TRACKING ANALYSIS OF ϵ-NLMS AND LEAKY ϵ-NLMS ALGORITHMS FOR COLORED GAUSSIAN INPUTS

BY

MAHMOUD JAMAL ANKEER

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

TELECOMMUNICATION ENGINEERING

JANUARY 2011

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN 31261, SAUDI ARABIA

DEANSHIP OF GRADUATE STUDIES

This thesis, written by **MAHMOUD JAMAL M. ANKEER** under the direction of his thesis adviser and approved by his thesis committee, has been presented to and accepted by the Dean of Graduate Studies, in partial fulfillment of the requirements for the degree of **MASTER OF SCIENCE IN TELECOMMUNICATION ENGINEERING**.

Thesis Committee

azzedine zerguine

Dr. Azzedine Zerguine (Adviser)

Dr. Tareq Y. Al Naffouri (Co-adviser)

Dr. Samir H. Abdul Jauwad (Member)

Dr. Abdelmalek Zidouri (Member)

Dr. Wajih Abu-Al-Saud (Member)

Dr. Ali Al-Shaikhi
Department Chairman

Dr. Salam A. Zummo
Dean of Graduate Studies

17/4/11

Date

عمادة الدراسات العليا
DEANSHIP OF GRADUATE STUDIES

Dedicated to the one I give my all for, My Mother

to the one who supports me, My Loving Father

and Finally, to My Wife, My Coming Baby, and to my Brothers and Sisters.

ACKNOWLEDGMENTS

In the name of Allah, Most Gracious, Most Merciful

Allah, With Your prolific praise,

Owner of Honor, I desire to begin A limitless praise,

with which You are Pleased.

All Praise be to Allah, the Lord of the worlds, who says in His Glorious Book, "My Lord! Increase me in knowledge". All praise is due to Allah for His limitless blessings. Peace and blessings of Allah be upon the noblest of the Prophets and Messengers, our Prophet Mohammed who said "he who treads the path in search of knowledge, Allah would make that path easy, leading to Paradise for him". May peace be upon him, his family, his companions, and his followers.

It is my pleasure to work with such kind, respective, understandable, very helpful and remarkable persons such as Dr. Azzedine Zerquine, and Dr. Tareq Al Naffouri, Without their guidance, advices, patience, kindness and expertise, nothing of this work would see the light. I really enjoyed working with them, they gave me the opportunity to increase my knowledge and my research experience in the adaptive filtering field.

I would like to express my profound gratitude and appreciation to my commitee members, Dr. Samir Abdul-Jauwad, Dr. Abdelmalek Zidouri, and Dr. Wajih Abu Al-Saud for their helping and support in addition to their acceptence to be members of my committee.

Special thanks to my dearest mother and father for their emotional support, love, sacrifices, prayers and understanding throughout my academic career. I would not be where I am in life if it weren't for their love.

Thank you very much to my wife who gives me her private time just to give me the ability to finsih this work, she is my partner in this success, she encouarges me to do my best and to work hard to reach this success.

My grateful thanks to my friend, Omer Bin Saeed for his unlimited helping and support, I would like to thank him for every piece of time that he gave to me during the time of this research.

I also acknowledge my colleagues and friends as I had a pleasant, enjoyable and fruitful company with them especially Yousuf Al-Arian, Ahmad Al-Masri, Hesham Abdul-Khaliq, Babar Khan, Ahmad Saleem, Mohammed Sitouah, Sameh, Suhaib, and Abdurahman Shaheen. I thank all of them (and many more whom I forgot to mention) for their love and encouragement and for dragging me out of the academic work. All the good time we had together will remain a valuable part of my memory for the rest of my life.

Finally, I acknowledge that this research would not have been possible without the support provided by the King Fahd University of Petroleum & Minerals.

We praise Allah for His favor to us in completing this work, praying Him to accept this as a sincere service for His sake.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

THESIS ABSTRACT

NAME: Mahmoud Jamal M. Ankeer

TITLE OF STUDY: Tracking Analysis of ϵ-NLMS and Leaky ϵ-NLMS Algorithms for Colored Gaussian Inputs

MAJOR FIELD: Telecommunication Engineering

DATE OF DEGREE: January 2011

One of the main features of the adaptive filters is its ability to track the variations in the underlaying signal statistics. This is due to the fact that they are relying on the instantaneous data, the statistical properties of the error signal and the weight vector of the adaptive filter will react to these variations in the input signal properties. In this thesis, the tracking analysis of ϵ-NLMS algorithm and leaky ϵ-NLMS algorithm are investigated in a nonstationary environments, the environment is characterized by a first order autoregressive model and the filter input is assumed to be correlated gaussian, A closed form EMSE expression is derived for both algorithms by relying on energy conservation and weighted energy conservation relations, Finally, simulation results show a great match between the analytical and the simulation results especially in the slow changable environments.

ملخـص الرســـالــة

الاسم: محمود جمال محمود عنكير

عنوان الدراسة: تحليل خاصية التتبع لكل من خوارزمية (e-NLMS) و خوارزمية (Leaky e-NLMS) عندما تكون المدخلات على شرط ارتباط جاوس.

التخصص: هندسة الاتصالات.

تاريخ التخرج: يناير ٢٠١١

احدى اهم المميزات الخاصة بالمنقحات المتكيفة هو قدرتها على عمل تتبع للتغيرات الحاصلة في خصائص الاشارة الكهربائية. وهذا كله يرجع الى حقيقة ان هذه المنقحات تعتمد على المعلومات اللحظية، الخصائص الاحصائية لاشارة الخطأ بالاضافة الى متجه الوزن والتي بدورها تقوم بالاستجابة لهذه التغيرات الحاصلة في خصائص الاشارة الكهربائية.

في هذه الرسالة، سوف نقوم بدراسة تحليل قدرة التتبع لخوارزمية e-NLMS وخوارزمية leaky e-NLMS وذلك ضمن البيئة المتغيرة مع الزمن. قمنا باعطاء وصف لهذه البيئة وذلك اعتمادا على نموذج الانحدار الذاتي. مدخلات المنقحات تم اعتبارها على انها على شرط ارتباط جاوس، اضافة الى ذلك فان هذه الرسالة تقدم اشتقاقا بصيغة مغلقة لفائض متوسط مربع الخطأ (EMSE) المتعلق بكلا الخوارزميتين وذلك بالاعتماد على علاقة حفظ الطاقة في حالة خوارزمية e-NLMS وعلاقة حفظ الطاقة الموزونة في حالة خوارزمية leaky e-NLMS.

في ختام الرسالة كانت نتائج المحاكاة متطابقة بشكل كبير مع العلاقات التي تم اشتقاقها وخاصة في البيئة ذات التغير البطيء مع الزمن.

Nomenclature

Abbreviations

SNR : Signal to Noise Ratio

AWGN : Additive white Gaussian noise

Notations

$(.)^H$	:	Hermitian transpose
$\|.\|$	:	Euclidean Norm
$Tr(.)$	:	Trace of a matrix
$Tr_a(.)$	:	Normalized trace of a matrix
$\mathcal{N}(m, R)$	:	Gaussian random variable with a mean m and a auto-correlation matrix R

CHAPTER 1

INTRODUCTION

Adaptive systems are playing a very important role in the development of modern communications. The concept of adaptive filtering is considered as an significant part of the statistical signal processing. Actually, Adaptive filter provides us with the attractive solution to the problem of processing or detecting a signal which results from an unknown statistics of an environment. In addition to that, Adaptive filters and systems are considered to be extremely effective in achieving high reliability, high quality, and high efficiency of around-the-world ubiquitous telecommunication services. Thus, adaptive filters are successfully applied in such diverse fields as noise cancellation [2], linear prediction [3], equalization [1], and in system identification [4],[5].

The two well known algorithms for adaptive filters in the literature are the Least Mean Squares (LMS) [4], [6] and the Recursive Least Squares (RLS)[4] algorithms.

The RLS algorithm basically depends on updating the inverse of the covariance matrix which has a size $L \times L$ (L is the filter's length), therefore, its order of computational complexity is $O(L^2)$. In order to reduce the RLS algorithms complexity, researchers in the adaptive filtering field develop some fast versions of the RLS (FRLS) algorithms, namely, the fast Kalman [7], Fast a posteriori Error Sequential Technique (FAEST) [8], and Fast Transversal Filter (FTF) [9], which reduces the order of the complexity to $O(L)$. In the past, these fast RLS algorithms were not used practically especially in the real time applications, this is due to the divergence of these algorithms which appears as a result of the numerical error accumulation in the prediction parameters. Efficient stabilization techniques, with limited additional complexity that prevent numerical divergence without performance degradation, have been proposed in [10], [11].

The LMS algorithm represents a solution to the optimal Weiner Filter criterion which is minimizing the mean square value of the error in a stochastic approximation sense. The LMS algorithm belongs to the gradient types algorithmic family, this implies that LMS algorithm inherits its low computational complexity ,i.e., $O(L)$ operations, in addition to the slow convergence property, especially in highly correlated signals like speech. Hence, one must whiten the input signal. using one of the following methods: the projection algorithm or the affine projection algorithm [12], or the Normalized LMS (NLMS) algorithm [13].

There are many other algorithms derived from the LMS algorithm such as sign LMS [4], Leaky LMS [14], and Block LMS [15] algorithms, just to name a few. The performance of the LMS algorithm can be improved by employing a time varying step size in the standard LMS[16], [17]. The basic idea of this approach depends on using large step size when the algorithm is far from the optimal solution, the result is speeding up the convergence rate, and when the algorithm becomes close to the optimum solution, small step size is used in order to achieve a low level of misadjustment, As a result a great improvement in the overall performance of the LMS algorithm is acheived. To obtain this, The step size is adjusted in accordance to some criterion. Different criteria have been mentioned in the literature in order to adjust the step size, examples of these criteria are: squared instantaneous error [16], sign changes of successive samples of the gradient[18], attempting to reduce the squared error at each instant [19], and cross correlation of input and error [17]. In [20], a new variable step size LMS algorithm was proposed in which the step size is adjusted according to the square of the time averaged estimate of the correlation of the error.

The Least Mean Fourth (LMF) algorithm, another member of the steepest descent algorithm with 4th error norm was suggested in [21], is a special case of the more general family of the descent algorithms [5] with 2k error norms, where k is a positive integer. While the LMS algorithm is well established in adaptive filtering, the LMF algorithm has gained attention [22], [23], [24].

1.1 Adaptive Filters

When the filter is linear and all the pretinent statistics are known, then the Weiner filter [5] provides us with the optimal solution, which is optimum from the mean square point of view. However, when the filter is required to operate in an environment of unknown statisitics or nonstationary environment, an adaptive filter provides an elegant solution to such problems. Generally, Adaptive filters are defined as filters which modify or adapt its characterisitics automatically (i.e. without user intervention) in order to achieve the desired objectives.

Working mechanism of the adaptive filter basically depends on a recursive algorithm, this gives the filter the ability to perform sufficiently in an environment where a complete knowledge of the signal characteristics is not available. The algorithm starts from some predetermined set of initial conditions, the algorithm initially ignores completely the environment. In a stationary environment, after successive iterations, the algorithm tries to converge to the optimum Wiener solution in some statistical sense.

There are different ways of classifying adaptive filters, depending on the feature of interest. When the feature of interest is input-output mapping, adaptive filters can be classified into two main categories: *linear and nonlinear*. Linear adaptive filters compute an estimate of the desired response by using a linear combination of the available set of the observables applied to the input. This form of input-output mapping satisfied by having a single layer of computational units or simply a single computational unit as the output layer. On the other hand, when the input-output mapping is required to be nonlinear, we need to use nonlinear adapative filter. Typically, nonlinear adaptive filters involove the use of one or more layers of computational units in addition to the output layer [5].

1.2 Applications Of Adaptive Filters

Adaptive filtering has a number of applications in different fields. Despite of these applications are really quite different in nature, but they have one basic feature which is shared between all of them. All of these applications have an input signal and a desired response to compute the error, which is in turn used to control the values of a set of adjustable filter coefficients. However, the main difference among the various applications is the manner in which the desired response is extracted. On this basis, adaptive filters are classified into the following four categories.

1.2.1 Inverse Modelling or Equalization

As the title implies, the basic function of the adaptive filter in this application is to provide an inverse model which represents the best fit to the unknown system. Thus, when the adaptive filter converges, then the transfer function inverse of the unknown system is approximated by the adaptive filter. A delay is introduced to the desired response path as shown in Figure (1.1), so as to ensure that the input to the adaptive filter is minimum phase and suitable for equalization. The primary use of the inverse modeling is to reduce the intersymbol interference (ISI) in digital receivers. This is can be achieved by using the channel equalization for digital communications [1].

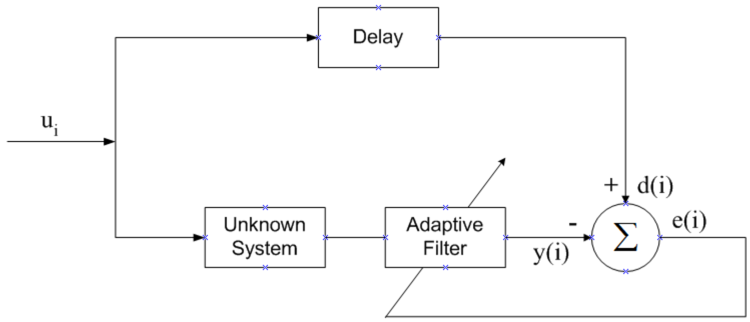

Figure 1.1: Inverse Modeling Scenario

1.2.2 Prediction

In this application, the adaptive filter is used to provide the best prediciton of the present value of the input signal from its previous values. The configuration shown in Figure (1.2) is used for this purpose, where the desired signal, d_n, is the instantaneous value and the input to the adaptive filter is the delayed version of the same signal. This application is used in linear predictive coding (LPC) of speech [3] and in adaptive differential pulse-code modulation (DPCM) [27].

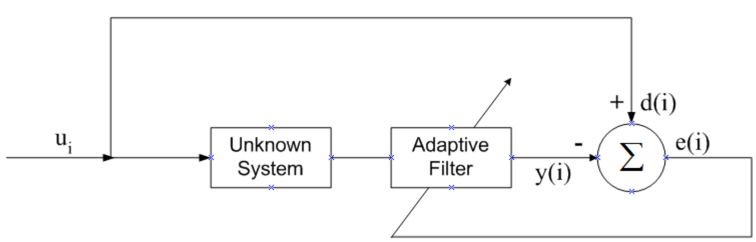

Figure 1.2: Prediction Scenario

1.2.3 Noise Cancellation

In this class of applications, the adaptive filter is used to cancel unknown interference contained in a primary signal, as shown in Figure (1.3). The primary signal serves as the desired response of the adaptive filter. This type of application is used in adaptive noise cancellation [2] , or adaptive beamforming or adaptive array processing [28].

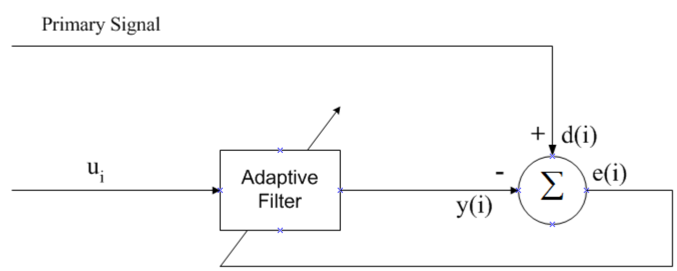

Figure 1.3: Noise Cancellation Scenario

1.2.4 System Identification

System Identification is the experimental approach to the modeling of a process or a plant. It involoves the following steps: experimental planning, the selection of a model of a structure, parameter estimation and model validation. The procedure of system identification, as pursued in practice, is iterative in nature in that we may have to go back and forth in these steps until a satisfactory model is built. The system to be identified is unknown which can be stationary or time varying. Figure (1.4) depicts the system identification scenario.

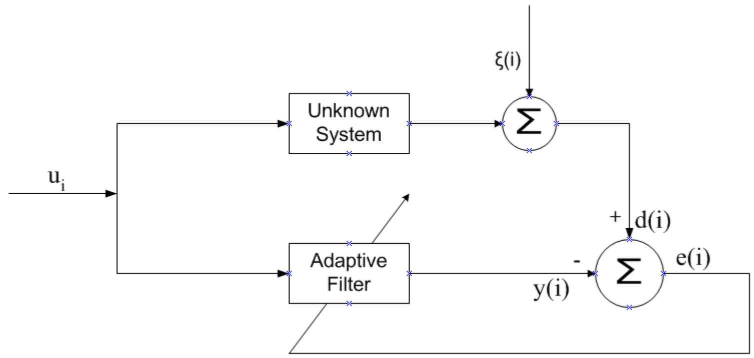

Figure 1.4: System Identification Scenario

Now we will discuss briefly the idea of adaptive filtering algorithms for estimating the parameters of an unknown system modeled as a transversal filter. Suppose that we have an unknown dynamic system that is linear and time varying. The system is characterized by a real valued set of discrete time measurements that describe the variations of the system output in response to a known stationary input.

The requirement is to develop an on-line transversal filter model for this plant. The model conists of finite number of unit delay elements and a corresponding set of adjustable parameters (tap weights). Let the available input sequence at time i be denoted by the set of samples: $u_i, u_{i-1}, \ldots, u_{i-M+1}$, where M is the number of adjustable parameters in the model. The input sequence is applied at the same time to the system and to the model. Let their outputs be denoted by d_i and

y_i respectively. The system output d_i serves as the purpose of desired response for the adaptive filtering algorithm employed to adjust the model paramters. The model output is given by:

$$y_i = \sum_{N=0}^{M-1} w_{ni} u_{i-n}, \tag{1.1}$$

where $w_{0i}, w_{1i}, \ldots, w_{(M-1)i}$ are the estimated model parameters at the ith iteration. The model output y_i is compared with the system output d_i. The difference between them defines the modeling (estimation) error. Let this error be denoted by e_i and defined as follows:

$$e_i = d_i - y_i. \tag{1.2}$$

Typically, at iteration i, the modeling error e_i is nonzero, implying that the model deviates from the system. In an attempt to account for this deviation, the error e_i is applied to an adaptive control algorithm. The samples of the input sequence $u_i, u_{i-1}, \ldots, u_{i-M+1}$, are also applied to the algorithm. The combination of the transversal filter and the adaptive control algorithm constitutes the adaptive filtering algorithm. The algorithm is designed to control the adjustment made in the values of the parameters of the model. As a result, the model parameter assumes a new set of values for use in the next iteration. Thus, at iteration $i + 1$, a new model output is computed, and with it a new value for the modeling error. The operation described is then repeated. This process is continued for a sufficiently large number of iterations (starting from $i = 0$), until the deviation of the model from the system, measured by the magnitude of the modeling error e_i, becomes sufficiently small in a statistical sense.

When the system is time varying, the system output is non-stationary, and so is the desired response presented to the adaptive filtering algorithm. In such a situation, the adaptive filter has the task of not only keeping the modeling error small but also continually tracking the time variations in the dynamics of the system.

1.3 Adaptive Filtering Algorithms

An adaptive algorithm refers to the criteria by which a filter is adapted in response to the outside environment. Let $\boldsymbol{w}_i$ be a vector of length L whose elements represent a time-varying finite impulse response of the adaptive filter. A general form for the algorithm that adapts the filter coefficient vector $\boldsymbol{w}_i$ is given by

$$\boldsymbol{w}_{i+1} = \boldsymbol{w}_i + f(\boldsymbol{u}_i, e(i), \mu), \tag{1.3}$$

where $\boldsymbol{u}_i$ is the input sequence, $e(i)$ is the adaptive error, and μ denotes the algorithm step-size which may be time varying, and f is a function of all these quantities. As mentioned previously, adaptive algorithms can be classified into two main categories: the least squares algorithms, and the least mean algorithms

[29].

1.3.1 Least Squares Algorithms

Algorithms based on the method of least squares can be classified into three major classes [5] as follows:

1. Recursive least squares (RLS) algorithm.
2. QR-decomposition based recursive least squares algorithm (QRD-RLS).
3. Fast algorithms:
    - Fast transversal filters algorithm (FTF),
    - Recursive least squares lattice algorithm (LSL), and
    - QR-decomposition based least squares lattice algorithm (QRD-LSL).

Among all the above mentioned adaptive least squares algorithm, the most popular and widely used one is the RLS algorithm. Thus, only RLS is discussed here.

Recursive Least-Squares (RLS) Algorithm

In this class of algorithm, the minimization cost function is defined as

$$J_n = \sum_{i=1}^{n} \beta_{(n,i)} [e(i)]^2, \tag{1.4}$$

where the sequence $\beta_{(n,i)}$ represents a weighting function to gradually fade out the effect of previous data and to make the finite dimension arithmetic possible [30]. One possible choice for $\beta_{(n,i)}$ is the exponential weighting defined as [5]

$$\beta_{(n,i)} = \lambda^{n-i}, \tag{1.5}$$

where λ is a positive constant close to, but less than one [5]. The adaptation algorithm should read [5]

$$\boldsymbol{w}_{i+1} = \boldsymbol{w}_i + \boldsymbol{k}_{i+1}\alpha^*_{\mathbf{i+1}} \tag{1.6}$$

where $\boldsymbol{k}_i$ is the gain vector, while $\alpha_{\mathbf{i}}$ is the innovation. They are defined in terms of the autocorrelation matrix, $\mathbf{R} = E\boldsymbol{u}_i^*\boldsymbol{u}_i$, and the desired response $d(i)$ as follows [5]:

$$k_i = \mathbf{R}_i^{-1}\boldsymbol{u}_i, \tag{1.7}$$

and

$$\alpha_i = \boldsymbol{d}(i) - \boldsymbol{w}_i^*\boldsymbol{u}_i. \tag{1.8}$$

The main advantage of the RLS algorithm is its faster convergence irrespective of the input statistics. But the main problem with RLS algorithm is its computational complexity.

1.4 Least Mean Algorithms

Instead of minimizing the time average of the error signal $e(i)$ as in the RLS algorithm, the least mean class of algorithms minimizes a statistical average of the error, i.e., $E[f(e(i))]$, a convex function of the filter coefficients $\boldsymbol{w}_i$. Thus, $\boldsymbol{w}_i$ can be adapted using the steepest-descent algorithm as follows:

$$\boldsymbol{w}_{i+1} = \boldsymbol{w}_i - \mu \Delta E[f(e(i))], \tag{1.9}$$

where $\Delta E[f(e(i))]$ represents the gradient of $E[f(e(i))]$ with respect to $\boldsymbol{w}_i$. Since, a closed form of this gradient is not available, it is replaced by its stochastic approximation, i.e., $\Delta E[f(e(i))]$. Thus, different least mean algorithms are obtained for each choice of the function $f(e(i))$. The well known of them are discussed next.

Least Mean Square (LMS) Algorithm

If $f(e(i)) = e(i)$, the least mean squares (LMS) algorithm is obtained which is one of the most popular algorithms in adaptive filtering. According to LMS algorithm, the filter coefficients are adapted according to the following recursion:

$$\boldsymbol{w}_{i+1} = \boldsymbol{w}_i + \mu \boldsymbol{u}_i^* e(i), \tag{1.10}$$

where the error signal $e(i)$ is defined in (1.2), $\boldsymbol{u}_i$ is the tap input vector, and $\boldsymbol{w}_{i+1}$ represents the tap weights of the adaptive filter. The parameter μ is a positive constant called step size which is used to control the size of the incremental correction applied to the tap weights as we proceed from one iteration to the next.

