QUASI-NEWTON LEAST MEAN FOURTH ADAPTIVE ALGORITHM

BY

UMAIR BIN MANSOOR

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

June 2009

KING FAHD UNIVERSITY OF PETROLEUM AND MINERALS

DHAHRAN 31261, SAUDI ARABIA

DEANSHIP OF GRADUATE STUDIES

This thesis, written by

UMAIR BIN MANSOOR

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

MASTER OF SCIENCE IN TELECOMMUNICATION ENGINEERING

Thesis Committee

Dr. Zerguine, A. (Chairman)

Dr. Al-Naffouri, T.Y. (Member)

Dr. Abu-Saud, Wajih. (Member)

Dr. Samir H. Abdul-Jauwad
Department Chairman

Dr. Zummo, Salam A.
Dean of Graduate Studies

15/7/09
Date

Dedicated to

My

Beloved Parents, Brothers

and My Nephew

Yahya

Acknowledgment

All praise belongs to Allah, glorified is He and exalted. Who caused this work to be completed successfully. Who gave me the opportunity, strength and persistence to work on it. And Who helped me in the most difficult of times. I'm happy to have had an opportunity to glorify His name in the sincerest way through this small accomplishment and pray to Him to accept my efforts.

I would like to acknowledge the role of the King Fahd University of Petroleum & Minerals in extending a generous financial assistance to me and for providing a wonderful environment, academic and otherwise, which made my stay at KFUPM a memorable experience.

My deep appreciation goes to my thesis advisor Dr. Azzedine Zerguine for his constant help and guidance and for his compassionate attitude.

I'm greatful to my brother Saad bin Mansoor for helping me with the overall understanding of my thesis problem or else my work have been delayed.

Very special thanks to Dr. Tariq Y. Al-Naffouri, Dr. Moinuddin, Syed Ali Aamir Imam, Obaid Khattak, Muzhar Azeem and Kashif Paracha. They extended their help to me in very difficult times.

Thanks to my parents and brothers for their moral and financial support. Their emotional support was invaluable during my stay at KFUPM.

I would also like to recognize the support of my friends with whom I spent considerable time here at KFUPM, (in alphabetical order) Abdul Karim, Abubakar, Aftab, Ahmed, Akhlaq, Al-Tayyeb, A. Salim, Asif, Ayaz, Babar, Farhan, Haris, Junaid, Kashif, M. Amalou, Muzhar, Osama, Raheel, Sabih, Salman K., Salman Y., Saqib, Umer, Zeehsham and Zeeshan. I spent an excellent time with them.

Contents

List of Figures

List of Tables

Abstract

Name: Umair bin Mansoor

Title: Quasi-Newton Least-Mean Fourth Adaptive Algorithm

Degree: Master of Science

Major Field: Telecommunications

Date of Degree: June 2009

In this thesis, a novel Quasi-Newton Least-Mean Fourth (QNLMF) adaptive filtering algorithm is analyzed. The main aim of this research is to derive the QNLMF adaptive algorithm and assess its performance in different noise environments. More specifically, both the convergence analysis and the steady-state performance analysis were derived. Finally, a number of simulation results were carried out to corroborate the theoretical findings.

Master of Science Degree
King Fahd University of Petroleum and Minerals
Dhahran, Saudi Arabia
June 2009

ملخص الرسالة

الاسم: عمير بن منصور

العنوان: قانون حساب يشابه نيوتن لأقل المتوسطات الرابعة

الدرجة العلمية: ماجستير

التخصص: الاتصالات البعيدة

تاريخ الدرجة العلمية: June 2009

في هذه الرسالة تم تحليل نظام حساب جديد و هو "الفلتر التكييفي المشابة لنيوتن لأقل المتوسطات الرابعة " ، الهدف الأساسي للبحث هو اشتقاق نظام الحساب الجديد و تقييم أدائه في في بيئات تشويش مختلفة ، و بشكل خاص ، تم اشتقاق نظام حساب التحليل التقاربي و نظام حساب تحليل أداء الحالة المستقرة. أخيرا تم تحصيل بعض النتائج من خلال برامج المحاكاة لتدعيم المعلومات النظرية.

ماجستير العلوم

جامعة الملك فهد للبترول و المعادن

الظهران ، المملكة العربية السعودية

June 2009

Chapter 1

Introduction

One example of a digital signal processing system is called filtering. Filtering is a signal processing operation whose objective is to process a signal in order to manipulate the information contained in the signal. In other words, a filter is a device that maps its input signal to another output signal facilitating the extraction of the desired information contained in the input signal. In case of a time-invariant filter the internal parameters and the structure of the filter are fixed, and if the filter is linear the output signal is a linear function of the input signal.

1.1 Adaptive filters

Adaptive systems are playing a vital role in the development of modern communications. The concept of adaptive filtering constitutes an important part of the statistical signal processing. Whenever there is a requirement to process signals that result from an unknown statistics of an environment, the use of an adative filter offers an attractive solution to the problem.

An adaptive filter is required when either the fixed specifications are unknown or the specifications cannot be satisfied by time-invariant filters. To be specific, an adaptive filter is a nonlinear filter since its characteristics are dependent on the input signal and consequently the homogeneity and additivity conditions are not satisfied. In our case, however, the adaptive filter will be considered linear in the sense that its output signal is a linear function of its input signal.

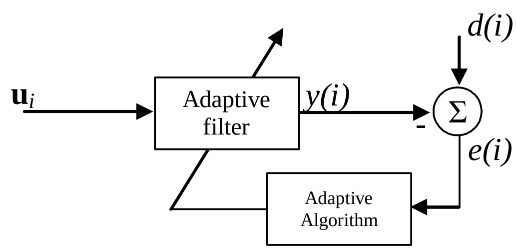

Figure 1.1: General adaptive filter configuration.

The complete specification of an adaptive system, as shown in Fig. 1.1, consists of two items:

1.1.1 Adaptive Filter Structure

The adaptive filter can be implemented in a number of different structures or realizations. The choice of the structure can influence the computational complexity (amount of airthmetic operations per iteration) of the process and also the necessary number of iterations to achive a desired performance level. Basically, there are two major classes of adaptive digital filter realizations, distinguished by the form of the impulse response, namely the finite duration impulse response (FIR) filter and the infinite-duration impulse response (IIR) filters. FIR filters are usually implemented with nonrecursive structures, whereas IIR filters utilize recursive realizations.

1.1.2 Algorithm

The algorithm is the procedure used to adjust the adaptive filter coefficients in order to minimize a prescribed criterion. The algorithm is determined by defining the search method (or minimization algorithm), the objective function, and the error signal nature. The choice of the algorithm determines several crucial aspects of the overall adaptive process, such as existance of sub-optimal solutions, biased optimal solution, and computational complexity.

1.2 Application

The type of application is defined by the choice of the signals acquired from the environment to be the input and desired output signals. The number of different applications in which adaptive techniques are being successfully used has increased enormously during the last decade. In the ensuing, some very common examples are discussed [3]:

1.2.1 System Identification

System Identification is the experimental approach to the modelling of a process or a plant. It involves the following steps: experimental planning, the selection of a model structure, parameter estimation and model validation. The procedure of system identification, as persued in practice, is iterative in nature in that we have to go back and fourth in these steps until a satisfactory model is built. The system to be identified is unknown which can be stationary or time varying. Figure 1.1.1 depicts a system identification scenario.

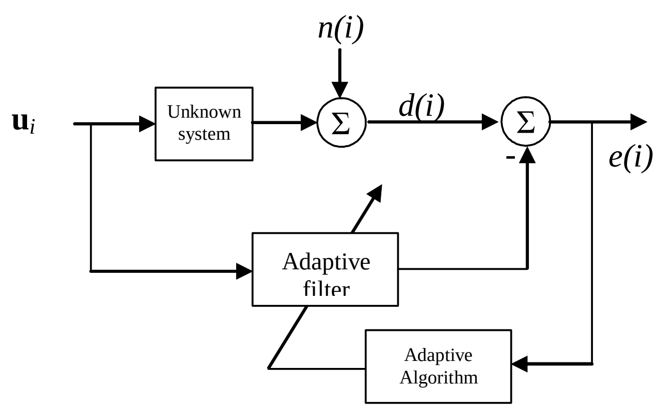

Figure 1.2: System identification scenario.

1.2.2 Inverse Modelling or Equalization

In this application, the adaptive filter is used to represent the best fit of an unknown noisy plant. Thus, at convergence, the inverse of the transfer function of the unknown system is approximated by the adaptive filter. A delay is introduced into the desired response path, as shown in Figure 1.1.2, so as to ensure that the input to the adaptive filter is minimum phase and suitable for equalization.

The primary use of the inverse modelling is to reduce inter-symbol interference (ISI) in digital receivers. This is achieved through the use of channel equalization in

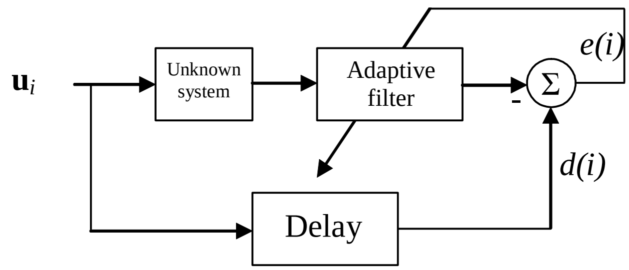

Figure 1.3: Equalization scenario.

digital communications [4].

1.2.3 Noise Cancellation

In this class of application, the adaptive filter is used to cancel unknown interference contained in a primary signal, as shown in Figure 1.1.3. The primary signal serves as the desired response of the adaptive filter. This type of application is used in adaptive beamforming or in adaptive noise cancellation [5].

1.2.4 Prediction

Finally, in this application, the adaptive filter is used to provide the best prediction of the present value of the input signal from its previous values. The desired signal, $d(i)$,

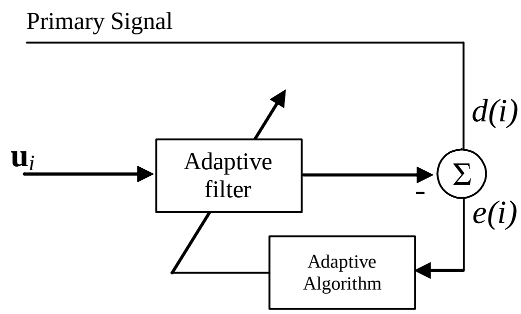

Figure 1.4: Noise Cancellation Scenario.

is the instantaneous value and the input of the adaptive filter is a delayed version of the same signal. This application is used in linear prediction coding (LPC) of speech [6] and in adaptive differential pulse-code modulation (DPCM) [7].

1.3 Adaptive Filtering Algorithms

As mentioned above, an adaptive algorithm refers to the criteria by which a filter is adapted in response to the outside environment. Let $\mathbf{w}_i$ 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 $\mathbf{w}_i$ is given by:

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu g\left(e(i)\right)\mathbf{u}_i, \tag{1.1}$$

where $\{\mathbf{u}_i\}$ is the input sequence, $e(i)$ is the adaptive error, $g\left(e(i)\right)$ is a function of the error and μ denotes the positive step-size which may be time varying. Some of the well known algorithms are discussed below.

1.3.1 Least Mean Squares (LMS) Algorithm

The above equation represents the method of Steepest Descent [14] that can be described as

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \left(-\nabla_{\mathbf{w}} J\left(\mathbf{w}_{i-1}\right)\right). \tag{1.2}$$

The next weight vector, $\mathbf{w}_{i+1}$, equals the present weight vector, $\mathbf{w}_i$, plus a change which is proportional to the negative gradient. The proportionality constant is μ, and this is a design parameter that controls stability and rate of convergence. The LMS algorithm is obtained from the equation (1.1) and using the error signal

$$e(i) = d(i) - y(i), \tag{1.3}$$

as

$$\mathbf{w}_i = \mathbf{w}_{i-1} + 2\mu e(i)\mathbf{u}_i, \tag{1.4}$$

$$e(i) = d(i) - \mathbf{u}_i^T \mathbf{w}_{i-1}. \tag{1.5}$$

The gradient of each iteration is instantaneous and is given by $-2e(i)\mathbf{u}_i$.

When using the steepest descent to find the minimum of a quadratic function of the weights, the weights progress geometrically towards the Wiener solution [14]. It has been shown [8] that there are as many distinct time constants as there are distinct eigenvalues of the input vector autocorrelation matrix $\mathbf{R}_{\mathbf{u}}$. These time constants depend on the eigenvalues of $\mathbf{R_u}$ and corresponds to natural modes of the adaptive algorithm. The relative amplitude of the modes are different from one weight to another and depend on the initial conditions of the weight vector, i.e., its initial value. Since one rarely has a priori knowledge of the orientation of the initial weight vector setting with respect to the eigenvectors of $\mathbf{R_u}$, it is difficult to predict the relative amplitudes of the modes and therefore difficult to predict the rate of convergence of the LMS algorithm. Inspite of this drawback, the LMS algorithm is very widely used.

Newton's Algorithm

A more predictable algorithm is Newton's method,

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \left(\nabla^2_{\mathbf{w}} J\left(\mathbf{w}_{i-1}\right)\right)^{-1} \left(-\nabla_{\mathbf{w}} J\left(\mathbf{w}_{i-1}\right)\right). \tag{1.6}$$

The second derivative of the cost function is a Hessian matrix which, in the case of the Mean-square-error criteria, is the input vector autocorrelation matrix $\mathbf{R_u}$. Assuming that $\mathbf{R_u}$ is not singular, the LMS-Newton algorithm can be written as

$$\mathbf{w}_i = \mathbf{w}_{i-1} + 2\mu \mathbf{R_u}^{-1} e(i)\mathbf{u}_i. \tag{1.7}$$

With Newton's method, there is only one natural mode, corresponding to one time constant. The learning rate of Newton's method is independent of the weight vector's initial condition.

