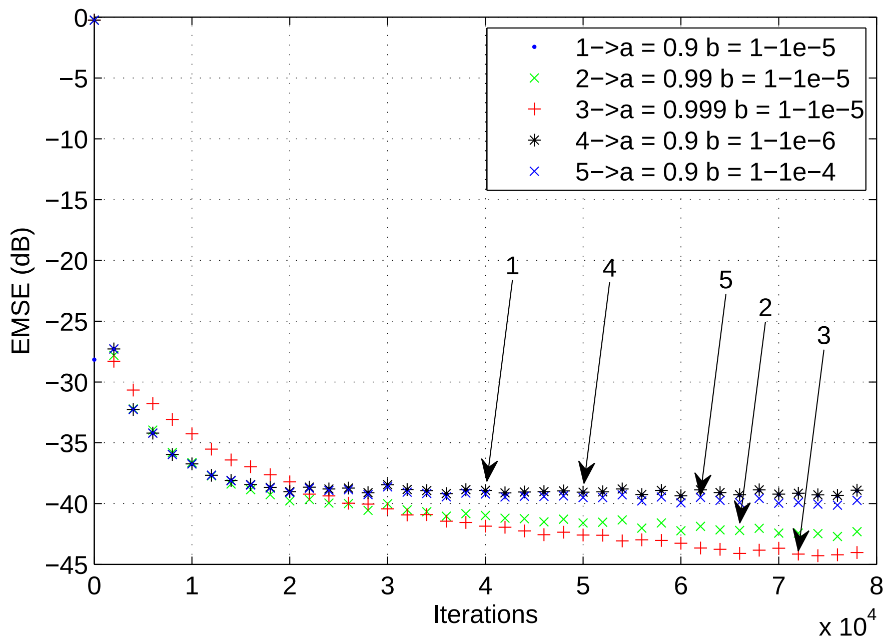

Figure 6.10: Comparison of the EMSE of the proposed VSSLMFQ algorithm for different values of parameter a and b in Gaussian environment with SNR $= 10$ dB.

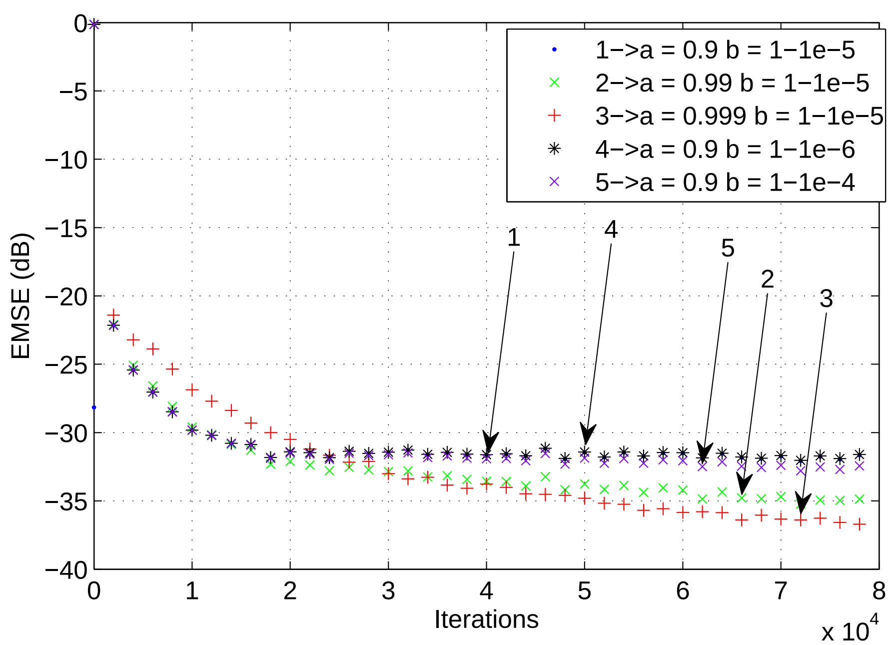

Figure 6.11: Comparison of the EMSE of the proposed VSSLMFQ algorithm for different values of parameter a and b in Laplacian environment with SNR $= 10$ dB.

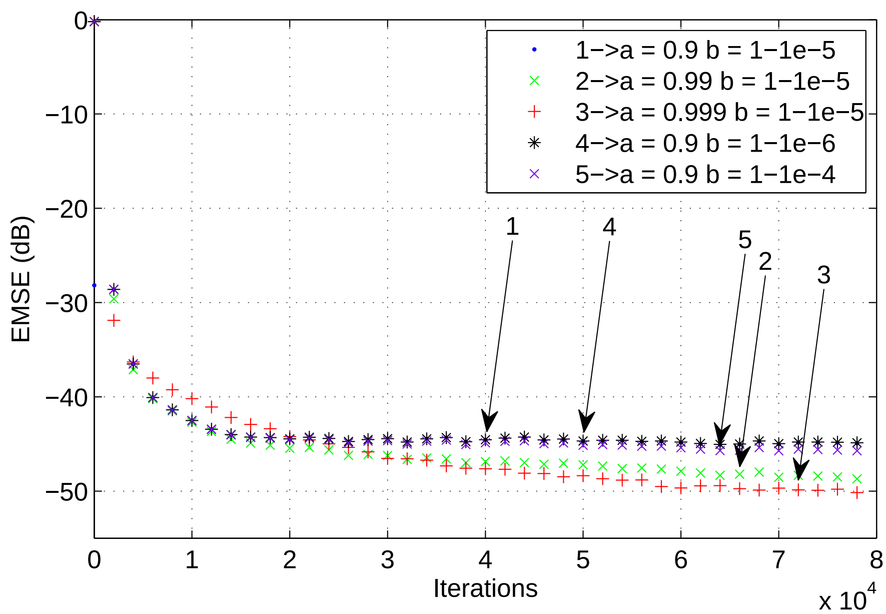

Figure 6.12: Comparison of the EMSE of the proposed VSSLMFQ algorithm for different values of parameter a and b in Uniform environment with SNR $= 10$ dB.

6.1.2 Comparison of the EMSE of LMF algorithm with the proposed VSSLMFQ algorithm

In this section, the LMF algorithm and the proposed VSSLMFQ algorithm are compared in terms of their excess mean-square error. In all the results provided here, we see that the proposed VSSLMFQ algorithm achieves a lower EMSE than the traditional LMF algorithm. Particularly, it can be seen in Figures 6.13, 6.14 and 6.15 that the proposed VSSLMFQ algorithm achieves an EMSE that is lower by 15 dB, 10 dB and 10 dB for Gaussian noise environments with SNR of 0 dB, 10 dB and 20 dB, respectively, as compared to the traditional LMF algorithm.

Similar behaviour is observed for the non-Gaussian environments where Laplacian and Uniform noises are used. The difference in the EMSE between LMF and the proposed VSSLMFQ for the Laplacian case as demonstrated in Figures 6.16, 6.17 and 6.18 is 15 dB, 10 dB and 10 dB with SNR of 0 dB, 10 dB and 20 dB, respectively. Figures 6.19, 6.20 and 6.21 illustrate the same performance enhancement for Uniform noise. Table 6.1 provides a summary of the results obtained in this section. In order to illustrate the respective algorithm's performance in all the noise environments i.e Gaussian, Laplacian and Uniform, Figures 6.22, 6.24 and 6.26 provide the EMSE for the LMF algorithm while Figures 6.23, 6.25 and 6.27 provide the EMSE for the proposed VSSLMFQ algorithm.

The point worth noting is the proposed algorithm's resilience exhibited in all the noise environments irrespective of the SNR level. This was predicted in the arguments presented while proposing the time-varying behaviour of the step-size. This has been clearly illustrated in the simulation results in the previous section.

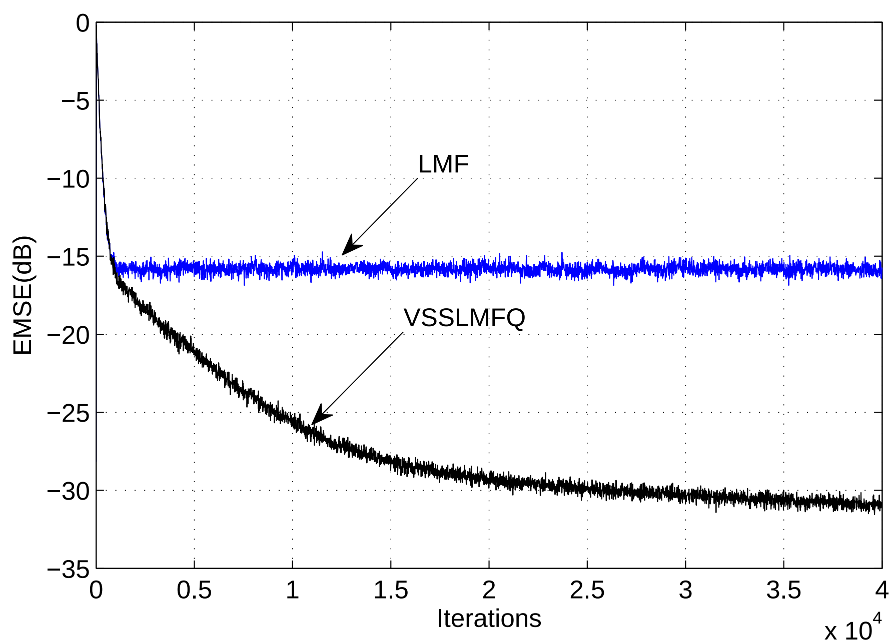

Figure 6.13: Comparison of the EMSE of LMF and the proposed VSSLMFQ in AWGN environment with SNR = 0 dB.

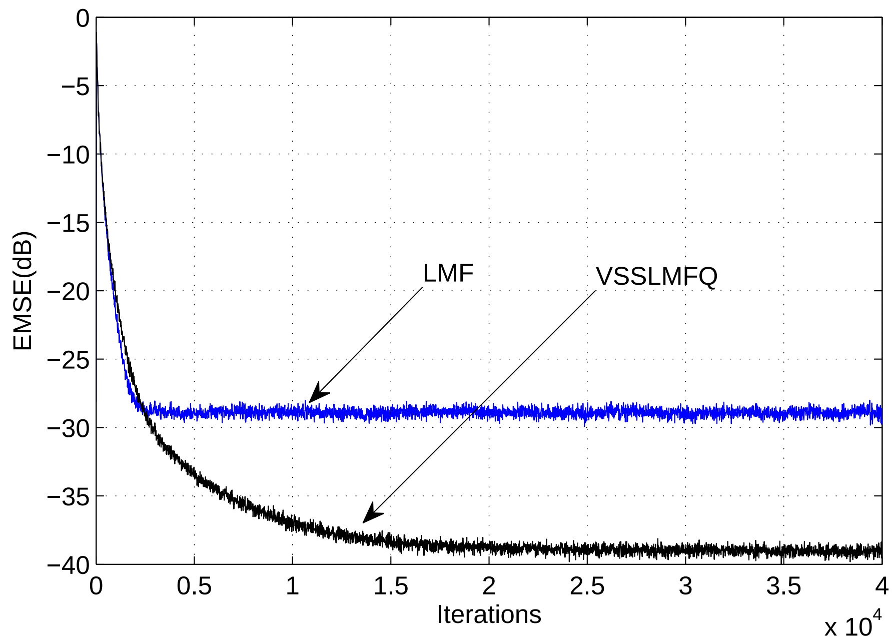

Figure 6.14: Comparison of the EMSE of LMF and the proposed VSSLMFQ in AWGN environment with SNR = 10 dB.