The LMS algorithm is simple to implement and yet capable of achieving satisfactory performance under the right conditions. Its major limitations are a relatively slow rate of convergence and a sensitivity to variations in the condition number of correlation matrix of the input signal[1].

In a non-stationary environment, the orientation of the error-performance surface varies continuously with time. In this case, the LMS algorithm has the added task of continually tracking the bottom of the error performance surface. Indeed, tracking will occur provided that the input data varies slowly compared to the learning rate of the LMS algorithm [5].

Least Mean Fourth (LMF) Algorithm

This algorithm consists of minimizing the fourth power of the error. In fact, it is a special case of the more general family of the steepest decent algorithms with $2k$ error norms [5], where k is a positive integer, and its weight update equation is defined as

$$\boldsymbol{w}_{i+1} = \boldsymbol{w}_i + \mu \boldsymbol{u}_i^* e^3(i). \tag{1.11}$$

The LMF algorithm has a faster convergence as compared to LMS but has higher steady state error. The complexity of the LMF algorithm is more as compared to the LMS algorithm because of the higher power of $e(i)$ involved in the

[1]The condition number of a hermitian matrix is defined as the ratio of its largest eigenvalue to its smallest eigenvalue.

adaptation of the weights.

Least Mean *p*-power Algorithm

Recently, other minimization criteria have emerged, in which adaptive structures are derived from the minimization of a class of functions of the form [21]

$$J_n = E[e^p(i)], \tag{1.12}$$

where p is an integer constant. It is seen from (1.12) that when $p = 2$, it is reduced to the cost function of LMS algorithm while LMF when $p = 4$.

Least Mean Mixed Norm (LMMN) Algorithm

It has been seen that the LMS algorithm can achieve the lower steady-state error floor than the LMF algorithm but its convergence speed is slower. So, knowing the fact that the convex addition of two convex function is also a convex function, a class of mixed norm algorithm [25] was developed with the following cost function:

$$J_n = \gamma E[e^2(i)] + (1 - \gamma)E[e^4(i)], \tag{1.13}$$

where γ is the mixing parameter. The adaptation algorithm of (1.4) reduces to the LMS algorithm and LMF algorithm, respectively, for $\gamma = 1$ and $\gamma = 0$. Thus, it represents the generalization of both the LMS and LMF algorithm.

1.5 Normalized LMS (NLMS) Adaptive Algorithm

The LMS algorithm performs badly with correlated input signal like speech signals. The reason is that LMS algorithm is directly dependent on the input vector $\boldsymbol{u}_i$. Therefore, when the eigen-value spread of the autocorrelation is large, LMS algorithm experiences a gradient noise amplification. In order to overcome this problem, in traditional normalized LMS algorithm, the input signal is normalized by the input signal power. Thus, the filter coefficients are adapted according to the following recursion [13]:

$$\boldsymbol{w}_{i+1} = \boldsymbol{w}_i + \mu \boldsymbol{u}_i^* \frac{e(i)}{\|\boldsymbol{u}_i\|^2}. \tag{1.14}$$

Due to normalization, The NLMS algorithm is made to have more stable behaviour for a known range ($0 < \mu < 2$), less sensitive to the colored input signal (as the effect of the eigen-value spread of the input vector is reduced) [31], and converges faster than the LMS algorithm [31].

1.6 Thesis Objectives and Organization

This thesis provides the tracking analysis for both ϵ - NLMS and leaky ϵ - NLMS algorithms in the presence of nonstationary environment and a colored gaussian input data.

In Chapter 2, the tracking analysis of the ϵ - NLMS is derived, the environment model is assumed to be a nonstationry in the leaky sense, and the filter input is assumed to be a colored gaussian regressors, The energy conservation relation is derived in order to derive the steady-state excess mean square error of this algorithm.

Chapter 3 provides the tracking analysis of the leaky ϵ - NLMS algorithm for the same environment and the same input data, the overall procedure which is used in this chapter is almost similar to that in chapter 2 (derivation of the energy conservation relation followed by the steady-state excess mean square error), but the main differece is the usage of the weighted energy conservation realtion in order to derive the EMSE of the leaky ϵ - NLMS.

Simulation results are shown in Chapter 4, The algorithms tested in the presence of a white gaussian noise, the analytical results are compared to the experimental results in order to show the accuracy of the obtained analytical expressions. Finally, Thesis conclusions and recommendation for the futer work is presented in Chapter 5.

CHAPTER 2

TRACKING ANALYSIS OF ϵ-NLMS ALGORITHM

Since adaptive filters rely on the instantaneous data, one of the most important features of them is their ability to track the underlying charateristics of the signal. This means that the statistical properties of the signal error in addition to the weight vector will react to the changes in the input signal properties.

In this chapter, the tracking ability of the ϵ-NLMS algorithm is derived in a special environemnt. First, the model are presented in the next section, after that the energy conservation relation is derived for the general form of adaptive filters. Second, the steady-state EMSE is derived also for the general form of adaptive filters, and finally the case of the ϵ-NLMS is derived out of that form.

2.1 Nonstationary Data Model

The data model which describes the nonstationary environment in our tracking analysis is described by the following:

(a) There exists a vector $\boldsymbol{w}_i^o$ such that $\boldsymbol{d}(i) = \boldsymbol{u}_i \boldsymbol{w}_i^o + \boldsymbol{v}(i)$.

(b) The weight vector varies according to $\boldsymbol{w}_i^o = w^o + \boldsymbol{\theta}_i$.

(c) The perturbation varies according to $\boldsymbol{\theta}_{i+1} = \beta \boldsymbol{\theta}_i + \boldsymbol{q}_i$.

(d) The noise sequence $\{\boldsymbol{v}(i)\}$ is i.i.d with variance $\sigma_v^2 = \mathrm{E}\,|\boldsymbol{v}(i)|^2$.

(e) The noise sequence $\{\boldsymbol{v}(i)\}$ is independent of $\boldsymbol{u}_j$ for all i, j.

(f) The sequence $\boldsymbol{q}_i$ has covariance $\boldsymbol{Q}$ and is independent of $\{\boldsymbol{v}(j), \boldsymbol{u}_j\}$ for all i, j.

(g) The initial conditions $\{\boldsymbol{w}_{-1}, \boldsymbol{\theta}_{-1}\}$ are independent of all $\{\boldsymbol{d}(j), \boldsymbol{u}_j, \boldsymbol{v}(j), \boldsymbol{q}_j\}$.

(h) The regressor covariance matrix is denoted by $\boldsymbol{R}_u = \mathrm{E}\,\boldsymbol{u}_i^* \boldsymbol{u}_i > 0$.

(i) The coefficient β satisifies $|\beta| \leq 1$.

(j) The random variables $\{\boldsymbol{d}(i), \boldsymbol{u}_i, \boldsymbol{v}(i), \boldsymbol{q}_i\}$ are zero mean.

(k) The weight vector $\boldsymbol{w}_i^o$ has constant mean w^o.

2.2 Energy Conservation Relation

Let us consider the adaptive algorithms whose update are of the form:

$$\boldsymbol{w}_{i+1} = \boldsymbol{w}_i + \mu \boldsymbol{u}_i^* g[\boldsymbol{e}(i)], \qquad \boldsymbol{w}_{-1} = \text{initial condition} \tag{2.1}$$

where g[.] denotes different error function. Table 2.1 list some of the error functions with their corresponding algorithms.

Table 2.1: The LMS algorithm and some of its derivatives.

Algorithm	Error Functions
LMS	$e(i)$
Normalized LMS	$\frac{e(i)}{\|\boldsymbol{u}_i\|^2}$
Leaky LMS	$e(i)$
Sign-Data LMS	$e(i)sign(\mathbf{u}_i)$
Sign-Error LMS	$sign(e(i))\mathbf{u}_i$
Sign-Sign LMS	$sign(e(i))sign(\mathbf{u}_i)$

Now in order to write the update recursion (2.1) in terms of the weight error vector $\tilde{\boldsymbol{w}}_{i+1} = \boldsymbol{w}^o_{i+1} - \boldsymbol{w}_{i+1}$ we have to subtract from both sides of (2.1) $\boldsymbol{w}^o_{i+1}$ which gives

$$\boldsymbol{w}^o_{i+1} - \boldsymbol{w}_{i+1} = \boldsymbol{w}^o_{i+1} - \boldsymbol{w}_i - \mu \boldsymbol{u}^*_i g[\boldsymbol{e}(i)]. \tag{2.2}$$

Further, On the right hand side we have to subtract and add the term $\boldsymbol{w}^o_i$, so (2.1) becomes

$$\begin{aligned}\boldsymbol{w}_{i+1}^{o} - \boldsymbol{w}_{i+1} &= \boldsymbol{w}_{i+1}^{o} - \boldsymbol{w}_{i}^{o} + \boldsymbol{w}_{i}^{o} - \boldsymbol{w}_{i} - \mu \boldsymbol{u}_{i}^{*} g[\boldsymbol{e}(i)] \\ \tilde{\boldsymbol{w}}_{i+1} &= \boldsymbol{w}_{i+1}^{o} - \boldsymbol{w}_{i}^{o} + \tilde{\boldsymbol{w}}_{i} - \mu \boldsymbol{u}_{i}^{*} g[\boldsymbol{e}(i)]. \end{aligned} \tag{2.3}$$

Depending on point(b) of the mentioned data model, the term $\boldsymbol{w}_{i+1}^{o} - \boldsymbol{w}_{i}^{o}$ can be rewritten as

$$\begin{aligned}\boldsymbol{c}_{i+1} &= \boldsymbol{\theta}_{i+1} - \boldsymbol{\theta}_{i} \\ &= \boldsymbol{w}_{i+1}^{o} - \boldsymbol{w}_{i}^{o}. \end{aligned} \tag{2.4}$$

Using (2.4), (2.3) can be written as

$$\tilde{\boldsymbol{w}}_{i+1} = \tilde{\boldsymbol{w}}_{i} + \boldsymbol{c}_{i+1} - \mu \boldsymbol{u}_{i}^{*} g[\boldsymbol{e}(i)] \tag{2.5}$$

$$\tilde{\boldsymbol{w}}_{i+1} - \boldsymbol{c}_{i+1} = \tilde{\boldsymbol{w}}_{i} - \mu \boldsymbol{u}_{i}^{*} g[\boldsymbol{e}(i)]. \tag{2.6}$$

Next,we will define a priori and a posteriori estimation errors as

$$\boxed{\boldsymbol{e}_{a}(i) = \boldsymbol{u}_{i} \tilde{\boldsymbol{w}}_{i}, \qquad \boldsymbol{e}_{p}(i) = \boldsymbol{u}_{i} (\tilde{\boldsymbol{w}}_{i+1} - \boldsymbol{c}_{i+1})} \tag{2.7}$$

By multiplying both sides of equation (2.6) by $\boldsymbol{u}_i$ gives

$$\boldsymbol{e}_p(i) = \boldsymbol{e}_a(i) - \mu\|\boldsymbol{u}_i\|^2 g[\boldsymbol{e}(i)]. \tag{2.8}$$

By rearranging (2.8) we get

$$g[\boldsymbol{e}(i)] = \frac{\boldsymbol{e}_a(i) - \boldsymbol{e}_p(i)}{\mu\|\boldsymbol{u}_i\|^2}. \tag{2.9}$$

Then substituting (2.9) in (2.6) and rearranging the terms we find that

$$\boxed{(\tilde{\boldsymbol{w}}_{i+1} - \boldsymbol{c}_{i+1}) + \tfrac{\boldsymbol{u}_i^*}{\|\boldsymbol{u}_i\|^2}\boldsymbol{e}_a(i) = \tilde{\boldsymbol{w}}_i + \tfrac{\boldsymbol{u}_i^*}{\|\boldsymbol{u}_i\|^2}\boldsymbol{e}_p(i)} \tag{2.10}$$

Taking the squared euclidean norms of both sides (i.e. evaluating the energies), we get

$$\boxed{\|\tilde{\boldsymbol{w}}_{i+1} - \boldsymbol{c}_{i+1}\|^2 + \tfrac{1}{\|\boldsymbol{u}_i\|^2}|\boldsymbol{e}_a(i)|^2 = \|\tilde{\boldsymbol{w}}_i\|^2 + \tfrac{1}{\|\boldsymbol{u}_i\|^2}|\boldsymbol{e}_p(i)|^2} \tag{2.11}$$

When $\boldsymbol{u}_i = \boldsymbol{0}$, it is obiviously true that

$$\|\tilde{\boldsymbol{w}}_{i+1} - \boldsymbol{c}_{i+1}\|^2 = \|\tilde{\boldsymbol{w}}_i\|^2. \tag{2.12}$$

Both results, (2.11) and (2.12), can be grouped together into a single equation by defining

$$\bar{\mu}(i) = \begin{cases} \frac{1}{\|\boldsymbol{u}_i\|^2} & \text{for} \quad \|\boldsymbol{u}_i\|^2 \neq \mathbf{0} \\ 0 & \text{for} \quad \|\boldsymbol{u}_i\|^2 = \mathbf{0} \end{cases}$$

Thus, using the above definition, (2.11) can be defined as

$$\boxed{\|\tilde{\boldsymbol{w}}_{i+1} - \boldsymbol{c}_{i+1}\|^2 + \bar{\mu}(i)|\boldsymbol{e}_a(i)|^2 = \|\tilde{\boldsymbol{w}}_i\|^2 + \bar{\mu}(i)|\boldsymbol{e}_p(i)|^2} \tag{2.13}$$

2.3 Variance Relation

Equation (2.13) can be used to derive the variance relation. To do so, we start by expanding the term $|\boldsymbol{e}_p(i)|^2$. Using (2.8) gives

$$\begin{aligned} |\boldsymbol{e}_p(i)|^2 &= |\boldsymbol{e}_a(i) - \mu\|\boldsymbol{u}_i\|^2 g[\boldsymbol{e}(i)]|^2 \\ &= |\boldsymbol{e}_a(i)|^2 + \mu^2\|\boldsymbol{u}_i\|^4 g[\boldsymbol{e}(i)]^2 - 2\mu\|\boldsymbol{u}_i\|^2 \mathrm{Re}(\boldsymbol{e}_{\mathrm{a}}^*(\mathrm{i})\mathrm{g}[\boldsymbol{e}(\mathrm{i})]). \end{aligned} \tag{2.14}$$

Multiplying both sides by $\bar{\mu}(i)$ and substituting in (2.14) will cancel out the term $\bar{\mu}(i)|\boldsymbol{e}_a(i)|^2$, so we will endup with

$$\|\tilde{\boldsymbol{w}}_{i+1} - \boldsymbol{c}_{i+1}\|^2 = \|\tilde{\boldsymbol{w}}_i\|^2 + \mu^2\|\boldsymbol{u}_i\|^2 g[\boldsymbol{e}(i)]^2 - 2\mu \mathrm{Re}(\boldsymbol{e}_{\mathrm{a}}^*(\mathrm{i})\mathrm{g}[\boldsymbol{e}(\mathrm{i})]). \tag{2.15}$$

Now, let us expand the term $\|\tilde{\boldsymbol{w}}_{i+1} - \boldsymbol{c}_{i+1}\|^2$, this will give us

$$\|\tilde{\boldsymbol{w}}_{i+1} - \boldsymbol{c}_{i+1}\|^2 = \|\tilde{\boldsymbol{w}}_{i+1}\|^2 + \|\boldsymbol{c}_{i+1}\|^2 - 2\mathrm{Re}(\boldsymbol{c}_{\mathrm{i+1}}^*\tilde{\boldsymbol{w}}_{\mathrm{i+1}}). \tag{2.16}$$

Substituting (2.16) in (2.15) and re-arranging the terms we get

$$\|\tilde{\boldsymbol{w}}_{i+1}\|^2 = \|\tilde{\boldsymbol{w}}_i\|^2 + \mu^2\|\boldsymbol{u}_i\|^2 g[\boldsymbol{e}(i)]^2 - 2\mu \mathrm{Re}(\boldsymbol{e}_{\mathrm{a}}^*(\mathrm{i})\mathrm{g}[\boldsymbol{e}(\mathrm{i})]) - \|\boldsymbol{c}_{i+1}\|^2 + 2\mathrm{Re}(\boldsymbol{c}_{\mathrm{i+1}}^*\tilde{\boldsymbol{w}}_{\mathrm{i+1}}). \tag{2.17}$$

In order to evaluate $\|\boldsymbol{c}_{i+1}\|^2$ we need to simplify (2.4) by using point (c) of the nonstationary data model as the following:

$$\begin{aligned} \boldsymbol{c}_{i+1} &= \boldsymbol{\theta}_{i+1} - \boldsymbol{\theta}_i \\ &= \beta\boldsymbol{\theta}_i - \boldsymbol{\theta}_i + \boldsymbol{q}_i \\ &= (\beta - 1)\boldsymbol{\theta}_i + \boldsymbol{q}_i \end{aligned} \tag{2.18}$$

Now the term $\|\boldsymbol{c}_{i+1}\|^2$ can be evaluated simply as

$$\begin{aligned}\|\boldsymbol{c}_{i+1}\|^2 &= \|(\beta-1)\boldsymbol{\theta}_i + \boldsymbol{q}_i\|^2 \\ &= \|(\beta-1)\boldsymbol{\theta}_i\|^2 + \|\boldsymbol{q}_i\|^2 + (\beta-1)\boldsymbol{q}_i^*\boldsymbol{\theta}_i + (\beta-1)^*\boldsymbol{\theta}_i{}^*\boldsymbol{q}_i \\ &= |\beta-1|^2\|\boldsymbol{\theta}_i\|^2 + \|\boldsymbol{q}_i\|^2 + (\beta-1)\boldsymbol{q}_i^*\boldsymbol{\theta}_i + (\beta-1)^*\boldsymbol{\theta}_i{}^*\boldsymbol{q}_i. \end{aligned} \tag{2.19}$$

In addation to the term $\|\boldsymbol{c}_{i+1}\|^2$, we need to evaluate the term $\mathrm{Re}(\boldsymbol{c}_{\mathrm{i+1}}^*\tilde{\boldsymbol{w}}_{\mathrm{i+1}})$, which can be done using (2.5) and (2.18) as follows:

$$\begin{aligned}\mathrm{Re}\{\boldsymbol{c}_{\mathrm{i+1}}^*\tilde{\boldsymbol{w}}_{\mathrm{i+1}}\} &= \mathrm{Re}\{((\beta-1)\boldsymbol{\theta}_{\mathrm{i}} + \boldsymbol{q}_{\mathrm{i}})^*(\tilde{\boldsymbol{w}}_{\mathrm{i}} + \boldsymbol{c}_{\mathrm{i+1}} - \mu\boldsymbol{u}_{\mathrm{i}}^*\mathrm{g}[\boldsymbol{e}(\mathrm{i})])\} \\ &= \mathrm{Re}\{((\beta-1)\boldsymbol{\theta}_{\mathrm{i}} + \boldsymbol{q}_{\mathrm{i}})^*(\tilde{\boldsymbol{w}}_{\mathrm{i}} + (\beta-1)\boldsymbol{\theta}_{\mathrm{i}} + \boldsymbol{q}_{\mathrm{i}} - \mu\boldsymbol{u}_{\mathrm{i}}^*\mathrm{g}[\boldsymbol{e}(\mathrm{i})])\} \\ &= |\beta-1|^2\|\boldsymbol{\theta}_i\|^2 + \|\boldsymbol{q}_i\|^2 + \mathrm{Re}\{((\beta-1)\boldsymbol{\theta}_{\mathrm{i}})^*(\tilde{\boldsymbol{w}}_{\mathrm{i}} + \boldsymbol{q}_{\mathrm{i}} - \mu\boldsymbol{u}_{\mathrm{i}}^*\mathrm{g}[\boldsymbol{e}(\mathrm{i})])\} \\ &\quad + \mathrm{Re}\{\boldsymbol{q}_{\mathrm{i}}{}^*(\tilde{\boldsymbol{w}}_{\mathrm{i}} + (\beta-1)\boldsymbol{\theta}_{\mathrm{i}} - \mu\boldsymbol{u}_{\mathrm{i}}^*\mathrm{g}[\boldsymbol{e}(\mathrm{i})])\}. \end{aligned} \tag{2.20}$$

In order to derive the variance relation, we take the expectaion operation on (2.17), To do so, let us take the expected value of (2.19) and (2.20). Starting with (2.19) gives

$$
\begin{aligned}
E\|\boldsymbol{c}_{i+1}\|^2 &= E|\beta-1|^2\|\boldsymbol{\theta}_i\|^2 + E\|\boldsymbol{q}_i\|^2 + E(\beta-1)\boldsymbol{q}_i^*\boldsymbol{\theta}_i + E(\beta-1)^*\boldsymbol{\theta}_i{}^*\boldsymbol{q}_i \\
&= |\beta-1|^2 E\|\boldsymbol{\theta}_i\|^2 + E\|\boldsymbol{q}_i\|^2 + (\beta-1)E\boldsymbol{q}_i^*\boldsymbol{\theta}_i + (\beta-1)E\boldsymbol{\theta}_i^*\boldsymbol{q}_i \\
&= |\beta-1|^2 E\|\boldsymbol{\theta}_i\|^2 + E\|\boldsymbol{q}_i\|^2 + (\beta-1)E\boldsymbol{q}_i^*\left(\beta^{i+1}\boldsymbol{\theta}_{-1} + \sum_{j=1}^{i+1}\beta^{(j-1)}\boldsymbol{q}_{(i-j)}\right) \\
&\quad +(\beta-1)E\left(\beta^{i+1}\boldsymbol{\theta}_{-1} + \sum_{j=1}^{i+1}\beta^{(j-1)}\boldsymbol{q}_{(i-j)}\right)^*\boldsymbol{q}_i. \qquad (2.21)
\end{aligned}
$$

But since $\boldsymbol{q}_i$ is independent of all previous $\boldsymbol{q}_j$ and of $\boldsymbol{\theta}_{-1}$, the last two terms will vanish and therfore (2.21) lookslike

$$
E\|\boldsymbol{c}_{i+1}\|^2 = |\beta-1|^2 E\|\boldsymbol{\theta}_i\|^2 + E\|\boldsymbol{q}_i\|^2. \qquad (2.22)
$$

Now, taking the expected value of (2.20) and using the same methodology used above gives

$$
\begin{aligned}
\mathrm{Re}\{\mathrm{E}\boldsymbol{c}_{\mathrm{i+1}}^*\tilde{\boldsymbol{w}}_{\mathrm{i+1}}\} &= |\beta-1|^2 E\|\boldsymbol{\theta}_i\|^2 + E\|\boldsymbol{q}_i\|^2 + \mathrm{Re}\{\mathrm{E}((\beta-1)\boldsymbol{\theta}_{\mathrm{i}})^*(\tilde{\boldsymbol{w}}_{\mathrm{i}} - \mu\boldsymbol{u}_{\mathrm{i}}^*\mathrm{g}[\boldsymbol{e}(\mathrm{i})])\} \\
&\quad +\mathrm{Re}\{\mathrm{E}\boldsymbol{q}_{\mathrm{i}}{}^*(\tilde{\boldsymbol{w}}_{\mathrm{i}} - \mu\boldsymbol{u}_{\mathrm{i}}^*\mathrm{g}[\boldsymbol{e}(\mathrm{i})])\}. \qquad (2.23)
\end{aligned}
$$

The last term of (2.23) will vanish due to the fact that the term $(\tilde{\boldsymbol{w}}_i - \mu\boldsymbol{u}_i^* g[\boldsymbol{e}(i)])$ is a function of $\{\boldsymbol{w}_{-1}, \boldsymbol{u}_{i-1}, \ldots, \boldsymbol{u}_0, \boldsymbol{d}(i-1), \ldots, \boldsymbol{d}(0)\}$ and $\boldsymbol{q}_i$ is independent of all these variables. Thus (2.23) becomes

$$\mathrm{Re}\{\mathrm{E}\boldsymbol{c}^*_{\mathrm{i+1}}\tilde{\boldsymbol{w}}_{\mathrm{i+1}}\} = |\beta-1|^2 E\|\boldsymbol{\theta}_i\|^2 + E\|\boldsymbol{q}_i\|^2 + \mathrm{Re}\{(\beta-1)^*\mathrm{E}\,(\boldsymbol{\theta}_{\mathrm{i}}{}^*(\tilde{\boldsymbol{w}}_{\mathrm{i}} - \mu\boldsymbol{u}^*_{\mathrm{i}}\mathrm{g}[\boldsymbol{e}(\mathrm{i})]))\} \tag{2.24}$$

Finally, the expected value of (2.17), substituting (2.22) and (2.24) and combining the common terms gives

$$\begin{aligned} E\|\tilde{\boldsymbol{w}}_{i+1}\|^2 &= E\|\tilde{\boldsymbol{w}}_i\|^2 + \mu^2 E(\|\boldsymbol{u}_i\|^2 g[\boldsymbol{e}(i)]^2) - 2\mu\mathrm{Re}\{\mathrm{E}(\boldsymbol{e}^*_{\mathrm{a}}(\mathrm{i})\mathrm{g}[\boldsymbol{e}(\mathrm{i})])\} \\ &\quad + |\beta-1|^2 E\|\boldsymbol{\theta}_i\|^2 + \mathrm{tr}(\boldsymbol{Q}) \\ &\quad + 2\mathrm{Re}\{(\beta-1)^*\mathrm{E}(\boldsymbol{\theta}_{\mathrm{i}}{}^*(\tilde{\boldsymbol{w}}_{\mathrm{i}} - \mu\boldsymbol{u}^*_{\mathrm{i}}\mathrm{g}[e(\mathrm{i})]))\}, \end{aligned} \tag{2.25}$$

where we have used the fact that $E\|\boldsymbol{q}_i\|^2 = \mathrm{tr}(\boldsymbol{Q})$.