The rate of convergence of the LMS-Newton is predictable and does not depend on initial conditions. The drawback lies in that one cannot implement this algorithm in practice because $\mathbf{R}_{\mathbf{u}}^{-1}$ is generally unknown. LMS, based on steepest descent, has disadvantages, but it is simple and easy to implement. It performs equivalently to LMS-Newton under many important conditions [14].

1.3.2 Least Mean Fourth (LMF) Algorithm

Adaptive algorithms based on higher order moments of the error signal have been shown to perform better mean square estimation than the well known Least Mean Square (LMS) algorithm in some important applications. The Least Mean Fourth (LMF) is one of such algorithms [8]. It seeks to minimize the mean fourth error, which is a convex function of the adaptive weight vector [2].The power of the LMF algorithm lies in its faster initial convergence and lower steady-state error relative to the LMS algorithm under sub-Gaussian noise environment [8]-[11],[13],[17].

It has been mentioned in [11] that the LMF algorithm can outperform LMS for non-Gaussian additive noise. In such a case, the LMF algorithm can lead to considerably smaller excess mean square error (MSE) for the same convergence speed. According to LMF algorithm, the filter coefficients are adapted according to the following recursion:

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu e^3(i)\mathbf{u}_i, \tag{1.8}$$

where $\{\mathbf{u}_i\}$ is the input sequence, $e(i)$ is the adaptive error, and μ is a positive constant called the step-size which is used to control the size of the incremental correction applied to the tap weights as it proceeds from one iteration to the next.

It has been shown in [25] that the LMF algorithm is never stable in the mean-square sense for gaussian regressors. Nevertheless, results based on standard mean-square stability analysis are useful for practical design purposes. This is because the probability of divergence as a function of the step-size value tends to rise abruptly only when it moves past a given threshold. Before that, the probability of divergence tends to be sufficiently small to grant the practical applicability of the LMF algorithm (in practical applications it may be of interest to include a re-initialization scheme in case, for instance, the error signal tends to increase without bound). Moreover, signal amplitudes are necessarily limited in practical applications, which contributes to reducing the probability of divergence for small step sizes smaller than the threshold mentioned above [25].

Another relevent aspect of the LMF algorithm behavior is its steady-state stability. Depending on the step-size and on the initial condition, the LMF probability of divergence may increase considerably with the number of iterations. However, if the algorithm is initialized close to the optimum solution and one chooses a large step-size, it may have a significant probability of divergence also after initial convergence [25].

1.4 Thesis Objectives and Organization

In this thesis work, Newton's method based least mean fourth adaptive algorithm namely Quasi-Newton Least Mean Fourth (QNLMF) is proposed. The basic update recursion for QNLMF is as follows

$$\mathbf{w}_i = \mathbf{w}_{i-1} - \mu \left[\bigtriangledown^2_{\mathbf{w}} J\left(\mathbf{w}_{i-1}\right)\right]^{-1} \left[\bigtriangledown_{\mathbf{w}} J\left(\mathbf{w}_{i-1}\right)\right]^T . \tag{1.9}$$

The main objective of this thesis are: First, to derive the basic algorithm, namely the QNLMF algorithm from the Newton's method. Second, to analyze the steady-state performance of the proposed algorithm and to derive the expression for the excess mean-square error. Third, to examine the convergence properties of the proposed algorithm. Fourth, to analyze the tracking performance of the proposed algorithm and to derive a mathematical expression for the tracking excess mean-square error. Fifth, to analyze the transient-state performance of the proposed algorithm and to derive a mathematical expression for mean-square error and mean-square deviation during the transient phase. Finally, to present simulation scenarios to support the analytical analysis.

This thesis is organised so as to achive all the above mentioned objectives. In Chapter 2, the proposed algorithm is derived using the concept of energy conservation [2]. This technique is used extensively to carry out the different analyses. In Chapter 3, the steady-state analysis is performed and expressions for excess mean-square error of the proposed algorithm is derived. A comparison of steady-state excess mean square error is carried out between the proposed algorithm and that of the conventional LMF. The tracking analysis of the proposed algorithms in a non-stationary enviornment is presented in Chapter 4 and the mathematical expression is derived for the tracking excess steady-state error for the algorithm. In Chapter 5, a complex mathematical model for the transient analysis of the algorithm is derived.

In support of the mathematical analysis listed above, the simulation scenario in different noise enviornments and the analytical results when compared with the experimental ones are presented in Chapter 6. Finally, thesis conclusions, contributions and recommendations for future work are presented in Chapter 7.

Chapter 2

Proposed Adaptive Algorithm

2.1 Introduction

There is a range of techniques available for trained (and decision directed) identification of linear FIR channels with additive white Gaussian noise (AWGN), which are broadly classified into two classes: adaptive and model-based. The adaptive algorithms do not explicitly use a model for the channel coefficients or noise and these include least mean squares (LMS) [1], recursive least squares (RLS) [1], and their derivatives. The model-based algorithms, however, use various type of models for the channel coefficients (e.g., random walk, autoregressive, or constant) and noise, where the model parameters are either known or jointly estimated with the channel. Many adaptive algorithms can be interpreted in a model based framework with data-dependent choice of model parameters. Also, some adaptive algorithms implicitly use model parameters to set the algorithm parameters, e.g., step-size and forgetting factors require partial knowledge of input statistics to guarantee stable behavior.

Iterative algorithms, in general, follow the basic update recursion,

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \mathbf{p}, \tag{2.1}$$

where, $\mathbf{w}_i$ is the adaptive weight vector which after convergence, in a system identification/channel estimation model, represents the unknown system, μ is the step-size which is responsible for the stability of the system and helps the system to gradually reach a minimum, and $\mathbf{p}$, is the direction vector that varies for varying adaptive algorithms. It is the main factor which is responsible for the algorithm to reach a global minimum, i.e., as $i \longrightarrow \infty$ it guarantees $\mathbf{w}_i \longrightarrow \mathbf{w}^o$ where, $\mathbf{w}^o$ represents the weights of the unknown system/channel. Convergence, in general, can be mathematically described as the condition when the last term on the right side of (2.1) becomes zero i.e., $\mathbf{w}_i = \mathbf{w}_{i-1}$, hence, the adaptive system has completely adapted the unknown system.

2.2 Proposed Algorithm

To develop the proposed algorithm a time invariant channel model is considered such that:

$$d(i) = \sum_{j=0}^{N-1} u(j) w^o(i-j) + n(i) = \mathbf{u}_i^T \mathbf{w}^o + n(i), \quad i = 0, 1, 2, 3..., \tag{2.2}$$

where $\{u(j)\}$ is a stationary input process with mean zero and variance σ_u^2, $\{n(i)\}$ is a stationary noise process with mean zero and variance σ_n^2, and $\mathbf{w}^o$ corresponds to a channel/impulse response with N taps.

Under the above model, minimizing the mean fourth error,

$$J(\mathbf{w}_i) = E\left[e^4(i)\right], \tag{2.3}$$

$$J(\mathbf{w}_i) = E\left[d(i) - \mathbf{u}_i^T \mathbf{w}_i\right]^4, \tag{2.4}$$

over $\mathbf{w}$ gives the optimal weight value $\mathbf{w}_i = \mathbf{w}^o$. Also, it should be noted that $J(\mathbf{w}^o) = E\left[n^4(i)\right]$.

The update direction vector $\mathbf{p}$ in equation (1.8), as derived in [2], can be written as follows,

$$\mathbf{p} = -\mathbf{B}\left[\bigtriangledown_{\mathbf{w}} J(\mathbf{w}_{i-1})\right]^T. \tag{2.5}$$

The special choice $\mathbf{B} = \mathbf{I}$ is very common and it corresponds to the update direction

$$\mathbf{p} = -\left[\bigtriangledown_{\mathbf{w}} J(\mathbf{w}_{i-1})\right]^T, \tag{2.6}$$

which leads us to the famous steepest-descent method

$$\mathbf{w}_i = \mathbf{w}_{i-1} - \mu\left[\bigtriangledown_{\mathbf{w}} J(\mathbf{w}_{i-1})\right]^T, \quad i \geq 0, \quad \mathbf{w}_{-1} = \text{initial guess}, \tag{2.7}$$

in which, the successive weight vectors $\{\mathbf{w}_i\}$ are obtained by descending along a path of decreasing cost values.

In this work, $\mathbf{B} = \{\bigtriangledown^2_{\mathbf{w}} J(\mathbf{w}_{i-1})\}^{-1}$, the inverse of the Hessian matrix, leads to the well known Newton's recursive update:

$$\mathbf{w}_i = \mathbf{w}_{i-1} - \mu \left[\bigtriangledown^2_{\mathbf{w}} J(\mathbf{w}_{i-1})\right]^{-1} \left[\bigtriangledown_{\mathbf{w}} J(\mathbf{w}_{i-1})\right]^T, \quad i \geq 0, \quad \mathbf{w}_{-1} = \text{initial guess.} \tag{2.8}$$

2.2.1 Quasi-Newton Least Mean Fourth (QNLMF)

In this work, the update recursion for the LMF will be derived using Newton's method. In order to achieve our desired relation we first need to evaluate the gradient vector and the Hessian matrix of the cost function (2.1), derived in the App. I and are, respectivly, given by

$$\bigtriangledown_{\mathbf{w}} J(\mathbf{w}_{i-1}) = -4E\left[e^3(i)\mathbf{u}_i^T\right], \tag{2.9}$$

and

$$\bigtriangledown^2_{\mathbf{w}} J(\mathbf{w}_{i-1}) = 12E\left[e^2(i)\mathbf{u}_i\mathbf{u}_i^T\right]. \tag{2.10}$$

These relations are extracted from the expanded Kronecker's form of the cost function (2.1), which can be shown to be set up as:

$$\begin{aligned} E\left[e^4(i)\right] \;=\;& E[d^4(i)] - 4E[d^3(i)\mathbf{u}_i^T]\mathbf{w} + 6\mathbf{w}^T E[d^2(i)\mathbf{u}_i\mathbf{u}_i^T]\mathbf{w} \\ & -4E\left[d(i)\mathbf{u}_i^T \otimes \mathbf{u}_i^T \otimes \mathbf{u}_i^T\right] \mathbf{w} \otimes \mathbf{w} \otimes \mathbf{w} \\ & +\mathbf{w}^T E\left[\mathbf{u}_i\mathbf{u}_i^T \otimes \mathbf{u}_i^T \otimes \mathbf{u}_i^T\right] \mathbf{w} \otimes \mathbf{w} \otimes \mathbf{w}, \end{aligned} \tag{2.11}$$

where $\otimes$ denotes the Kronecker's product and is defined as:

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{bmatrix}$$

Now, using the results from (2.9) and (2.10),we can rewrite (2.8) as,

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \left[\epsilon(i)\mathbf{I} + \hat{\mathbf{R}}_i\right]^{-1} \left[e^3(i)\mathbf{u}_i^T\right]^T, \quad i \geq 0, \quad \mathbf{w}_{-1} = \text{initial guess,} \tag{2.12}$$

where, $\epsilon(i)$ is a small positive scalar that will prevent the Hessian matrix from becoming singular and $\hat{\mathbf{R}}_i$ is the approximation for the actual Hessian which, is given as,

$$\hat{\mathbf{R}}_i = \alpha \sum_{j=0}^{i} (1-\alpha)^{i-j} e^2(j)\mathbf{u}_j\mathbf{u}_j^T. \tag{2.13}$$

Let $\mathbf{\Phi}_i$ be

$$\mathbf{\Phi}_i \triangleq \epsilon(i)\mathbf{I} + \hat{\mathbf{R}}_i, \tag{2.14}$$

with $\epsilon(i) = (1-\alpha)^{(i+1)}\epsilon$ and ϵ a small positive scalar.

After some algebric manipulation we can show that $\mathbf{\Phi}_i$ recursively follows the following recursion:

$$\mathbf{\Phi}_{i-1} = (1-\alpha)^i \epsilon\mathbf{I} + \alpha \sum_{j=0}^{i-1} (1-\alpha)^{i-j-1} e^2(j)\mathbf{u}_j\mathbf{u}_j^T. \tag{2.15}$$

Multiplying both sides of (2.15) by $(1-\alpha)$, one obtains

$$(1-\alpha)\,\mathbf{\Phi}_{i-1} = (1-\alpha)^{(i+1)}\,\epsilon\mathbf{I} + \alpha\sum_{j=0}^{i-1}(1-\alpha)^{i-j}\,e^2(j)\mathbf{u}_j\mathbf{u}_j^T. \tag{2.16}$$

Finally, $\mathbf{\Phi}_i$ in its recursion format:

$$\mathbf{\Phi}_i = (1-\alpha)\,\mathbf{\Phi}_{i-1} + \alpha e^2(i)\mathbf{u}_i\mathbf{u}_i^T. \tag{2.17}$$

The next step involves finding out the inverse of the Hessian matrix which will be evaluated using the matrix inversion lemma [1] on (2.17). This formula is a very useful matrix theory result. The result states that for arbitrary matrices $\{A, B, C, D\}$ of compatible dimensions, if A and C are invertible, then

$$(A+BCD)^{-1} = A^{-1} - A^{-1}B\left(C^{-1}+DA^{-1}B\right)^{-1}DA^{-1}. \tag{2.18}$$

Using the matrix inversion lemma, the inverse of (2.17), that is $\mathbf{\Phi}_i^{-1}$, is evaluated next. First, $\mathbf{\Phi}_i^{-1}$ looks like

$$\mathbf{\Phi}_i^{-1} = \left\{(1-\alpha)\,\mathbf{\Phi}_{i-1} + \alpha e^2(i)\mathbf{u}_i\mathbf{u}_i^T\right\}^{-1} \tag{2.19}$$

Applying the matrix inversion lemma, one gets:

$$\begin{aligned}\mathbf{\Phi}_i^{-1} &= \{(1-\alpha)\,\mathbf{\Phi}_{i-1}\}^{-1} \\ &- \{(1-\alpha)\,\mathbf{\Phi}_{i-1}\}^{-1}\mathbf{u}_i\left[\left(\alpha e^2(i)\right)^{-1} + \mathbf{u}_i^T\{(1-\alpha)\,\mathbf{\Phi}_{i-1}\}^{-1}\mathbf{u}_i\right]^{-1}\mathbf{u}_i^T\{(1-\alpha)\,\mathbf{\Phi}_{i-1}\}^{-1},\end{aligned} \tag{2.20}$$

where, we have used the following in the derivation: $A = (1-\alpha)\,\mathbf{\Phi}_{i-1}$, $B = \mathbf{u}_i$, $C = \alpha e^2(i)$ and $D = \mathbf{u}_i^T$.