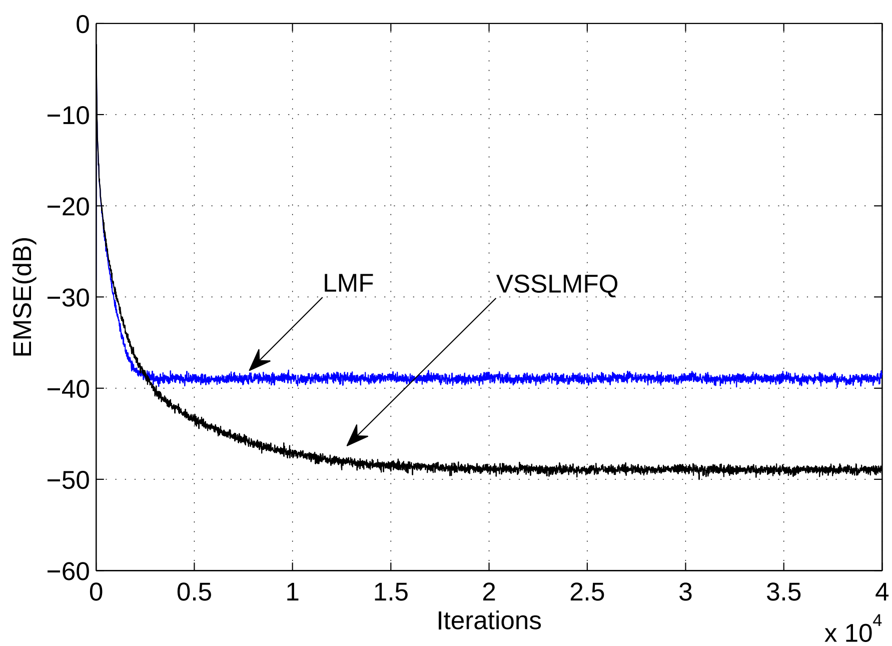

Figure 6.15: Comparison of the EMSE of LMF and the proposed VSSLMFQ in AWGN environment with SNR = 20 dB.

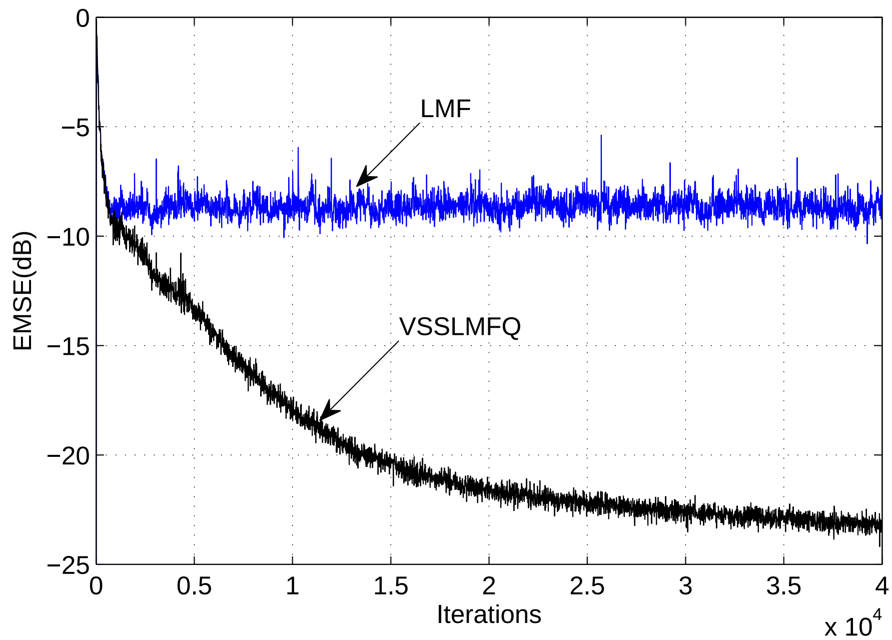

Figure 6.16: Comparison of the EMSE of LMF and the proposed VSSLMFQ in Laplacian environment with SNR = 0 dB.

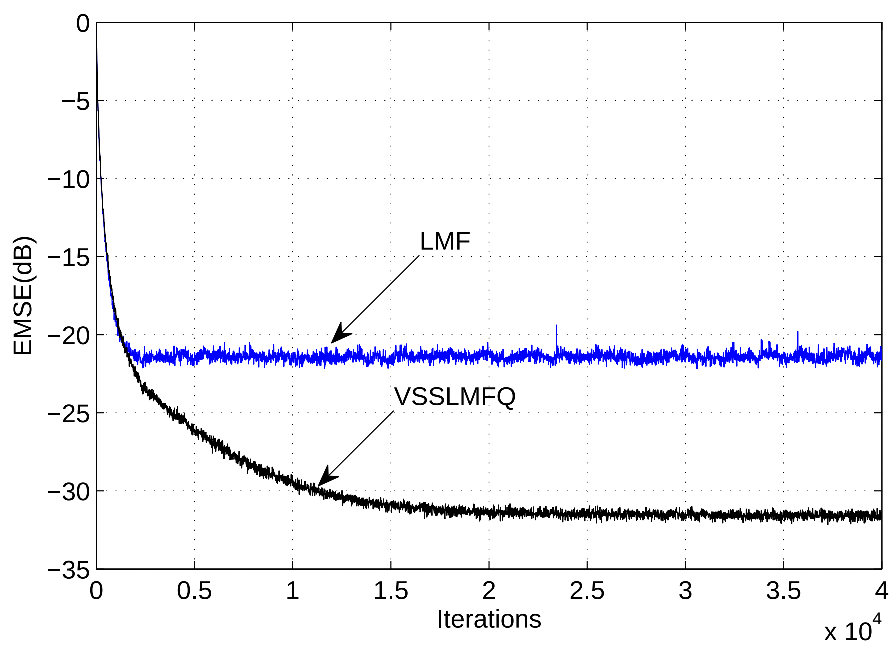

Figure 6.17: Comparison of the EMSE of LMF and the proposed VSSLMFQ in Laplacian environment with SNR = 10 dB.

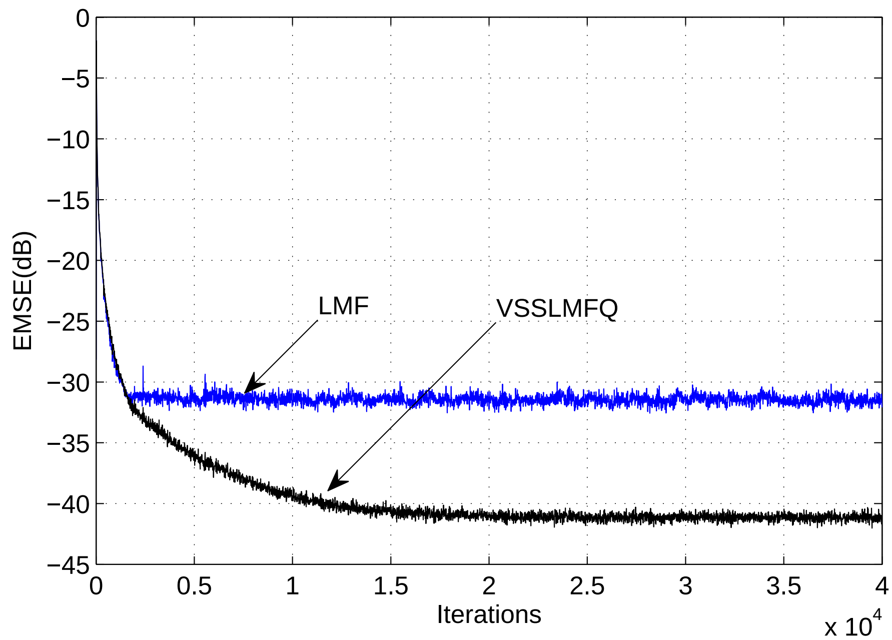

Figure 6.18: Comparison of the EMSE of LMF and the proposed VSSLMFQ in Laplacian environment with SNR = 20 dB.

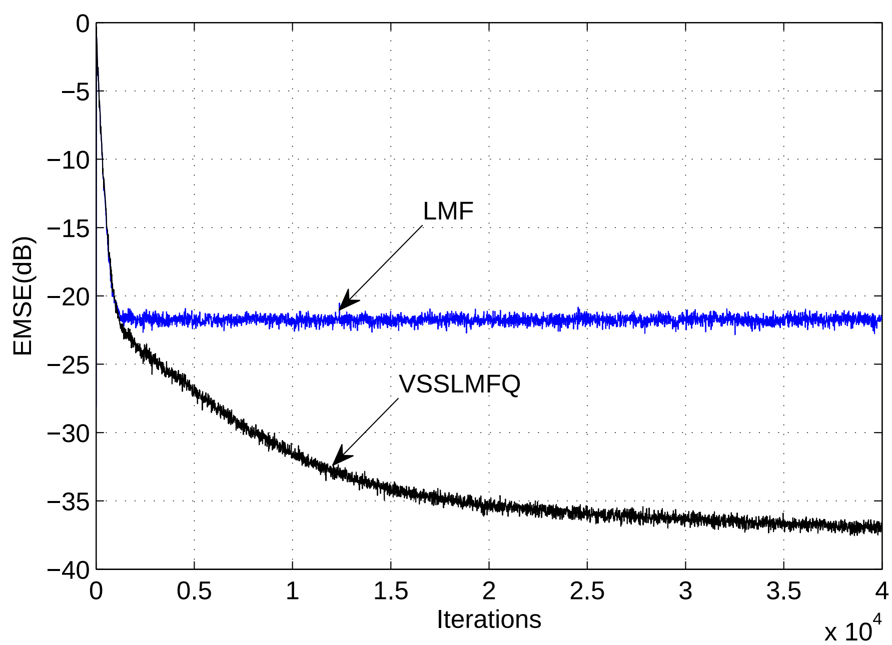

Figure 6.19: Comparison of the EMSE of LMF and the proposed VSSLMFQ in Uniform environment with SNR = 0 dB.

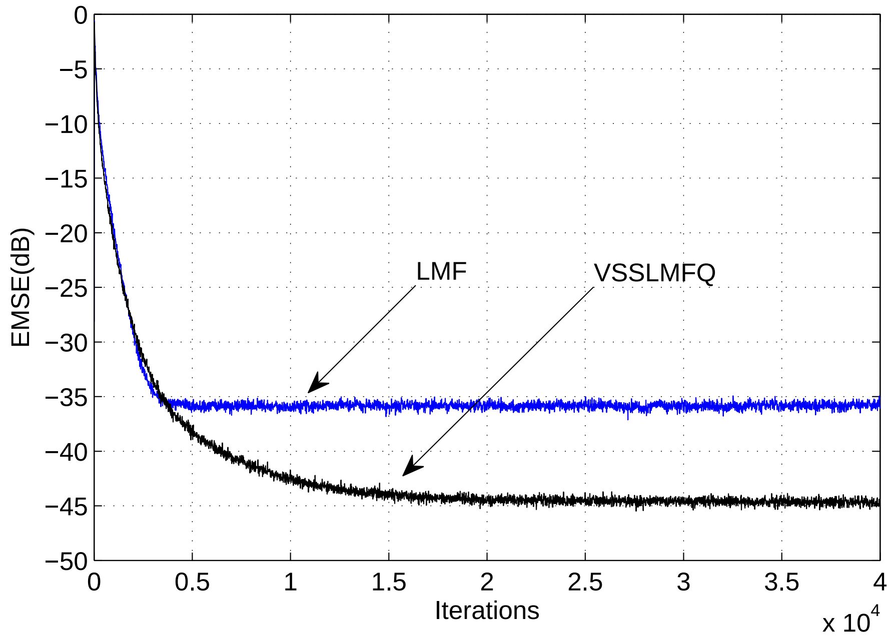

Figure 6.20: Comparison of the EMSE of LMF and the proposed VSSLMFQ in Uniform environment with SNR = 10 dB.

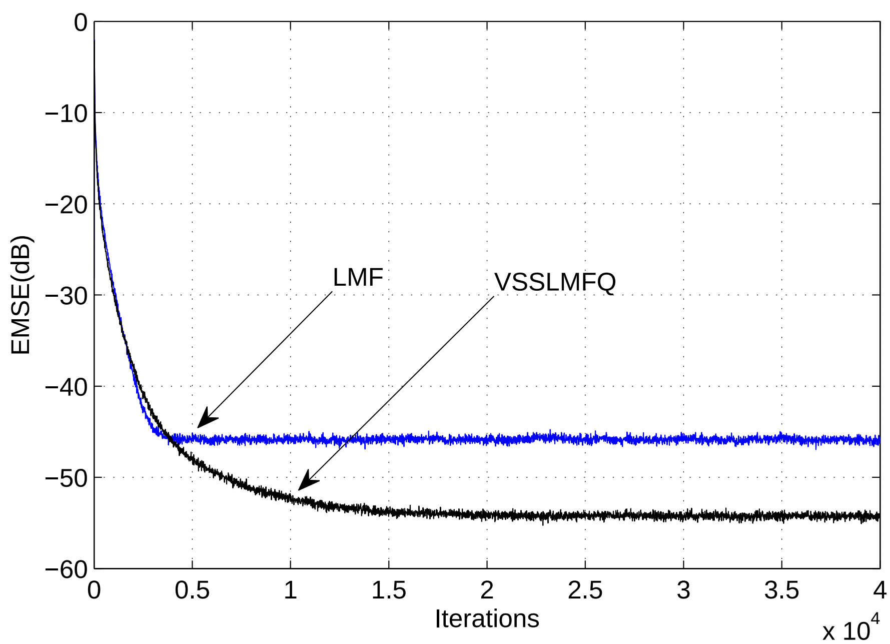

Figure 6.21: Comparison of the EMSE of LMF and the proposed VSSLMFQ in Uniform environment with SNR = 20 dB.