2.4 Steady State Performance

We are interested in using the variance relation (2.25) to evaluate the excess mean square error (EMSE) of an adaptive filter at steady state. It is known that $E\|\tilde{\boldsymbol{w}}_{i+1}\|^2 = E\|\tilde{\boldsymbol{w}}_i\|^2$ at steady state. By applying this fact to (2.25) and rearranging the terms we get

$$\begin{aligned} 2\mu \mathrm{Re}\{\mathrm{E}(\boldsymbol{e}_{\mathrm{a}}^{*}(\mathrm{i})\mathrm{g}[\boldsymbol{e}(\mathrm{i})])\} &= \mu^2 E(\|\boldsymbol{u}_i\|^2 g[\boldsymbol{e}(i)]^2) + |\beta-1|^2 E\|\boldsymbol{\theta}_i\|^2 + \mathrm{tr}(\boldsymbol{Q}) \\ &\quad +2\mathrm{Re}\{(\beta-1)^* \mathrm{E}(\boldsymbol{\theta}_{\mathrm{i}}{}^{*}(\tilde{\boldsymbol{w}}_{\mathrm{i}} - \mu \boldsymbol{u}_{\mathrm{i}}^{*}\mathrm{g}[\boldsymbol{e}(\mathrm{i})]))\}. \end{aligned} \tag{2.26}$$

At steady state, it can be easily verified that

$$\begin{aligned} \lim_{i\to\infty} E(\boldsymbol{\theta}_i \boldsymbol{\theta}_i^*) &= \frac{\boldsymbol{Q}}{1-|\beta|^2} \\ &= \boldsymbol{\Theta}. \end{aligned} \tag{2.27}$$

Using (2.27), (2.26) becomes

$$\begin{aligned} 2\mu \mathrm{Re}\{\mathrm{E}(\boldsymbol{e}_{\mathrm{a}}^{*}(\mathrm{i})\mathrm{g}[\boldsymbol{e}(\mathrm{i})])\} &= \mu^2 E(\|\boldsymbol{u}_i\|^2 g[\boldsymbol{e}(i)]^2) + |\beta-1|^2 \mathrm{tr}(\boldsymbol{\Theta}) + \mathrm{tr}(\boldsymbol{Q}) \\ &\quad +2\mathrm{Re}\{(\beta-1)^* \mathrm{E}(\boldsymbol{\theta}_{\mathrm{i}}{}^{*}(\tilde{\boldsymbol{w}}_{\mathrm{i}} - \mu \boldsymbol{u}_{\mathrm{i}}^{*}\mathrm{g}[\boldsymbol{e}(\mathrm{i})]))\}. \end{aligned} \tag{2.28}$$

2.5 ϵ-NLMS Tracking Analysis

In the previous sections, we derived the variance relation for any adaptive filter that statisfies the recursion (2.1). In this section, we will derive an expression for the EMSE of the ϵ-NLMS algorithm.

For ϵ-NLMS algorithm we have:

$$\begin{aligned} g[\boldsymbol{e}(i)] &= \frac{\boldsymbol{e}(i)}{\left(\varepsilon + \|\boldsymbol{u}_i\|^2\right)} \\ &= \frac{\boldsymbol{d}(i) - \boldsymbol{u}_i \boldsymbol{w}_i}{\left(\varepsilon + \|\boldsymbol{u}_i\|^2\right)} \\ &= \frac{\boldsymbol{u}_i \boldsymbol{w}_i^o - \boldsymbol{u}_i \boldsymbol{w}_i + \boldsymbol{v}(i)}{\left(\varepsilon + \|\boldsymbol{u}_i\|^2\right)} \\ &= \frac{\boldsymbol{u}_i \tilde{\boldsymbol{w}}_i + \boldsymbol{v}(i)}{\left(\varepsilon + \|\boldsymbol{u}_i\|^2\right)} \\ &= \frac{\boldsymbol{e}_a(i) + \boldsymbol{v}(i)}{\left(\varepsilon + \|\boldsymbol{u}_i\|^2\right)}. \end{aligned} \tag{2.29}$$

Since $\{\boldsymbol{v}(i)\}$ is i.i.d with zero mean and statistically independent of the regressor sequence $\boldsymbol{u}_i$ then it can be shown that $\{\boldsymbol{v}(i)\}$ is also independent of $\boldsymbol{e}_a(i)$. Using this fact and by substituting (2.29) in the relation(2.28), we get

$$\begin{aligned} 2\mu E\left\{\frac{|\boldsymbol{e}_a(i)|^2}{\left(\varepsilon + \|\boldsymbol{u}_i\|^2\right)}\right\} &= \mu^2 E\left(\|\boldsymbol{u}_i\|^2 \frac{[|\boldsymbol{e}_a(i)|^2 + |\boldsymbol{v}(i)|^2]}{\left(\varepsilon + \|\boldsymbol{u}_i\|^2\right)^2}\right) + |\beta - 1|^2 \mathrm{tr}(\boldsymbol{\Theta}) + \mathrm{tr}(\mathbf{Q}) \\ &\quad + 2\mathrm{Re}\left\{(\beta - 1)^* \mathrm{E}\left(\boldsymbol{\theta}_{\mathrm{i}}{}^*(\tilde{\boldsymbol{w}}_{\mathrm{i}} - \mu \frac{\boldsymbol{u}_{\mathrm{i}}^* \boldsymbol{u}_{\mathrm{i}}}{\left(\varepsilon + \|\boldsymbol{u}_{\mathrm{i}}\|^2\right)} \tilde{\boldsymbol{w}}_{\mathrm{i}})\right)\right\} \\ &= \mu^2 E\left(\frac{\|\boldsymbol{u}_i\|^2 |\boldsymbol{e}_a(i)|^2}{\left(\varepsilon + \|\boldsymbol{u}_i\|^2\right)^2}\right) + \mu^2 \sigma_v^2 E\left(\frac{\|\boldsymbol{u}_i\|^2}{\left(\varepsilon + \|\boldsymbol{u}_i\|^2\right)^2}\right) + |\beta - 1|^2 \mathrm{tr}(\boldsymbol{\Theta}) \\ &\quad + \mathrm{tr}(\mathbf{Q}) + 2\mathrm{Re}\left\{(\beta - 1)^* \mathrm{E}\left(\boldsymbol{\theta}_{\mathrm{i}}{}^*(\mathbf{I} - \mu \frac{\boldsymbol{u}_{\mathrm{i}}^* \boldsymbol{u}_{\mathrm{i}}}{\left(\varepsilon + \|\boldsymbol{u}_{\mathrm{i}}\|^2\right)}) \tilde{\boldsymbol{w}}_{\mathrm{i}}\right)\right\} \\ &= \mu^2 E\left(\frac{\|\boldsymbol{u}_i\|^2 |\boldsymbol{e}_a(i)|^2}{\left(\varepsilon + \|\boldsymbol{u}_i\|^2\right)^2}\right) + \mu^2 \sigma_v^2 E\left(\frac{\|\boldsymbol{u}_i\|^2}{\left(\varepsilon + \|\boldsymbol{u}_i\|^2\right)^2}\right) + |\beta - 1|^2 \mathrm{tr}(\boldsymbol{\Theta}) \\ &\quad + \mathrm{tr}(\mathbf{Q}) + 2\mathrm{Re}\left\{(\beta - 1)^* \mathrm{E}\left(\boldsymbol{\theta}_{\mathrm{i}}{}^*(\mathbf{I} - \mu \mathbf{R}_{\mathrm{n}}) \tilde{\boldsymbol{w}}_{\mathrm{i}}\right)\right\} \end{aligned} \tag{2.30}$$

where

$$\mathbf{R}_n \triangleq \ \mathbf{\Gamma} \triangleq \ E\left(\frac{\boldsymbol{u}_i^* \boldsymbol{u}_i}{\left(\varepsilon + \|\boldsymbol{u}_i\|^2\right)}\right). \tag{2.31}$$

2.5.1 Separation Principile

Separation Principle states that

$$\boxed{\text{At steady state } \|\boldsymbol{u}_i\|^2 \text{ is statistically independent of } |\boldsymbol{e}_a(i)|^2} \tag{2.32}$$

By imposing this principle to (2.30) we get

$$\begin{aligned} 2\mu E\left\{\frac{1}{\left(\varepsilon + \|\boldsymbol{u}_i\|^2\right)}\right\} E|\boldsymbol{e}_a(i)|^2 &= \mu^2 E\left(\frac{\|\boldsymbol{u}_i\|^2}{\left(\varepsilon + \|\boldsymbol{u}_i\|^2\right)^2}\right) E|\boldsymbol{e}_a(i)|^2 \\ &+\mu^2\sigma_v^2 E\left(\frac{\|\boldsymbol{u}_i\|^2}{\left(\varepsilon + \|\boldsymbol{u}_i\|^2\right)^2}\right) + |\beta-1|^2 \mathrm{tr}(\mathbf{\Theta}) \\ &+\mathrm{tr}(\mathbf{Q}) + 2\mathrm{Re}\left\{(\beta-1)^* \mathrm{E}\left(\boldsymbol{\theta}_{\mathrm{i}}{}^*(\mathbf{I}-\mu\mathbf{R}_{\mathrm{n}})\tilde{\boldsymbol{w}}_{\mathrm{i}}\right)\right\} \end{aligned} \tag{2.33}$$

$$
\begin{aligned}
2\mu\eta_u E|\boldsymbol{e}_a(i)|^2 &= \mu^2\alpha_u E|\boldsymbol{e}_a(i)|^2 + \mu^2\sigma_v^2\alpha_u + |\beta-1|^2\mathrm{tr}(\boldsymbol{\Theta}) \\
&\quad + \mathrm{tr}(\mathbf{Q}) + 2\mathrm{Re}\left\{(\beta-1)^*\mathrm{E}\left(\boldsymbol{\theta}_{\mathrm{i}}{}^*(\mathbf{I}-\mu\mathbf{R}_{\mathrm{n}})\tilde{\boldsymbol{w}}_{\mathrm{i}}\right)\right\} \\
(2\mu\eta_u - \mu^2\alpha_u)E|\boldsymbol{e}_a(i)|^2 &= \mu^2\sigma_v^2\alpha_u + |\beta-1|^2\mathrm{tr}(\boldsymbol{\Theta}) + \mathrm{tr}(\mathbf{Q}) \\
&\quad + 2\mathrm{Re}\left\{(\beta-1)^*\mathrm{E}\left(\boldsymbol{\theta}_{\mathrm{i}}{}^*(\mathbf{I}-\mu\mathbf{R}_{\mathrm{n}})\tilde{\boldsymbol{w}}_{\mathrm{i}}\right)\right\} \\
\zeta^{\varepsilon-NLMS} &= \left(\frac{1}{2\mu\eta_u - \mu^2\alpha_u}\right)\Bigg(\mu^2\sigma_v^2\alpha_u + |\beta-1|^2\mathrm{tr}(\boldsymbol{\Theta}) + \mathrm{tr}(\mathbf{Q}) \\
&\quad + 2\mathrm{Re}\left\{(\beta-1)^*\mathrm{E}\left(\boldsymbol{\theta}_{\mathrm{i}}{}^*(\mathbf{I}-\mu\mathbf{R}_{\mathrm{n}})\tilde{\boldsymbol{w}}_{\mathrm{i}}\right)\right\}\Bigg), \qquad (2.34)
\end{aligned}
$$

where

$$
\eta_u \triangleq E\left(\frac{1}{(\varepsilon + \|\boldsymbol{u}_i\|^2)}\right), \qquad \alpha_u \triangleq E\left(\frac{\|\boldsymbol{u}_i\|^2}{(\varepsilon + \|\boldsymbol{u}_i\|^2)^2}\right) \qquad (2.35)
$$

The cross correlation $E(\boldsymbol{\theta}_i{}^*\tilde{\boldsymbol{w}}_i)$ which appears in (2.34) can be computed by the same procedure which was followed by [32] (especially Lemma 2), this gives us the following result:

$$
\begin{aligned}
\zeta^{\varepsilon-NLMS} &= \left(\frac{1}{2\mu\eta_u - \mu^2\alpha_u}\right)\Bigg(\mu^2\sigma_v^2\alpha_u + |\beta-1|^2\mathrm{tr}(\boldsymbol{\Theta}) + \mathrm{tr}(\mathbf{Q}) \\
&\quad + 2\mathrm{Re}\left\{\mathrm{tr}\left[(\beta-1)^*(\mathbf{I}-\mu\mathbf{R}_{\mathrm{n}})\mathbf{W}\right]\right\}\Bigg), \qquad (2.36)
\end{aligned}
$$

where

$$\mathbf{W} = [\beta^*(\mathbf{I} - \mu\mathbf{\Gamma}) - \mathbf{I}]^{-1}\mathbf{C}$$

$$\mathbf{C} = \beta^*(1-\beta)\mathbf{\Theta} - \mathbf{Q}. \tag{2.37}$$

Equation (2.36) can be further simplified to be

$$\zeta^{\varepsilon-NLMS} = \frac{\mu\sigma_v^2\alpha_u - \mu^{-1}\psi}{2\eta_u - \mu\alpha_u}, \tag{2.38}$$

where

$$\begin{aligned}\psi = {} & |\beta-1|^2\mathrm{Re}\Big[\mathrm{tr}(\mathbf{\Theta}(\mathbf{I} - 2\beta^*\mathbf{X}_\alpha))\Big] \\ & +\mathrm{Re}\Big[\mathrm{tr}(\mathbf{Q}(\mathbf{I} - 2(\beta-1)^*\mathbf{X}_\alpha))\Big],\end{aligned} \tag{2.39}$$

where

$$\mathbf{X}_\alpha \triangleq (\mathbf{I} - \mu\mathbf{R_n})[\beta^*(\mathbf{I} - \mu\mathbf{\Gamma}) - \mathbf{I}]^{\mathbf{-1}}. \tag{2.40}$$

It is clear from (2.38) that the tracking analysis of the ϵ - NLMS algorithm depends on the moments α_u and η_u in addition to the matrix identified by (2.31). These moments are given in the Appendix of [38].

2.6 Summary

In this chapter, the tracking analysis of the ϵ - NLMS algorithm in a nonstationary environment was investigated. Here, it was assumed that the environment is time varying according to a first order auto regressive model and the filter input is correlated gaussian, Finally an exact expression for the steady state EMSE of the ϵ-NLMS algorithm is derived based on the energy conservation relation.

CHAPTER 3

TRACKING ANALYSIS OF LEAKY ϵ-NLMS ALGORITHM

As it was done in the previous chapter, here we derive the steady-state performance of the leaky ϵ-NLMS algorithm in the presence of the white gaussian noise in addition to the changing environment. In order to do that, we start by mentionining the nonstationary model, then the weighted energy conservation relation is derived in order to derive a general leaky adaptive filters formula for the steady state excess mean square error, finallay, the case of the leaky ϵ-NLMS is considered as a special case of the general formula and the EMSE of the leaky ϵ-NLMS is derived.

3.1 Nonstationary Data Model

Since there are no changes in our assumptions about the environment in which the adaptive filter will operate on, we will use the same data model which was used in Chapter 2.

3.2 Energy Conservation Relation

Let us consider adaptive algorithms whose update equations are of the form

$$\boldsymbol{w}_{i+1} = (1-\alpha\mu)\boldsymbol{w}_i + \mu\boldsymbol{u}_i^* \frac{\boldsymbol{e}(i)}{g[\boldsymbol{u}_i]}, \qquad \boldsymbol{w}_{-1} = \text{initial condition} \tag{3.1}$$

where g[.] denotes different error function, which differs according to the algorithm itself.

By using the same methodology which was used in the previous chapter, we will subtract from both sides of (3.1) $\boldsymbol{w}_{i+1}^o$, this will give us the same recursion but in terms of the weight error vector $\tilde{\boldsymbol{w}}_{i+1}$

$$\begin{aligned}
\boldsymbol{w}_{i+1}^o - \boldsymbol{w}_{i+1} &= \boldsymbol{w}_{i+1}^o - (1-\alpha\mu)\boldsymbol{w}_i - \mu\boldsymbol{u}_i^* \frac{\boldsymbol{e}(i)}{g[\boldsymbol{u}_i]} \\
\tilde{\boldsymbol{w}}_{i+1} &= \alpha\mu\boldsymbol{w}_{i+1}^o + (1-\alpha\mu)\boldsymbol{w}_{i+1}^o - (1-\alpha\mu)\boldsymbol{w}_i - \mu\boldsymbol{u}_i^* \frac{\boldsymbol{e}(i)}{g[\boldsymbol{u}_i]} \\
\tilde{\boldsymbol{w}}_{i+1} &= \alpha\mu\boldsymbol{w}_{i+1}^o + (1-\alpha\mu)(\tilde{\boldsymbol{w}}_i + \boldsymbol{c}_{i+1}) - \mu\boldsymbol{u}_i^* \frac{\boldsymbol{e}(i)}{g[\boldsymbol{u}_i]}
\end{aligned} \tag{3.2}$$

where $\boldsymbol{c}_i$ is defined according to (1.4).

Now, we will take the weighted euclidean norms of both sides, we have:

$$\|\tilde{\boldsymbol{w}}_{i+1}\|_\Sigma^2 = \left\|\alpha\mu\boldsymbol{w}_{i+1}^o + (1-\alpha\mu)(\tilde{\boldsymbol{w}}_i + \boldsymbol{c}_{i+1}) - \mu\boldsymbol{u}_i^*\frac{\boldsymbol{e}(i)}{g[\boldsymbol{u}_i]}\right\|_\Sigma^2. \tag{3.3}$$

In order to solve the above equation, we will use the fact that $\boldsymbol{e}(i) = \boldsymbol{e}_a(i) + \boldsymbol{v}(i)$, so (3.3) becomes

$$\begin{aligned} \|\tilde{\boldsymbol{w}}_{i+1}\|_\Sigma^2 &= \left\|\alpha\mu\boldsymbol{w}_{i+1}^o + (1-\alpha\mu)(\tilde{\boldsymbol{w}}_i + \boldsymbol{c}_{i+1}) - \mu\boldsymbol{u}_i^*\frac{\boldsymbol{e}_a(i)}{g[\boldsymbol{u}_i]} - \mu\boldsymbol{u}_i^*\frac{\boldsymbol{v}(i)}{g[\boldsymbol{u}_i]}\right\|_\Sigma^2 \\ &= \left\|\alpha\mu\boldsymbol{w}_{i+1}^o + (1-\alpha\mu)(\tilde{\boldsymbol{w}}_i + \boldsymbol{c}_{i+1}) - \mu\boldsymbol{u}_i^*\frac{\boldsymbol{e}_a(i)}{g[\boldsymbol{u}_i]}\right\|_\Sigma^2 + \mu^2\frac{\|\boldsymbol{u}_i\|_\Sigma^2}{g^2[\boldsymbol{u}_i]}|\boldsymbol{v}(i)|^2 \\ &\quad -\frac{2\mu}{g[\boldsymbol{u}_i]}\mathrm{Re}\left\{\boldsymbol{u}_\mathrm{i}\boldsymbol{v}^*(\mathrm{i})\Sigma(\alpha\mu\boldsymbol{w}_{\mathrm{i+1}}^\mathrm{o} + (1-\alpha\mu)(\tilde{\boldsymbol{w}}_\mathrm{i} + \boldsymbol{c}_{\mathrm{i+1}}) - \mu\boldsymbol{u}_\mathrm{i}^*\frac{\boldsymbol{e}_\mathrm{a}(\mathrm{i})}{\mathrm{g}[\boldsymbol{u}_\mathrm{i}]})\right\}. \end{aligned} \tag{3.4}$$

3.3 Variance Relation

In order to calculate the variance relation, we apply the expectation operator to both sides of (3.4), to get

$$E\|\tilde{\boldsymbol{w}}_{i+1}\|_\Sigma^2 = \overbrace{E\left\|\alpha\mu\boldsymbol{w}_{i+1}^o + (1-\alpha\mu)(\tilde{\boldsymbol{w}}_i + \boldsymbol{c}_{i+1}) - \mu\boldsymbol{u}_i^*\frac{\boldsymbol{e}_a(i)}{g[\boldsymbol{u}_i]}\right\|_\Sigma^2}^{①} + \mu^2\sigma_v^2 E\frac{\|\boldsymbol{u}_i\|_\Sigma^2}{g^2[\boldsymbol{u}_i]}, \tag{3.5}$$

where the last term in (3.4) vanishes due to the following reasons:

- Since $\boldsymbol{w}_i^o = w^o + \boldsymbol{\theta}_i$ and $\boldsymbol{\theta}_i$ is a function of $\{\boldsymbol{\theta}_{-1}, \boldsymbol{q}_{-1}, \boldsymbol{q}_0, \ldots, \boldsymbol{q}_i\}$ and $\boldsymbol{v}(i)$ is independent of all of these variables, then $E\boldsymbol{u}_i\boldsymbol{v}^*(i)\Sigma\alpha\mu\boldsymbol{w}_{i+1}^o = 0$.