Eventually, after some arrangements one can set up $\mathbf{\Phi}_i^{-1}$ in the following format:

$$\mathbf{\Phi}_i^{-1} = \frac{1}{(1-\alpha)}\left[\mathbf{\Phi}_{i-1}^{-1} - \frac{\mathbf{\Phi}_{i-1}^{-1}\mathbf{u}_i\mathbf{u}_i^T\mathbf{\Phi}_{i-1}^{-1}}{\frac{(1-\alpha)}{\alpha e^2(i)} + \mathbf{u}_i^T\mathbf{\Phi}_{i-1}^{-1}\mathbf{u}_i}\right]. \tag{2.21}$$

Finally, substituting $\mathbf{P}_i$ for $\mathbf{\Phi}_i^{-1}$, (2.21) looks like the following:

$$\mathbf{P}_i = \frac{1}{(1-\alpha)}\left[\mathbf{P}_{i-1} - \frac{\mathbf{P}_{i-1}\mathbf{u}_i\mathbf{u}_i^T\mathbf{P}_{i-1}}{\frac{(1-\alpha)}{\alpha e^2(i)} + \mathbf{u}_i^T\mathbf{P}_{i-1}\mathbf{u}_i}\right]. \tag{2.22}$$

Ultimately, the final update recursion of QNLMF algorithm can be written as

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu\mathbf{P}_i e^3(i)\mathbf{u}_i, \quad i \geq 0, \quad \mathbf{w}_{-1} = \text{initial guess}, \tag{2.23}$$

which updates the adaptive weights $\mathbf{w}_i$, iteratively.

Before proceeding to the next stage of studying the convergence analysis of the proposed algorithm, in the following section the concept of fundamental energy conservation is introduced [2]. This is very fundamental for this research as it leads very trackable and easy analysis when compared to its counterpart derived through the concept of ensemble averaging [1].

2.3 Fundamental Energy Conservation Method

The generic form of an adaptive filter update is given by:

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \mathbf{P}_i g\left[e(i)\right] \mathbf{u}_i, \tag{2.24}$$

where $\mathbf{u}_i$ is the input sequence, μ is a positive constant called the step-size, and $g\left[e(i)\right]$ denotes some function of the error signal.

The above generic update recursion in terms of the weight-error vector can be shown to be:

$$\mathbf{v}_i = \mathbf{v}_{i-1} - \mu \mathbf{P}_i g\left[e(i)\right] \mathbf{u}_i. \tag{2.25}$$

where $\mathbf{v}_i = \mathbf{w}^o - \mathbf{w}_i$ represents the weight error vector.

Now, let us define two kind of errors known as a-priori estimation error, $e_a(i)$, and a-psteriori estimation error, $e_p(i)$, respectively, as follows:

$$e_a(i) \triangleq \mathbf{u}_i^T \mathbf{v}_{i-1}, \tag{2.26}$$

and

$$e_p(i) \triangleq \mathbf{u}_i^T \mathbf{v}_i. \tag{2.27}$$

Both of these estimation errors $e_a(i)$ and $e_p(i)$ are related according to:

$$e_p(i) = e_a(i) - \mu \left\| \mathbf{u}_i \right\|_{\mathbf{P}_i}^2 g\left[e(i)\right]. \tag{2.28}$$

The above equation provides an alternative description of an adaptive filter in terms of the error quantities $e_a(i)$, $e_p(i)$, $\mathbf{v}_{i-1}$, $\mathbf{v}_i$ and $g\left[e(i)\right]$. This description is useful as we are often interested in questions related to the behavior of these errors, such as:

1. **Steady-state behavior**: which relates to determining the steady-state values of $E\left\|\mathbf{v}_{i-1}\right\|^2$, $E\left[\left|e_a(i)\right|^2\right]$ and $E\left[\left|e(i)\right|\right]^2$.

2. **Stability**, which relates to determining the range of values of the step-size over which the variance $E\left[\left|e_a(i)\right|^2\right]$ and $E\left\|\mathbf{v}_{i-1}\right\|^2$ remain bounded.

3. **Transient behavior**, which relates to studying the time evolution of the curves $E\left[\left|e_a(i)\right|^2\right]$, $E\left[\mathbf{v}_{i-1}\right]$ and $E\left\|\mathbf{v}_{i-1}\right\|^2$.

Now in order to answer the above questions, we look forward to an energy equality [2] that relates all the squared norms of the errors. To derive the energy relation, all the above equations are combined together to eliminate the error non-linearity function $g\left[\cdot\right]$, this means that the resulting energy relation will hold irrespective of the error nonlinearity. Hence, two cases will be considered:

1. $\underline{\mathbf{u}_i = 0}$. This is a degenerate situation. In this case, it is obvious from the above equations that $\mathbf{v}_i = \mathbf{v}_{i-1}$ and $e_p(i) = e_a(i)$ so that $E\left\|\mathbf{v}_i\right\|^2 = E\left\|\mathbf{v}_{i-1}\right\|^2$ and $E\left|e_a(i)\right|^2 = E\left|e_p(i)\right|^2$.

2. $\underline{\mathbf{u}_i \neq 0}$. In this case, the relation between a-priori and a-posteriori have been used to solve for $g\left[e(i)\right]$,

$$g\left[e(i)\right] = \frac{1}{\mu\left\|\mathbf{u}_i\right\|^2_{\mathbf{P}_i}}\left[e_a(i) - e_p(i)\right]. \tag{2.29}$$

Now the error non-linearity function $g\left[\cdot\right]$ is substituted in the weight-error vector equation to obtain:

$$\mathbf{v}_i = \mathbf{v}_{i-1} - \frac{\mathbf{P}_i\mathbf{u}_i}{\left\|\mathbf{u}_i\right\|^2_{\mathbf{P}_i}}\left[e_a(i) - e_p(i)\right]. \tag{2.30}$$

It is clear that the above relation involves the four errors $\mathbf{v}_{i-1}$, $\mathbf{v}_i$, $e_a(i)$, $e_p(i)$; it is also observed that even the step-size parameter is not present. Rearranging the above equation we get:

$$\mathbf{v}_i + \frac{\mathbf{P}_i\mathbf{u}_i}{\left\|\mathbf{u}_i\right\|^2_{\mathbf{P}_i}}e_a(i) = \mathbf{v}_{i-1} + \frac{\mathbf{P}_i\mathbf{u}_i}{\left\|\mathbf{u}_i\right\|^2_{\mathbf{P}_i}}e_p(i). \tag{2.31}$$

By evaluating the energies (i.e., the squared-weighted Euclidean norms) of both sides and after some straight forward calculation, it was found that the following energy equality holds:

$$\left\|\mathbf{v}_i\right\|^2_{\mathbf{P}_i^{-1}} + \frac{1}{\left\|\mathbf{u}_i\right\|^2_{\mathbf{P}_i}}\left|e_a(i)\right|^2 = \left\|\mathbf{v}_{i-1}\right\|^2_{\mathbf{P}_i^{-1}} + \frac{1}{\left\|\mathbf{u}_i\right\|^2_{\mathbf{P}_i}}\left|e_p(i)\right|^2. \tag{2.32}$$

This result has proven very useful in the study of the performance of adaptive filters [2].

2.4 Summary

In this chapter we described the basic development of the QNLMF algorithm by explaining all the steps clearly. We also introduced the method, the fundamental energy conservation method, using which we will analyze the proposed algorithm.

Chapter 3

Steady-State Analysis of the proposed QNLMF Adaptive Algorithm

3.1 Introduction

In this Chapter, the steady-state analysis of the proposed algorithm is carried out. The following assumptions [2] were used during the analysis of the proposed algorithm:

A1 The input process $\{\mathbf{u}_i\}$ is a sequence of independent and identically distributed (i.i.d) Gaussian random vectors with zero-mean and auto-correlation matrix $\mathbf{R}_{\mathbf{u}}$. Moreover, $\mathbf{u}_i$ and $\mathbf{u}_j$ are uncorrelated for $i \neq j$.

A2 The noise process $\{n(i)\}$ is a zero-mean independent and identically distributed (i.i.d) random process with variance σ_n^2, and is independent of the input process.

A1 is not true in practice but it is very common in literature and a large number of work has shown that the analytical results obtained under this assumption agree closely with the simulation results under general conditions. A2 is very common in literature and is termed as the independence assumption [1] which can also be justified in several practical instances. The gaussian assumption is used to simplify the analysis and simulation results show that analytical results derived based on this assumption are well matched.

3.2 Mean Square Analysis of QNLMF

The proposed algorithm (QNLMF) update recursion obtained in Chapter 2 can be written as:

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \mathbf{P}_i e^3(i)\mathbf{u}_i, \quad i \geq 0, \quad \mathbf{w}_{-1} = \text{initial guess}, \tag{3.1}$$

$$\mathbf{P}_i = \frac{1}{(1-\alpha)}\left[\mathbf{P}_{i-1} - \frac{\mathbf{P}_{i-1}\mathbf{u}_i\mathbf{u}_i^T\mathbf{P}_{i-1}}{\frac{(1-\alpha)}{\alpha e^2(i)} + \mathbf{u}_i^T\mathbf{P}_{i-1}\mathbf{u}_i}\right], \quad 0 < \alpha \leq 0.1, \quad \mathbf{P}_{-1} = \epsilon^{-1}\mathbf{I}. \tag{3.2}$$

The weight error vector is defined as $\mathbf{v}_i = \mathbf{w}^o - \mathbf{w}_i$. Therefore, equation (3.1) can be expressed in terms of weight-error vector as:

$$\mathbf{v}_i = \mathbf{v}_{i-1} - \mu \mathbf{P}_i e^3(i)\mathbf{u}_i \tag{3.3}$$

Now, we can easily derive equations for the evolution of the weight-error mean vector, $\bar{\mathbf{v}}_i = E[\mathbf{v}_i]$ and the step-size μ. Considering the above mentioned assumptions, we can write:

$$\bar{\mathbf{v}}_i = \left\{\mathbf{I} - 3\mu\left(\sigma_e^2\mathbf{R_u} + 2\mathbf{R_u}\bar{\mathbf{v}}_{i-1}\bar{\mathbf{v}}_{i-1}^T\mathbf{R_u}\right)^{-1}\sigma_e^2\mathbf{R_u}\right\}\bar{\mathbf{v}}_{i-1}. \tag{3.4}$$

When we multiply (3.3) by $\mathbf{u}_i^T$ from the left, we get a new equation employing a-priori and a-posteriori estimation error as:

$$e_p(i) = e_a(i) - \mu\left\|\mathbf{u}_i\right\|_{\mathbf{P}_i}^2 e^3(i), \tag{3.5}$$

where, $\left\|\mathbf{u}_i\right\|_{\mathbf{P}_i}^2 = \mathbf{u}_i^T\mathbf{P}_i\mathbf{u}_i$ is the weighted-squared Euclidean norm. Using this result, (3.3) can be written as:

$$\mathbf{v}_i = \mathbf{v}_{i-1} - \bar{\mu}(i)\mathbf{u}_i\left[e_a(i) - e_p(i)\right]. \tag{3.6}$$

where $\bar{\mu}(i) = \frac{1}{\left\|\mathbf{u}_i\right\|_{\mathbf{P}_i}^2}$.