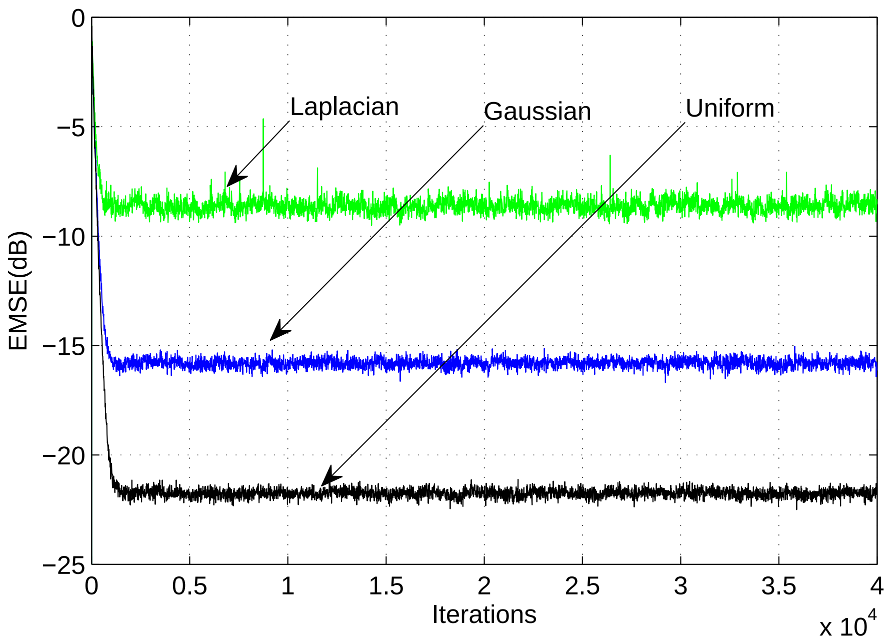

Figure 6.22: EMSE of the LMF algorithm in the presence of Gaussian, Laplacian and Uniform environment with SNR = 0 dB.

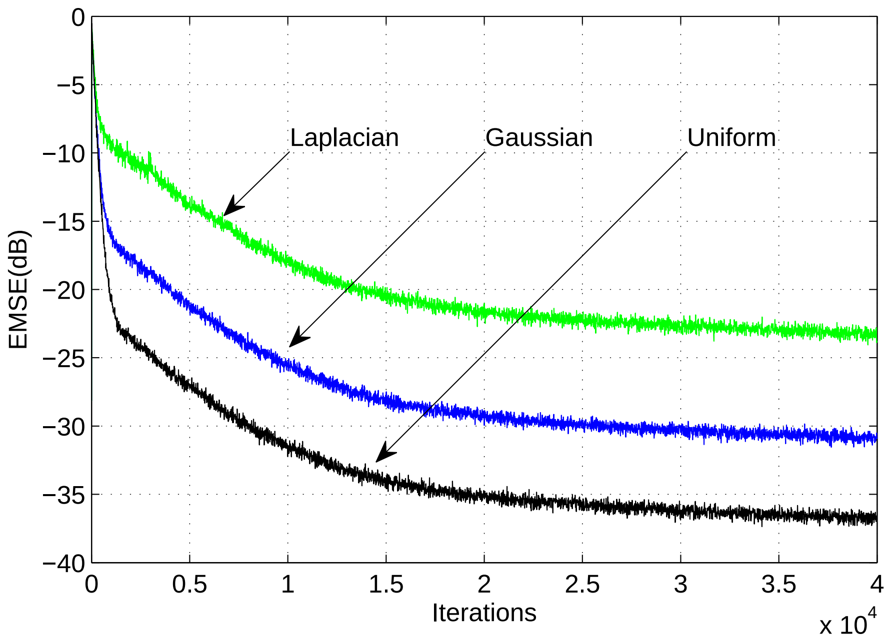

Figure 6.23: EMSE of the proposed VSSLMFQ algorithm in the presence of Gaussian, Laplacian and Uniform environment with SNR = 0 dB.

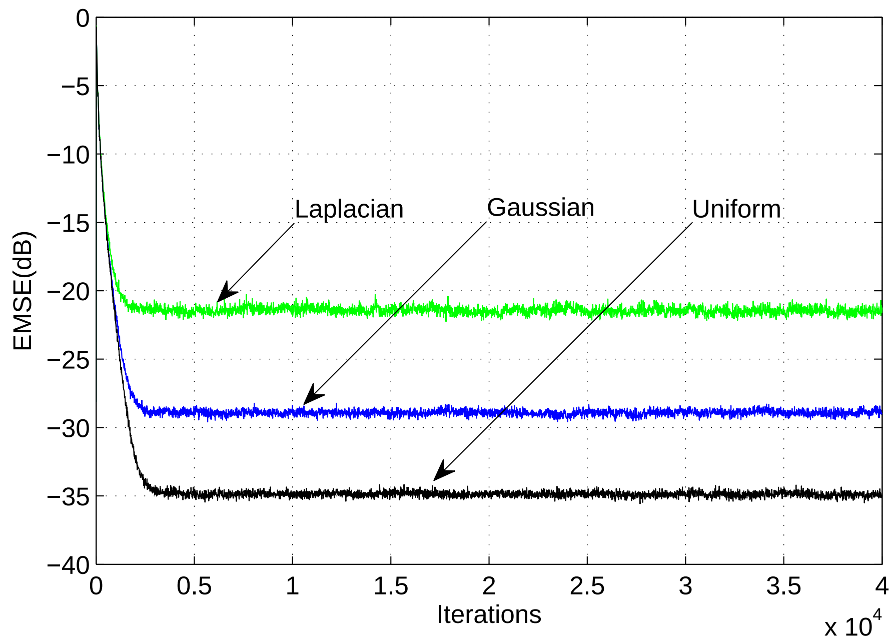

Figure 6.24: EMSE of the LMF algorithm in the presence of Gaussian, Laplacian and Uniform environment with SNR = 10 dB.

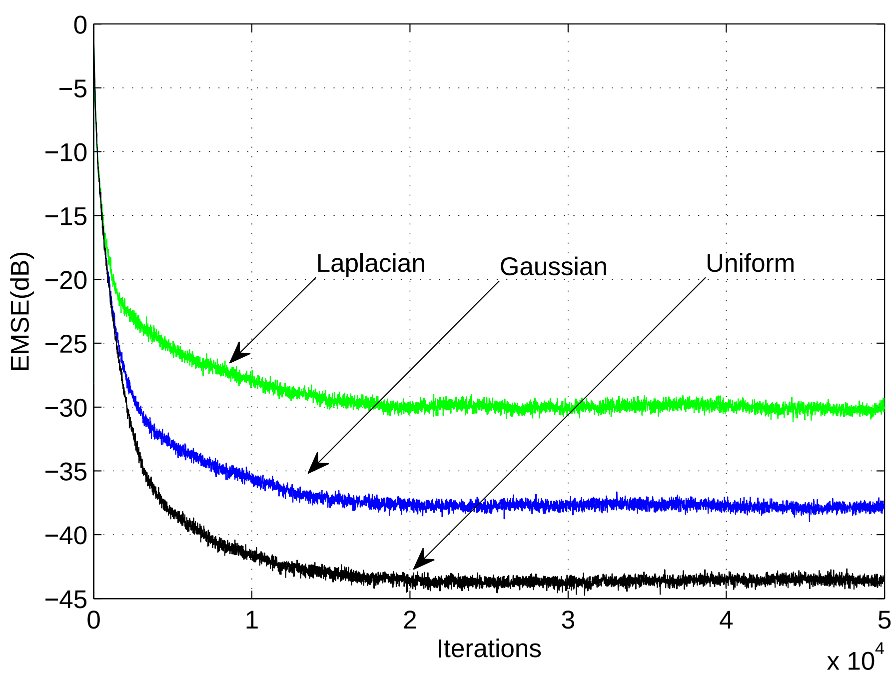

Figure 6.25: EMSE of the proposed VSSLMFQ algorithm in the presence of Gaussian, Laplacian and Uniform environment with SNR = 10 dB.

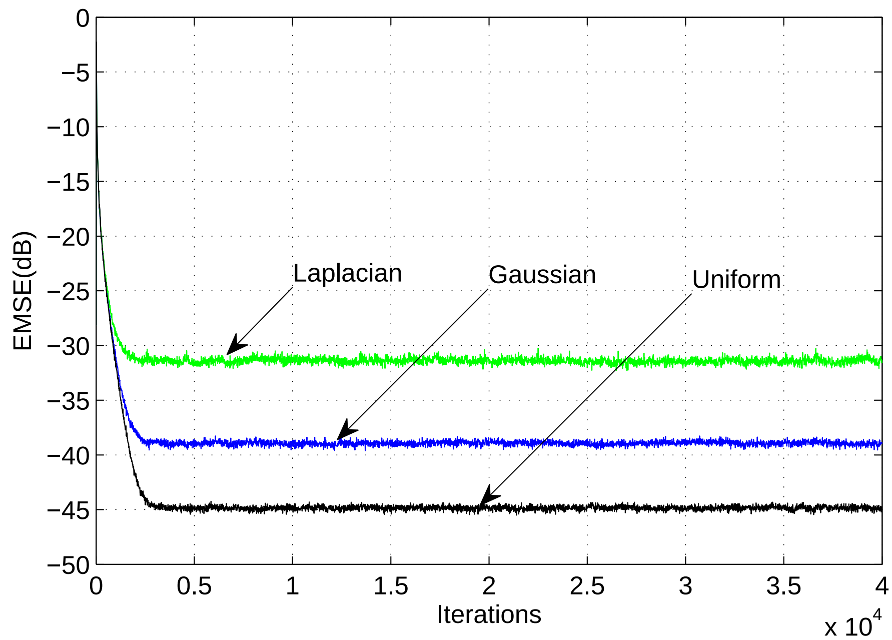

Figure 6.26: EMSE of the LMF algorithm in the presence of Gaussian, Laplacian and Uniform environment with SNR = 20 dB.

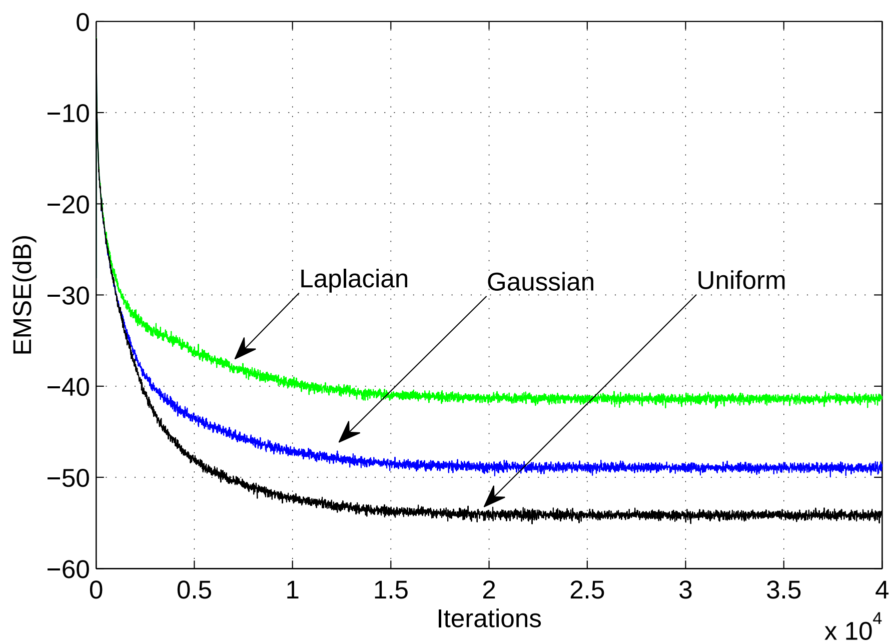

Figure 6.27: EMSE of the proposed VSSLMFQ algorithm in the presence of Gaussian, Laplacian and Uniform environment with SNR = 20 dB.