- Since $\boldsymbol{c}_i$ is a function of $\{\boldsymbol{\theta}_{-1}, \boldsymbol{q}_{-1}, \boldsymbol{q}_0, \ldots, \boldsymbol{q}_i\}$, and for the same reason mentioned in the previous point, then $E\boldsymbol{u}_i\boldsymbol{v}^*(i)\Sigma(1-\alpha\mu)\boldsymbol{c}_{i+1} = 0$.

- Finally, $\boldsymbol{v}(i)$ is independent of $\tilde{\boldsymbol{w}}_i$ and it's also independent of $\boldsymbol{e}_a(i)$, then $E\boldsymbol{u}_i\boldsymbol{v}^*(i)\Sigma(1-\alpha\mu)\tilde{\boldsymbol{w}}_i = 0$ and $E\boldsymbol{u}_i\boldsymbol{v}^*(i)\Sigma\mu\boldsymbol{u}_i^*\frac{e_a(i)}{g[\boldsymbol{u}_i]} = 0$.

Equation (3.5) can be further simplified by expanding the first term on the right hand side as follows:

$$
\begin{aligned}
① \quad = \quad & E\|\alpha\mu \boldsymbol{w}_{i+1}^{o}\|_{\Sigma}^{2} + E\left\|(1-\alpha\mu)(\tilde{\boldsymbol{w}}_i + \boldsymbol{c}_{i+1}) - \mu \boldsymbol{u}_i^{*}\frac{\boldsymbol{e}_a(i)}{g[\boldsymbol{u}_i]}\right\|_{\Sigma}^{2} + \\
& 2\alpha\mu\mathrm{Re}\left\{\mathrm{E}\boldsymbol{w}^{\mathrm{o}*}_{\mathrm{i+1}}\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_{\mathrm{i}} + (1-\alpha\mu)\boldsymbol{c}_{\mathrm{i+1}} - \mu\boldsymbol{u}_{\mathrm{i}}^{*}\frac{\boldsymbol{e}_{\mathrm{a}}(\mathrm{i})}{\mathrm{g}[\boldsymbol{u}_{\mathrm{i}}]}\right]\right\}
\end{aligned}
$$

$$
\begin{aligned}
= \quad & E\|\alpha\mu \boldsymbol{w}_{i+1}^{o}\|_{\Sigma}^{2} + E\left\|(1-\alpha\mu)\tilde{\boldsymbol{w}}_i - \mu \boldsymbol{u}_i^{*}\frac{\boldsymbol{e}_a(i)}{g[\boldsymbol{u}_i]}\right\|_{\Sigma}^{2} + E\left\|(1-\alpha\mu)\boldsymbol{c}_{i+1}\right\|_{\Sigma}^{2} \\
& +2(1-\alpha\mu)\mathrm{Re}\left\{\mathrm{E}\boldsymbol{c}_{\mathrm{i+1}}^{*}\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_{\mathrm{i}} - \mu\boldsymbol{u}_{\mathrm{i}}^{*}\frac{\boldsymbol{e}_{\mathrm{a}}(\mathrm{i})}{\mathrm{g}[\boldsymbol{u}_{\mathrm{i}}]}\right]\right\} \\
& +2\alpha\mu\mathrm{Re}\left\{\mathrm{E}\boldsymbol{w}^{\mathrm{o}*}_{\mathrm{i+1}}\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_{\mathrm{i}} + (1-\alpha\mu)\boldsymbol{c}_{\mathrm{i+1}} - \mu\boldsymbol{u}_{\mathrm{i}}^{*}\frac{\boldsymbol{e}_{\mathrm{a}}(\mathrm{i})}{\mathrm{g}[\boldsymbol{u}_{\mathrm{i}}]}\right]\right\}
\end{aligned}
$$

$$
\begin{aligned}
= \quad & E\|\alpha\mu \boldsymbol{w}_{i+1}^{o}\|_{\Sigma}^{2} + E\left\|(1-\alpha\mu)\tilde{\boldsymbol{w}}_i\right\|_{\Sigma}^{2} + \mu^2 E\left\|\boldsymbol{u}_i^{*}\frac{\boldsymbol{e}_a(i)}{g[\boldsymbol{u}_i]}\right\|_{\Sigma}^{2} \\
& -\mu(1-\alpha\mu)E\left[\tilde{\boldsymbol{w}}_i^{*}\Sigma\boldsymbol{u}_i^{*}\frac{\boldsymbol{e}_a(i)}{g[\boldsymbol{u}_i]}\right] - \mu(1-\alpha\mu)E\left[\frac{\boldsymbol{e}_a{}^{*}(i)}{g[\boldsymbol{u}_i]}\boldsymbol{u}_i\Sigma\tilde{\boldsymbol{w}}_i\right] \\
& +E\left\|(1-\alpha\mu)\boldsymbol{c}_{i+1}\right\|_{\Sigma}^{2} + 2(1-\alpha\mu)\mathrm{Re}\left\{\mathrm{E}\boldsymbol{c}_{\mathrm{i+1}}^{*}\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_{\mathrm{i}} - \mu\boldsymbol{u}_{\mathrm{i}}^{*}\frac{\boldsymbol{e}_{\mathrm{a}}(\mathrm{i})}{\mathrm{g}[\boldsymbol{u}_{\mathrm{i}}]}\right]\right\} \\
& +2\alpha\mu\mathrm{Re}\left\{\mathrm{E}\boldsymbol{w}^{\mathrm{o}*}_{\mathrm{i+1}}\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_{\mathrm{i}} + (1-\alpha\mu)\boldsymbol{c}_{\mathrm{i+1}} - \mu\boldsymbol{u}_{\mathrm{i}}^{*}\frac{\boldsymbol{e}_{\mathrm{a}}(\mathrm{i})}{\mathrm{g}[\boldsymbol{u}_{\mathrm{i}}]}\right]\right\}. \qquad (3.6)
\end{aligned}
$$

Using the weighted norm algebra, some terms of (3.6) can be further simplified as

- $E\left\|\boldsymbol{u}_i^{*}\frac{\boldsymbol{e}_a(i)}{g[\boldsymbol{u}_i]}\right\|_{\Sigma}^{2} = E\left\|\tilde{\boldsymbol{w}}_i\right\|^{2}_{\boldsymbol{u}_i^{*}\frac{\|\boldsymbol{u}_i\|_{\Sigma}^{2}}{g^{2}[\boldsymbol{u}_i]}\boldsymbol{u}_i}$,
- $E\left[\tilde{\boldsymbol{w}}_i^{*}\Sigma\boldsymbol{u}_i^{*}\frac{\boldsymbol{e}_a(i)}{g[\boldsymbol{u}_i]}\right] = E\|\tilde{\boldsymbol{w}}_i\|^{2}_{\Sigma\frac{\boldsymbol{u}_i^{*}\boldsymbol{u}_i}{g[\boldsymbol{u}_i]}}$,
- $E\left[\frac{\boldsymbol{e}_a{}^{*}(i)}{g[\boldsymbol{u}_i]}\boldsymbol{u}_i\Sigma\tilde{\boldsymbol{w}}_i\right] = E\|\tilde{\boldsymbol{w}}_i\|^{2}_{\frac{\boldsymbol{u}_i^{*}\boldsymbol{u}_i}{g[\boldsymbol{u}_i]}\Sigma}$,

In addition to the above terms, we can simplify the term $E\left\|(1-\alpha\mu)\boldsymbol{c}_{i+1}\right\|_{\Sigma}^{2}$ to become

$$E\left\|(1-\alpha\mu)\boldsymbol{c}_{i+1}\right\|_{\Sigma}^{2}=(1-\alpha\mu)^{2}\Big[|\beta-1|^{2}E\|\boldsymbol{\theta}_{i}\|_{\Sigma}^{2}+E\|\boldsymbol{q}_{i}\|_{\Sigma}^{2}\Big].$$

Substituting all the above in (3.6) gives

$$\begin{aligned}
① \;=\; & \alpha^{2}\mu^{2}E\|\boldsymbol{w}_{i+1}^{o}\|_{\Sigma}^{2}+(1-\alpha\mu)^{2}E\|\tilde{\boldsymbol{w}}_{i}\|_{\Sigma}^{2}+\mu^{2}E\,\|\tilde{\boldsymbol{w}}_{i}\|^{2}_{\boldsymbol{u}_i^{*}\frac{\|\boldsymbol{u}_i\|_{\Sigma}^{2}}{g^{2}[\boldsymbol{u}_i]}\boldsymbol{u}_i} \\
& -\mu(1-\alpha\mu)E\|\tilde{\boldsymbol{w}}_{i}\|^{2}_{\Sigma\frac{\boldsymbol{u}_i^{*}\boldsymbol{u}_i}{g[\boldsymbol{u}_i]}}-\mu(1-\alpha\mu)E\|\tilde{\boldsymbol{w}}_{i}\|^{2}_{\frac{\boldsymbol{u}_i^{*}\boldsymbol{u}_i}{g[\boldsymbol{u}_i]}\Sigma} \\
& +(1-\alpha\mu)^{2}\Big[|\beta-1|^{2}E\|\boldsymbol{\theta}_{i}\|_{\Sigma}^{2}+E\|\boldsymbol{q}_{i}\|_{\Sigma}^{2}\Big] \\
& +2(1-\alpha\mu)\mathrm{Re}\left\{\mathrm{E}\boldsymbol{c}_{\mathrm{i+1}}^{*}\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_{\mathrm{i}}-\mu\boldsymbol{u}_{\mathrm{i}}^{*}\frac{\boldsymbol{e}_{\mathrm{a}}(\mathrm{i})}{\mathrm{g}[\boldsymbol{u}_{\mathrm{i}}]}\right]\right\} \\
& +2\alpha\mu\mathrm{Re}\left\{\mathrm{E}\boldsymbol{w}^{\mathrm{o}*}_{\mathrm{i+1}}\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_{\mathrm{i}}+(1-\alpha\mu)\boldsymbol{c}_{\mathrm{i+1}}-\mu\boldsymbol{u}_{\mathrm{i}}^{*}\frac{\boldsymbol{e}_{\mathrm{a}}(\mathrm{i})}{\mathrm{g}[\boldsymbol{u}_{\mathrm{i}}]}\right]\right\}
\end{aligned}$$

$$\begin{aligned}
=\; & \alpha^{2}\mu^{2}\overbrace{E\|\boldsymbol{w}_{i+1}^{o}\|_{\Sigma}^{2}}^{②}+E\|\tilde{\boldsymbol{w}}_{i}\|_{\boldsymbol{\Sigma}'}^{2}+(1-\alpha\mu)^{2}\Big[|\beta-1|^{2}E\|\boldsymbol{\theta}_{i}\|_{\Sigma}^{2}+E\|\boldsymbol{q}_{i}\|_{\Sigma}^{2}\Big] \\
& +2(1-\alpha\mu)\mathrm{Re}\left\{\mathrm{E}\boldsymbol{c}_{\mathrm{i+1}}^{*}\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_{\mathrm{i}}-\mu\boldsymbol{u}_{\mathrm{i}}^{*}\frac{\boldsymbol{e}_{\mathrm{a}}(\mathrm{i})}{\mathrm{g}[\boldsymbol{u}_{\mathrm{i}}]}\right]\right\} \\
& +2\alpha\mu\mathrm{Re}\left\{\mathrm{E}\boldsymbol{w}^{\mathrm{o}*}_{\mathrm{i+1}}\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_{\mathrm{i}}+(1-\alpha\mu)\boldsymbol{c}_{\mathrm{i+1}}-\mu\boldsymbol{u}_{\mathrm{i}}^{*}\frac{\boldsymbol{e}_{\mathrm{a}}(\mathrm{i})}{\mathrm{g}[\boldsymbol{u}_{\mathrm{i}}]}\right]\right\}. \qquad (3.7)
\end{aligned}$$

Where $\boldsymbol{\Sigma}'$ is defined as

$$\boxed{\boldsymbol{\Sigma}'=(1-\alpha\mu)^{2}\Sigma-\mu(1-\alpha\mu)\Sigma\frac{\boldsymbol{u}_i^{*}\boldsymbol{u}_i}{g[\boldsymbol{u}_i]}-\mu(1-\alpha\mu)\frac{\boldsymbol{u}_i^{*}\boldsymbol{u}_i}{g[\boldsymbol{u}_i]}\Sigma+\mu^{2}\boldsymbol{u}_i^{*}\frac{\|\boldsymbol{u}_i\|_{\Sigma}^{2}}{g^{2}[\boldsymbol{u}_i]}\boldsymbol{u}_i}. \qquad (3.8)$$

② can be evaluated using point (b) of our model which is mentioned in section 1 as the following:

$$
\begin{aligned}
E\|\boldsymbol{w}^o_{i+1}\|^2_\Sigma &= E\|w^o + \boldsymbol{\theta}_{i+1}\|^2_\Sigma \\
&= E\|w^o\|^2_\Sigma + E\|\boldsymbol{\theta}_{i+1}\|^2_\Sigma \\
&= E\|w^o\|^2_\Sigma + |\beta|^2 E\|\boldsymbol{\theta}_i\|^2_\Sigma + E\|\boldsymbol{q}_i\|^2_\Sigma \qquad (3.9)
\end{aligned}
$$

Substituting (3.9) in (3.7) we conclude

$$
\begin{aligned}
① \; = \; & \alpha^2\mu^2(E\|w^o\|^2_\Sigma + |\beta|^2 E\|\boldsymbol{\theta}_i\|^2_\Sigma + E\|\boldsymbol{q}_i\|^2_\Sigma) + E\|\tilde{\boldsymbol{w}}_i\|^2_{\boldsymbol{\Sigma}'} \\
& + (1-\alpha\mu)^2\left[|\beta-1|^2 E\|\boldsymbol{\theta}_i\|^2_\Sigma + E\|\boldsymbol{q}_i\|^2_\Sigma\right] \\
& + 2(1-\alpha\mu)\,\mathrm{Re}\overbrace{\left\{\mathrm{E}\boldsymbol{c}^*_{\mathrm{i+1}}\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_{\mathrm{i}} - \mu\boldsymbol{u}^*_{\mathrm{i}}\frac{\boldsymbol{e}_{\mathrm{a}}(\mathrm{i})}{\mathrm{g}[\boldsymbol{u}_{\mathrm{i}}]}\right]\right\}}^{③} \\
& + 2\alpha\mu\mathrm{Re}\left\{\mathrm{E}\boldsymbol{w}^{\mathrm{o}*}_{\mathrm{i+1}}\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_{\mathrm{i}} + (1-\alpha\mu)\boldsymbol{c}_{\mathrm{i+1}} - \mu\boldsymbol{u}^*_{\mathrm{i}}\frac{\boldsymbol{e}_{\mathrm{a}}(\mathrm{i})}{\mathrm{g}[\boldsymbol{u}_{\mathrm{i}}]}\right]\right\} \qquad (3.10)
\end{aligned}
$$

③ in (3.10) can be evaluated by using equation (2.18) which gives

$$
\begin{aligned}
E\boldsymbol{c}^*_{i+1}\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_i - \mu\boldsymbol{u}^*_i\frac{\boldsymbol{e}_a(i)}{g[\boldsymbol{u}_i]}\right] &= E\left\{((\beta-1)\boldsymbol{\theta}_i + \boldsymbol{q}_i)^*\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_i - \mu\boldsymbol{u}^*_i\frac{\boldsymbol{e}_a(i)}{g[\boldsymbol{u}_i]}\right]\right\} \\
&= E\left\{((\beta-1)\boldsymbol{\theta}_i)^*\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_i - \mu\boldsymbol{u}^*_i\frac{\boldsymbol{e}_a(i)}{g[\boldsymbol{u}_i]}\right]\right\}
\end{aligned}
\qquad (3.11)
$$

where we used the fact that $E\left\{\boldsymbol{q}^*_i\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_i - \mu\boldsymbol{u}^*_i\frac{\boldsymbol{e}_a(i)}{g[\boldsymbol{u}_i]}\right]\right\} = 0$. By substituting (3.11) in (3.10), (3.10) becomes

$$
\begin{aligned}
① \;=\;& \alpha^2\mu^2(E\|w^o\|^2_\Sigma + |\beta|^2 E\|\boldsymbol{\theta}_i\|^2_\Sigma + E\|\boldsymbol{q}_i\|^2_\Sigma) + E\|\tilde{\boldsymbol{w}}_i\|^2_{\boldsymbol{\Sigma}'} \\
&+(1-\alpha\mu)^2\left[|\beta-1|^2 E\|\boldsymbol{\theta}_i\|^2_\Sigma + E\|\boldsymbol{q}_i\|^2_\Sigma\right] \\
&+2(1-\alpha\mu)\mathrm{Re}\left\{\mathrm{E}(\beta-1)^*\boldsymbol{\theta}^*_\mathrm{i}\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_\mathrm{i} - \mu\boldsymbol{u}^*_\mathrm{i}\frac{\boldsymbol{e}_\mathrm{a}(\mathrm{i})}{\mathrm{g}[\boldsymbol{u}_\mathrm{i}]}\right]\right\} \\
&+2\alpha\mu\overbrace{\mathrm{Re}\left\{\mathrm{E}\boldsymbol{w}^{\mathrm{o}*}_{\mathrm{i+1}}\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_\mathrm{i} + (1-\alpha\mu)\boldsymbol{c}_\mathrm{i+1} - \mu\boldsymbol{u}^*_\mathrm{i}\frac{\boldsymbol{e}_\mathrm{a}(\mathrm{i})}{\mathrm{g}[\boldsymbol{u}_\mathrm{i}]}\right]\right\}}^{④} \quad (3.12)
\end{aligned}
$$

④ can be evaluated as

$$
\begin{aligned}
④ \;=\;& E(w^o + \boldsymbol{\theta}_{i+1})^*\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_i + (1-\alpha\mu)\boldsymbol{c}_{i+1} - \mu\boldsymbol{u}^*_i\frac{\boldsymbol{e}_a(i)}{g[\boldsymbol{u}_i]}\right] \\
=\;& E(w^o + \boldsymbol{\theta}_{i+1})^*\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_i - \mu\boldsymbol{u}^*_i\frac{\boldsymbol{e}_a(i)}{g[\boldsymbol{u}_i]}\right] + (1-\alpha\mu)E(w^o + \boldsymbol{\theta}_{i+1})^*\Sigma\boldsymbol{c}_{i+1} \\
=\;& Ew^{o*}\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_i - \mu\boldsymbol{u}^*_i\frac{\boldsymbol{e}_a(i)}{g[\boldsymbol{u}_i]}\right] + \beta^* E\boldsymbol{\theta}^*_i\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_i - \mu\boldsymbol{u}^*_i\frac{\boldsymbol{e}_a(i)}{g[\boldsymbol{u}_i]}\right] \\
&+\beta^*(1-\alpha\mu)(\beta-1)E\|\boldsymbol{\theta}_i\|^2_\Sigma + (1-\alpha\mu)E\|\boldsymbol{q}_i\|^2_\Sigma \quad (3.13)
\end{aligned}
$$

By substituting (3.13) in (3.12), and re-arranging its terms we get

$$
\begin{aligned}
① \quad = \quad & E\|\tilde{\boldsymbol{w}}_i\|^2_{\boldsymbol{\Sigma}'} + \alpha^2\mu^2(E\|w^o\|^2_\Sigma + |\beta|^2 E\|\boldsymbol{\theta}_i\|^2_\Sigma + E\|\boldsymbol{q}_i\|^2_\Sigma) \\
& +(1-\alpha\mu)^2\left[|\beta-1|^2 E\|\boldsymbol{\theta}_i\|^2_\Sigma + E\|\boldsymbol{q}_i\|^2_\Sigma\right] \\
& +2(1-\alpha\mu)\mathrm{Re}\left\{\mathrm{E}(\beta-1)^*\boldsymbol{\theta}^*_\mathrm{i}\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_\mathrm{i} - \mu\boldsymbol{u}^*_\mathrm{i}\frac{\boldsymbol{e}_\mathrm{a}(\mathrm{i})}{\mathrm{g}[\boldsymbol{u}_\mathrm{i}]}\right]\right\} \\
& +2\alpha\mu\mathrm{Re}\left\{\mathrm{E}w^{\mathrm{o}*}\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_\mathrm{i} - \mu\boldsymbol{u}^*_\mathrm{i}\frac{\boldsymbol{e}_\mathrm{a}(\mathrm{i})}{\mathrm{g}[\boldsymbol{u}_\mathrm{i}]}\right]\right. \\
& +\beta^* E\boldsymbol{\theta}^*_i\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_i - \mu\boldsymbol{u}^*_i\frac{\boldsymbol{e}_a(i)}{g[\boldsymbol{u}_\mathrm{i}]}\right] \\
& \left. +\beta^*(1-\alpha\mu)(\beta-1)E\|\boldsymbol{\theta}_i\|^2_\Sigma + (1-\alpha\mu)E\|\boldsymbol{q}_i\|^2_\Sigma\right\}. \qquad (3.14)
\end{aligned}
$$

By combining similar terms

$$
\begin{aligned}
① \quad = \quad & E\|\tilde{\boldsymbol{w}}_i\|^2_{\boldsymbol{\Sigma}'} + E\|\boldsymbol{q}_i\|^2_\Sigma + \alpha^2\mu^2(E\|w^o\|^2_\Sigma + |\beta|^2 E\|\boldsymbol{\theta}_i\|^2_\Sigma) \\
& +(1-\alpha\mu)^2|\beta-1|^2 E\|\boldsymbol{\theta}_i\|^2_\Sigma \\
& +2\mathrm{Re}\left\{(\alpha\mu-1+\beta^*)\mathrm{E}\boldsymbol{\theta}^*_\mathrm{i}\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_\mathrm{i} - \mu\boldsymbol{u}^*_\mathrm{i}\frac{\boldsymbol{e}_\mathrm{a}(\mathrm{i})}{\mathrm{g}[\boldsymbol{u}_\mathrm{i}]}\right]\right\} \\
& +2\alpha\mu\mathrm{Re}\left\{\mathrm{E}w^{\mathrm{o}*}\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_\mathrm{i} - \mu\boldsymbol{u}^*_\mathrm{i}\frac{\boldsymbol{e}_\mathrm{a}(\mathrm{i})}{\mathrm{g}[\boldsymbol{u}_\mathrm{i}]}\right]\right. \\
& \left. +\beta^*(1-\alpha\mu)(\beta-1)E\|\boldsymbol{\theta}_i\|^2_\Sigma\right\} \qquad (3.15)
\end{aligned}
$$

Finally, substituting this in (3.5) we get

$$
\begin{aligned}
E\|\tilde{\boldsymbol{w}}_{i+1}\|^2_{\Sigma} \;=\;& E\|\tilde{\boldsymbol{w}}_i\|^2_{\boldsymbol{\Sigma}'} + \mathrm{tr}(\Sigma\mathbf{Q}) + \alpha^2\mu^2(\mathrm{tr}(\Sigma\mathbf{W}) + |\beta|^2\mathrm{E}\|\boldsymbol{\theta}_{\mathrm{i}}\|^2_{\Sigma}) \\
&+(1-\alpha\mu)^2|\beta-1|^2E\|\boldsymbol{\theta}_i\|^2_{\Sigma} \\
&+2\mathrm{Re}\Bigg\{(\alpha\mu-1+\beta^*)\mathrm{E}\boldsymbol{\theta}^*_{\mathrm{i}}\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_{\mathrm{i}} - \mu\boldsymbol{u}^*_{\mathrm{i}}\frac{\boldsymbol{e}_{\mathrm{a}}(\mathrm{i})}{\mathrm{g}[\boldsymbol{u}_{\mathrm{i}}]}\right]\Bigg\} \\
&+2\alpha\mu\mathrm{Re}\Bigg\{\mathrm{E}w^{o*}\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_{\mathrm{i}} - \mu\boldsymbol{u}^*_{\mathrm{i}}\frac{\boldsymbol{e}_{\mathrm{a}}(\mathrm{i})}{\mathrm{g}[\boldsymbol{u}_{\mathrm{i}}]}\right] \\
&+\beta^*(1-\alpha\mu)(\beta-1)E\|\boldsymbol{\theta}_i\|^2_{\Sigma}\Bigg\} + \mu^2\sigma_v^2 E\frac{\|\boldsymbol{u}_i\|^2_{\Sigma}}{g^2[\boldsymbol{u}_i]},
\end{aligned}
\qquad (3.16)
$$

where we have the fact that $E\|\boldsymbol{q}_i\|^2_{\Sigma} = \mathrm{tr}(\Sigma\mathbf{Q})$ and $E\|w^o\|^2_{\Sigma} = \mathrm{tr}(\Sigma\mathbf{W})$, and $\mathbf{W} = w^o w^{o*}$.