Evaluating the energies of both sides of the equation results in what is known as the energy relation:

$$\left\|\mathbf{v}_i\right\|_{\mathbf{P}_i^{-1}}^2 + \bar{\mu}(i)\left|e_a(i)\right|^2 = \left\|\mathbf{v}_{i-1}\right\|_{\mathbf{P}_i^{-1}}^2 + \bar{\mu}(i)\left|e_p(i)\right|^2. \tag{3.7}$$

This important fundamental energy relation will now be used to evaluate the steady-state relation for the Excess Mean-Sqaure Error (EMSE) of the proposed QNLMF algorithm at steady state. As we know, an adaptive filter is said to operate in steady state iff:

$$\lim_{i\to\infty} E\left[\mathbf{v}_i\right] = \lim_{i\to\infty} E\left[\mathbf{v}_{i-1}\right], \tag{3.8}$$

and

$$E\left[\|\mathbf{v}_i\|^2_{\mathbf{P}_i^{-1}}\right] = E\left[\|\mathbf{v}_{i-1}\|^2_{\mathbf{P}_i^{-1}}\right] = c < \infty \quad as \quad i \to \infty. \tag{3.9}$$

Now, we take expectation of the fundamental energy relation (3.7) just derived above and also apply this condition $i \to \infty$, to get:

$$E\left[\bar{\mu}(i)\left|e_a(i)\right|^2\right] = E\left[\bar{\mu}(i)\left|e_a(i) - \frac{\mu}{\bar{\mu}(i)}e^3(i)\right|^2\right]. \tag{3.10}$$

Since, $e(i) = e_a(i) + n(i)$, then substituting this in the above equation results in:

$$E\left[\frac{|e_a(i)|^2}{\|\mathbf{u}_i\|^2_{\mathbf{P}_i}}\right] = E\left[\frac{|e_a(i)|^2}{\|\mathbf{u}_i\|^2_{\mathbf{P}_i}}\right] - 2\mu E\left(|e_a(i)|\left[e_a(i)+n(i)\right]^3\right) + \mu^2 E\left(\|\mathbf{u}_i\|^2_{\mathbf{P}_i}\left[e_a(i)+n(i)\right]^6\right), \tag{3.11}$$

or equivalently:

$$\mu E\left(\|\mathbf{u}_i\|^2_{\mathbf{P}_i}\left[e_a(i)+n(i)\right]^6\right) = 2E\left(|e_a(i)|\left[e_a(i)+n(i)\right]^3\right). \tag{3.12}$$

To be able to proceed further we need to expand $e^6(i)$ and $e^3(i)$ (we are omitting the time index i):

$$e^3 = e_a^3 + 3e_a^2 n + 3e_a n^2 + n^3 \tag{3.13}$$

$$e^6 = e_a^6 + 6e_a^5 n + 6e_a n^5 + 15e_a^4 n^2 + 15e_a^2 n^4 + 20e_a^3 n^3 + n^6 \tag{3.14}$$

A3 Assume that the a-priori estimation error $e_a(i)$ and $\mathbf{u}_i$ are independent of the noise process $n(i)$.

Using A3 and ignoring third and higher-order terms in $e_a(i)$, since at steady-state these terms become small enough, we now evaluate each side of (3.12) separately and are given by

$$\begin{aligned} E\left(\|\mathbf{u}_i\|^2_{\mathbf{P}_i}\left[e_a(i)+n(i)\right]^6\right) &= 15\delta_n^4 E\left[\|\mathbf{u}_i\|^2_{\mathbf{P}_i}\right] e_a^2(i) + 6\delta_n^5 E\left[\|\mathbf{u}_i\|^2_{\mathbf{P}_i}\right] e_a(i) \\ &\quad + \delta_n^6 E\left[\|\mathbf{u}_i\|^2_{\mathbf{P}_i}\right], \qquad (3.15) \\ E\left(|e_a(i)|\left[e_a(i)+n(i)\right]^3\right) &= 3\sigma_n^2 E\left[e_a^2(i)\right]. \qquad (3.16) \end{aligned}$$

Let $E\left[n^m(i)\right] = \delta_n^m$ and using the above expanded forms, (3.12) looks like the following:

$$15\mu\delta_n^4 E\left[\|\mathbf{u}_i\|^2_{\mathbf{P}_i} e_a^2(i)\right] + 6\mu\delta_n^5 E\left[\|\mathbf{u}_i\|^2_{\mathbf{P}_i} e_a(i)\right] + \mu\delta_n^6 E\left[\|\mathbf{u}_i\|^2_{\mathbf{P}_i}\right] = 6\sigma_n^2 E\left[e_a^2(i)\right]. \tag{3.17}$$

To be able to proceed further the following assumption is used:

A4 Assume that at steady state $\|\mathbf{u}_i\|^2_{\mathbf{P}_i}$ is independent of $e_a^2(i)$

This assumption will provide us the power to solve (3.17) and get a relation for Excess Mean-Square Error at steady state,

$$EMSE = \lim_{i\to\infty} E\left|e_a^2(i)\right| = \varsigma_* \tag{3.18}$$

After some straight forward calculations we arrive at the following relation

$$\varsigma(i) = \frac{\mu\delta_n^6 E\left[\|\mathbf{u}_i\|^2_{\mathbf{P}_i}\right]}{6\sigma_n^2 - 15\mu\delta_n^4 E\left[\|\mathbf{u}_i\|^2_{\mathbf{P}_i}\right]}, \tag{3.19}$$

and which, for smaller values of μ can be written as

$$\varsigma(i) = \frac{\mu\delta_n^6 E\left[\|\mathbf{u}_i\|^2_{\mathbf{P}_i}\right]}{6\sigma_n^2}. \tag{3.20}$$

3.2.1 Steady-State Approximation

To begin with, note that (2.10) can be setup to look like the following:

$$\begin{aligned}\mathbf{P}_i^{-1} &= (1-\alpha)^{(i+1)}\epsilon\mathbf{I} \quad (3.21)\\ &\quad +\alpha\left[e^2(i)\mathbf{u}_i\mathbf{u}_i^T + (1-\alpha)e^2(i-1)\mathbf{u}_{i-1}\mathbf{u}_{i-1}^T + \cdots + (1-\alpha)^i e^2(0)\mathbf{u}_0\mathbf{u}_0^T\right],\end{aligned}$$

so that, as $i \to \infty$, and since $\alpha < 1$, the steady-state mean value of P_i^{-1} is given by

$$\lim_{i\to\infty} E\left(\mathbf{P}_i^{-1}\right) = \lim_{i\to\infty} \frac{\alpha E\left[e^2(i)\mathbf{u}_i\mathbf{u}_i^T\right]}{1-(1-\alpha)}, \tag{3.22}$$

$$= \lim_{i\to\infty} E\left[e^2(i)\mathbf{u}_i\mathbf{u}_i^T\right]. \tag{3.23}$$

Applying the Gaussian moment factoring theorem, in (3.23), we obtain:

$$\begin{aligned} \lim_{i\to\infty} E\left(\mathbf{P}_i^{-1}\right) &= \lim_{i\to\infty} E\left[e^2(i)\right] E\left[\mathbf{u}_i\mathbf{u}_i^T\right] + 2\lim_{i\to\infty} E\left[e(i)\mathbf{u}_i\right] E\left[e(i)\mathbf{u}_i^T\right], \quad (3.24) \\ &= \sigma_e^2 \mathbf{R}_\mathbf{u} \\ &= \left(\sigma_n^2 + \varsigma_*\right)\mathbf{R}_\mathbf{u} \approx \mathbf{P}^{-1}. \end{aligned}$$

We denote the result by $\mathbf{P}^{-1}$. The mean value of $\mathbf{P}_i$, on the other hand, is considerably harder to evaluate. So we shall satisfy ourselves with the approximation

$$E\left(\mathbf{P}_i\right) \approx \left[E\left(\mathbf{P}_i^{-1}\right)\right]^{-1} = \frac{\mathbf{R}_\mathbf{u}^{-1}}{\left(\sigma_n^2 + \varsigma_*\right)} \approx \mathbf{P}, \quad as \quad i \to \infty. \tag{3.25}$$

This is an approximation, of course, because even though $\mathbf{P}_i$ and $\mathbf{P}_i^{-1}$ are the inverses of one another, it does not hold that their expected values will have the same inverse relation. Still, approximation (3.25) is reasonable for Gaussian regressors.

Now, as $i \rightarrow \infty$,

$$\begin{aligned} \lim_{i\rightarrow\infty} \varsigma(i) &= \varsigma_* \\ &= \frac{\mu\delta_n^6 E\left[\|\mathbf{u}_i\|^2_{\mathbf{P}_i}\right]}{6\sigma_n^2 - 15\mu\delta_n^4 E\left[\|\mathbf{u}_i\|^2_{\mathbf{P}_i}\right]}, \\ &= \frac{\mu\delta_n^6 Tr\left(\mathbf{R_u P}\right)}{6\sigma_n^2 - 15\mu\delta_n^4 Tr\left(\mathbf{R_u P}\right)}, \end{aligned} \tag{3.26}$$

with $E\|\mathbf{u}_i\|^2_{\mathbf{P}} = E\left(\mathbf{u}_i^T \mathbf{P}_i \mathbf{u}_i\right) = Tr\left(\mathbf{R_u P}\right)$ by using the independence assumption.

Now, with the $E\left(\mathbf{P}_i\right)$ solved we can write a new relation for $Tr\left(\mathbf{R_u P}\right)$ at steady state as

$$Tr\left(\mathbf{R_u P}\right) = Tr\left(\frac{\mathbf{R_u R_u^{-1}}}{\sigma_n^2 + \varsigma_*}\right) \tag{3.27}$$

$$= \frac{Tr\left(\mathbf{I}\right)}{\sigma_n^2 + \varsigma_*} \tag{3.28}$$

$$= \frac{M}{\sigma_n^2 + \varsigma_*} \tag{3.29}$$

Finally, after evaluation we come across a $2nd$ equation for the Excess MSE ς_* of the form:

$$A\varsigma_*^2 + B\varsigma_* + C = 0 \tag{3.30}$$

where,

$$A = 6\sigma_n^2 \tag{3.31}$$

$$B = 6\sigma_n^4 - \mu 15\delta_n^4 M \tag{3.32}$$

$$C = -\mu\delta_n^6 M \tag{3.33}$$

Ignoring higher powers of ς_* we can rewrite (3.30) as follows:

$$B\varsigma_* + C \approx 0 \tag{3.34}$$

Hence, an asymptotic approximation for the excess MSE of QNLMF can be written as:

$$\varsigma_{QNLMF} \approx -\frac{C}{B} \tag{3.35}$$

$$\varsigma_{QNLMF} = \varsigma_* \approx \frac{\mu\delta_n^6 M}{6\sigma_n^4 - \mu 15\delta_n^4 M} \tag{3.36}$$

and for smaller values of μ it can be further approximated to

$$\varsigma_{QNLMF} = \varsigma_* \approx \frac{\mu\delta_n^6 M}{6\sigma_n^4} \tag{3.37}$$

The value of σ_n^2 varies according to the chosen noise. Therefore, we can conclude that for Gaussian noise the EMSE of the QNLMF is given by:

$$\varsigma_{QNLMF} = \varsigma_* \approx \frac{15\mu\sigma_n^6 M}{6\sigma_n^4} \tag{3.38}$$

From the above expression, one can see that the EMSE does not dependent on the input statistics, e.g. the input autocorrelation matrix; however, they depend upon the length of the FIR filter M.

3.3 Mean Convergence of the Step-Size

Proposition 1 :

From (2.8), under the assumption that convergence has taken place, μ can be approximated as

$$0 < \mu < \frac{2}{3} \tag{3.39}$$

Proof.

Subtracting (2.8) from $\mathbf{w}^o$ and taking expected value on both sides, we come across

$$E(\mathbf{v}_i) = E(\mathbf{v}_{i-1}) + \mu \left\{E\left[e^2(i)\mathbf{u}_i\mathbf{u}_i^T\right]\right\}^{-1} E\left[e^3(i)\mathbf{u}_i\right] \tag{3.40}$$

We can show that the expectation term $E\left[e^2(i)\mathbf{u}_i\mathbf{u}_i^T\right]$ can be expanded as

$$E\left[e^2(i)\mathbf{u}_i\mathbf{u}_i^T\right] = E\left[n^2(i)\mathbf{u}_i\mathbf{u}_i^T\right] + 2E\left[n(i)e_a(i)\mathbf{u}_i\mathbf{u}_i^T\right] + E\left[e_a^2(i)\mathbf{u}_i\mathbf{u}_i^T\right], \tag{3.41}$$

which, at steady-state, can be approximated to

$$E\left[e^2(i)\mathbf{u}_i\mathbf{u}_i^T\right] \approx \sigma_e^2(i)\mathbf{R}_\mathbf{u} \tag{3.42}$$

Also, from [13], it can be shown that

$$E\left[e^3(i)\mathbf{u}_i\right] = -3\sigma_e^2(i)\mathbf{R}_\mathbf{u}E\left(\mathbf{v}_{i-1}\right) \tag{3.43}$$

Therefore, (3.40) becomes

$$E\left(\mathbf{v}_i\right) = E\left(\mathbf{v}_{i-1}\right) - 3\mu\left\{\sigma_e^2(i)\mathbf{R}_\mathbf{u}\right\}^{-1}\sigma_e^2(i)\mathbf{R}_\mathbf{u}E\left(\mathbf{v}_{i-1}\right), \tag{3.44}$$

$$E\left(\mathbf{v}_i\right) = \left[1-3\mu\right]E\left(\mathbf{v}_{i-1}\right). \tag{3.45}$$

From the above equation it is easy to show that the mean behavior of the weight error vector $E\left(\mathbf{v}_{i-1}\right)$ converges to the zero vector if the convergence parameter μ is selected to be (3.39). ■

3.4 Adaptation Time Constants and Comparisons

If the unknown channel is time invariant, the usual way to compare the performance of different algorithms is to set the parameters such that all algorithms under test have the same misadjustment and then compare their convergence rates.

Recall that the (steady-state) misadjustment is defined as:

$$\mathcal{M}_{adj} = \frac{\text{excess MSE}}{\text{minimum MSE}} = \frac{\varsigma_*}{\sigma_n^2} \tag{3.46}$$

In our case, suitable choise of the parameter μ is responsible for the variable time constants but keeping μ constant for a certain misadjustment will give us the same time constant everytime, since, EMSE is not dependent on the eigenvalue spread of the input autocorrelation matrix. Hence, from (3.39), the time constant relation for the QNLMF algorithm can be shown to be:

$$\tau_{QNLMF} \approx \frac{1}{6\mu} \tag{3.47}$$

3.5 Computational Cost

Each step of the algorithm requires a handful of straightforward computations. Below are tables that explains and compares the computations of QNLMF with RLS for real data.