Table 6.1: Steady-State EMSE of the LMF algorithm and the proposed VSSLMFQ algorithm

	Gaussian			Laplacian			Uniform		
	0dB	10dB	20dB	0dB	10dB	20dB	0dB	10dB	20dB
LMF	-15.99	-29	-38.8	-8.82	-21.31	-31.59	-21.68	-35.78	-45.88
VSSLMFQ	-31	-39.14	-49	-23.3	-31.64	-41.2	-37.03	-44.7	-54.33

6.2 Tracking Performance of the VSSLMFQ Algorithm

In this section, we demonstrate the tracking capabilities of the proposed VSSLMFQ algorithm for time-varying channel. The two channel model used are Random-Walk Channel and Rayleigh Fading Channel (both single and multipath). The input regressor $\{x_n\}$ is taken to be BPSK $\{\pm 1\}$ signal. A Gaussian noise environment is used with variance to achieve an SNR of 10 dB.

6.2.1 Random-Walk Model

The Random-Walk model as discussed in Section 4.2 provides the time-varying channel as:

$$\mathbf{w}^o_{n+1} = \mathbf{w}^o_n + \mathbf{q}_n, \tag{6.1}$$

where vector $\mathbf{q}_n$ is an iid zero mean Gaussian sequence with variance $\sigma_q^2 = 10^{-10}$. This corresponds to a *degree of nonstationarity* (DN) of 3.16×10^{-4}. Thus it satisfies the condition that $DN \ll 1$ for the adaptive filter to track the variations successfully. The parameters of the proposed VSSLMFQ algorithms are set as in the mean-square analysis. The initial condition of the channel is also taken to be the same as in the mean-square analysis, i.e., normalised Hanning window. The results given in Figures 6.28 and 6.29 are the experimental and analytical MSE of the proposed VSSLMFQ algorithm for different values of parameters a and b. It can be observed that the experimental and analytical findings are in close agreement. Also the dependence of the MSE on a and b, as observed in the stationary case is also exhibited here.

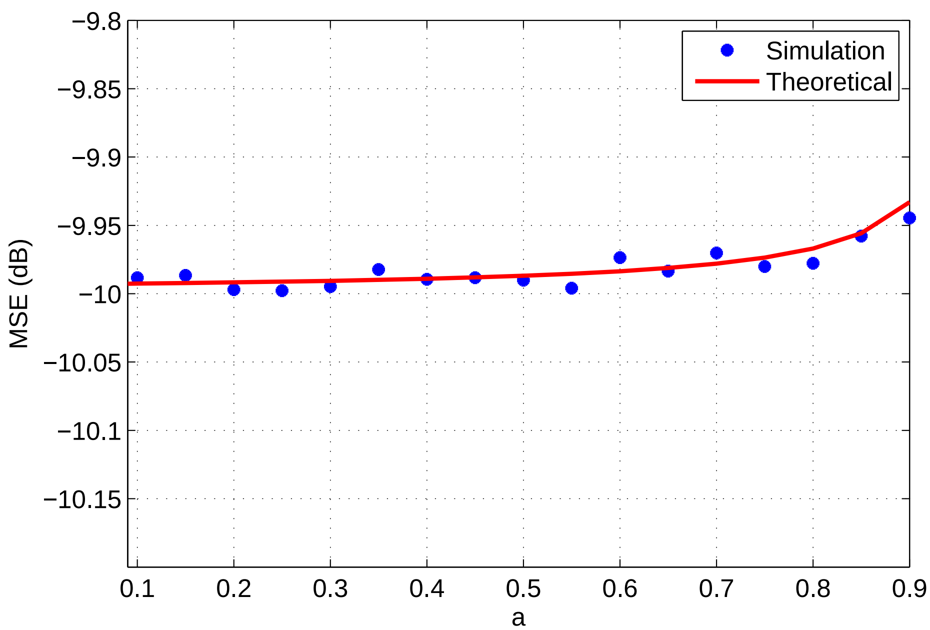

Figure 6.28: Comparison of Experimental and Analytical MSE of VSSLMFQ algorithm for Random-Walk Channel for different values of parameter a.

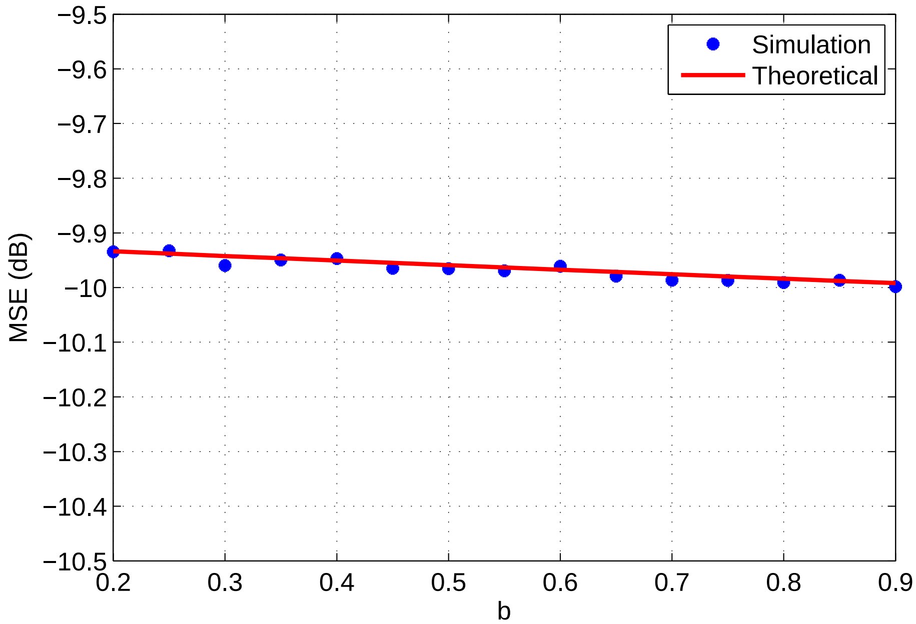

Figure 6.29: Comparison of Experimental and Analytical MSE of VSSLMFQ algorithm for Random-Walk Channel for different values of parameter b.

6.2.2 Rayleigh Fading Channel Model

In the case of Rayleigh fading channel model, the two types of channels considered are single-path and multipath channels. Both channels to be estimated are give as:

$$[0 \quad 0 \quad s_{1n} \quad 0 \quad 0], \tag{6.2}$$

$$[0 \quad 0 \quad s_{1n} \quad 0 \quad s_{2n}], \tag{6.3}$$

where $\{s_{1n}\}$ and $\{s_{2n}\}$ are the absolute values of a Rayleigh distributed sequence that represent Rayleigh fading channel. The Doppler frequency was chosen to be $10Hz$ with a channel sampling period of $T_s = 0.8\mu s$. This corresponds to a *degree of non-stationarity* (DN) of 1.1×10^{-3}. Thus also satisfies the condition that $DN \ll 1$ for the adaptive filter to track the fading channel successfully. The experimental and analytical MSE of the proposed VSSLMFQ algorithm for the single-path fading channel for different values of parameters a and b is shown in Figures 6.30 and 6.31. It is observed that both the experimental and analytical results are in very close agreement. Figures 6.32 and 6.33 provide the experimental and analytical MSE of the proposed VSSLMFQ algorithm for a multipath fading channel for the same parameters. Again the close agreement between theory and simulation is observed. Although not visible but the dependency of the MSE on parameters a and b, seen in the Random-Walk case, is still valid here. This behaviour would be evident for either very higher value of a and lower value of b.

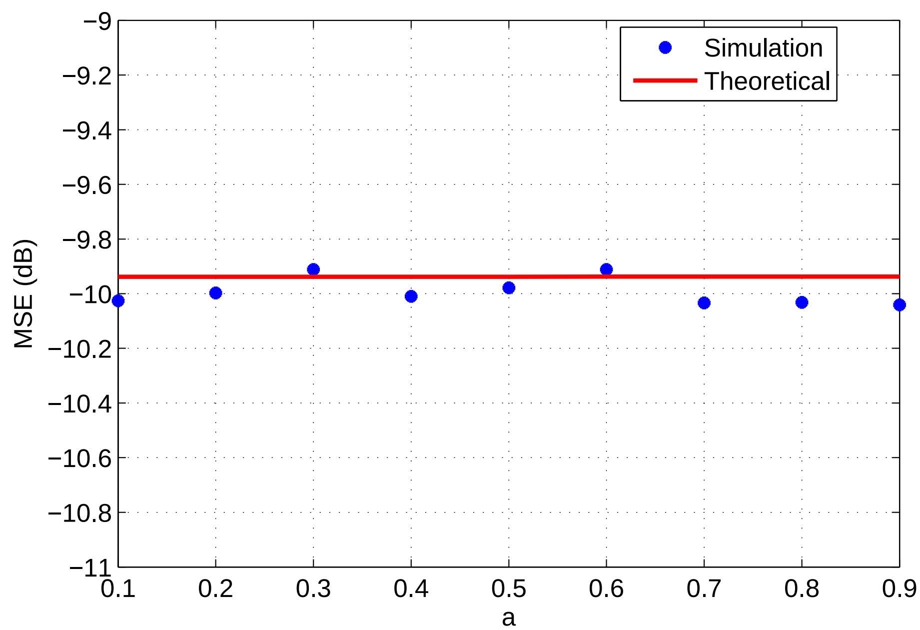

Figure 6.30: Comparison of Experimental and Analytical MSE of VSSLMFQ algorithm for single-path Rayleigh Fading Channel for different values of parameter a.

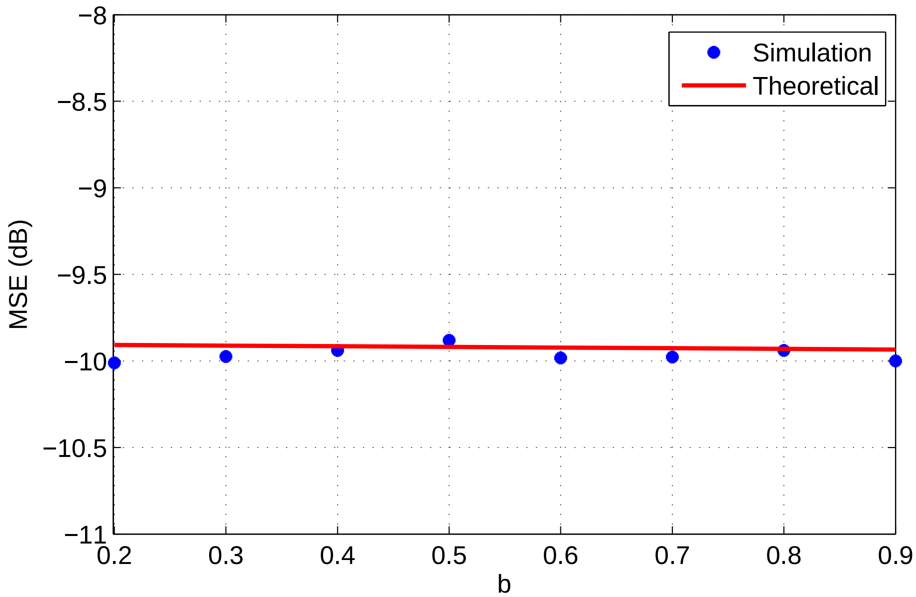

Figure 6.31: Comparison of Experimental and Analytical MSE of VSSLMFQ algorithm for single-path Rayleigh Fading Channel for different values of parameter b.