3.3.1 Independence Assumption

Independence assumption states

$$
\boxed{\text{The sequence } \tilde{\boldsymbol{u}}_i \text{ is independent and identically distributed}} \qquad (3.17)
$$

This assumption will allow us to replace $\boldsymbol{\Sigma}'$ by Σ' where $\Sigma' = E[\boldsymbol{\Sigma}']$, therefore (3.16) becomes

$$
\begin{aligned}
E\|\tilde{\boldsymbol{w}}_{i+1}\|^2_{\Sigma} \quad = \quad & E\|\tilde{\boldsymbol{w}}_i\|^2_{\Sigma'} + \mathrm{tr}(\Sigma\mathbf{Q}) + \alpha^2\mu^2(\mathrm{tr}(\Sigma\mathbf{W}) + |\beta|^2\mathrm{E}\|\boldsymbol{\theta}_{\mathrm{i}}\|^2_{\Sigma}) \\
& +(1-\alpha\mu)^2|\beta-1|^2 E\|\boldsymbol{\theta}_i\|^2_{\Sigma} \\
& +2\mathrm{Re}\left\{(\alpha\mu - 1 + \beta^*)\mathrm{E}\boldsymbol{\theta}^*_{\mathrm{i}}\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_{\mathrm{i}} - \mu\boldsymbol{u}^*_{\mathrm{i}}\frac{\boldsymbol{e}_{\mathrm{a}}(\mathrm{i})}{\mathrm{g}[\boldsymbol{u}_{\mathrm{i}}]}\right]\right\} \\
& +2\alpha\mu\mathrm{Re}\left\{\mathrm{E}w^{o*}\Sigma\left[(1-\alpha\mu)\tilde{\boldsymbol{w}}_{\mathrm{i}} - \mu\boldsymbol{u}^*_{\mathrm{i}}\frac{\boldsymbol{e}_{\mathrm{a}}(\mathrm{i})}{\mathrm{g}[\boldsymbol{u}_{\mathrm{i}}]}\right]\right. \\
& \left. +\beta^*(1-\alpha\mu)(\beta-1)E\|\boldsymbol{\theta}_i\|^2_{\Sigma}\right\} + \mu^2\sigma_v^2 E\frac{\|\boldsymbol{u}_i\|^2_{\Sigma}}{g^2[\boldsymbol{u}_i]},
\end{aligned} \tag{3.18}
$$

where Σ' is defined as

$$
\boxed{\Sigma' = (1-\alpha\mu)^2\Sigma - \mu(1-\alpha\mu)\Sigma E\left[\frac{\boldsymbol{u}_i^*\boldsymbol{u}_i}{g[\boldsymbol{u}_i]}\right] - \mu(1-\alpha\mu)E\left[\frac{\boldsymbol{u}_i^*\boldsymbol{u}_i}{g[\boldsymbol{u}_i]}\right]\Sigma + \mu^2 E\left[\boldsymbol{u}_i^*\frac{\|\boldsymbol{u}_i\|^2_{\Sigma}}{g^2[\boldsymbol{u}_i]}\boldsymbol{u}_i\right]} \tag{3.19}
$$

3.3.2 Steady State Performance

Initially we will start by substituting (2.27) in (3.16) which gives

$$
\begin{aligned}
E\|\tilde{\boldsymbol{w}}_{i+1}\|^2_{\Sigma} \;=\;& E\|\tilde{\boldsymbol{w}}_i\|^2_{\Sigma'} + \mathrm{tr}(\Sigma\mathbf{Q}) + \alpha^2\mu^2(\mathrm{tr}(\Sigma\mathbf{W}) + |\beta|^2\mathrm{tr}(\Sigma\boldsymbol{\Theta})) \\
& +(1-\alpha\mu)^2|\beta-1|^2\mathrm{tr}(\Sigma\boldsymbol{\Theta}) \\
& +2\mathrm{Re}\Bigg\{(\alpha\mu - 1 + \beta^*)\mathrm{E}\boldsymbol{\theta}^*_{\mathrm{i}}\Sigma\Bigg[\Bigg((1-\alpha\mu)\mathbf{I} - \mu\frac{\boldsymbol{u}^*_{\mathrm{i}}\boldsymbol{u}_{\mathrm{i}}}{\mathrm{g}[\boldsymbol{u}_{\mathrm{i}}]}\Bigg)\tilde{\boldsymbol{w}}_{\mathrm{i}}\Bigg]\Bigg\} \\
& +2\alpha\mu\mathrm{Re}\Bigg\{\mathrm{E}w^{\mathrm{o}*}\Sigma\Bigg[\Bigg((1-\alpha\mu) - \mu\frac{\boldsymbol{u}^*_{\mathrm{i}}\boldsymbol{u}_{\mathrm{i}}}{\mathrm{g}[\boldsymbol{u}_{\mathrm{i}}]}\Bigg)\tilde{\boldsymbol{w}}_{\mathrm{i}}\Bigg] \\
& +\beta^*(1-\alpha\mu)(\beta-1)\mathrm{tr}(\Sigma\boldsymbol{\Theta})\Bigg\} + \mu^2\sigma^2_v E\frac{\|\boldsymbol{u}_i\|^2_{\Sigma}}{g^2[\boldsymbol{u}_i]},
\end{aligned} \tag{3.20}
$$

Or

$$
\begin{aligned}
E\|\tilde{\boldsymbol{w}}_{i+1}\|^2_{\Sigma} \;=\;& E\|\tilde{\boldsymbol{w}}_i\|^2_{\Sigma'} + \mathrm{tr}(\Sigma\mathbf{Q}) + \alpha^2\mu^2(\mathrm{tr}(\Sigma\mathbf{W}) + |\beta|^2\mathrm{tr}(\Sigma\boldsymbol{\Theta})) \\
& +(1-\alpha\mu)^2|\beta-1|^2\mathrm{tr}(\Sigma\boldsymbol{\Theta}) \\
& +2\mathrm{Re}\Big\{(\alpha\mu - 1 + \beta^*)\mathrm{tr}(\Sigma[(\mathbf{I} - \mu\mathbf{R}_{\mathrm{n}})\mathbf{W}])\Big\} \\
& +2\alpha\mu\mathrm{Re}\Big\{w^{\mathrm{o}*}\Sigma\boldsymbol{J}\mathrm{E}\tilde{\boldsymbol{w}}_{\mathrm{i}} + \beta^*(1-\alpha\mu)(\beta-1)\mathrm{tr}(\Sigma\boldsymbol{\Theta})\Big\} \\
& +\mu^2\sigma^2_v E\frac{\|\boldsymbol{u}_i\|^2_{\Sigma}}{g^2[\boldsymbol{u}_i]},
\end{aligned} \tag{3.21}
$$

where

$$
\boxed{\mathbf{R}_n = \boldsymbol{U}_i \;\triangleq\; \alpha\mathbf{I} + \tfrac{\boldsymbol{u}^*_i\boldsymbol{u}_i}{g[\boldsymbol{u}_i]}, \quad \boldsymbol{J} \;\triangleq\; \mathbf{I} - \mu E(\boldsymbol{U}_i)} \tag{3.22}
$$

also (3.19) will be

$$\Sigma' = \Sigma - \mu(E\boldsymbol{U}_i)\Sigma - \mu\Sigma(E\boldsymbol{U}_i) + \mu^2 E(\boldsymbol{U}_i\Sigma\boldsymbol{U}_i) \tag{3.23}$$

Vector Notation

From [33], we can see many useful definitions which can be used in our case, which are

$$\boxed{\sigma = \text{vec}(\Sigma), \sigma' = \text{vec}(\Sigma'),\ \sigma' = F\sigma, \quad F \stackrel{\Delta}{=} E\{(I_M - \mu\boldsymbol{U}_i)^* \otimes (I_M - \mu\boldsymbol{U}_i)\}} \tag{3.24}$$

At steady state it is known that

$$\lim_{i\to\infty} E\|\tilde{\boldsymbol{w}}_{i+1}\|^2_\sigma = \lim_{i\to\infty} E\|\tilde{\boldsymbol{w}}_i\|^2_\sigma, \quad \lim_{i\to\infty} E\tilde{\boldsymbol{w}}_{i+1} = \lim_{i\to\infty} E\tilde{\boldsymbol{w}}_i.$$

Using the above, it can be shown that

$$\boxed{\lim_{i\to\infty} E\tilde{\boldsymbol{w}}_{i+1} = \alpha\mu(\mathbf{I} - \boldsymbol{J})^{-1} w^o} \tag{3.25}$$

Applying all the mentioned above to (3.21), we get

$$
\begin{aligned}
\lim_{i\to\infty} E\|\tilde{\boldsymbol{w}}_i\|^2_{(\mathbf{I}-F)\sigma} &= \mathrm{tr}(\Sigma\mathbf{Q}) + \alpha^2\mu^2\|w^{\mathrm{o}}\|^2_{\mathrm{T}\sigma} + \alpha^2\mu^2|\beta|^2\mathrm{tr}(\Sigma\boldsymbol{\Theta}) \\
&\quad +(1-\alpha\mu)^2|\beta-1|^2\mathrm{tr}(\Sigma\boldsymbol{\Theta}) + \mu^2\sigma_v^2 E\frac{\|\boldsymbol{u}_i\|^2_\sigma}{g^2[\boldsymbol{u}_i]} \\
&\quad +2\mathrm{Re}\Big\{(\alpha\mu-1+\beta^*)\mathrm{tr}(\Sigma[(\mathbf{I}-\mu\mathbf{R}_{\mathrm{n}})\mathbf{W}])\Big\} \\
&\quad +2\alpha\mu\mathrm{Re}\Big\{\beta^*(1-\alpha\mu)(\beta-1)\mathrm{tr}(\Sigma\boldsymbol{\Theta})\Big\}
\end{aligned} \tag{3.26}
$$

where $\mathbf{T}$ is defined as

$$
\boxed{\mathbf{T} \triangleq \mathbf{I} + ((\mathbf{I}-\boldsymbol{J}^T)^{-1}\boldsymbol{J}^T \otimes \mathbf{I}) + (\mathbf{I}\otimes(\mathbf{I}-\boldsymbol{J})^{-1}\boldsymbol{J})} \tag{3.27}
$$

Now, we already know that

$$
\begin{aligned}
EMSE &= E|\boldsymbol{e}_a(i)|^2 \\
&= E\|\tilde{\boldsymbol{w}}_i\|^2_{R_u}
\end{aligned} \tag{3.28}
$$

if we define

$$
r = vec(R_u) \tag{3.29}
$$

then the EMSE of the leaky ϵ-NLMS can be evaluated by substituting

$\sigma = (\mathbf{I} - F)^{-1} r$ as the following:

$$\begin{aligned}
\lim_{i\to\infty} E\|\tilde{\boldsymbol{w}}_i\|_r^2 \;\; = \;\; & \mathrm{tr}(\Sigma\mathbf{Q}) + \alpha^2\mu^2\|w^o\|^2_{\mathrm{T}(\mathbf{I}-\mathrm{F})^{-1}\mathrm{r}} + \alpha^2\mu^2|\beta|^2\mathrm{tr}(\Sigma\boldsymbol{\Theta}) \\
& +(1-\alpha\mu)^2|\beta-1|^2\mathrm{tr}(\Sigma\boldsymbol{\Theta}) + \mu^2\sigma_v^2 E\frac{\|\boldsymbol{u}_i\|^2_{(\mathbf{I}-F)^{-1}r}}{g^2[\boldsymbol{u}_i]} \\
& +2\mathrm{Re}\Big\{(\alpha\mu - 1 + \beta^*)\mathrm{tr}(\Sigma[(\mathbf{I}-\mu\mathbf{R}_{\mathrm{n}})\mathbf{W}])\Big\} \\
& +2\alpha\mu\mathrm{Re}\Big\{\beta^*(1-\alpha\mu)(\beta-1)\mathrm{tr}(\Sigma\boldsymbol{\Theta})\Big\}
\end{aligned} \tag{3.30}$$

3.4 Tracking Analysis of Leaky ϵ-NLMS Algorithm

In this section we will apply the relations which were derived in the previous section to Leaky ϵ-NLMS algorithm.

The recursion of Leaky ϵ-NLMS algorithm is given by

$$\boldsymbol{w}_{i+1} = (1-\alpha\mu)\boldsymbol{w}_i + \mu\boldsymbol{u}_i^*\frac{\boldsymbol{e}(i)}{\left(\varepsilon + \|\boldsymbol{u}_i\|^2\right)}. \tag{3.31}$$

Therefore, from this recursion we can see that $g[\boldsymbol{u}_i]$ is replaced by $\left(\varepsilon + \|\boldsymbol{u}_i\|^2\right)$, substituting this in (3.22) and (3.26) produces

$$\boxed{\mathbf{R}_n = \boldsymbol{U}_i \;\triangleq\; \alpha\mathbf{I} + \frac{\boldsymbol{u}_i^*\boldsymbol{u}_i}{\left(\varepsilon+\|\boldsymbol{u}_i\|^2\right)}, \quad \boldsymbol{J} \triangleq \mathbf{I} - \mu E(\boldsymbol{U}_i)} \tag{3.32}$$

$$\lim_{i\to\infty} E\|\tilde{\boldsymbol{w}}_i\|_r^2 = \mathrm{tr}(\Sigma\mathbf{Q}) + \alpha^2\mu^2\|w^o\|^2_{\mathrm{T}(\mathbf{I}-\mathrm{F})^{-1}\mathrm{r}} + \alpha^2\mu^2|\beta|^2\mathrm{tr}(\Sigma\boldsymbol{\Theta})$$
$$+(1-\alpha\mu)^2|\beta-1|^2\mathrm{tr}(\Sigma\boldsymbol{\Theta}) + \mu^2\sigma_v^2 E\left[\frac{\|\boldsymbol{u}_i\|^2_{(\mathbf{I}-F)^{-1}r}}{\left(\varepsilon+\|\boldsymbol{u}_i\|^2\right)^2}\right]$$
$$+2\mathrm{Re}\left\{(\alpha\mu-1+\beta^*)\mathrm{tr}(\Sigma[(\mathbf{I}-\mu\mathbf{R}_\mathrm{n})\mathbf{W}])\right\}$$
$$+2\alpha\mu\mathrm{Re}\left\{\beta^*(1-\alpha\mu)(\beta-1)\mathrm{tr}(\Sigma\boldsymbol{\Theta})\right\} \quad (3.33)$$

3.5 Summary

The tracking analysis of the leaky ϵ - NLMS algorithm in a nonstationary environments was investigated in this chapter, the environment is assumed to be time varying according to a first order auto regressive model. The filter input is assumed to be correlated gaussian. Finally an expression for the steady state EMSE is derived based on the weighted energy conservation relation thanks to the independence assumption principle. The derived EMSE of the leaky ϵ - NLMS algorithm depends mainly on two moments, and an exact expression for the steady state EMSE of the leaky ϵ-NLMS algorithm is derived using the results of [38] defined by (3.33).

CHAPTER 4

SIMULATION RESULTS

In this chapter the results of the computer simulations are presented which are made to investigate the validity of the derived steady-state EMSE for both ϵ - NLMS and leaky ϵ - NLMS algorithms.

In order to test the performance of these algorithms, an unknown real valued system identification problem is considered. The system noise is taken as zero mean i.i.d with variance 0.01. The length of the adaptive filter is taken to be equal to that of the unknown system, i.e., 5. The correlation matrix of the correlated complex Gaussian input to the adaptive filter and unknown system is

$$\mathbf{R} = \begin{pmatrix} 1 & \rho_c & \rho_c^2 & \cdots & \rho_c^{M-1} \\ \rho_c & 1 & \rho_c & \cdots & \rho_c^{M-2} \\ \rho_c^2 & \rho_c & 1 & \cdots & \rho_c^{M-3} \\ \vdots & & & & \\ \rho_c^{M-1} & \rho_c^{M-2} & \rho_c^{M-3} & \cdots & 1 \end{pmatrix}$$

where $0 < \rho_c < 1$ is the factor that controls the correlation between the regressor elements, large value of ρ_c results in a highly correlated input.

4.1 Simulation Results for ϵ - NLMS algorithm

In this section we are going to test the accuracy of the derived EMSE for the ϵ-NLMS algorithm. As it was mentioned in the introduction of this chapter, the filter length is M = 5, the SNR is set to be 20 dB which is equivalent to $\sigma_v^2 = 0.01$. The nonstationary environement controlled by choosing different values of σ_d^2 and the correlation of the adaptive filter input is controlloed by ρ_c. The simulated results are obtained by averaging over 100 experiments while the analytical results are obtained by plotting the steady state tracking expression given by(2.36).

Tables below show a comparison between the analytical and simulated EMSE of the ϵ - NLMS algorithm for different values of ϵ (and the same value of the correlation factor $\rho_c = 0.5$). Table 4.1 represents the case when $\sigma_d^2 = 10^{-6}$, Table 4.2 when $\sigma_d^2 = 10^{-5}$, and Table 4.3 when $\sigma_d^2 = 10^{-4}$. The tables show the analytical results match the experimental results whith a notable error in the third case (Table 4.3) and with a great accuracy in the first two tables. This implies that when rapid variations occured in the environemnt, the tracking becomes difficult.

EMSE in 10^{-3}									
$\epsilon = 0.1$		$\epsilon = 0.3$		$\epsilon = 0.5$		$\epsilon = 0.7$		$\epsilon = 0.9$	
Anal.	Sim.	Anal.	Sim.	Anal.	Sim.	Anal.	Sim.	Anal.	Sim.
4.2	4.33	3.8	3.82	3.5	3.56	3.3	3.34	3.0	3.08

Table 4.1: EMSE for ϵ-NLMS when Q is a multiple of 10^{-6}.

EMSE in 10^{-3}									
$\epsilon = 0.1$		$\epsilon = 0.3$		$\epsilon = 0.5$		$\epsilon = 0.7$		$\epsilon = 0.9$	
Anal.	Sim.	Anal.	Sim.	Anal.	Sim.	Anal.	Sim.	Anal.	Sim.
4.4	4.44	4.1	3.93	3.7	3.66	3.5	3.45	3.3	3.14

Table 4.2: EMSE for ϵ-NLMS when Q is a multiple of 10^{-5}.

EMSE in 10^{-3}									
$\epsilon = 0.1$		$\epsilon = 0.3$		$\epsilon = 0.5$		$\epsilon = 0.7$		$\epsilon = 0.9$	
Anal.	Sim.	Anal.	Sim.	Anal.	Sim.	Anal.	Sim.	Anal.	Sim.
6.3	5.5	5.9	5.31	5.6	4.98	4.6	4.28	4.3	3.96

Table 4.3: EMSE for ϵ-NLMS when Q is a multiple of 10^{-4}.

Figures {4.1 - 4.4} show how ϵ-NLMS EMSE against ϵ in 4 nonstationary environments with different σ_d^2. The figures show that the EMSE of the ϵ-NLMS is monotically decreasing as ϵ is increasing (inversely proportional to ϵ) which means that choosing ϵ near 1 will ensure a low EMSE. In addition to that, the figures show that in a slow changeble environment the simulation and analytical results are almost the same, this clear in Figures {4.2 - 4.4}, while in rapid changeble environments, a gap appears between the analytical and experimental results which again implies the tracking difficulty in a rapid nonstationary environements.

A study for the effect of the step size on the EMSE of the ϵ-NLMS algorithm at different values of the ratio of $\frac{\sigma_d^2}{\sigma_v^2}$ are shown in the Figures {4.5 - 4.9}. The ratio $\frac{\sigma_d^2}{\sigma_v^2}$ gives us an information about which is the dominant factor of the algorithm if

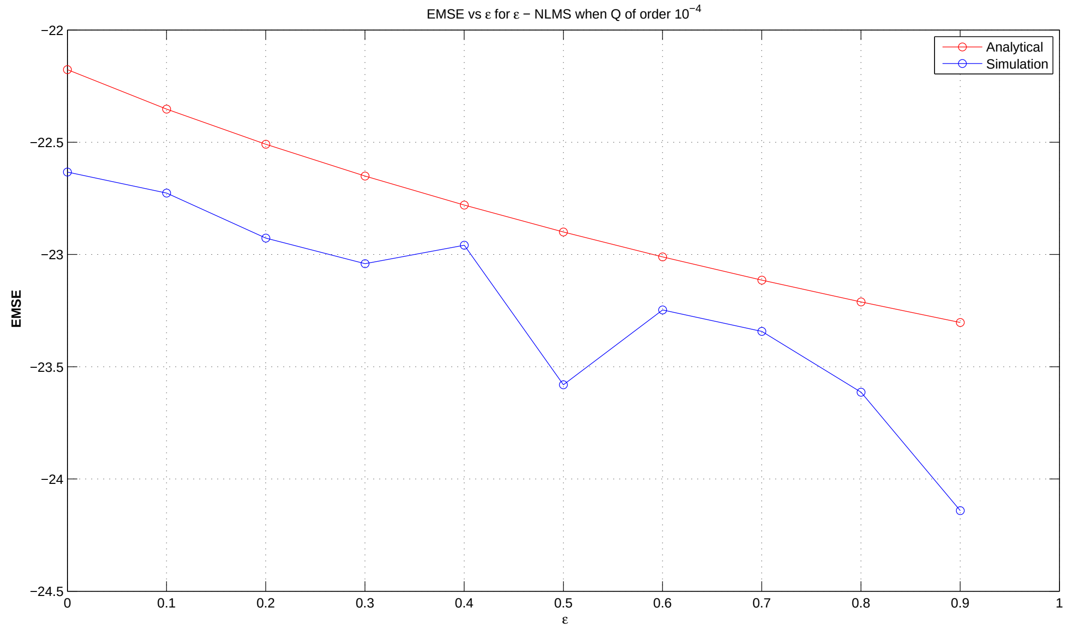

Figure 4.1: EMSE for ϵ - NLMS Algorithm for varying ϵ when Q of order 10^{-4}.

it was the noise or the rapid changes in the environment. The figures show that when this ratio is small, the noise is the dominant factor and the environment is slowly changing, the step size is monotically increasing and the EMSE is very small, when this ratio increases until it reaches 1, the analytical curve becomes a parabola while the experimental curve still monotically increasing and the EMSE becomes very large. These resutls imply that small step size is better for EMSE performance but its bad for the convergence speed of the ϵ - NLMS algorithm.