3.6 Summary

In this chapter, a thorough steady-state analysis is carried out. The relation for the steady-state EMSE is derived. Expressions for mean convergence of the step-size, adaptation time constant are also derived. Finally, the computational cost of the proposed algorithm is compared with the well known RLS algorithm.

RLS Algorithm			
Term	$\times$	$+$	$/$
$\mathbf{u}_i\mathbf{w}_{i-1}$	M	$M-1$	
$d(i) - \mathbf{u}_i\mathbf{w}_{i-1}$		1	
$\lambda^{-1}\mathbf{u}_i$	M		
$\mathbf{P}_{i-1}\left(\lambda^{-1}\mathbf{u}_i\right)$	M^2	$M(M-1)$	
$\mathbf{u}_i^T\mathbf{P}_{i-1}\left(\lambda^{-1}\mathbf{u}_i\right)$	M	$M-1$	
$1+\mathbf{u}_i^T\mathbf{P}_{i-1}\left(\lambda^{-1}\mathbf{u}_i\right)$		1	
$1/\left[1+\mathbf{u}_i^T\mathbf{P}_{i-1}\left(\lambda^{-1}\mathbf{u}_i\right)\right]$			1
$\left(\lambda^{-1}\mathbf{u}_i^T\mathbf{P}_{i-1}\mathbf{u}_i\right)\cdot\frac{1}{1+\mathbf{u}_i^T\mathbf{P}_{i-1}\left(\lambda^{-1}\mathbf{u}_i\right)}$	1		
$\left(\lambda^{-1}\mathbf{P}_{i-1}\mathbf{u}_i\right)\times\frac{\lambda^{-1}\mathbf{u}_i^T\mathbf{P}_{i-1}\mathbf{u}_i}{1+\mathbf{u}_i^T\mathbf{P}_{i-1}\left(\lambda^{-1}\mathbf{u}_i\right)}$	M		
$\mathbf{P}_i\mathbf{u}_i$		M	
$\mathbf{P}_i\mathbf{u}_i\left[d(i)-\mathbf{u}_i\mathbf{w}_{i-1}\right]$	M		
$\mathbf{w}_i$		M	
TOTAL per iteration	M^2+5M+1	M^2+3M	1

Table 3.1: Computational cost of RLS

QNLMF Algorithm			
Term	$\times$	$+$	$/$
$\mathbf{u}_i\mathbf{w}_{i-1}$	M	$M-1$	
$d(i)-\mathbf{u}_i\mathbf{w}_{i-1}$		1	
$\mathbf{P}_{i-1}\mathbf{u}_i$	M^2	$M(M-1)$	
$\mathbf{u}_i^T\mathbf{P}_{i-1}\mathbf{u}_i$	M	$M-1$	
$\alpha e^2(i)$	2		
$\frac{1-\alpha}{\alpha e^2(i)}$		1	1
$\frac{1-\alpha}{\alpha e^2(i)}+\mathbf{u}_i^T\mathbf{P}_{i-1}\mathbf{u}_i$		1	
$1/\left[\frac{1-\alpha}{\alpha e^2(i)}+\mathbf{u}_i^T\mathbf{P}_{i-1}\mathbf{u}_i\right]$			1
$\left(\mathbf{u}_i^T\mathbf{P}_{i-1}\mathbf{u}_i\right)\cdot\frac{1}{\frac{1-\alpha}{\alpha e^2(i)}+\mathbf{u}_i^T\mathbf{P}_{i-1}\mathbf{u}_i}$	1		
$\left(\mathbf{P}_{i-1}\mathbf{u}_i\right)\times\frac{\mathbf{u}_i^T\mathbf{P}_{i-1}\mathbf{u}_i}{\frac{1-\alpha}{\alpha e^2(i)}+\mathbf{u}_i^T\mathbf{P}_{i-1}\mathbf{u}_i}$	M		
$\frac{1}{1-\alpha}\cdot\left(\mathbf{P}_{i-1}\mathbf{u}_i\right)\times\frac{\mathbf{u}_i^T\mathbf{P}_{i-1}\mathbf{u}_i}{\frac{1-\alpha}{\alpha e^2(i)}+\mathbf{u}_i^T\mathbf{P}_{i-1}\mathbf{u}_i}$			1
$\mu\mathbf{P}_i\mathbf{u}_i$	M	M	
$\mathbf{P}_i\mathbf{u}_i\left[d(i)-\mathbf{u}_i\mathbf{w}_{i-1}\right]^3$	$M+1$		
$\mathbf{w}_i$		M	
TOTAL per iteration	M^2+5M+4	M^2+3M+1	3

Table 3.2: Computational cost of QNLMF

Chapter 4

Tracking Analysis of the Proposed QNLMF Adaptive Algorithm

4.1 Introduction

The main aim of tracking analysis of an adaptive filter is to quantify its ability to track time variations in the channel. In this Chapter, the tracking analysis of the proposed QNLMF algorithms are carried out. We have considered Random Walk model and Rayleigh fading model (both single and multiple path) to model the time varying channels and the analysis is carried out using the energy relation in the same way as described in [2].

4.2 Random Walk Model

The model that is widely used in the adaptive filtering literature is a first order Random-Walk model. The model assumes that $\mathbf{w}_i^o$ undergoes random variations of the form

$$\mathbf{w}_i^o = \mathbf{w}_{i-1}^o + \mathbf{q}_i \tag{4.1}$$

with $\mathbf{q}_i$ denoting some random perturbation that is independent of $\{\mathbf{u}_j, n(j)\}$ for all i, j. Here, $\mathbf{w}_i^o$ is a random variable now due to the presence of the random quantity $\mathbf{q}_i$. The sequence $\{\mathbf{q}_i\}$ is assumed to be i.i.d., zero-mean, with covariance matrix

$$E\left(\mathbf{q}_i \mathbf{q}_i^T\right) = \mathbf{Q} \tag{4.2}$$

It is easy to see from (4.1) that

$$E\left(\mathbf{w}_i^o\right) = E\left(\mathbf{w}_{i-1}^o\right) \tag{4.3}$$

so that the $\{\mathbf{w}_i^o\}$ have a constant mean, which we shall denote by $\bar{\mathbf{w}}^o$,

$$E\left(\mathbf{w}_i^o\right) \triangleq \bar{\mathbf{w}}^o \tag{4.4}$$

The initial condition for model (4.1) is modeled as a random variable $\mathbf{w}_{-1}^o$, with mean $\bar{\mathbf{w}}^o$ and independent of all other variables, $\{\mathbf{q}_i, n(i), \mathbf{u}_i\}$ for all i.

It has been observed in [2], that the covariance matrix of $\mathbf{w}_i^o$:

$$\mathbf{w}_i^o - \bar{\mathbf{w}}^o = \mathbf{w}_{i-1}^o - \bar{\mathbf{w}}^o + \mathbf{q}_i \tag{4.5}$$

given by,

$$E\left[\left(\mathbf{w}_i^o - \bar{\mathbf{w}}^o\right)\left(\mathbf{w}_i^o - \bar{\mathbf{w}}^o\right)^T\right] = E\left[\left(\mathbf{w}_{i-1}^o - \bar{\mathbf{w}}^o\right)\left(\mathbf{w}_{i-1}^o - \bar{\mathbf{w}}^o\right)^T\right] + \mathbf{Q} \tag{4.6}$$

grows unbounded, which means that at each time instant i, a nonnegative-definite matrix $\mathbf{Q}$ is added to the covariance matrix of $\mathbf{w}_{i-1}^o$ in order to obtain a covariance matrix of $\mathbf{w}_i^o$. So, as time progresses the covariance matrix of $\mathbf{w}_i^o$ becomes unbounded. A more practical model can be developed by replacing (4.5) by:

$$\mathbf{w}_i^o - \bar{\mathbf{w}}^o = \gamma\left(\mathbf{w}_{i-1}^o - \bar{\mathbf{w}}^o\right) + \mathbf{q}_i \tag{4.7}$$

for some scalar $|\gamma| < 1$. In this case the covariance matrix of $\mathbf{w}_i^o$ would tend to a finite steady-state value given by:

$$\lim_{i\to\infty} E\left[\left(\mathbf{w}_i^o - \bar{\mathbf{w}}^o\right)\left(\mathbf{w}_i^o - \bar{\mathbf{w}}^o\right)^T\right] = \frac{\mathbf{Q}}{1-|\gamma|^2} \tag{4.8}$$

But the tracking analysis for this kind of model is very demanding. As mentioned in [2], it was found that in the literature it is customary to assume the value of γ

is sufficiently close to one and use the model described by (4.5) which simplifies our analysis tremendously. For this reason, we have used the random walk model described by (4.1) for the tracking analysis of the proposed algorithms.

The algorithms' update recursion obtained in Chapter 2 can be written as:

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \mathbf{P}_i e^3(i)\mathbf{u}_i, \quad i \geq 0, \quad \mathbf{w}_{-1} = \text{initial guess} \tag{4.9}$$

$$\mathbf{P}_i = \frac{1}{(1-\alpha)}\left[\mathbf{P}_{i-1} - \frac{\mathbf{P}_{i-1}\mathbf{u}_i\mathbf{u}_i^T\mathbf{P}_{i-1}}{\frac{(1-\alpha)}{\alpha e^2(i)} + \mathbf{u}_i^T\mathbf{P}_{i-1}\mathbf{u}_i}\right], \quad 0 < \alpha \leq 0.1,\ \mathbf{P}_{-1} = \epsilon^{-1}\mathbf{I} \tag{4.10}$$

The update recursion can be written in terms of weight error vector $\mathbf{v}_i = \mathbf{w}_i^o - \mathbf{w}_i$, a-priori estimation error $e_a(i) = \mathbf{u}_i^T\left(\mathbf{w}_i^o - \mathbf{w}_{i-1}\right)$, and a-posteriori estimation error $e_p(i) = \mathbf{u}_i^T\left(\mathbf{w}_i^o - \mathbf{w}_i\right)$ as:

$$e_p(i) = e_a(i) - \mu \left\|\mathbf{u}_i\right\|^2_{\mathbf{P}_i} e^3(i) \tag{4.11}$$

Equation (4.11) has the same form as (3.5) in the previous Chapter. Therefore, following the same exact arguments that we presented in that section, we arrive at the following energy relation:

$$\left\|\mathbf{w}_i^o - \mathbf{w}_i\right\|^2_{\mathbf{P}_i^{-1}} + \bar{\mu}(i)\left|e_a(i)\right|^2 = \left\|\mathbf{w}_i^o - \mathbf{w}_{i-1}\right\|^2_{\mathbf{P}_i^{-1}} + \bar{\mu}(i)\left|e_p(i)\right|^2 \tag{4.12}$$

where $\bar{\mu}(i) = \frac{1}{\|\mathbf{u}_i\|^2_{\mathbf{P}_i}}$. Now, the first term on the left hand side of the above equation is the weight-error vector $\mathbf{v}_i$ but same cannot be said for the first term on the right hand side. Now we take expectation on both sides of (4.12) and use all the assumptions of Chapter 3 and the following assumption,

A5 $\{\mathbf{q}_i\}$ is a zero-mean stationary random process with a positive definite covariance matrix $\mathbf{Q}$ and is statistically independent of both the input regressor vector $\{\mathbf{u}_i\}$ and the noise sequence $\{n(i)\}$,

to conclude that:

$$E\left[\|\mathbf{v}_i\|^2_{\mathbf{P}_i^{-1}}\right] + E\bar{\mu}(i)\left|e_a(i)\right|^2 = E\left[\|\mathbf{v}_{i-1}\|^2_{\mathbf{P}_i^{-1}}\right] + Tr\left(\mathbf{Q}\mathbf{P}_i^{-1}\right) + E\bar{\mu}(i)\left|e_p(i)\right|^2 \quad (4.13)$$

Comparing the above equation with its counterpart (3.7), we notice that the only difference is the extra term $Tr\left(\mathbf{Q}\mathbf{P}_i^{-1}\right)$ on the right hand side of the above. Hence, following the same arguments that were presented in Chapter 3, we notice that :

$$6\sigma_n^2\varsigma_* = \mu^{-1}Tr\left(\mathbf{Q}\mathbf{P}^{-1}\right) + \mu Tr\left(\mathbf{R_u}\mathbf{P}\right)\left[15\delta_n^4\varsigma_* + \mu\delta_n^6\right] \quad (4.14)$$

where,

$$Tr\left(\mathbf{Q}\mathbf{P}^{-1}\right) = \left(\sigma_n^2 + \varsigma_*\right) Tr\left(\mathbf{Q}\mathbf{R_u}\right) \quad (4.15)$$

Now, all the above terms have been defined in Chapter 3 and holds good here also, upon following the similar procedure, tracking steady-state EMSE for the proposed

QNLMF is given by:

$$\varsigma_{QNLMF} = \varsigma_* \approx \frac{\mu\delta_n^6 M + \mu^{-1}\sigma_n^4 Tr\left(\mathbf{Q}\mathbf{R}_\mathbf{u}\right)}{6\sigma_n^4 - 2\mu^{-1}\sigma_n^2 Tr\left(\mathbf{Q}\mathbf{R}_\mathbf{u}\right)} \tag{4.16}$$

The value of σ_n^2 varies depending upon the different noise enviornments. Under the assumption that the random non-stationarity $\{\mathbf{q}_i\}$ is i.i.d., and further assuming that the noise is Gaussian, the above equation becomes:

$$\varsigma_* \approx \frac{15\mu\sigma_n^6 M + \mu^{-1}\sigma_n^4 Tr\left(\mathbf{Q}\mathbf{R}_\mathbf{u}\right)}{6\sigma_n^4 - 2\mu^{-1}\sigma_n^2 Tr\left(\mathbf{Q}\mathbf{R}_\mathbf{u}\right)} \tag{4.17}$$

Remark 2 *It is clear from the above expression that a time varying channel results in an increse in the EMSE of the proposed QNLMF algorithm as compared to the stationary model. The additional terms reflects the effect of non-stationarity on the filter performance. Observe in particular that* $Tr\left(\mathbf{Q}\mathbf{R}_\mathbf{u}\right)$ *is multipled by* μ^{-1}*, so that the larger the step-size the smaller the effect of non-stationarity on the EMSE. This behavior is intuitive since a larger step-size (usually) signifies faster adaptation [2], in which case our proposed QNLMF willl have a better chance at "learning" and at "following" the data statistics. A small step-size, on the other hand, leads to a smaller EMSE under stationary conditions, but it may also lead to poor tracking performance. So, there must be an optimum choice for the step-size, which is obtained by minimizing* (4.16) *with respect to* μ*. Taking the derivative of* (4.16) *and equating it to zero gives:*

$$\mu_{opt} = \frac{Tr\left(\mathbf{QR_u}\right)}{3\sigma_n^2} \pm \frac{\sqrt{\delta_n^6 MTr\left(\mathbf{QR_u}\right)\left\{36\sigma_n^8 + 4\delta_n^6 MTr\left(\mathbf{QR_u}\right)\right\}}}{6\sigma_n^2\delta_n^6 M} \tag{4.18}$$

Substituting the above optimal value for μ into (4.16) we get the minimum EMSE.