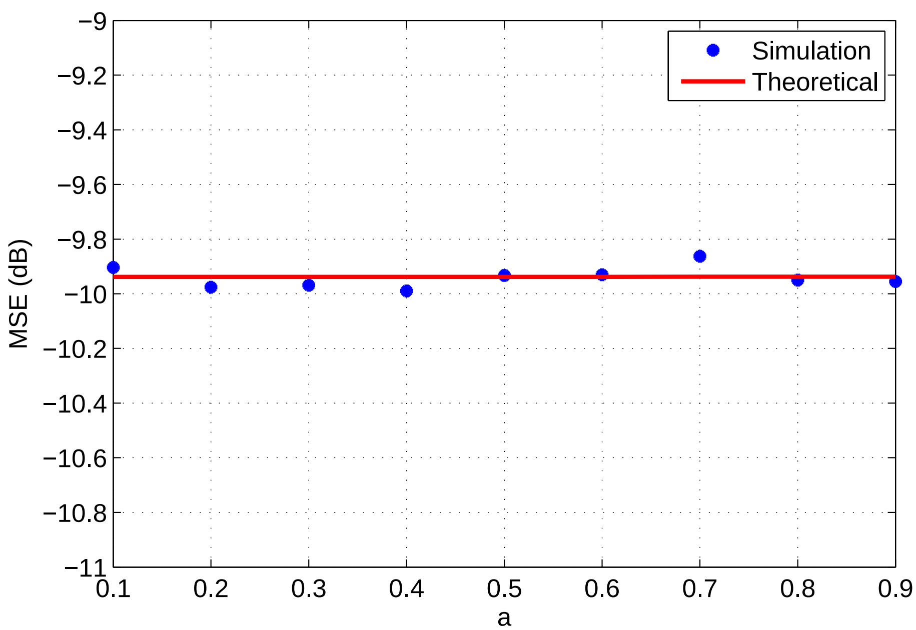

Figure 6.32: Comparison of Experimental and Analytical MSE of VSSLMFQ algorithm for Multipath Rayleigh Fading Channel for different values of parameter a.

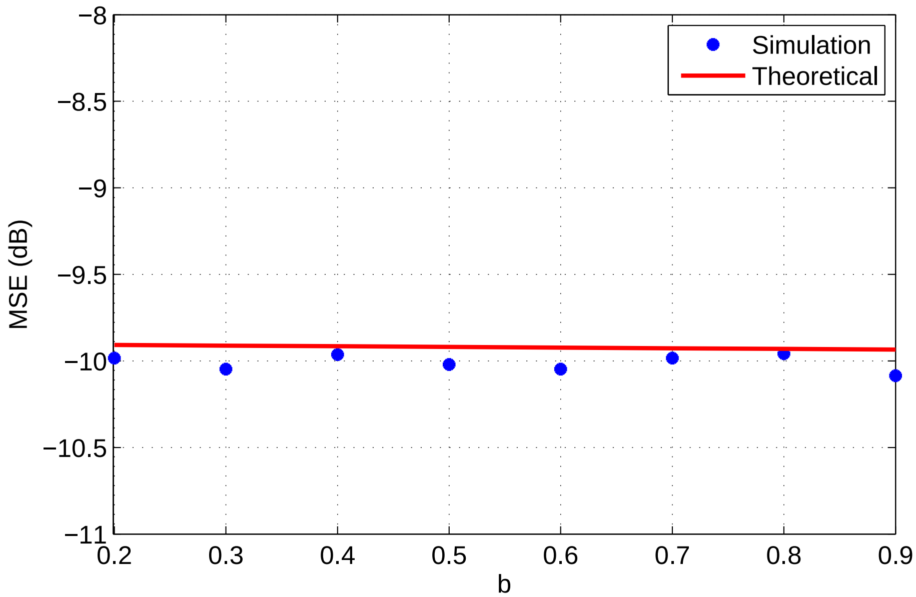

Figure 6.33: Comparison of Experimental and Analytical MSE of VSSLMFQ algorithm for Multipath Rayleigh Fading Channel for different values of parameter b.

6.3 Transient Performance of the VSSLMFQ Algorithm

In this section, we illustrate the transient behaviour of the proposed VSSLMFQ algorithm to investigate the time-evolution of both the weight-error vector as well as the time-varying step-size. The investigation for the time-varying step-size is under taken to corroborate the theoretical assumptions done during its analysis. Mainly we had discussed 3 cases for the transient behaviour of the variable step-size. Experiments to verify all the three cases have been carried out. Figure 6.34 illustrates the experimental and theoretical mean step-size behaviour of the proposed algorithm for all the three cases discussed. It can be seen that all the three cases demonstrate very close agreement with the theoretical result. Therefore all the three cases provide a very accurate model for the evolution of the time-varying step-size. Also the time-evolution of the weight-error vector is demonstrated through the MSD (mean-square deviation) and MSE. The theoretical values of the MSD are evaluated using the first entry in the state-space vector given by (5.36) and the MSE is evaluated using the second entry in (5.40) and subsequently (5.40).

For the case of time-evolution of the weight-error vector, we have considered both the uncorrelated and correlated input signals.

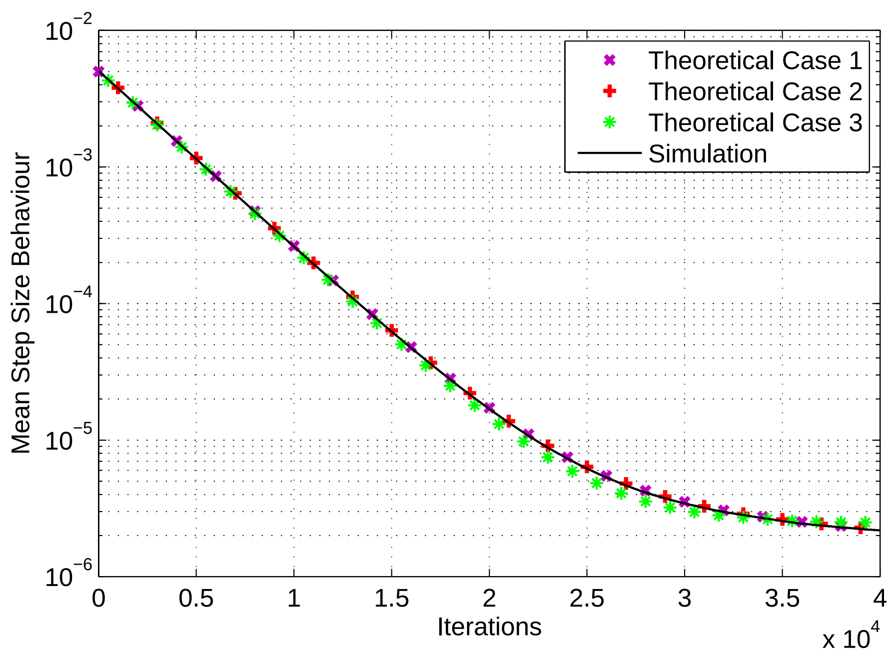

Figure 6.34: Comparison of Experimental and Analytical mean step-size behaviour of the proposed VSSLMFQ algorithm for all the three cases with SNR 10dB.

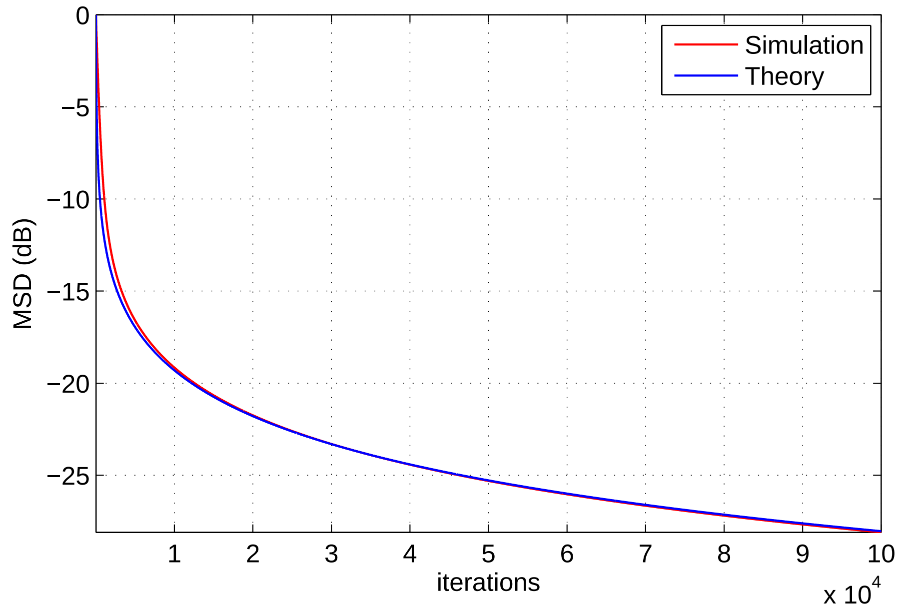

Figure 6.35: Experimental and Analytical MSD of the LMF algorithm for Correlated input.

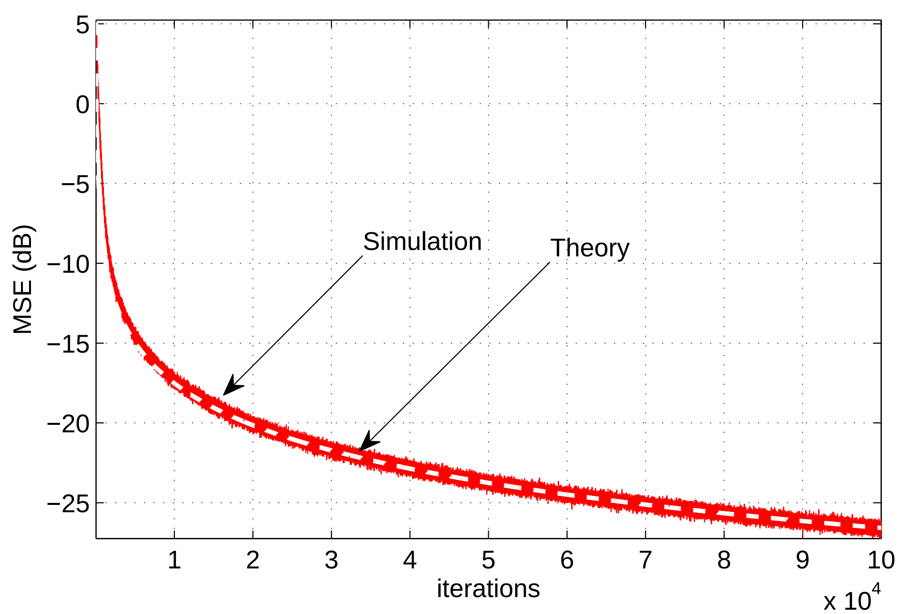

Figure 6.36: Experimental and Analytical MSE of the LMF algorithm for Correlated input.

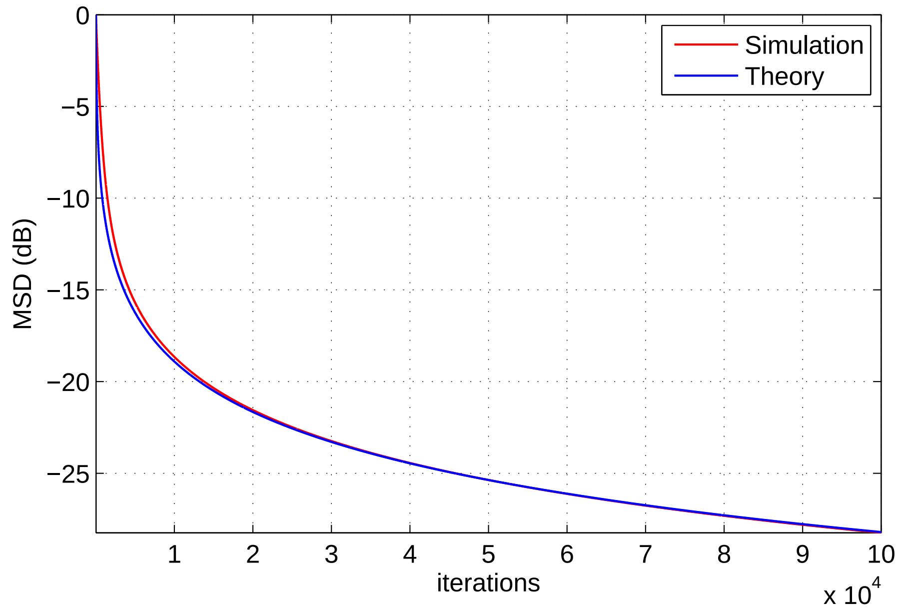

Figure 6.37: Experimental and Analytical MSD of the LMF algorithm for white input data.