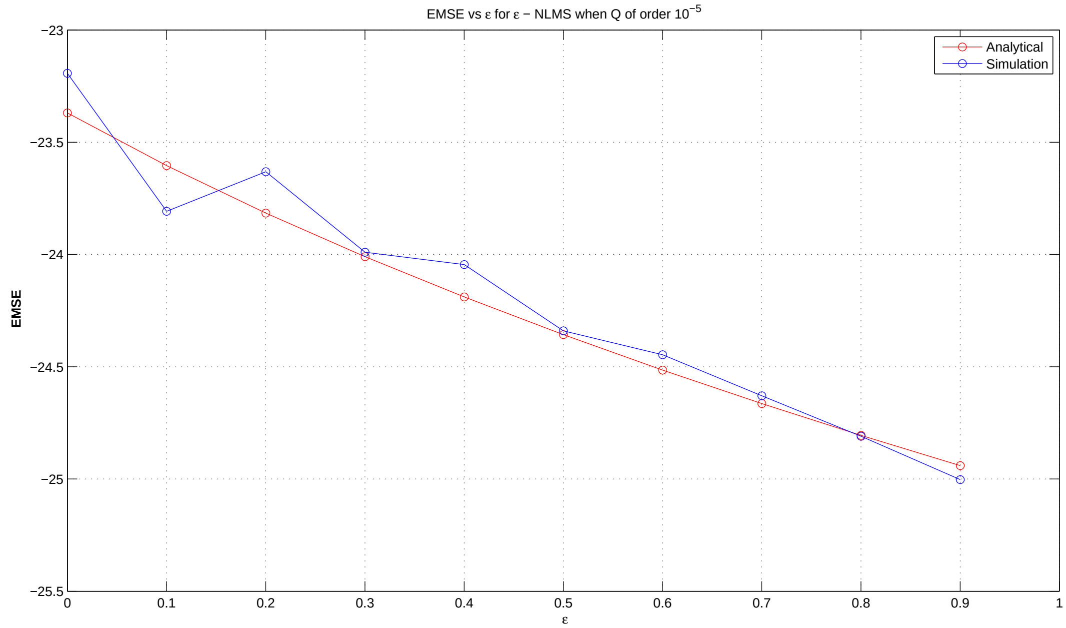

Figure 4.2: EMSE for ϵ - NLMS Algorithm for varying ϵ when Q of order 10^{-5}.

4.2 Simulation Results for Leaky ϵ - NLMS algorithm

The derivations of the Leaky ϵ - NLMS Algorithm which is shown in Chapter 3 is tested here. The simulation parameters are simialr to that in the previous section, i.e., the filter length is M = 5, SNR is 20 dB ($\sigma_v^2 = 0.01$). The nonstationary environement controlled by choosing different values of σ_d^2 and the correlation of the adaptive filter input is controlloed by ρ_c. The simulation results are obtained by averaging over 100 experiments while the analytical results are obtained by plotting the steady state tracking expression given in (3.33).

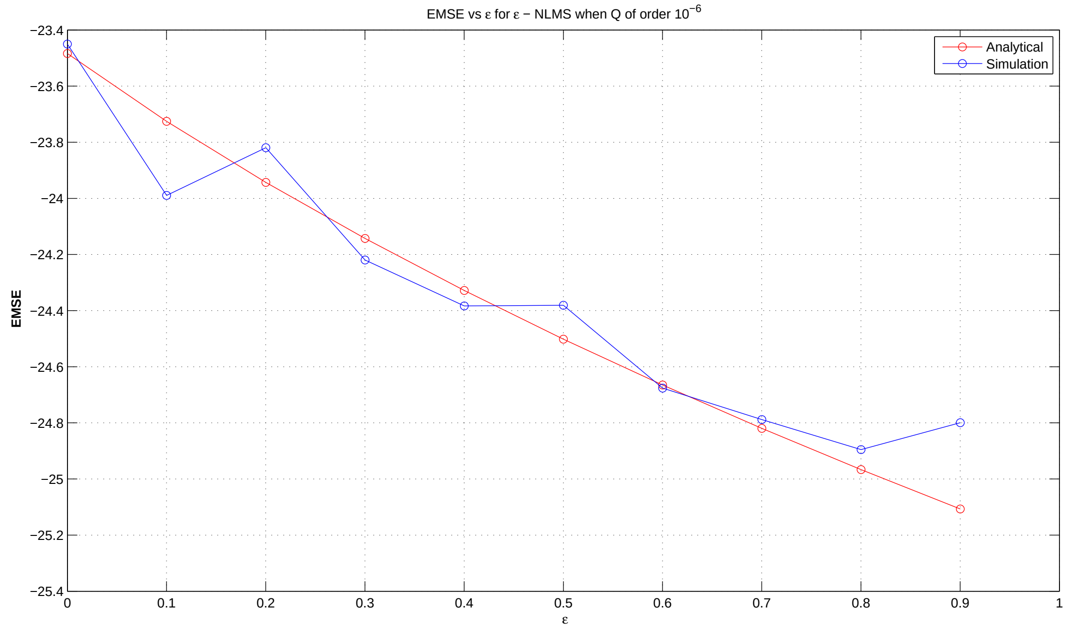

Figure 4.3: EMSE for ϵ - NLMS Algorithm for varying ϵ when Q of order 10^{-6}.

Tables {4.4 and 4.5} show how the EMSE of the Leaky ϵ - NLMS algorithm behaves for different values of ϵ. Table 4.4 represents the case when $\sigma_d^2 = 10^{-6}$. Table 4.5 when $\sigma_d^2 = 10^{-5}$ and Table 4.6 when $\sigma_d^2 = 10^{-4}$. The tables show the analytical results match the experimental results whith a notable error in the third case (Table 4.3), and with a great accuracy in the first two tables. This implies that when a rapid variations occured in the environemnt the tracking becomes difficult.

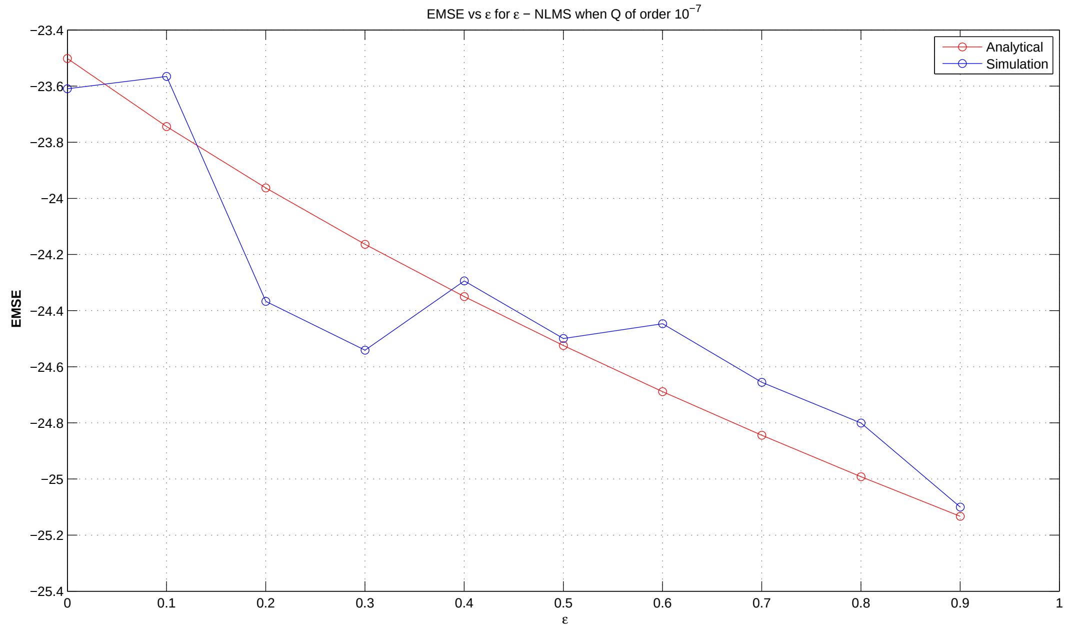

Figure 4.4: EMSE for ϵ - NLMS Algorithm for varying ϵ when Q of order 10^{-7}.

EMSE in 10^{-3}									
$\epsilon = 0.1$		$\epsilon = 0.3$		$\epsilon = 0.5$		$\epsilon = 0.7$		$\epsilon = 0.9$	
Anal.	Sim.	Anal.	Sim.	Anal.	Sim.	Anal.	Sim.	Anal.	Sim.
4.2	4.12	3.9	3.89	3.7	3.66	3.3	3.23	3.1	3.15

Table 4.4: EMSE for Leaky ϵ-NLMS when Q is a multiple of 10^{-6}

EMSE in 10^{-3}									
$\epsilon = 0.1$		$\epsilon = 0.3$		$\epsilon = 0.5$		$\epsilon = 0.7$		$\epsilon = 0.9$	
Anal.	Sim.	Anal.	Sim.	Anal.	Sim.	Anal.	Sim.	Anal.	Sim.
4.4	4.43	4.1	4.07	3.8	3.6	3.5	3.42	3.2	3.17

Table 4.5: EMSE for Leaky ϵ-NLMS when Q is a multiple of 10^{-5}

Figures {4.10 - 4.13} show how leaky ϵ - NLMS EMSE against ϵ in 4 non-stationary environments with different σ_d^2. The figures show that the EMSE of the ϵ - NLMS is monotically decreasing as ϵ increases (inversely proportional to ϵ) which means that choosing ϵ near 1 will ensure a low EMSE. In addition to that, the figures show that in a slow changeble environment the simulation and

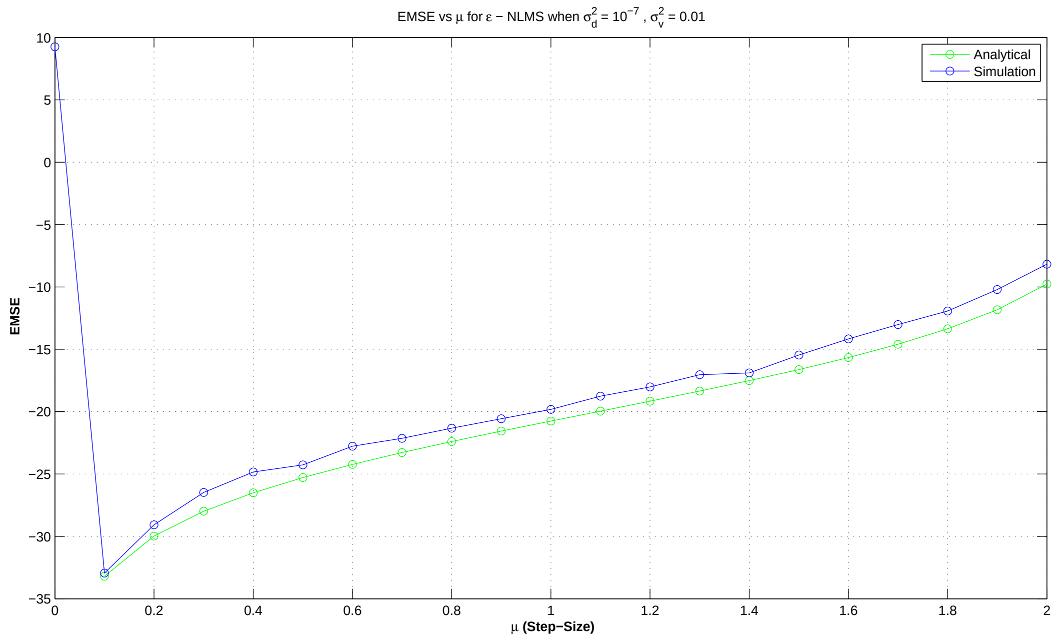

Figure 4.5: EMSE for ϵ - NLMS Algorithm when $\frac{\sigma_d^2}{\sigma_v^2} = 10^{-5}$.

EMSE in 10^{-3}									
$\epsilon = 0.1$		$\epsilon = 0.3$		$\epsilon = 0.5$		$\epsilon = 0.7$		$\epsilon = 0.9$	
Anal.	Sim.	Anal.	Sim.	Anal.	Sim.	Anal.	Sim.	Anal.	Sim.
7.3	6.96	6.0	5.72	5.6	4.75	5.1	4.46	4.7	4.27

Table 4.6: EMSE for Leaky ϵ-NLMS when Q is a multiple of 10^{-4}

analytical results are almost the same, this clear in Figures {4.11 - 4.13}, while in rapid changeble environments, a gap appears between the analytical and experimental results which again implies the tracking difficulty in a rapid nonstationary environements.

A study for the effect of the step size on the EMSE of the leaky ϵ-NLMS algorithm at different ratios of $\frac{\sigma_d^2}{\sigma_v^2}$ are shown in Figures {4.14 - 4.18}. The ratio $\frac{\sigma_d^2}{\sigma_v^2}$ gives us an information about which is the dominant factor of the algorithm if it was the noise or the rapid changes in the environment. The figures show that when this ratio is small the noise is the dominant factor and the environment

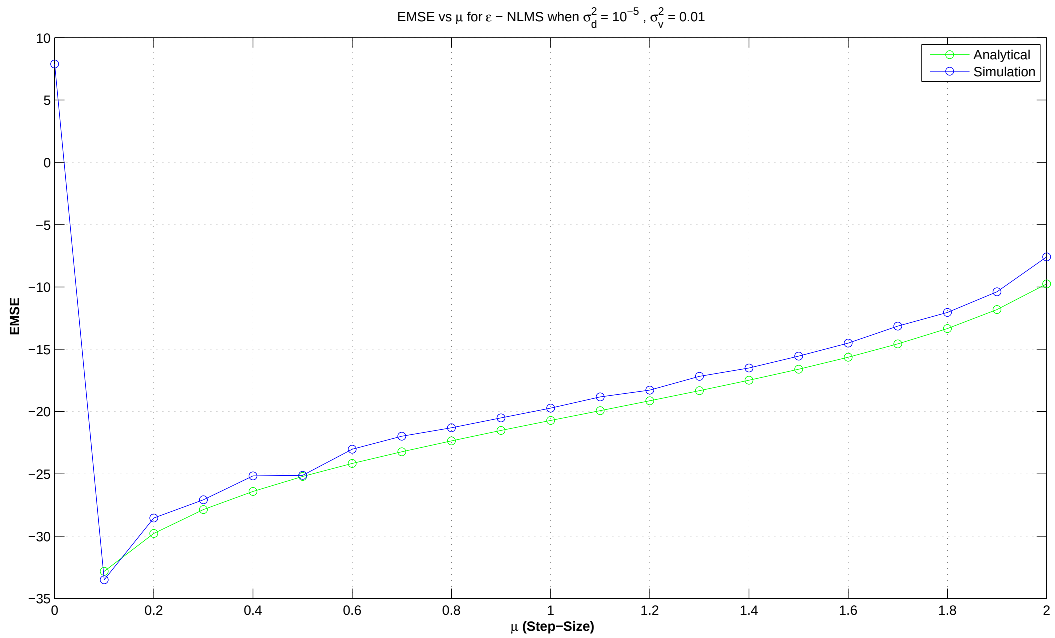

Figure 4.6: EMSE for ϵ - NLMS Algorithm when $\frac{\sigma_d^2}{\sigma_v^2} = 10^{-3}$.

is slowly changing. The step size is monotically increasing and the EMSE is very small, when this ratio increases until it reaches 1. The analytical curve becomes a parabola while the experimental curve still monotically increasing and the EMSE becomes very large. These resutls imply that small step size is better for EMSE performance but it is bad for the convergence speed of the leaky ϵ - NLMS algorithm.

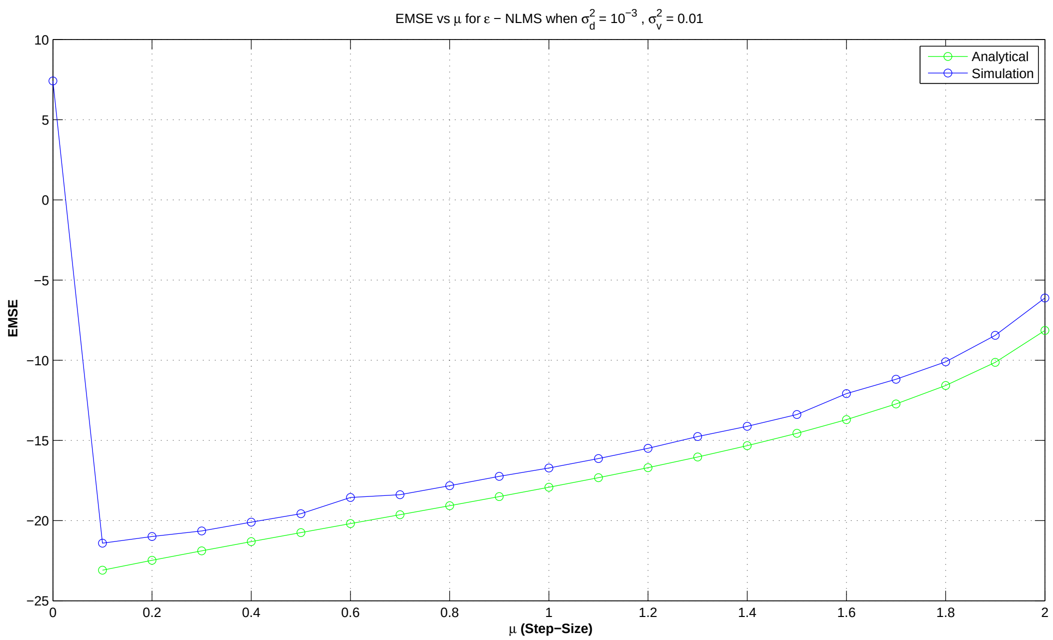

Figure 4.7: EMSE for ϵ - NLMS Algorithm when $\frac{\sigma_d^2}{\sigma_v^2} = 10^{-1}$.

4.3 Comparison of the ϵ - NLMS and leaky ϵ - NLMS ($\alpha = 0$, $\beta = 1$)

The anlaytical EMSE of the ϵ - NLMS algorithm can be compared to the equivalent one of the leaky ϵ - NLMS algorithm by setting $\alpha = 0$ (leaky factor of the update recursion (3.1)) and $\beta = 1$ (leaky factor of the environment mentioned in the nonstaionary data model point (c)).

The comparison performed for different values of σ_d^2 and for two different values of the correlation factor of the input data $\rho_c = 0.2$, and $\rho_c = 0.6$.

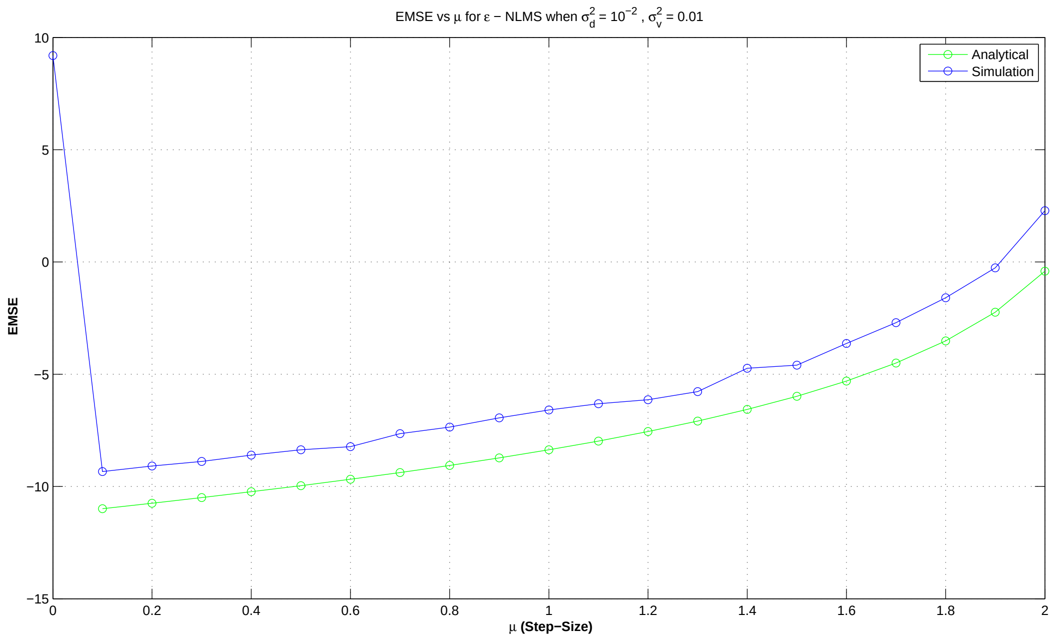

Figure 4.8: EMSE for ϵ - NLMS Algorithm when $\frac{\sigma_d^2}{\sigma_v^2} = 1$.

The results show that the equation (3.33) is more accurate than equation (2.36) for all values of μ and for different values of the correlation factor of the input data ρ_c.

4.4 Summary

In this chapter, the accuracy of the derived theoretical expressions was examined. The simulation results for both cases (i.e., ϵ-NLMS and Leaky ϵ-NLMS) show that the analytical expressions matching the simulation results especially in slow changing environment, while some notable error appear when the adaptive filter works in a rapid changing environment. Moreover, simulation shows that the derived expression for the Leaky ϵ-NLMS algorithm is more accurate than the

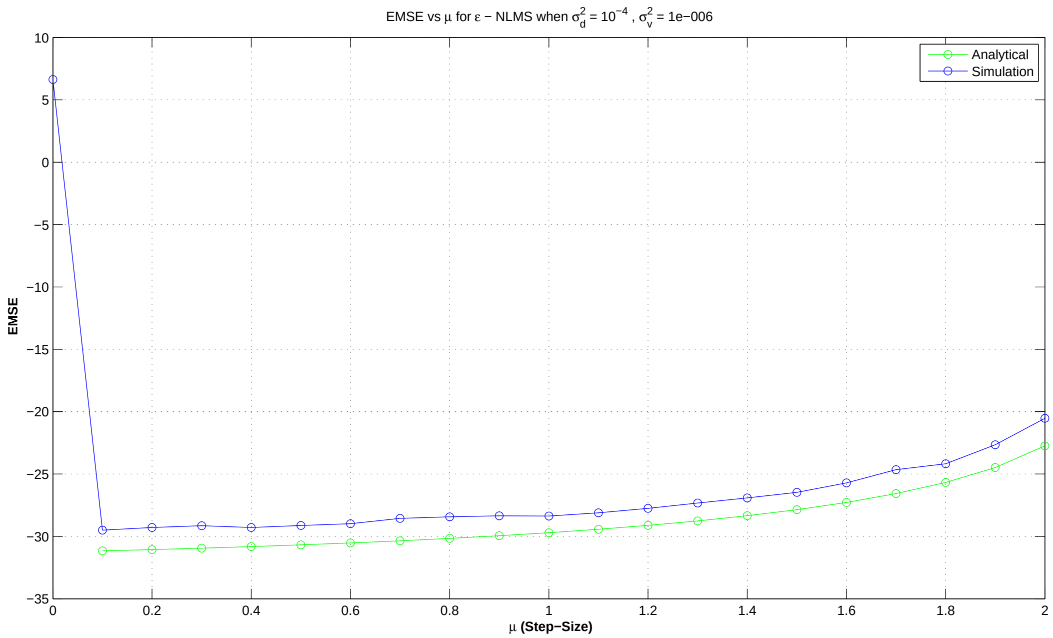

Figure 4.9: EMSE for ϵ - NLMS Algorithm when $\frac{\sigma_d^2}{\sigma_v^2} = 10^2$.

one derived for the ϵ-NLMS algorithm. This is due to the assumptions used to derive the EMSE expression for the ϵ-NLMS algorithm are weaker than the assumptions used to derive the same expression for the leaky ϵ-NLMS algorithm.