4.3 Rayleigh Fading Channel

As we know in a wireless communication enviornment, signal suffers from multiple reflections while travelling from the transmitter to the receiver so that the receiver ends up getting several (almost simultaneous) replicas of the transmitted signal. The reflections are received with different amplitude and phase distortions , and the overall received signal is the resultant of all these reflections. Based on the relative phases of the reflections, the signal may add up constructively or destructively at the receiver. Furthermore, if the transmitter is moving with respect to the receiver, then these interferences will vary with time. This phenomenon is known as channel fading [7].

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

$$h(i) = \vartheta x(i)\delta(i - i_o) \tag{4.19}$$

where $\{x(i)\}$ is a time-variant complex sequence that models the time-variations in the channel, and i_o is the channel delay. The sequence $\{x(i)\}$ is assumed to have unit variance , and the scalar ϑ is then used to model the actual path loss that is introduced by the channel. That is, ϑ^2 is equal to the power attenuation that a signal will undergo when it travels through the channel.

Several mathematical models can be used to characterize the fading properties of $\{x(i)\}$, and consequently those of the channel. A widely used model is known as Rayleigh fading model. In this case, for each i, the amplitude $|x(i)|$ is assumed to have a rayleigh distribution, i.e.,

$$f_{|x(i)|}\left(|x(i)|\right) = |x(i)|\, e^{-|x(i)|^2/2}, \quad |x(i)| \geq 0 \tag{4.20}$$

while the phase $\angle x(i)$ is assumed to be uniformly distributed within $[-\pi, \pi]$:

$$f\left(\angle x(i)\right) = \frac{1}{2\pi}, \qquad -\pi \leq \angle x(i) \leq \pi \tag{4.21}$$

Zeroth-order Bessel function of the first kind has been widely used in the lierature to model the auto-correlation function of the sequence $\{x(i)\}$. It is based on the fact that all the scatterers are uniformly distributed circularly around the receiver, so that its power spectral density has a U-shaped spectrum. This function is defined as:

$$r(i) \cong E\left[x(j)x(j-i)\right] = \Im_o\left(2\pi f_D T_s n\right), \quad n = \cdots, -1, 0, 1, \cdots \tag{4.22}$$

where T_s is the sampling period, f_D is called the maximum Doppler frequency of the rayleigh fading channel, and the function $\Im_o$ is defined by:

$$\Im_o(y) \cong \frac{1}{\pi} \int_0^\pi \cos(y \sin\theta) d\theta \tag{4.23}$$

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

$$f_D = \frac{\nu f_c}{c}, \tag{4.24}$$

where c is speed of the light, $c = 3 \times 10^8 m/s$. Therefore, the weight vector that we wish to estimate looks like:

$$\begin{bmatrix} 0 & 0 & x_2(i) & 0 & 0 \end{bmatrix} \tag{4.25}$$

When we dig more in to the fading phenomenon it was mentioned in [2] that in some instance the reflections might have originated from a far away object like a mountain or tall buildings. These reflections arrive at the receiver with longer delay than the first group of reflections. In this case, a single-path Rayleigh channel is not sufficient, therefore a multi-path model is preferred which is governed by this finite-impulse response:

$$h(i) = \sum_{j=1}^{M-1} \vartheta_j x_j(i) \delta(i - j + 1) \tag{4.26}$$

where $\{\vartheta_j\}$ and $\{x_j(i)\}$ are, respectively, the path loss and fading sequence of the $i - th$ cluster of reflectors. In our analysis we have considered a wireless channel with two Rayleigh fading rays; furthermore, both rays are assumed to fade at the same Doppler frequency. The channel impulse response sequence consists of an initial delay of two samples, followed by a Rayleigh fading ray, then another zero sample, which is finally followed by another Rayleigh fading ray; so the 5-tap weight-vector that we wish to be estimated looks like:

$$\begin{bmatrix} 0 & 0 & x_2(i) & 0 & x_4(i) \end{bmatrix} \tag{4.27}$$

As mentioned in [2], we came to know that a first-order approximation for the variation of a Rayleigh fading coefficient $\{x_j(i)\}$ is to assume that $\{x_j(i)\}$ varies according too the auto-regressive model:

$$x(i) = r(1)x(i-1) + \sqrt{1 - |r(1)|^2}\eta(i), \tag{4.28}$$

where $r(1) = \Im_o(2\pi f_D T_s)$ and $\eta(i)$ denotes a white noise process with unit-variance. Now, since the multi-path rays of the channel (4.27) are assumed to fade at the same rate, the above approximation indicates that the variations in the channel weight vector could be approximated as:

$$\mathbf{w}_i^o = \tau \mathbf{w}_{i-1}^o + \mathbf{q}_i \tag{4.29}$$

where the covariance matrix of $\{\mathbf{q}_i\}$ is $\mathbf{Q} = (1 - \tau^2)\,\mathbf{I}$ with $\tau = r(1)$. It is clear, that the value of τ depends upon the doppler frequency and if the value of τ is chose to be approximated equal to one then the results of analysis that we have done in the previous section for Random Walk model is applicable in this case also.

4.4 Summary

In this chapter we studied and analyzed the effect of time variation of the enviornment over the tracking ability of the adaptive filter. Variations in the environment were introduced using the Random Walk model and the Rayleigh fading model. Consequently, the tracking ability of the filter is hampered more as the variation in the environment increases.

Chapter 5

Transient Analysis of the Proposed QNLMF Algorithm

5.1 Introduction

Adaptive filters are time-variant and non-linear stochastic systems with inherent learning and tracking abilities.The transient analysis of adaptive algorithms is very important because the success of their learning mechanism depends on how fast and how stable they adapt to changes in the signal statistics. Transient performance is concerned with the stability and convergence rate of an adaptive scheme.

In the previous chapters we focused on the steady-state performance of adaptive filters. In this chapter we turn our attention to the transient performance of adaptive filters. In this section we have used the methodology of [30] (where a unified approach to the transient analysis of adaptive filters with error non-linearities is discussed) to carry out transient analysis of QNLMF. This approach is also described in [26]-[30] does not restrict the regression data to be Gaussian and avoids the need for explicit recursions for the covariance matrix of the weight-error vector. Before we start we will specify a data model which will be for the stationary environment.

A1 There exists a vector $\mathbf{w}^o$ such that $d(i) = \mathbf{u}_i\mathbf{w}^o + n(i)$.

A2 The noise sequence $\{n(i)\}$ is i.i.d. with variance $\sigma_n^2 = E\left|n(i)\right|^2$

A3 The sequence $n(i)$ is independent of $\mathbf{u}_j$ for all i,j.

A4 The initial condition $\mathbf{w}_{-1}$ is independent of all $\{d(j), \mathbf{u}_j, n(j)\}$.

A5 The regressor covariance matrix is $\mathbf{R}_\mathbf{u} = E\mathbf{u}_i\mathbf{u}_i^T > 0$

A6 The random variables $\{d(i), \mathbf{u}_i, n(i)\}$ have zero means.

5.2 Transient Analysis of the QNLMF

The upadte recursion obtained in Chapter 2 can be written as:

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu\mathbf{P}_i e^3(i)\mathbf{u}_i, \quad i \geq 0, \quad \mathbf{w}_{-1} = \text{initial guess} \tag{5.1}$$

$$\mathbf{P}_i = \frac{1}{(1-\alpha)}\left[\mathbf{P}_{i-1} - \frac{\mathbf{P}_{i-1}\mathbf{u}_i\mathbf{u}_i^T\mathbf{P}_{i-1}}{\frac{(1-\alpha)}{\alpha e^2(i)} + \mathbf{u}_i^T\mathbf{P}_{i-1}\mathbf{u}_i}\right], \quad 0 < \alpha \leq 0.1, \quad \mathbf{P}_{-1} = \epsilon^{-1}\mathbf{I} \tag{5.2}$$

The weight error vector is defined as $\mathbf{v}_i = \mathbf{w}^o - \mathbf{w}_i$. Therefore, the above equation can be expressed in terms of weight-error vector as:

$$\mathbf{v}_i = \mathbf{v}_{i-1} - \mu \mathbf{P}_i e^3(i) \mathbf{u}_i \tag{5.3}$$

Now, we will define two kinds of weighted errors known as weighted a priori and a posteriori error signals.

$$e_a^{\mathbf{P}_i \Sigma}(i) = \mathbf{u}_i^T \mathbf{P}_i \Sigma \mathbf{v}_{i-1}, \qquad e_p^{\mathbf{P}_i \Sigma}(i) = \mathbf{u}_i^T \mathbf{P}_i \Sigma \mathbf{v}_i \tag{5.4}$$

If we multiply both sides of (5.3) by $\mathbf{u}_i^T \mathbf{P}_i \boldsymbol{\Sigma}$ from the left we find that the a priori and a posteriori estimation errors $\left\{e_a^{\mathbf{P}_i \Sigma}(i), e_p^{\mathbf{P}_i \Sigma}(i)\right\}$ are related via

$$e_p^{\mathbf{P}_i \Sigma}(i) = e_a^{\mathbf{P}_i \Sigma}(i) - \mu \left\|\mathbf{u}_i\right\|^2_{\mathbf{P}_i \Sigma \mathbf{P}_i} e^3(i) \tag{5.5}$$

The above equation provides an alternative description of an adaptive filter in terms of error quantities, $e_a^{\mathbf{P}_i \Sigma}(i), e_p^{\mathbf{P}_i \Sigma}(i), \mathbf{v}_{i-1}, \mathbf{v}_i$ and $e^3(i)$. Now, by substituting above equation in (5.3) we get:

$$\mathbf{v}_i + \frac{\mathbf{P}_i \mathbf{u}_i e_a^{\mathbf{P}_i \Sigma}(i)}{\|\mathbf{u}_i\|^2_{\mathbf{P}_i \Sigma \mathbf{P}_i}} = \mathbf{v}_{i-1} + \frac{\mathbf{P}_i \mathbf{u}_i e_p^{\mathbf{P}_i \Sigma}(i)}{\|\mathbf{u}_i\|^2_{\mathbf{P}_i \Sigma \mathbf{P}_i}} \tag{5.6}$$

On each side of this identity we have a combination of a priori and a posteriori errors. By equating the above weighted Euclidean norms of both sides of the equation, i.e., by setting

$$\left\| \mathbf{v}_i + \frac{\mathbf{P}_i \mathbf{u}_i e_a^{\mathbf{P}_i \Sigma}(i)}{\|\mathbf{u}_i\|^2_{\mathbf{P}_i \Sigma \mathbf{P}_i}} \right\|^2_{\Sigma} = \left\| \mathbf{v}_{i-1} + \frac{\mathbf{P}_i \mathbf{u}_i e_p^{\mathbf{P}_i \Sigma}(i)}{\|\mathbf{u}_i\|^2_{\mathbf{P}_i \Sigma \mathbf{P}_i}} \right\|^2_{\Sigma} \tag{5.7}$$

we find, after a straight forward calculation, that the following energy equality holds:

$$\|\mathbf{u}_i\|^2_{\mathbf{P}_i \Sigma \mathbf{P}_i} \cdot \|\mathbf{v}_i\|^2_{\Sigma} + e_a^{\mathbf{P}_i \Sigma}(i)^2 = \|\mathbf{u}_i\|^2_{\mathbf{P}_i \Sigma \mathbf{P}_i} \cdot \|\mathbf{v}_{i-1}\|^2_{\Sigma} + e_p^{\mathbf{P}_i \Sigma}(i)^2 \tag{5.8}$$

Observe that the equality simply amounts to adding the weighted energies of the individual terms of (5.6); the cross-terms cancel out. It has been shown that different choices of $\mathbf{\Sigma}$ allow us to evaluate different performance measures of an adaptive filter [2]. In our analysis two measures of performance indices are used: The steady-state EMSE and steady-state MSD. The steady state MSE is defined as follows:

$$MSE = \lim_{i \to \infty} E\left[e^2(i)\right], \tag{5.9}$$

$$\lim_{i \to \infty} E\left[\left(\mathbf{u}_i^T \mathbf{v}_{i-1} + n(i)\right)^2\right] \tag{5.10}$$