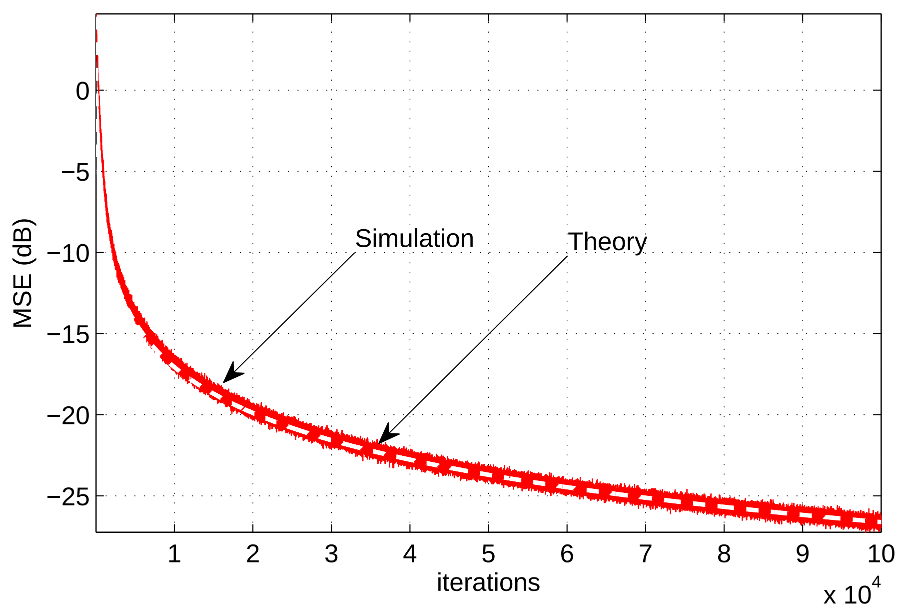

Figure 6.38: Experimental and Analytical MSE of the LMF algorithm for white input data.

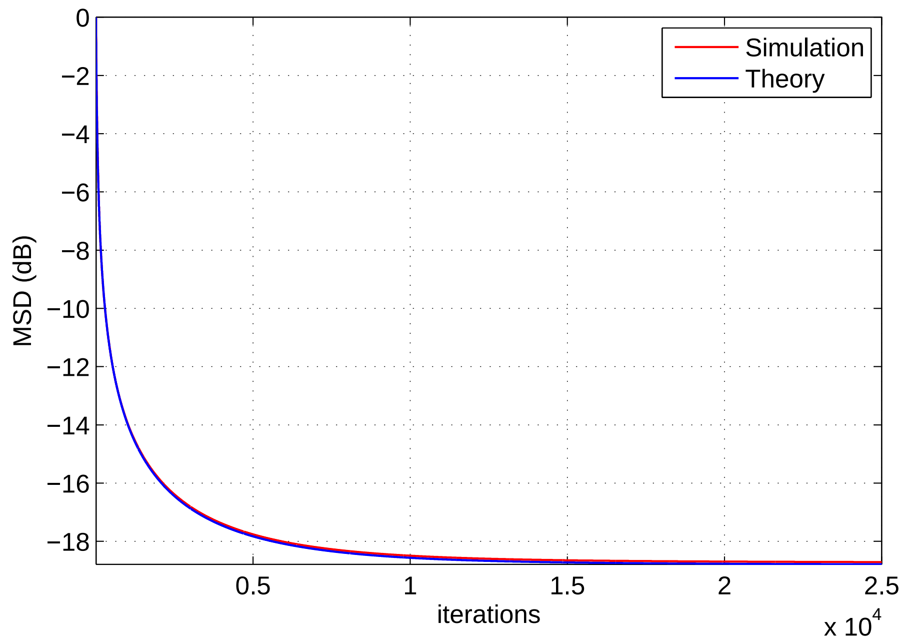

Figure 6.39: Experimental and Analytical MSD of the VSSLMFQ algorithm for correlated data.

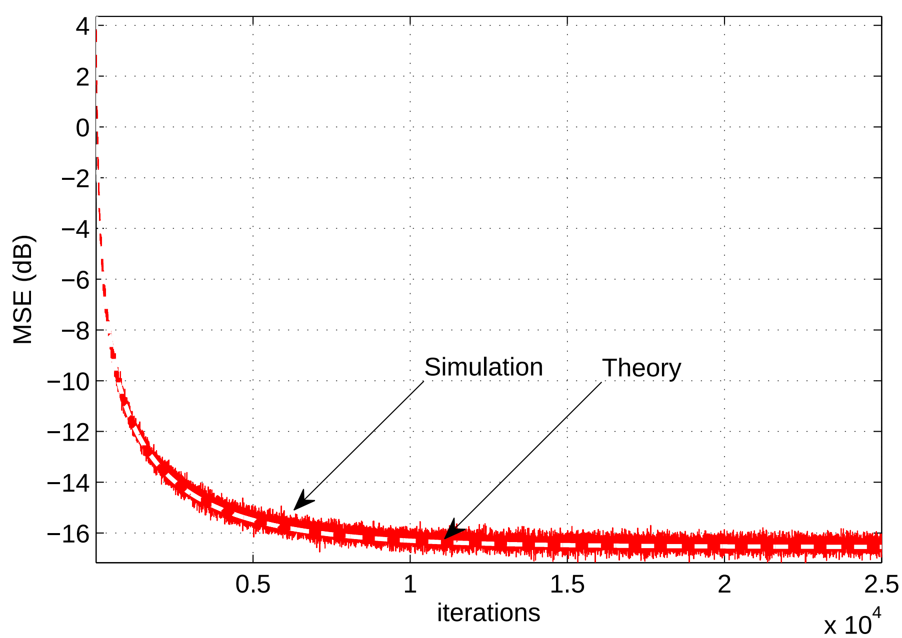

Figure 6.40: Experimental and Analytical MSE of the VSSLMFQ algorithm for correlated data.

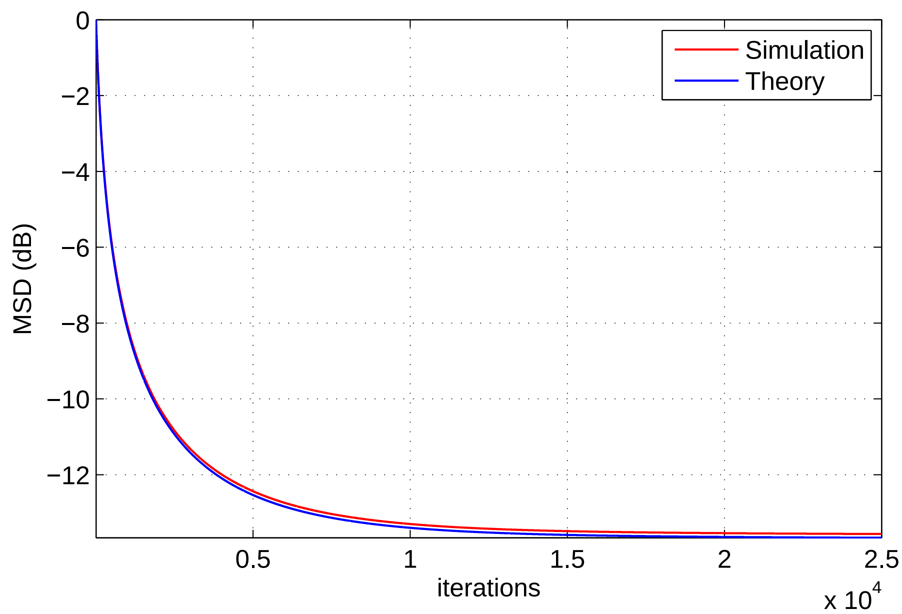

Figure 6.41: Experimental and Analytical MSD of the VSSLMFQ algorithm for white input data.

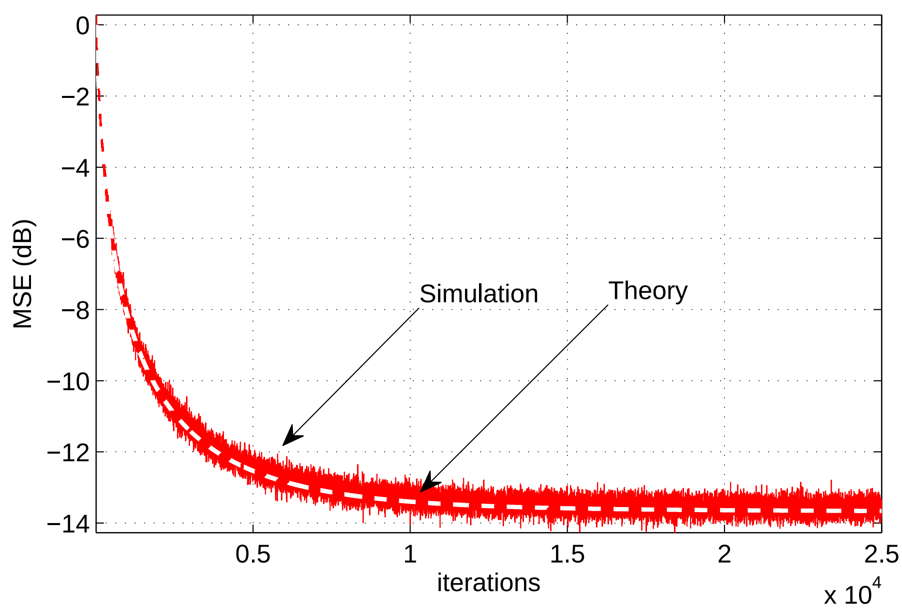

Figure 6.42: Experimental and Analytical MSE of the VSSLMFQ algorithm for white input data.

Chapter 7

Conclusions and Future Recommendations

7.1 Thesis Contributions

In this work, we have successfully presented a variable step-size LMF algorithm, namely the Variable Step-Size Least-Mean Fourth Algorithm of the Quotient Form (VSSLMFQ). The algorithm has been thoroughly analysed for its steady-state performance, tracking properties, transient performance, stability and convergence rate. These analyses have been supported by experimental results. The major contributions of this thesis can be summarised as follows:

1. A new variable step-size LMF (VSSLMFQ) algorithm has been presented that takes into account the measurement noise as a major degrading factor in achieving a lower EMSE performance and proposes a new methodology to mitigate

it.

2. The steady-state analysis of (VSSLMFQ) algorithm is presented and carried out in the mean square sense to derive an expression for the excess mean-square error using the fundamental energy relation.
3. The tracking properties of (VSSLMFQ) algorithm are analysed and an expression for the tracking excess mean-square error is derived.
4. Transient behaviour of (VSSLMFQ) algorithm is analysed to derive the time-evolution of the algorithm.
5. Lastly, the analytical results are experimentally verified to be consistent.

7.2 Conclusions

In this work, we have presented a variable step-size LMF algorithm (VSSLMFQ) to be used in applications for digital or wireless communication. Hence, the algorithm has been analysed for such applications and environments. The study included a comprehensive comparison of the VSSLMFQ algorithm with the traditional LMF algorithm and demonstrated its dominance over it in terms of the excess mean-square error for different noise environments.

Although the performance enhancement was achieved with a slight increase in the complexity of the algorithm, it was stated earlier that with the advance in digital electronics, this algorithm can have real time applications. With the demand in wireless communication on the increase, the proposed algorithm, with its ability to mitigate the measurement noise to such a degree, can be used in wireless communication receivers where accurate channel estimation is required that result in lower bit-error rates.

7.3 Future Recommendations

With every research work, there is always room for improvement. There are a few suggestions regarding any future work that can be undertaken. It was seen that the performance of the algorithm was highly dependent upon the design parameters of the variable step-size, i.e., α, γ, a and b. With regards to parameters α and γ, they can be optimised to achieve a lower EMSE than achieved in the experiments. For the case of variable step-size LMS, these parameters were analysed in [18]. A similar approach can be used for the VSSLMFQ case. For the case of parameters a and b, it was seen that they also play a key role in the MSE of the algorithm. These parameters of the quotient form can be optimised to get a minimum MSE.