The simulation also shows some facts ablout the relation between the derived expressions for the EMSE and some paramters of the both algorithms ϵ-NLMS and Leaky ϵ-NLMS. The simulations show that the EMSE decreases monotically against the ϵ parameter for both algorithms, which means that choosing ϵ close to 1 will give us lower EMSE. Also, the simulation tries to detect the optimal step size for both algorithms by testing the EMSE against the step size of the choosen algorithm. Finally, the simulations show that the algorithm will be inefficient to perform the tracking task in a very rapid changing environment but it will bo so efficient for a slow changing one.

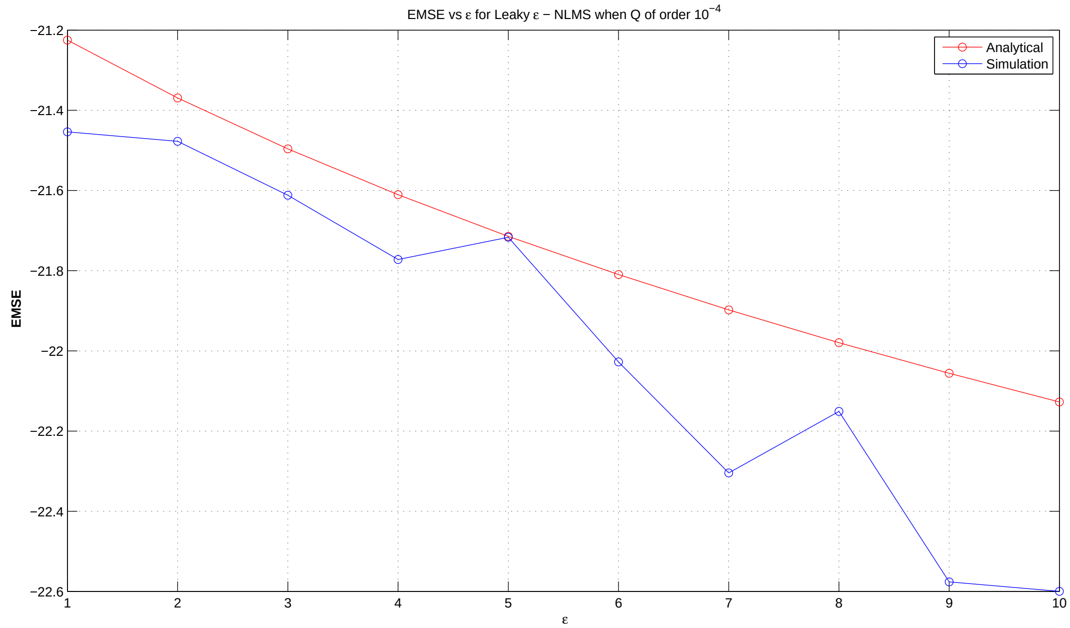

Figure 4.10: EMSE for Leaky ϵ - NLMS Algorithm for varying ϵ when Q of order 10^{-4}

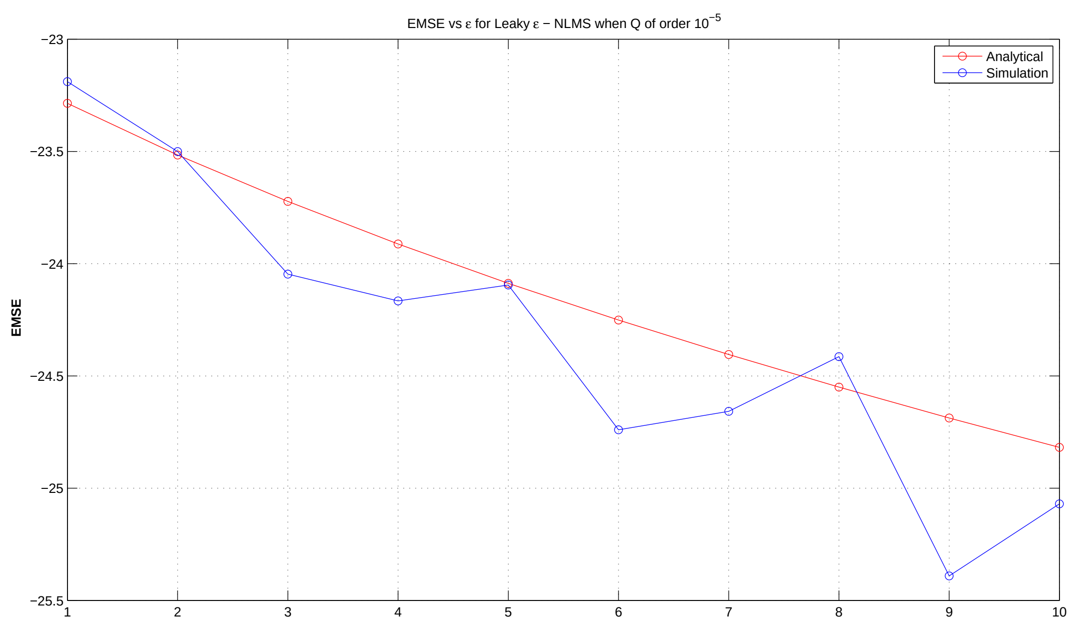

Figure 4.11: EMSE for Leaky ϵ - NLMS Algorithm for varying ϵ when Q of order 10^{-5}

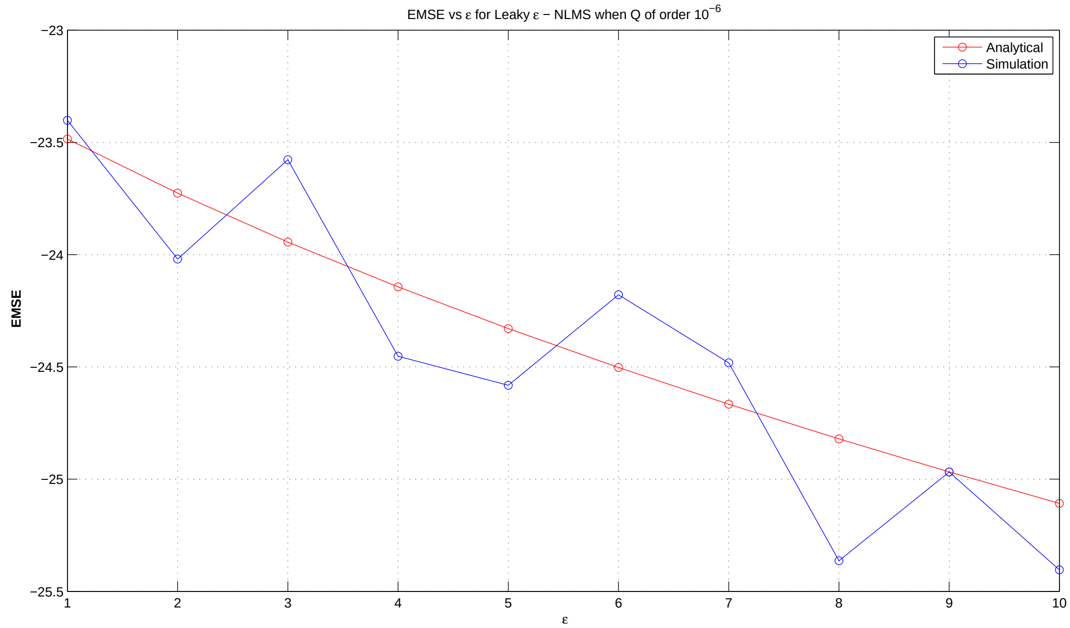

Figure 4.12: EMSE for Leaky ϵ - NLMS Algorithm for varying ϵ when Q of order 10^{-6}

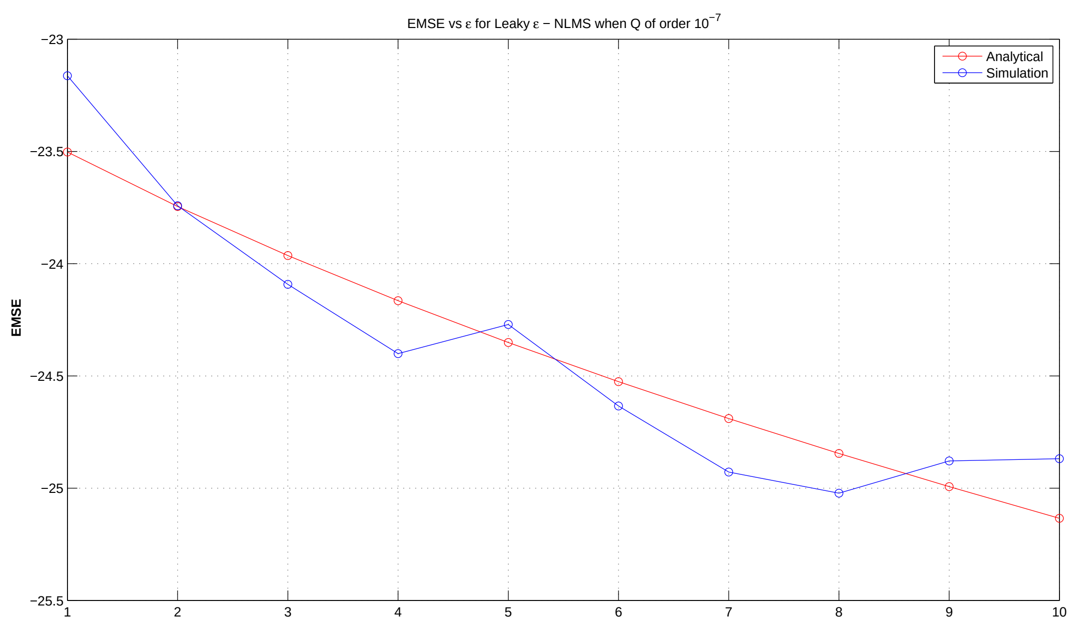

Figure 4.13: EMSE for Leaky ϵ - NLMS Algorithm for varying ϵ when Q of order 10^{-7}

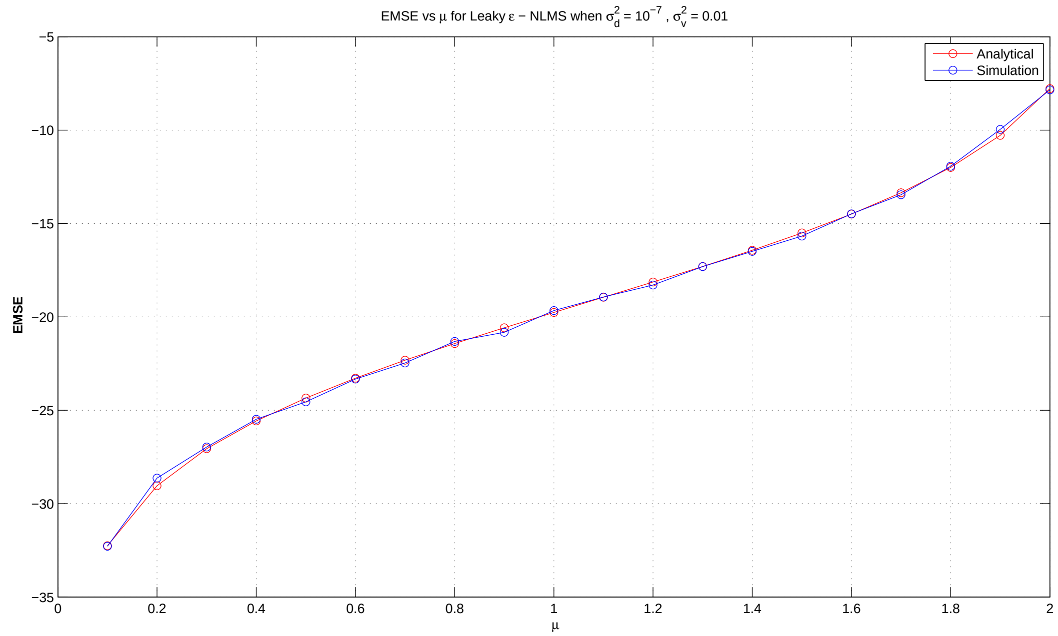

Figure 4.14: EMSE for Leaky ϵ - NLMS Algorithm when $\frac{\sigma_d^2}{\sigma_v^2} = 10^{-5}$.

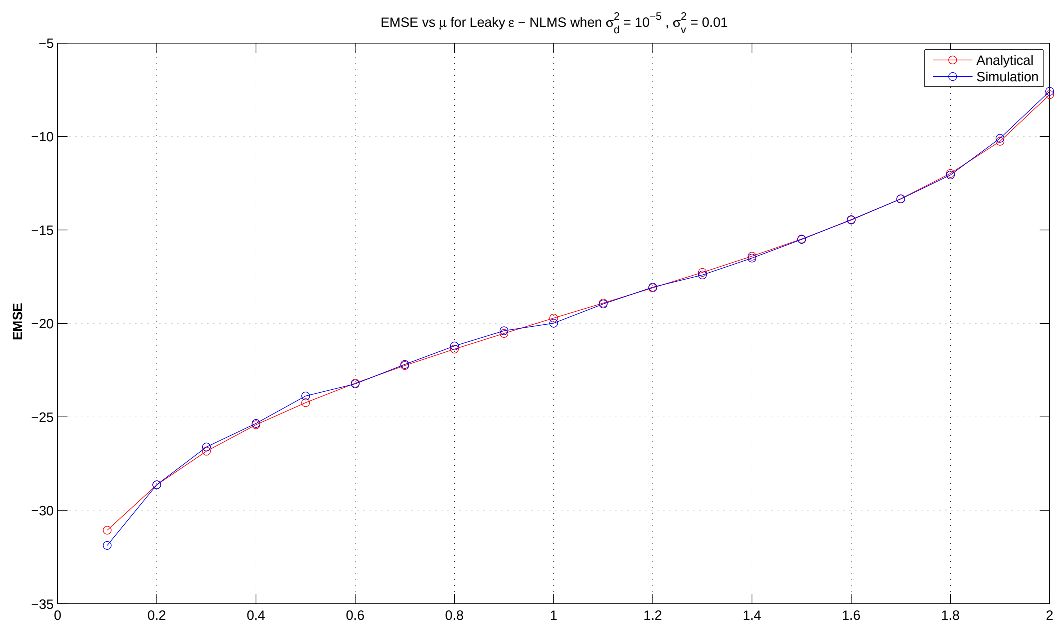

Figure 4.15: EMSE for Leaky ϵ - NLMS Algorithm when $\frac{\sigma_d^2}{\sigma_v^2} = 10^{-3}$.

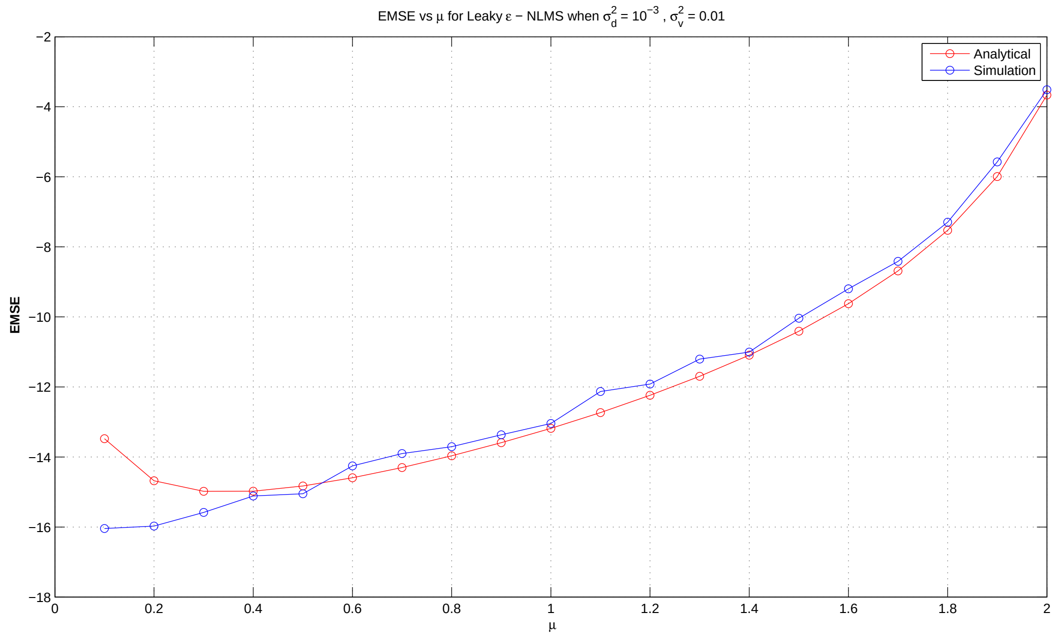

Figure 4.16: EMSE for Leaky ϵ - NLMS Algorithm when $\frac{\sigma_d^2}{\sigma_v^2} = 10^{-1}$.

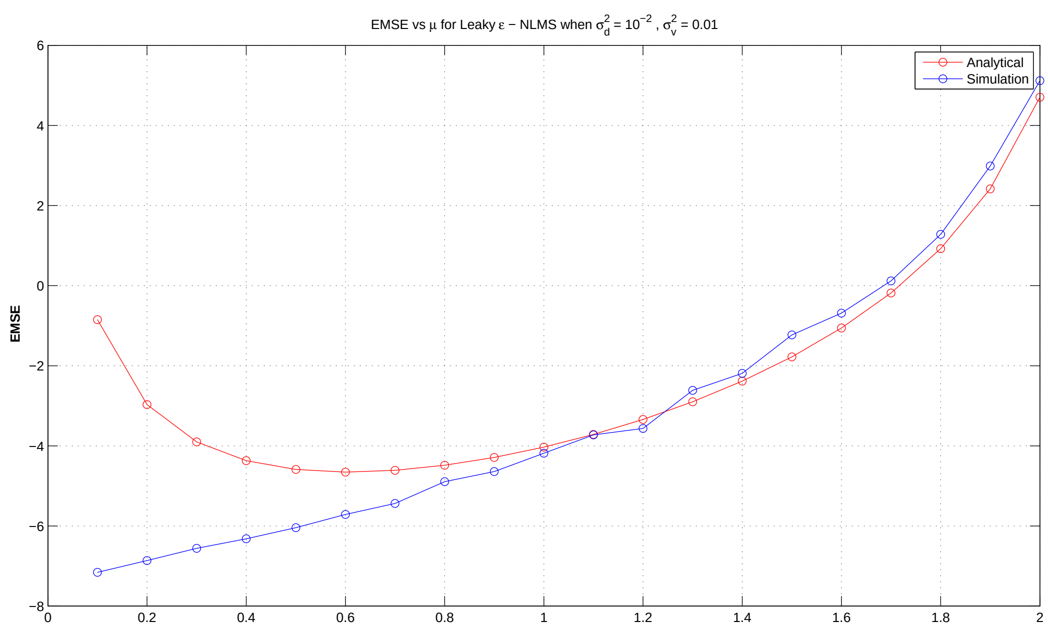

Figure 4.17: EMSE for Leaky ϵ - NLMS Algorithm when $\frac{\sigma_d^2}{\sigma_v^2} = 1$.

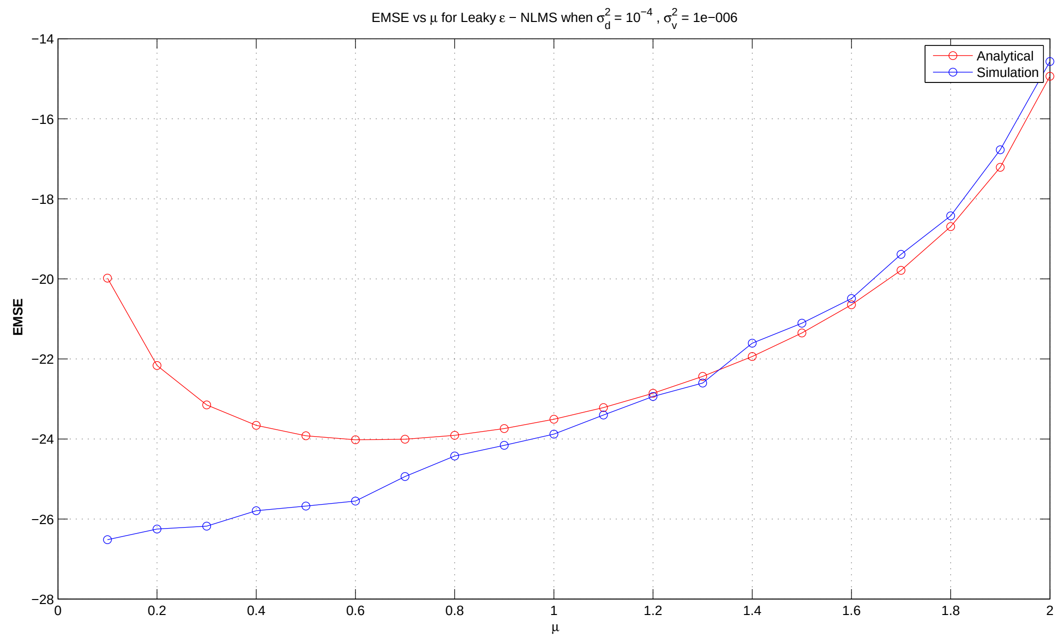

Figure 4.18: EMSE for Leaky ϵ - NLMS Algorithm when $\frac{\sigma_d^2}{\sigma_v^2} = 10^2$.

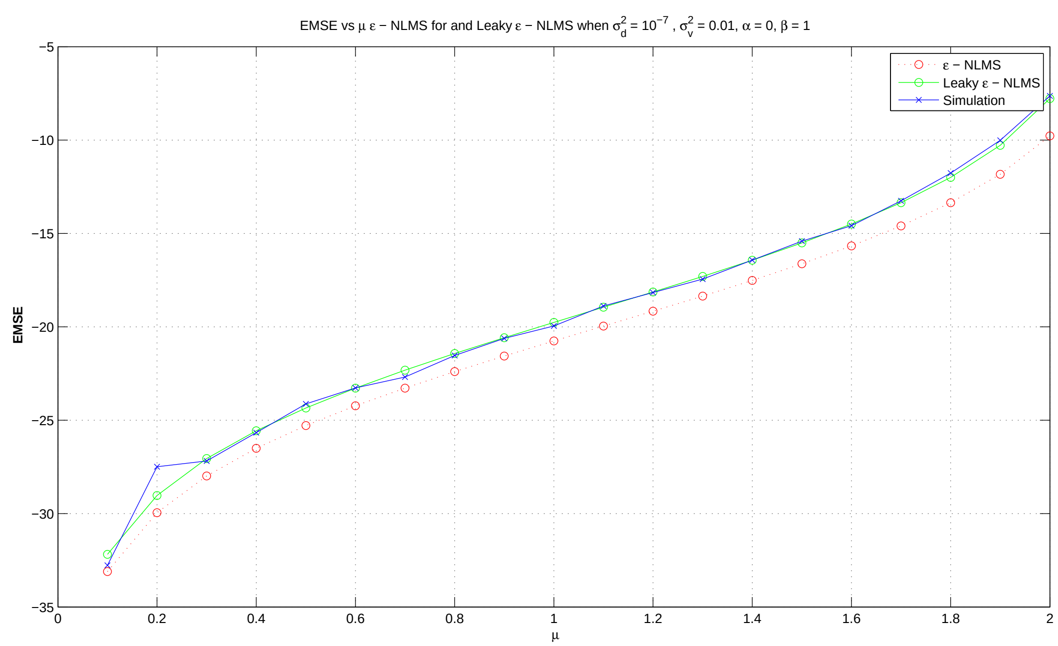

Figure 4.19: EMSE vs μ for ϵ - NLMS , Leaky ϵ - NLMS when $\sigma_d^2 = 10^{-7}$ $\sigma_v^2 = 10^{-2}$, and $\rho = 0.2$.