Using assumptions A2-A3, the MSE reduces to :

$$
\begin{aligned}
MSE &= \sigma_n^2 + \lim_{i\to\infty} E\left[\left(\mathbf{u}_i^T \mathbf{v}_{i-1}\right)^2\right], && (5.11)\\
&= \sigma_n^2 + \lim_{i\to\infty} Tr\left\{\mathbf{R_u}\mathbf{K}(i)\right\}, && (5.12)\\
&= \sigma_n^2 + \lim_{i\to\infty} E\left[e_a^2(i)\right], && (5.13)\\
&= \sigma_n^2 + \varsigma_* && (5.14)
\end{aligned}
$$

while, the steady-state MSD is the steady-state value of the weight-error variance, i.e.,

$$
MSD = \lim_{i\to\infty} E\left\|\mathbf{v}_i\right\|^2 \tag{5.15}
$$

Thus returning to (5.8), and replacing $e_p^{\mathbf{P}_i\Sigma}(i)$ by its equivalent expression (5.5) in terms of $e_a^{\mathbf{P}_i\Sigma}(i)$ and $e(i)$ we get

$$
\left\|\mathbf{u}_i\right\|^2_{\mathbf{P}_i\Sigma\mathbf{P}_i} \cdot \left\|\mathbf{v}_i\right\|^2_{\Sigma} + e_a^{\mathbf{P}_i\Sigma}(i)^2 = \left\|\mathbf{u}_i\right\|^2_{\mathbf{P}_i\Sigma\mathbf{P}_i} \cdot \left\|\mathbf{v}_{i-1}\right\|^2_{\Sigma} + \left[e_a^{\mathbf{P}_i\Sigma}(i) - \mu\left\|\mathbf{u}_i\right\|^2_{\mathbf{P}_i\Sigma\mathbf{P}_i} e^3(i)\right]^2 \tag{5.16}
$$

Furthermore, upon expanding the rightmost term in the above equation and using the fact that the event $\left\|\mathbf{u}_i\right\|^2_{\mathbf{P}_i\Sigma\mathbf{P}_i} = 0$ has probability zero, we can eliminate it and then after taking the expectation on both sides, leads to:

$$
E\left\|\mathbf{v}_i\right\|^2_{\Sigma} = E\left\|\mathbf{v}_{i-1}\right\|^2_{\Sigma} + \mu^2 E\left(\left\|\mathbf{u}_i\right\|^2_{\mathbf{P}_i\Sigma\mathbf{P}_i} e^6(i)\right) - 2\mu E\left(e_a^{\mathbf{P}_i\Sigma}(i) e^3(i)\right) \tag{5.17}
$$

We shall now proceed by evaluating the expectations

$$E\left(\|\mathbf{u}_i\|^2_{\mathbf{P}_i\Sigma\mathbf{P}_i}\, e^6(i)\right) \quad \text{and} \quad E\left(e_a^{\mathbf{P}_i\Sigma}(i)e^3(i)\right) \tag{5.18}$$

in terms of a weighted norm of $\mathbf{v}_{i-1}$. These expectations are hard to compute. In-order to facilitate their evaluation, we shall rely on the following assumption on the distribution of the a priori estimation errors [2].

A7 The a priori estimation errors $\left\{e_a(i), e_a^{\mathbf{P}_i\Sigma}(i)\right\}$ are jointly Gaussian.

A8 We also assume that a priori estimation error $e_a(i)$ and the noise process $\{n(i)\}$ are independent.

We will also rely on one very important assumption [35] in order to simplify our expectations and that is

$$\hat{\mathbf{R}}_i \approx \alpha \sum_{j=0}^{i-1} (1-\alpha)^{i-j}\, e^2(j)\mathbf{u}_j\mathbf{u}_j^T, \tag{5.19}$$

ignoring the term

$$\alpha e^2(i)\mathbf{u}_i\mathbf{u}_i^T, \tag{5.20}$$

by noticing that

$$\alpha e^2(i) \ll \alpha \sum_{j=0}^{i-1} (1-\alpha)^{i-j}\, e^2(j). \tag{5.21}$$

This approximation may be not good for initial iterations but will be applicable as i increases, so, $\hat{\mathbf{P}}_i = \epsilon(i)\mathbf{I} + \hat{\mathbf{R}}_i$.

These assumptions are resonable for long adaptive filters, for instance, that since $e_a(i) = \mathbf{u}_i\mathbf{v}_{i-1}$, it can be regarded as the sum of $(M-1)$ random variables. As its length increases its distribution can be approximated by a Gaussian distribution in view of central limit theorem [31]-[32]. A similar remark hold for $\left\{e_a^{\hat{\mathbf{P}}_i\Sigma}(i)\right\}$.

Hence, we can simplify the expectation $E\left[e_a^{\hat{\mathbf{P}}_i\Sigma}(i)e^3(i)\right]$ using the Price's theorem [33], A2 and the fact that $e(i) = e_a(i) + n(i)$, we get

$$E\mathbf{x}g(\mathbf{y}+z) = \frac{E\mathbf{x}\mathbf{y}}{E\mathbf{y}^2}E\mathbf{y}^T g(\mathbf{y}+z) \tag{5.22}$$

$$E\left[e_a^{\hat{\mathbf{P}}_i\Sigma}(i)e^3(i)\right] = E\left[e_a^{\hat{\mathbf{P}}_i\Sigma}(i)e_a(i)\right]\cdot\left(\frac{E\left[e_a(i)e^3(i)\right]}{E\left[e_a^2(i)\right]}\right) \tag{5.23}$$

The point now is that in view of the Gaussian A7, the expectation $E\left[e_a^{\hat{\mathbf{P}}_i\Sigma}(i)e^3(i)\right]$ depends on $e_a(i)$ only through its second moment, $E\left[e_a^2(i)\right]$. It is also well known from [32] that the expectation of a function of a Gaussian random variable will only depend on the variance of this variable and not on higher-order moments of it. Consequently, along with equality (5.23), we can introduce the following function of $E\left[e_a^2(i)\right]$,

$$H_G = \frac{E\left[e_a(i)e^3(i)\right]}{E\left[e_a^2(i)\right]} \tag{5.24}$$

Now in the case of our algorithm, QNLMF, after some straight forward calculation the value of h_g for real data was found to be:

$$H_G = 3\left(E\left[e_a^2(i)\right] + \sigma_n^2\right) \tag{5.25}$$

So, therefore, the expectation $E\left[e_a^{\hat{\mathbf{P}}_i\Sigma}(i)e^3(i)\right]$ can then be expressed as:

$$E\left[e_a^{\hat{\mathbf{P}}_i\Sigma}(i)e^3(i)\right] = H_G \cdot E\left[e_a^{\hat{\mathbf{P}}_i\Sigma}(i)e_a(i)\right]. \tag{5.26}$$

The left over expectation term $E\left[\|\mathbf{u}_i\|^2_{\hat{\mathbf{P}}_i\Sigma\hat{\mathbf{P}}_i} e^6(i)\right]$, from (5.17), which is still to be evaluated in order to facilitate the study of transient analysis of our algorithm, QNLMF, will be dealt as follows. In order to do that we shall rely on the below assumption and it is also assumed that the filter is long enough,

A9 The weighted norm of input $\|\mathbf{u}_i\|^2_{\hat{\mathbf{P}}_i\Sigma\hat{\mathbf{P}}_i}$ is independent of $e(i)$.

The A9 allows us to split this above expectation as:

$$E\left[\|\mathbf{u}_i\|^2_{\hat{\mathbf{P}}_i\Sigma\hat{\mathbf{P}}_i} e^6(i)\right] = E\left[\|\mathbf{u}_i\|^2_{\hat{\mathbf{P}}_i\Sigma\hat{\mathbf{P}}_i}\right] \cdot E\left[e^6(i)\right] \tag{5.27}$$

Now, gaining knowledge from the logic that since $e_a(i)$ is Gaussian and independent of the noise, it is worth to argue that $E\left[e^6(i)\right]$ depends on $e_a(i)$ through its second moment only, so, another term that can come up as defined in [2]:

$$H_U = E\left[e^6(i)\right] \tag{5.28}$$

Now in the case of our algorithm, QNLMF i.e. in the real case, after some straight forward calculation the value of H_U was found to be:

$$H_U = 15\left(E\left[e_a^2(i)\right]\right)^3 + 45\sigma_n^2\left(E\left[e_a^2(i)\right]\right)^2 + 15\delta_n^4 E\left[e_a^2(i)\right] + \delta_n^6 \tag{5.29}$$

Hence, the expectation $E\left[\|\mathbf{u}_i\|^2_{\hat{\mathbf{P}}_i\Sigma\hat{\mathbf{P}}_i} e^6(i)\right]$ can then be expressed as:

$$E\left[\|\mathbf{u}_i\|^2_{\hat{\mathbf{P}}_i\Sigma\hat{\mathbf{P}}_i} e^6(i)\right] = H_U \cdot TrE\left(\mathbf{u}_i^T\hat{\mathbf{P}}_i\boldsymbol{\Sigma}\hat{\mathbf{P}}_i\mathbf{u}_i\right) \tag{5.30}$$

By substituting (5.25) and (5.29) in (5.17), we come up to this version of the weighted-variance relation as follows:

$$E\left\|\mathbf{v}_i\right\|^2_\Sigma = E\left\|\mathbf{v}_{i-1}\right\|^2_\Sigma + \mu^2 H_U ETr\left(\mathbf{u}_i^T\hat{\mathbf{P}}_i\boldsymbol{\Sigma}\hat{\mathbf{P}}_i\mathbf{u}_i\right) - 2\mu H_G E\left[e_a^{\hat{\mathbf{P}}_i\Sigma}(i)e_a(i)\right]. \tag{5.31}$$

Due to the dependency among regressors $\{\mathbf{u}_i\}$ evaluation of the expectation term $E\left[e_a(i)e_a^{\hat{\mathbf{P}}_i\Sigma}(i)\right]$ is made difficult, so, in order to make the transient analysis more tractable we shall rely on the assumption below.

A10 The sequence of vectors $\{\mathbf{u}_i\}$ are i.i.d., i.e. independent and identically distributed.

This above assumption helped us to modify the above weighted-variance relation as follows:

$$E\left\|\mathbf{v}_i\right\|_{\Sigma}^2 = E\left\|\mathbf{v}_{i-1}\right\|_{\Sigma}^2 + \mu^2 H_U Tr\left[\mathbf{R_u} E\left(\hat{\mathbf{P}}_i \boldsymbol{\Sigma} \hat{\mathbf{P}}_i\right)\right] - 2\mu H_G E\left\|\mathbf{v}_{i-1}\right\|_{\mathbf{R_u} E\left(\hat{\mathbf{P}}_i\right)\boldsymbol{\Sigma}}^2. \quad (5.32)$$

Thus, we conclude that by evaluating H_G, H_U and by following the resulting variance relation, the transient behavior of our QNLMF algorithm can be studied. Now, depending upon the correlation of the input we can divide the analysis further.

5.2.1 Transient Analysis of the QNLMF for White Input Data

We start with the case that the input data is white for which the individual entries of $\{\mathbf{u}_i\}$ are i.i.d. i.e. $\mathbf{R_u}$ is diagonal. Lets say $\mathbf{R_u} = \sigma_{\mathbf{u}}^2 \mathbf{I}$ and $E\left[e_a^2(i)\right] = \sigma_{\mathbf{u}}^2 E\left\|\mathbf{v}_i\right\|^2$. Hence, when $\boldsymbol{\Sigma} = \mathbf{I}$, the variance relation in (5.32) becomes:

$$E\left\|\mathbf{v}_i\right\|^2 = E\left\|\mathbf{v}_{i-1}\right\|^2 + \mu^2 H_U \sigma_{\mathbf{u}}^2 Tr\left[E\left(\hat{\mathbf{P}}_i^2\right)\right] - 2\mu\sigma_{\mathbf{u}}^2 H_G E\left\|\mathbf{v}_{i-1}\right\|_{E\left(\hat{\mathbf{P}}_i\right)}^2 \quad (5.33)$$

where H_G and H_U are also functions of $E\left[\left\|\mathbf{v}_{i-1}\right\|^2\right]$. Here, in order to evaluate the second moment of $\mathbf{P}_i$, we will use some experimental results. Over a large simulation run we found out that

$$Tr\left[E\left(\hat{\mathbf{P}}_i^2\right)\right] > Tr\left[E\left(\hat{\mathbf{P}}_i\right)^2\right]. \quad (5.34)$$

Using this information and for smaller step-sizes we can write equation (5.33) as

$$E\left\|\mathbf{v}_i\right\|^2 \geq E\left\|\mathbf{v}_{i-1}\right\|^2 + \mu^2 H_U \sigma_{\mathbf{u}}^2 Tr\left(\hat{\mathbf{P}}^2\right) - 2\mu\sigma_{\mathbf{u}}^2 H_G E\left\|\mathbf{v}_{i-1}\right\|_{\hat{\mathbf{P}}}^2, \tag{5.35}$$

where $\hat{\mathbf{P}} = E\left(\hat{\mathbf{P}}_i\right)$. To simplify further, we will resort to the steady state assumption i.e.,

$$E\left(\hat{\mathbf{P}}_i\right) \approx \left[E\left(\hat{\mathbf{P}}_i^{-1}\right)\right]^{-1} = \left\{E\left[e^2(i)\mathbf{u}_i\mathbf{u}_i^T\right]\right\}^{-1}, \tag{5.36}$$

and with the help of App. II, we can finally write the variance relation for white input after some algebra as,

$$E\left\|\mathbf{v}_i\right\|^2 \geq \mathbf{f}E\left\|\mathbf{v}_{i-1}\right\|^2 + 5\mu^2\left(\sigma_{\mathbf{u}}^2\right)^{-1} H_G M, \tag{5.37}$$

where

$$\mathbf{f} = 1 - 6\mu \tag{5.38}$$

and M is the filter length.