Appendix A

Approximation in Equation (3.23)

The assumption taken in (3.23) is not valid under normal circumstances. Even if we assume independence, then the ratio, $E\left[\frac{A_\infty}{B_\infty}\right]$, can be written as

$$E\left[\frac{A_\infty}{B_\infty}\right] = E\left[A_\infty\right] E\left[\frac{1}{B_\infty}\right].$$

where

$$E\left[A_\infty\right] E\left[\frac{1}{B_\infty}\right] \neq \frac{E\left[A_\infty\right]}{E\left[B_\infty\right]}, \tag{A.1}$$

because

$$E\left[\frac{1}{B_\infty}\right] \neq \frac{1}{E\left[B_\infty\right]}. \tag{A.2}$$

Since our assumption is taken for the limiting case when $n \to \infty$, it has been shown in [40] and [41] that for vectors of the form

$$\frac{\mathbf{x}'\mathbf{A}\mathbf{x}}{\mathbf{x}'\mathbf{B}\mathbf{x}} = \frac{\mathbf{x}'\mathbf{A}\mathbf{x}}{\Sigma b_i x_i^2}, \tag{A.3}$$

the ratio is independent of its denominator if and only if the $b's$ are all equal and $\mathbf{x}$ has zero mean. In our case, the values of b, for $n \rightarrow \infty$, are almost identical. Also the vector $\mathbf{x}$ which in our case is the *a priori* estimation error, e_{an}, is zero mean, hence the expectation of the ratio can be written as the ratio of expectations.

Appendix B

Derivation of Equation (3.33)

This equation suggests that we can express $E\left[\|\mathbf{x}_n\|^2 e_{an}^2\right]$ and a scaled multiple of $E\left[e_{an}^2\right]$ [5]. In order to do this, lets first expand the expectation $E\left[\|\mathbf{x}_n\|^2 e_{an}^2\right]$ as:

$$
\begin{aligned}
E\left[\|\mathbf{x}_n\|^2 e_{an}^2\right] &= E\left[\mathbf{x}_n^T\mathbf{x}_n\left(\mathbf{x}_n^T\mathbf{w}_{n+1}^o - \mathbf{x}_n^T\mathbf{w}_n\right)\left(\mathbf{w}_{n+1}^{oT}\mathbf{x}_n - \mathbf{w}_n^T\mathbf{x}_n\right)\right], \\
&= E\left[tr\left(\mathbf{x}_n^T\mathbf{x}_n\mathbf{x}_n^T\left(\mathbf{w}_{n+1}^o - \mathbf{w}_n\right)\left(\mathbf{w}_{n+1}^{oT} - \mathbf{w}_n^T\right)\mathbf{x}_n\right)\right], \\
&= E\left[tr\left(\mathbf{x}_n^T\mathbf{x}_n\left(\mathbf{w}_{n+1}^o - \mathbf{w}_n\right)\left(\mathbf{w}_{n+1}^{oT} - \mathbf{w}_n^T\right)\mathbf{x}_n^T\mathbf{x}_n\right)\right], \\
&= tr\left(E\left[\mathbf{x}_n^T\mathbf{x}_n\left(\mathbf{w}_{n+1}^o - \mathbf{w}_n\right)\left(\mathbf{w}_{n+1}^{oT} - \mathbf{w}_n^T\right)\mathbf{x}_n^T\mathbf{x}_n\right]\right). \qquad \text{(B.1)}
\end{aligned}
$$

As the term $E\left[\mathbf{x}_n^T\mathbf{x}_n\left(\mathbf{w}_{n+1}^o - \mathbf{w}_n\right)\left(\mathbf{w}_{n+1}^{oT} - \mathbf{w}_n^T\right)\mathbf{x}_n^T\mathbf{x}_n\right]$ is a covariance matrix, then under the property of conditional expectation for any two random variable x and y, it states that $E[x] = E[E[x|y]]$. So we have

$$E\left[\mathbf{x}_n^T\mathbf{x}_n\left(\mathbf{w}_{n+1}^o-\mathbf{w}_n\right)\left(\mathbf{w}_{n+1}^{oT}-\mathbf{w}_n^T\right)\mathbf{x}_n^T\mathbf{x}_n\right]$$

$$= E\left[E\left[\mathbf{x}_n^T\mathbf{x}_n\left(\mathbf{w}_{n+1}^o-\mathbf{w}_n\right)\left(\mathbf{w}_{n+1}^{oT}-\mathbf{w}_n^T\right)\mathbf{x}_n^T\mathbf{x}_n|\mathbf{x}_n\right]\right],$$

$$= E\left[\mathbf{x}_n^T\mathbf{x}_n E\left[\left(\mathbf{w}_{n+1}^o-\mathbf{w}_n\right)\left(\mathbf{w}_{n+1}^{oT}-\mathbf{w}_n^T\right)|\mathbf{x}_n\right]\mathbf{x}_n^T\mathbf{x}_n\right],$$

$$E\left[\mathbf{x}_n^T\mathbf{x}_n\left(\mathbf{w}_{n+1}^o-\mathbf{w}_n\right)\left(\mathbf{w}_{n+1}^{oT}-\mathbf{w}_n^T\right)\mathbf{x}_n^T\mathbf{x}_n\right] = E\left[\mathbf{x}_n^T\mathbf{x}_n\mathbf{C}_n\mathbf{x}_n^T\mathbf{x}_n\right], \tag{B.2}$$

where

$$E\left[\left(\mathbf{w}_{n+1}^o-\mathbf{w}_n\right)\left(\mathbf{w}_{n+1}^{oT}-\mathbf{w}_n^T\right)|\mathbf{x}_n\right] = E\left[\left(\mathbf{w}_{n+1}^o-\mathbf{w}_n\right)\left(\mathbf{w}_{n+1}^{oT}-\mathbf{w}_n^T\right)\right] \triangleq \mathbf{C}_n.$$

In order to solve $E\left[\mathbf{x}_n^T\mathbf{x}_n\mathbf{C}_n\mathbf{x}_n^T\mathbf{x}_n\right]$ which involves fourth order moments, for any real-valued Gaussian random variable $\mathbf{g}$ with zero mean and diagonal covariance matrix $E\left[\mathbf{g}\mathbf{g}^T\right] = \Lambda$, and any symmetric matrix $\mathbf{K}$ we have:

$$E\left[\mathbf{g}\mathbf{g}^T\mathbf{K}\mathbf{g}\mathbf{g}^T\right] = \Lambda tr(\mathbf{K}\Lambda) + 2\Lambda\mathbf{K}\Lambda,$$

which in this case will be:

$$\mathbf{g} \leftarrow \mathbf{x}_n, \quad \mathbf{K} \leftarrow \mathbf{C}_n, \quad \Lambda \leftarrow \sigma_x^2\mathbf{I},$$

so that we can write:

$$E\left[\mathbf{x}_n^T\mathbf{x}_n\mathbf{C}_n\mathbf{x}_n^T\mathbf{x}_n\right] = \sigma_x^4\left[tr\left(\mathbf{C}_n\right)\mathbf{I}+2\mathbf{C}_n\right]. \tag{B.3}$$

Substitute in (B.1), we get:

$$E\left[\|\mathbf{x}_n\|^2 e_{an}^2\right] = tr\left[\sigma_x^4\left[tr\left(\mathbf{C}_n\right)\mathbf{I}+2\mathbf{C}_n\right]\right] = \left(N+2\right)\sigma_x^4 tr\left(\mathbf{C}_n\right). \tag{B.4}$$

Moreover we can express $E\left[e_{an}^2\right]$ in terms of $tr\left(\mathbf{C}_n\right)$ as follows

$$\begin{aligned}
E\left[e_{an}^2\right] &= E\left[\mathbf{x}_n^T\left(\mathbf{w}_{n+1}^o - \mathbf{w}_n\right)\left(\mathbf{w}_{n+1}^{oT} - \mathbf{w}_n^T\right)\mathbf{x}_n\right], \\
&= E\left[\mathbf{x}_n^T\mathbf{C}_n\mathbf{x}_n\right], \\
&= tr\left(E\left[\mathbf{x}_n\mathbf{x}_n^T\mathbf{C}_n\right]\right), \\
&= tr\left(E\left[\mathbf{R}\mathbf{C}_n\right]\right), \\
&= \sigma_x^2 tr\left(E\left[\mathbf{C}_n\right]\right), \\
&= \sigma_x^2 tr\left(\mathbf{C}_n\right).
\end{aligned}$$

Substituting the above result in (B.4) we get

$$E\left[\|\mathbf{x}_n\|^2 e_{an}^2\right] = \left(N+2\right)\sigma_x^2 E\left[e_{an}^2\right]. \tag{B.5}$$

Appendix C

Exact Moments of Ratio of Quadratic Forms in Normal Variables

Ratio of quadratic forms are often encountered in estimation problems. But their characterisation is difficult due to the complex expectation that have to be evaluated in order to get the desired result. Here, we are going to lay out an approach to find the exact moments of ratio of quadratic form in normal variables based on [38].

The problem can be motivated by stating that for Gaussian random variable, often a situation is encountered when we need to perform an expectation of the form:

$$E\left[\frac{\mathbf{x}'\mathbf{A}\mathbf{x}}{\mathbf{x}'\mathbf{B}\mathbf{x}}\right] \tag{C.1}$$

where $\mathbf{A}$ is symmetric, $\mathbf{B}$ positive semi-definite and $\mathbf{x}$ is normally distributed with some non zero mean and positive definite covariance matrix. The main idea in finding the moments of the ratio is that the moments of any random variable can be expressed as polynomial of its cumulants (cumulants are logs of moment-generating function).

This result is then extended to the case of quadratic form.

The moments of the ratio of two random variables is dealt with the s-fold differentiation of the moment-generating function (provided it exists) under the integral sign to obtain the sth moment. Then this result is extended to the quadratic form. In the end, we are able to combine all these results to get the expression for the exact moments of the ratio of quadratic form.

In the case of our proposed algorithm, we encountered this problem when trying to evaluate the expectation:

$$\theta_n = \frac{\sum_{i=0}^{n} a^i e_{n-i}^2}{\sum_{j=0}^{n} b^j e_{n-j}^2}. \tag{C.2}$$

We can reformulate the problem as:

$$E\left[\theta_n\right] = E\left[\frac{\mathbf{e}^T\mathbf{A}\mathbf{e}}{\mathbf{e}^T\mathbf{B}\mathbf{e}}\right], \tag{C.3}$$

where $\mathbf{e}$ is a vector of length $(N+1)\times 1$ with elements $\mathbf{e} = (e_n \, e_{n-1} \, e_{n-2} \, \ldots \, e_0)$, $\mathbf{A}$ and $\mathbf{B}$ are $(N+1)\times(N+1)$ diagonal matrices given by:

$$diag\left(\mathbf{A}\right) = \left\{a^0 \, a^1 \, \ldots \, a^n\right\}, \quad diag\left(\mathbf{B}\right) = \left\{b^0 \, b^1 \, \ldots \, b^n\right\}. \tag{C.4}$$

Now we suppose $\mathbf{e}$ to be a normally distributed vector with mean $\mathbf{m}$ and positive-definite covariance matrix $\mathbf{K}$. Let $\mathbf{P}$ be an orthogonal $(N+1)\times(N+1)$ and Λ a diagonal $(N+1)\times(N+1)$ matrix such that

$$\mathbf{P}^T\mathbf{L}^T\mathbf{B}\mathbf{L}\mathbf{P} = \Lambda, \quad \mathbf{P}^T\mathbf{P} = \mathbf{I}. \tag{C.5}$$

This type of decomposition can be achieved by Cholesky decomposition. Also define

$$\overline{\mathbf{A}} = \mathbf{P}^T\mathbf{L}^T\mathbf{A}\mathbf{L}\mathbf{P}, \quad \overline{\mathbf{m}} = \mathbf{P}^T\mathbf{L}^{-1}\mathbf{m}. \tag{C.6}$$