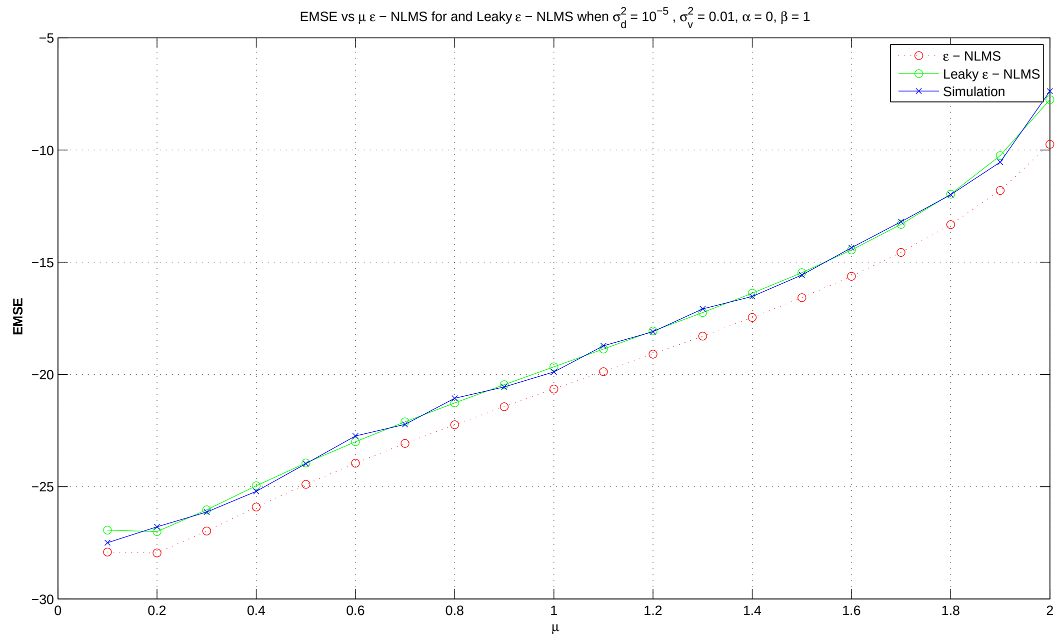

Figure 4.20: EMSE vs μ for ϵ - NLMS , Leaky ϵ - NLMS when $\sigma_d^2 = 10^{-5}$ $\sigma_v^2 = 10^{-2}$, and $\rho = 0.2$.

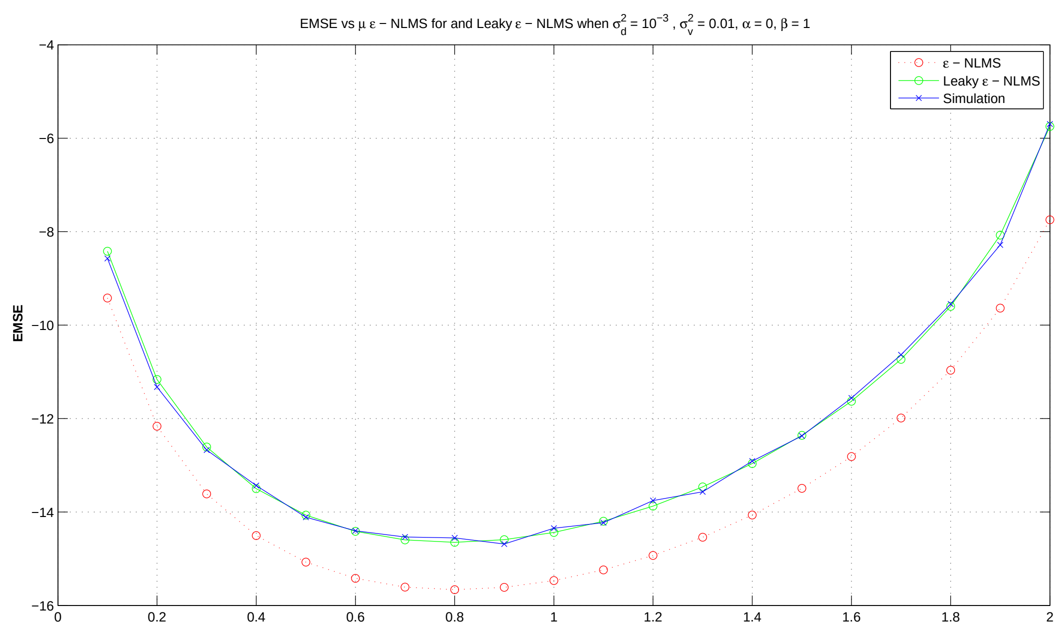

Figure 4.21: EMSE vs μ for ϵ - NLMS , Leaky ϵ - NLMS when $\sigma_d^2 = 10^{-3}$ $\sigma_v^2 = 10^{-2}$, and $\rho = 0.2$.

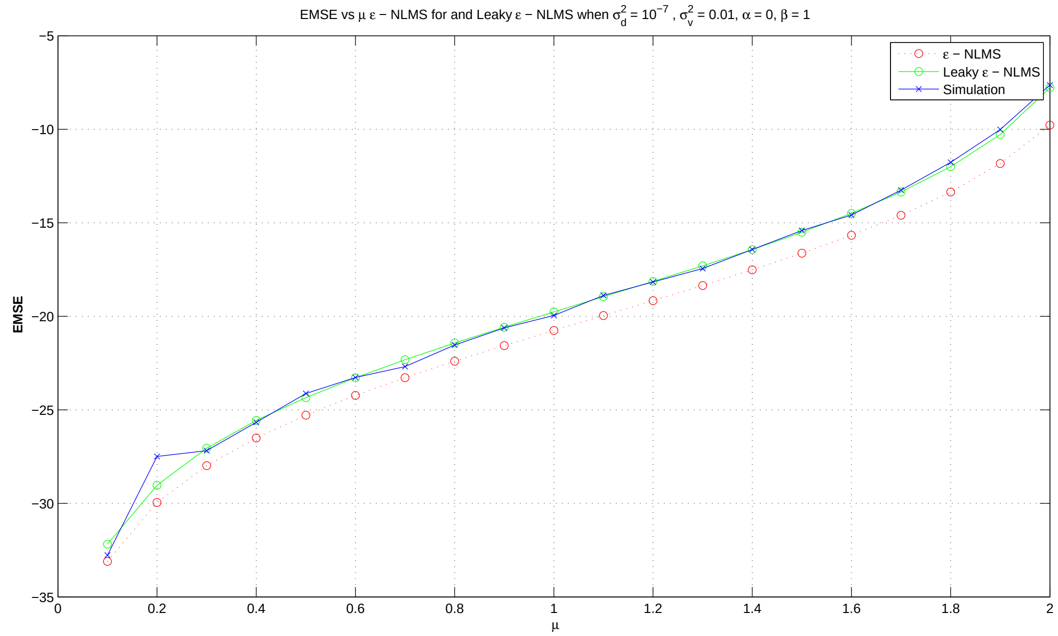

Figure 4.22: EMSE vs μ for ϵ - NLMS , Leaky ϵ - NLMS when $\sigma_d^2 = 10^{-7}$ $\sigma_v^2 = 10^{-2}$, and $\rho = 0.6$.

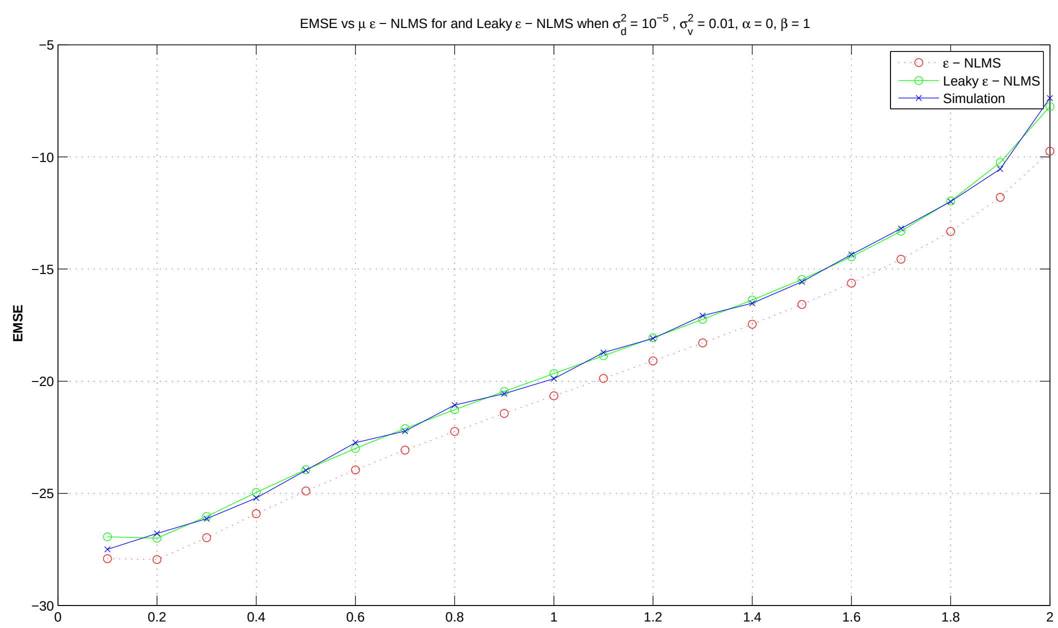

Figure 4.23: EMSE vs μ for ϵ - NLMS , Leaky ϵ - NLMS when $\sigma_d^2 = 10^{-5}$ $\sigma_v^2 = 10^{-2}$, and $\rho = 0.6$.

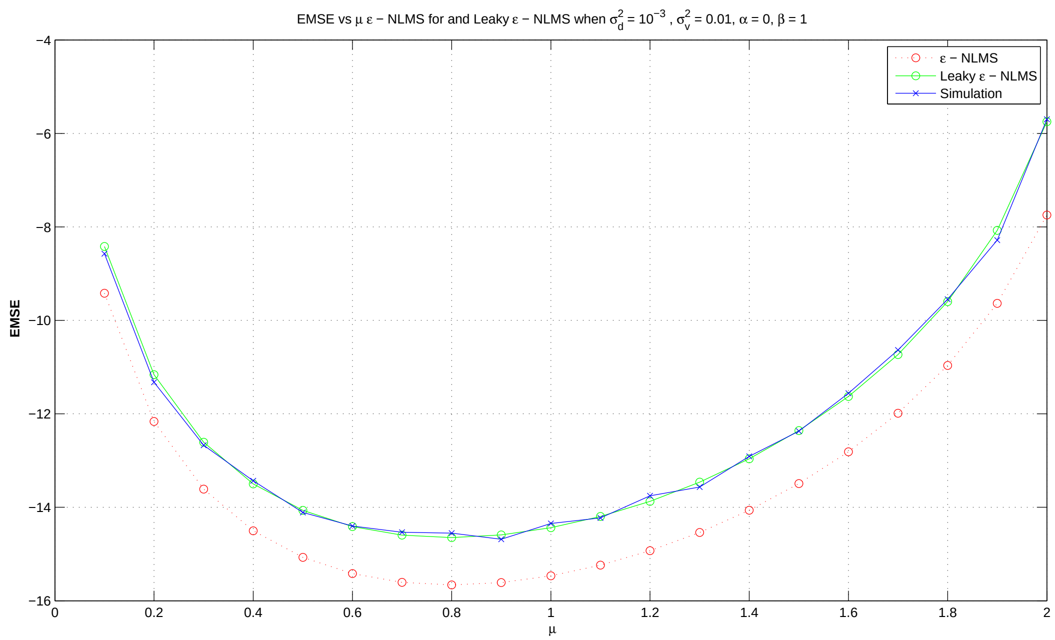

Figure 4.24: EMSE vs μ for ϵ - NLMS , Leaky ϵ - NLMS when $\sigma_d^2 = 10^{-3}$ $\sigma_v^2 = 10^{-2}$, and $\rho = 0.6$.

CHAPTER 5

CONCLUSIONS AND FUTURE WORK

5.1 Conclusions

In this thesis, the tracking analysis in a non-stationary environment of two adaptive filtering algorithms, the ϵ-NLMS algorithm and Leaky ϵ-NLMS algorithm, have been considered. In the first part of the thesis, the tracking analysis for the ϵ-NLMS algorithm was derived based on the energy conservation relation. The environment assumed to be nonstationary and it is changing with time according to a first order autoregrissive model. Finally, this part provides us with an expression for the steady state EMSE of the ϵ-NLMS algorithm where the input data is assumed to be correlated gaussian. This expression depends mainly on two moments (represented by α_u and η_u), previously, there are no close form expressions for these two moments, recently [38] derives the exact expressions for these moments, which give us the ability to get the exact expression for the steady state EMSE of the ϵ-NLMS algorithm.

The second part of the thesis provides us with an expression for the steady state EMSE of the leaky ϵ-NLMS algorithm. As in the first part of the thesis, the environment is assumed to be time varying environment and its variation model assumed to be a first order auto regressive model. The derivation of the steady state EMSE of the leaky ϵ-NLMS algorithm relies on the weighted energy conservation relation with the help of the independence assumption principle. The final expression of the EMSE also depends on some moments, again, these moments can't be expressed in a closed form expression, but with the help of [38], a closed form expressions for these moments are derived, so a closed form expression for the steady state EMSE is also available.

In the last part of this thesis, the accuracy of the derived theoretical expressions was examined. The simulation results for both cases (i.e., ϵ-NLMS and Leaky ϵ-NLMS) show that the analytical expressions matching the simulation results especially in slow changing environment, while some notable error appears when the adaptive filter works in a rapid changing environment, in addition to that, simulation shows that the derived expression for the Leaky ϵ-NLMS algorithm is more accurate than the one which was derived for the ϵ-NLMS algorithm, this is due to the fact that assumptions used to derive the EMSE expression for the ϵ-NLMS algorithm is weaker than the assumptions used to derive the same expression for the leaky ϵ-NLMS algorithm.

The simulation also shows some facts ablout the relation between the derived expressions for the EMSE and some paramters of the both algorithms ϵ-NLMS and Leaky ϵ-NLMS. The simulations show that the EMSE will decrease monotically against the ϵ parameter for both algorithms, this means that choosing ϵ close to 1 will give us lower EMSE, also the simulations try to detect the optimal step size for both algorithms by testing the EMSE against the step size of the choosen algorithm. At the end, the simulation shows that the algorithm will be inefficient to perform the tracking task in a very rapid changing environment but it will bo so efficient for slow changing one.

5.2 Future Work

The work applied here in this thesis to derive the EMSE of ϵ-NLMS and leaky ϵ-NLMS algorithms can be carried out for other algorithms, for example the tracking analysis of the least mean fourth LMF algorithm can be done using the same methodology which is used in this thesis. In addition to that, the time average of these two algorithms can be done in a nonstationary environment with a colored gaussian input.

REFERENCES

[1] S. U. H. Qureshi "Adaptive equalization" *IEEE Proceedings,* 53:1349-1387, 1985.

[2] B. Widrow, J. R. Glover, Jr., J. M. McCool, J. Kaunitz, C. S. Williams, J. R. Zeidler, E. Dong, Jr., and R. C. Goodlin "Adaptive noise cancelling: Principles and applications" *IEEE Proceedings,* 63(12):1692-1716, 1975.

[3] L. Makhoul "Linear predition a tutorail review" *IEEE Proceedings,* 63(4):561-580, 1975.

[4] C. F. N. Cowan and P. M. Grant "Adaptive Filters," *Englewood Cliffs, NJ:Prentice Hall,* 1985.

[5] S. Haykin "Adaptive Filter Theory," *Englewood Cliffs, NJ:Prentice Hall,* 1996.

[6] B. Widrow, J. M. McCool, M. G. Larimore, and C. R. Johnson "Stationary and Nonstationary Learning Characteristics of the LMS Adaptive Filter," *IEEE Proceedings,* 64(8):1151-1162, 1976.

[7] L. Ljung, M. Morf and D. Falconer "Fast calculation of gain matrices for recursive estimation schemes," *Int. J. Contr,* 27:1-19, 1978.

[8] D. Manolakis G. Carayannis and N. Kalouptsidis "A fast sequential algorithm for least squares filtering and prediction," *IEEE transactions on Acoustic, Speech and Signal Processing,* 31:1394-1402, 1983.

[9] J. M. Cioffi and T. Kailath "Fast recursive least squares transversal filters for adaptive filtering," *IEEE Transaction, Acoustic, Speech, Signal Processing,* 32:304-337, 1984.

[10] A. Benallal and A. Gillorie "A new method to stabilize Fast RLS algorithms based on a First Order Model of the propagation of numerical errors," *Proc. ICASSP88,* pages 1373-1376, 1988.

[11] D. Mansour and A. H. Gray "Unconstrained frequency domain adaptive filters," *IEEE Transaction Acoustic Speech and Signal Processing,* 30(5):726-734, 1982.

[12] K. Ozeki and T. Umeda "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties," *IEICE Transactions,* Japan, J67-A: 126-132, Feb 1984.

[13] J. I., Nagumo and A. Noda "A learning method for system identification," *IEEE Transactions, Automatic control,* AC-12:282-287, 1967.

[14] R. D. Gitlin, H. C. Meadors, and S. B. Weinstein "The Tap-Leakage Algorithm: An Algorithm for the stable operation of a digitally implemented, Fractionally Spaced adaptive equalizer," *Bell Syst. Tech. J.,* 61:1817-1839, 1982.

[15] G. A. Clark, S. K. Mitra, and S. R. Parker "Block implementatiom of adaptive digital filters," *IEEE Transaction Acoustic Speech and Signal Processing,* 29(3):744-752, 1981.

[16] R. H. Kwong and E. W. Johnston "A Variable Stepsize LMS algorithm," *IEEE transactions on signal procssing,* 40:1633-1642, 1992.

[17] T. J. Shan and T. Kailath "Adaptive algorithms with an automatic gain control feature," *IEEE transactions on Acoustic Speech and Signal Processing,* 35:122- 127, 1988.

[18] R. W. Harris, D. M. Chabries, and F. A. Bishop "A Variable Stepsize (VS) algorithm," *IEEE transactions on Acoustic, Speech and Signal Processing,* 34:499- 510, 1986.

[19] V. J. Methews and Z. Xie "A stochastic gradient adaptive filter with gradient adaptive step size," *IEEE transactions on Signal Processing,* 41:2075-2087,1993.

[20] T. Aboulnasr and K. Mayyas "A Robust Variable Step-Size LMS Type Algorithm: Analysis and Simulation," *IEEE transactions on Signal Processing,* 45:631-639, 1997.

[21] E. Walach and B. Widrow "The Least Mean Fourth(LMF) adaptive algorithm and its family," *IEEE transactions, Information Theory,* 30(2):275-283, 1984.

[22] A. Zerguine, C. F. N. Cowan, and M. Bettayeb "Novel LMS-LMF adaptive scheme for echo cancellation," *Electronic Letters.,* 32(19): 1776-1778, 1996.

[23] A. Zerguine, M. Bettayeb, and C. F. N. Cowan "Hybrid LMS-LMF scheme for echo cancellation," *IEE Proceedings-VIS Image Signal Processing.,* 146(4):173-180, 1999.

[24] A. Zerguine, C. F. N. Cowan, and M. Bettayeb "Adaptive echo cancellation using least mean mixed-norm," *IEEE Transactions Signal Processing.,* SP-45(5):1340-1343, 1997.

[25] J. A. Chambers, O. Tanrikulu, and A. G. Constantinides "Least mean mixed-norm adaptive filtering," *Electronic Letters,* 30(19):1574-1575, 1994.

[26] T. Aboulnasr and A. Zerguine "Variable Weight Mixed-Norm LMS-LMF Adaptive Algorithm," *IEEE conference on signals, systems and computers,* 1:791-794, 1999.

[27] J. G. Proakis "Digital Communications," *Singapore: McGraw-Hill,* 1983.

[28] B. Widrow, P.E. Mantey, L. J. Griffiths, and B.B. Goode "Adaptive Antenna Systems Array," *IEEE Proceedings,* 55:2143-2159, 1967.

[29] C. W. K. Gritton and D. W. Lin "Echo Cancellation Algorithms," *IEEE ASSP Magazine,* pages 30-37, 1984.

[30] K. Murano, S. Unagami, and F. Amano "IEEE Communication Magazine," *IEEE ASSP Magazine,* 28(l):49-55, 1990.

[31] N. J. Bershad "Analysis of the normalized LMS algorithm with gaussian inputs," *IEEE transactions on Acoustic, Speech and Signal Processing,* ASSP-34:793- 806, 1986.

[32] N.R. Yousef and A.H. Sayed "Ability of adaptive filters to track carrier offsets and channel nonstationarities," *IEEE Transactions on Signal Processing,* Vol.50, Issue 7, pp. 1533-1544, 2002.

[33] A.H. Sayed and T.Y. Al-Naffouri "Mean-square analysis of normalized leaky adaptive filters," *Proc, ICASSP, Salt Lake City, UT,* vol.6, pp. 3873-3876, 2001.

[34] A.H. Sayed "Fundamentals of Adaptive Filtering," *Wiley, NJ,* , 2008.

[35] T.Y. Al-Naffouri and A.H. Sayed "Transient Analysis of Data-Normalized Adaptive Filters," *IEEE Transactions on Signal Processing,* vol.51, no. 3, pp. 639-652, 2003.

[36] T.Y. Al-Naffouri and A.H. Sayed "Transient Analysis of Adaptive Filters -PARTI: The Data Nonlinearity case ," *Proc, IEEE-EURASIP, Baltimore, Maryland,* 2001.

[37] T.Y. Al-Naffouri and A.H. Sayed "Transient Analysis of Adaptive Filters -PARTII: The Error Nonlinearity case," *Proc, IEEE-EURASIP, Baltimore, Maryland,* 2001.

[38] T.Y. Al-Naffouri and M. Moinuddin "Exact performance analysis of the e-NLMS algorithm for colored circular Gaussian inputs ," *IEEE Transactions on Signal Processing,* vol.58, no. 10, pp. 5080-5090, Oct.2010.

Vitae

- Mahmoud Jamal Ankeer
- Born in Amman, Jordan on February 17, 1981
- Jordanian Nationality.
- Received Bachelor of Science (B.Sc) in Information Technology from Al-Balqa Applied University, Jordan in July 2002.
- Joined King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia as a Research Assistant in February 2007.
- Completed Master of Science (M.S.) in Electrical Engineering in January 2011.
- Email: *mahmoud_ankeer@hotmail.com*
- Mobile: +962 - 796903278
- Present And Permanent Address: Naur, Amman, Jordan.