Then if the expectations exist, we can write:

$$\begin{aligned} E\left[\frac{\mathbf{e}^T\mathbf{A}\mathbf{e}}{\mathbf{e}^T\mathbf{B}\mathbf{e}}\right]^s &= \frac{e^{\left(-(1/2)\mathbf{m}'\mathbf{K}^{-1}\mathbf{m}\right)}}{(s-1)!}\sum_{\mathbf{l}}\delta_s\left(\mathbf{l}\right) \\ &\quad \times \int_0^\infty t^{s-1}\left|\triangle\right| e^{1/2\zeta'\zeta}\prod_{j=1}^{s}\left[Tr\left(\mathbf{G}^j\right)+\zeta'\mathbf{G}^j\zeta\right]^{r_j}dt. \end{aligned} \tag{C.7}$$

We are going to explain the terms one by one. Firstly the summation is over all possible $1\times s$ vectors $\mathbf{l} = (r_1, r_2, \ldots, r_s)$ whose components r_j are non-negative integers satisfying $\sum_{j=1}^{s} jr_j = s$, so that:

$$\delta_s\left(\mathbf{l}\right) = s!2^s\prod_{j=1}^{s}\left[r_j!\left(2j\right)^{r_j}\right]^{-1}. \tag{C.8}$$

The matrix Δ is diagonal positive definite, $\mathbf{G}$ is a symmetric $(N+1)\times(N+1)$ matrix and ζ is an $(N+1)\times 1$ vector given:

$$\Delta = \left(\mathbf{I}+2t\Lambda\right)^{-1/2}, \quad \mathbf{G} = \Delta\overline{\mathbf{A}}\Delta, \quad \zeta = \Delta\overline{\mathbf{m}}. \tag{C.9}$$

For our case, we require the first two moments of θ_n, so $s = 1, 2$. The corresponding vectors, **l**, for this case are:

$$
\begin{aligned}
s &= 1 : r_1 = 1 \\
s &= 2 : (r_1, r_2) = (2,0), (0,1)
\end{aligned}
$$

so

$$
\begin{aligned}
\delta_1 (\mathbf{l}_1) &= 1, \\
\delta_2 (\mathbf{l}_1) &= 1, \\
\delta_2 (\mathbf{l}_2) &= 2.
\end{aligned}
$$

Hence substituting this data to evaluate for $s = 1, 2$ we get:

$$
E[\theta_n] = e^{\left(-(1/2)\mathbf{m}'\mathbf{K}^{-1}\mathbf{m}\right)} \times \int_0^1 |\triangle| \, e^{1/2\zeta'\zeta} \left[Tr(\mathbf{G}) + \zeta'\mathbf{G}\zeta\right] dt, \tag{C.10}
$$

and

$$
E\left[\theta_n^2\right] = e^{\left(-(1/2)\mathbf{m}'\mathbf{K}^{-1}\mathbf{m}\right)} \sum_{\mathbf{l}} \delta_2(\mathbf{l}) \times \int_0^1 |\triangle| \, e^{1/2\zeta'\zeta} \prod_{j=1}^{2} \left[Tr\left(\mathbf{G}^j\right) + \zeta'\mathbf{G}^j\zeta\right]^{r_j} dt. \tag{C.11}
$$

Notice the upper limit is set to 1. This is because θ_n is a decreasing function and bounded between $[0, 1]$. Lastly, both the equations have integrals in them as there is no close form solution to them. So they have to be evaluated through numerical integration.

Bibliography

[1] S. Qureshi, "Adaptive equalization," *Proceedings of the IEEE*, vol. 73, no. 9, pp. 1349–1387, 1985.

[2] S. Haykin, *Adaptive filter theory.* Prentice-Hall, Inc. Upper Saddle River, NJ, USA, 4th ed., 2002.

[3] B. Widrow, J. Glover Jr, J. McCool, J. Kaunitz, C. Williams, R. Hearn, J. Zeidler, E. Dong Jr, and R. Goodlin, "Adaptive noise cancelling: Principles and applications," *Proceedings of the IEEE*, vol. 63, no. 12, pp. 1692–1716, 1975.

[4] J. Makhoul, "Linear prediction: A tutorial review," *Proceedings of the IEEE*, vol. 63, no. 4, pp. 561–580, 1975.

[5] A. Sayed, *Adaptive Filters.* IEEE Press, Wiley-Interscience, 2008.

[6] T. Aboulnasr and K. Mayyas, "A robust variable step-size LMS-type algorithm: analysis andsimulations," *Signal Processing, IEEE Transactions on [see also Acoustics, Speech, and Signal Processing, IEEE Transactions on]*, vol. 45, no. 3, pp. 631–639, 1997.

[7] W. Chen, R. Haddad, B. Res, and N. Morristown, "A variable step size LMS algorithm," *Circuits and Systems, 1990., Proceedings of the 33rd Midwest Symposium on*, pp. 423–426, 1990.

[8] M. Costa and J. Bermudez, "A noise resilient variable step-size LMS algorithm," *Signal Processing*, vol. 88, no. 3, pp. 733–748, 2008.

[9] M. Costa and J. Bermudez, "A Robust Variable Step Size Algorithm for LMS Adaptive Filters," *IEEE Conference on Acoustics, Speech and Signal Processing, ICASSP*, vol. 3, 2006.

[10] R. Harris, D. Chabries, and F. Bishop, "A variable step (VS) adaptive filter algorithm," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 34, no. 2, pp. 309–316, 1986.

[11] T. Haweel, "A simple variable step size LMS adaptive algorithm," *International Journal of Circuit Theory and Applications*, vol. 32, pp. 523–536, 2004.

[12] R. Kwong and E. Johnston, "A variable step size LMS algorithm," *Signal Processing, IEEE Transactions on [see also Acoustics, Speech, and Signal Processing, IEEE Transactions on]*, vol. 40, no. 7, pp. 1633–1642, 1992.

[13] S. Zhao, Z. Man, S. Khoo, and H. Wu, "Variable step-size LMS algorithm with a quotient form," *Signal Processing*, vol. 89, pp. 67–76, 2008.

[14] W. Ang and B. Farhang-Boroujeny, "A new class of gradient adaptive step-size LMS algorithms," *IEEE transactions on signal processing*, vol. 49, no. 4, pp. 805–810, 2001.

[15] V. Mathews and Z. Xie, "A stochastic gradient adaptive filter with gradient adaptive stepsize," *IEEE Transactions on Signal Processing*, vol. 41, no. 6, pp. 2075–2087, 1993.

[16] J. Evans, P. Xue, and B. Liu, "Analysis and implementation of variable step size adaptive algorithms," *IEEE Transactions on Signal Processing*, vol. 41, no. 8, pp. 2517–2535, 1993.

[17] S. Gelfand, Y. Wei, and J. Krogmeier, "The stability of variable step-size LMS algorithms," *IEEE Transactions on Signal Processing*, vol. 47, no. 12, pp. 3277–3288, 1999.

[18] C. Lopes and J. Bermudez, "Evaluation and design of variable step size adaptive algorithms," *Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP '01). 2001 IEEE International Conference on*, vol. 6, pp. 3845–3848, 2001.

[19] E. Walach and B. Widrow, "The least mean fourth (LMF) adaptive algorithm and its family," *IEEE Transactions on Information Theory*, vol. 30, pp. 275–283, Feb 1984.

[20] A. Zerguine, "Convergence behavior of the normalized least mean fourth algorithm," *Signals, Systems and Computers, 2000. Conference Record of the Thirty-Fourth Asilomar Conference on*, vol. 1, 2000.

[21] A. Zerguine, "Convergence and steady-state analysis of the normalized least mean fourth algorithm," *Digital Signal Processing*, vol. 17, no. 1, pp. 17–31, 2007.

[22] M. Moinuddin and A. Zerguine, "Tracking analysis of normalized adaptive algorithms," *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03)*, vol. 6, pp. 637–640, 2003.

[23] M. Moinuddin, A. Zerguine, and A. Sheikh, "Tracking analysis of the NLMF algorithm in the presence of both random and cyclic nonstationarities," *Signal Processing and Its Applications, 2005. Proceedings of the Eighth International Symposium on*, vol. 2, pp. 755–758, 2005.

[24] C. C. M.K. Chan, "Using a normalised lmf algorithm for channel equalisation with co-channel interference," *EUSIPCO-2002 Toulouse, France*, pp. 48–51, September 3-6 2002.

[25] M. Chan, A. Zerguine, and C. Cowan, "An optimised normalised LMF algorithm for sub-Gaussian noise," *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03)*, vol. 6, pp. 377–380, 2003.

[26] A. Zerguine, M. Chan, T. Al-Naffouri, M. Moinuddin, and C. Cowan, "Convergence and tracking analysis of a variable normalised LMF (XE-NLMF) algorithm," *Signal Processing*, vol. 89, pp. 778–780, 2008.

[27] S. Cho, S. Kim, and K. Jeon, "Statistical convergence of the adaptive least mean fourth algorithm," *Signal Processing, 1996., 3rd International Conference on*, vol. 1, pp. 610–613, 1996.

[28] P. I. Hubscher and J. Bermudez, "An improved statistical analysis of the least mean fourth (LMF) adaptive algorithm," *Signal Processing, IEEE Transactions on [see also Acoustics, Speech, and Signal Processing, IEEE Transactions on]*, vol. 51, no. 3, pp. 664–671, 2003.

[29] P. Hubscher, V. Nascimento, and J. Bermudez, "New results on the stability analysis of the LMF (least mean fourth) adaptive algorithm," *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03)*, vol. 6, pp. 369–372, 2003.

[30] P. I. Hubscher, J. C. M. Bermudez, and V. H. Nascimento, "A Mean-Square Stability Analysis of the Least Mean Fourth Adaptive Algorithm," *Signal Processing, IEEE Transactions on [see also Acoustics, Speech, and Signal Processing, IEEE Transactions on]*, vol. 55, no. 8, pp. 4018–4028, 2007.

[31] V. Nascimento and J. Bermudez, "Probability of divergence for the least-mean fourth algorithm," *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. 1376–1385, 2006.

[32] T. Al-Naffouri and A. Sayed, "Adaptive filters with error nonlinearities: Mean-square analysis and optimum design," *EURASIP Journal on Applied Signal Processing*, vol. 2001, no. 4, pp. 192–205, 2001.

[33] T. Al-Naffouri and A. Sayed, "Transient analysis of adaptive filters with error nonlinearities," *IEEE Transactions on Signal Processing*, vol. 51, no. 3, pp. 653–663, 2003.

[34] A. Leon-Garcia, *Probability, statistics, and random processes for electrical engineering*. Prentice Hall, 2007.

[35] A. Papoulis, S. Pillai, P. A, and P. SU, *Probability, random variables, and stochastic processes*. McGraw-Hill New York, 1965.

[36] R. Price, "A useful theorem for nonlinear devices having Gaussian inputs," *Information Theory, IRE Transactions on*, vol. 4, no. 2, pp. 69–72, 1958.

[37] T. Al-Naffouri and A. Sayed, "Transient analysis of data-normalized adaptive filters," *IEEE Transactions on Signal Processing*, vol. 51, no. 3, pp. 639–652, 2003.

[38] J. Magnus, "The exact moments of a ratio of quadratic forms in Normal variables," *Annales d'Economie et de Statistique*, pp. 95–109, 1986.

[39] H. Cramer, *Mathematical methods of statistics*. Princeton university press, 1946.

[40] M. Jones, "On moments of ratios of quadratic forms in normal variables.," *STAT. PROB. LETT.*, vol. 6, no. 2, pp. 129–136, 1987.

[41] D. Conniffe and J. Spencer, "When moments of ratios are ratios of moments," *The Statistician*, pp. 161–168, 2001.

Vitae

– Syed Muhammad Asad.

– Nationality : Pakistani.

– Received Bachelor of Science (BS) degree in Electronics Engineering from Sir Syed University of Engineering & Technology, Karachi, Pakistan in 2001.

– Worked at Saudi International Trading & Marketing Ltd. in Jeddah, Saudi Arabia.

– Joined King Fahd University of Petroleum & Minerals in February 2006.

– Email: syed.m.asad@gmail.com

– Present Address : P. O. Box 8219, Dhahran 31261, Saudi Arabia.

– Permanent Address : Flat # A-12, Noman Terrace, Gulshan-e-Iqbal, Block 11, Karachi, Pakistan.