Mean-square Stability

Observe that we can write the relation (5.33), for MSD as

$$E\left\|\mathbf{v}_i\right\|^2 = E\left\|\mathbf{v}_{i-1}\right\|^2 + \mu^2 H_U Tr\left[\mathbf{R}_{\mathbf{u}} E\left(\hat{\mathbf{P}}_i^2\right)\right] - 2\mu H_G E\left\|\mathbf{v}_{i-1}\right\|_{\mathbf{R}_{\mathbf{u}}E\left(\hat{\mathbf{P}}_i\right)}^2 \tag{5.39}$$

from which it follows that $E\left\|\mathbf{v}_i\right\|^2$ converges for step-sizes satisfying

$$\mu^2 H_U Tr\left(\mathbf{R_u} E\left(\hat{\mathbf{P}}_i^2\right)\right) - 2\mu H_G E\left\|\mathbf{v}_{i-1}\right\|^2_{\mathbf{R_u}E\left(\hat{\mathbf{P}}_i\right)} < 0 \tag{5.40}$$

or, equivalently,

$$0 < \mu < \frac{2H_G E\left\|\mathbf{v}_{i-1}\right\|^2_{\mathbf{R_u}E\left(\hat{\mathbf{P}}_i\right)}}{H_U Tr\left(\mathbf{R}^2_{\mathbf{u}} E\left(\hat{\mathbf{P}}_i^2\right)\right)} \tag{5.41}$$

where

$$H_G = 3\left(E\left[e_a^2(i)\right] + \sigma_n^2\right) \tag{5.42}$$

$$H_U = 15\left(E\left[e_a^2(i)\right]\right)^3 + 45\sigma_n^2\left(E\left[e_a^2(i)\right]\right)^2 + 15\delta_n^4 E\left[e_a^2(i)\right] + \delta_n^6 \tag{5.43}$$

The step-size μ can be further bounded [34] by applying time independent lower and upper bounds on $E\left[e_a^2(i)\right]$ as,

$$\zeta \leq E\left[e_a^2\right] \leq \frac{1}{4} Tr(\mathbf{R_u}) E\left\|\mathbf{v}_0\right\|^2 \tag{5.44}$$

where ζ represents the Cramer-Rao bound and $E\left\|\mathbf{v}_0\right\|^2$ represents the mean weight-error vector for $i = 0$.

In sum, in this chapter we have carried out transient analysis of our proposed QNLMF algorithm under white input. The analytical results obtained are further compared with the experimental ones and the results obtained are presented in Chapter 6.

5.3 Summary

In this chapter we studied the time evolution of the Mean-Square Deviation (MSD) and MSE for White input. Derivation has been carried out for the relations governing the time evolution. Also, relation for the mean-square stability has been derived.

Chapter 6

Performance Analysis of the Proposed Algorithm

In this chapter the performance of the proposed algorithm (QNLMF) using computer simulations is presented. QNLMF is compared with the traditional LMF algorithm in an unknown system identification problem. A number of simulations are carried out to corroborate the theoretical findings and, as expected, better results are obtained over the traditional LMF algorithm.

The objective of designing QNLMF algorithm is to expand the area of research involving Newton's method based adaptive algorithms that is still unexplored particularly in the case of LMF. The performance analysis has been divided into three sections. The first section deals with the mean-square analysis of the proposed algorithms in stationary enviornment, the second section deals with the tracking analysis of the proposed algorithm under non-stationary enviornment and the final section presents a clearer picture about the transient analysis of the two proposed algorithms.

6.1 Mean-Square Performance Analysis of the Proposed Algorithm

In order to compare the convergence rates of the proposed algorithms in the presence of different noise enviornments, the usual way is to set the parameters such that algorithms under observation have same misadjustments and then compare their convergence rates. We know that the steady-state misadjustment is defined by the ratio:

$$\tilde{M} = \frac{\text{excess MSE}}{\text{minimum MSE}} = \frac{\varsigma_*}{\sigma_n^2} \tag{6.1}$$

In order to do a fare comparison we first find the step-size of the LMF μ_{LMF} to obtain a specified misadjustment (say, M_{LMF}), then we set the step-size of the QNLMF algorithm such that:

$$M_{LMF} \approx M_{QNLMF} \tag{6.2}$$

The simulations reported here are based on the FIR channel estimation/system identification. Furthermore, we have considered the following channel:

$$\mathbf{w}^o = \begin{bmatrix} 0.227 & 0.460 & 0.688 & 0.460 & 0.227 \end{bmatrix}^T \tag{6.3}$$

The input vector $\{\mathbf{u}_i\}$ is a BPSK $\{\pm 1\}$ signal. Three different noise enviornments have been considered namely Gaussian, Uniform and Laplacian. The variance of the noise is initially chosen as 1, leading to SNR of 0 dB, and the experiments are repeated for SNRs of 10 and 20 dB. The length of the adaptive filter is chosen equal to the length of the unknown system and the results obtained are averaged over 1000 independent runs. This section has been further categorized into two parts.

- Comparison of the LMF algorithm with the proposed QNLMF algorithm.

  1. Comparison of the learning curves of MSE and third tap in the presence of Gaussian noise for SNR 0 dB, 10 dB and 20 dB.
  2. Comparison of the learning curves of MSE and third tap when there is a sudden burst in the AWGN channel for SNR 20 dB.
  3. Comparison of the learning curves of MSE and third tap in the presence of Uniform noise for SNR 0 dB, 10 dB and 20 dB.
  4. Comparison of the learning curves of MSE and third tap in the presence of Laplacian noise for SNR 0 dB, 10 dB and 20 dB.
  5. Comparison of the learning curves of MSE for LMF and QNLMF in the presence of all three noise processes namely Gaussian, Uniform and Laplacian with SNR 20 dB.
  6. Sensitivity analysis of $\mathbf{\Phi}_i$, the approximation of the inverse of the Hessian matrix by variation in α.

6.1.1 Comparison of the LMF Algorithm with the Proposed QNLMF Algorithm

In this section the LMF algorithm and the proposed Quasi-Newton Least Mean Fourth (QNLMF) algorithm are compared in terms of the convergence time. It has been shown that the proposed QNLMF algorithm has achieved the same noise floor in a much lesser number of iterations as compared to the traditional LMF algorithm. It is observed that in Figure 6.2, 6.4 and 6.6 for gaussian noise with SNR 0 dB, 10 dB and 20 dB the proposed QNLMF algorithm achieved the same steady-state in approximately 5500, 6500 and 8500 iterations earlier than the traditional LMF algorithm, respectively. Also, the behavior of learning curves in Figure 6.3, 6.5 and 6.7 for the third-tap of QNLMF is better than that of LMF for SNR 0 dB, 10 dB and 20 dB in Gaussian noise enviornment. The behavior of the proposed QNLMF algorithm was also observed when there is a sudden burst in the enviornment in Figures 6.8, 6.9 we came to a conclusion that the convergence speed and the third tap behavior of the proposed QNLMF algorithm does not degrade.

In the case of uniform noise, it is also observed in Figure 6.10, 6.12 and 6.14 that the proposed QNLMF algorithm for SNR 0 dB, 10 dB and 20 dB achieved the same steady-state in approximately 5500, 7500 and 10500 iterations earlier than the traditional LMF algorithm.

When the noise is Laplacian, it is observed in Figure 6.16, 6.18 and 6.20 that the proposed QNLMF algorithm for SNR 0 dB, 10 dB and 20 dB achieved the same steady-state in approximately 3500, 4500 and 11000 iterations earlier than the traditional LMF algorithm.

Also, the behavior of learning curves for SNR 0 dB, 10 dB and 20 dB in Figure 6.11, 6.13, 6.15, 6.17, 6.19 and 6.21 for the third-tap of QNLMF algorithm is better than that of LMF in Uniform as well as Laplacian noise enviornments. In addition, in Figure 6.22 we have also shown the learning curve of MSE for QNLMF algorithm in the presence of all three enviornments with SNR 10 dB.

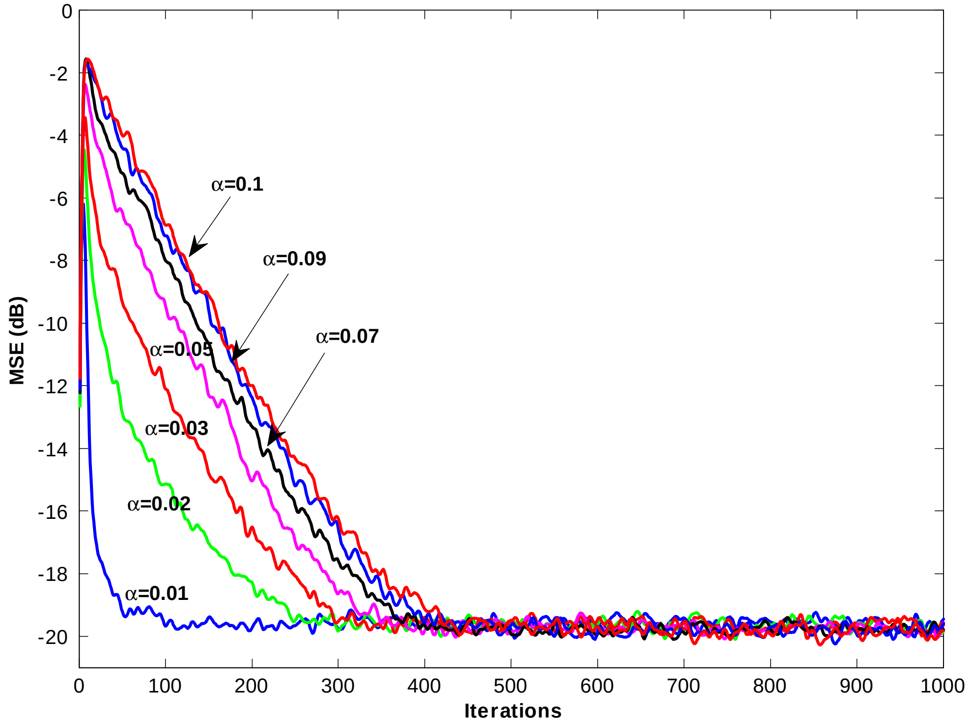

Figure 6.1: Sensitivity analysis of $\mathbf{\Phi}_i$ by varying α of the proposed QNLMF in an AWGN environment.

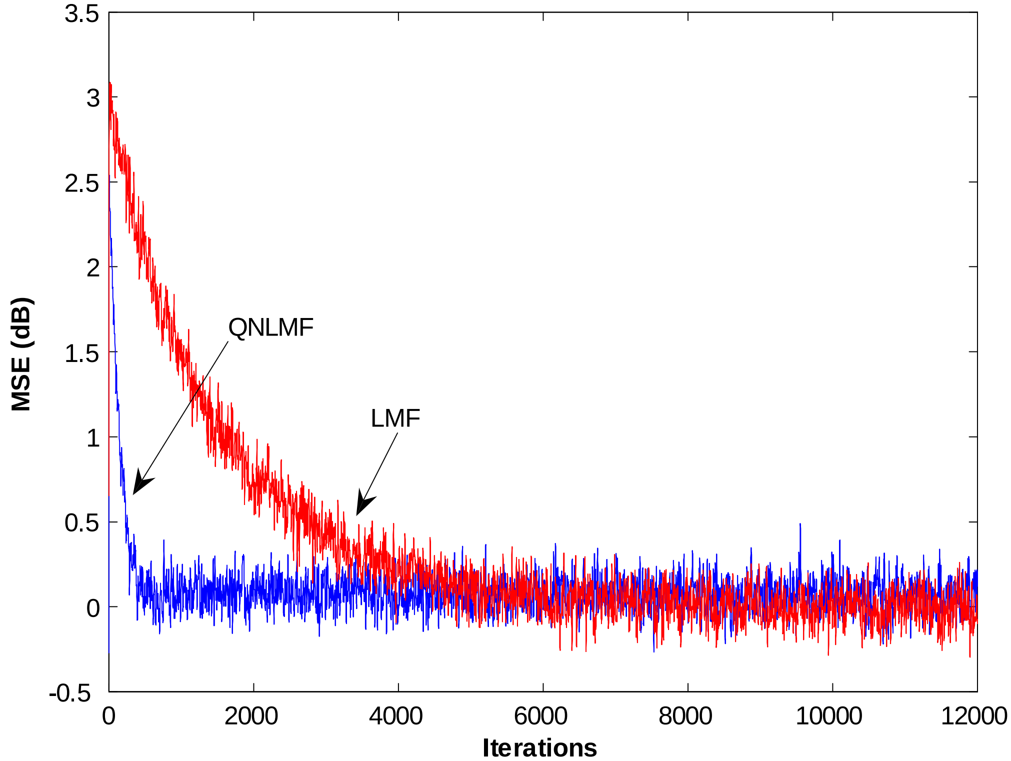

Figure 6.2: Comparison of the convergence speed of the LMF and the proposed QNLMF in AWGN environment with SNR = 0 dB.

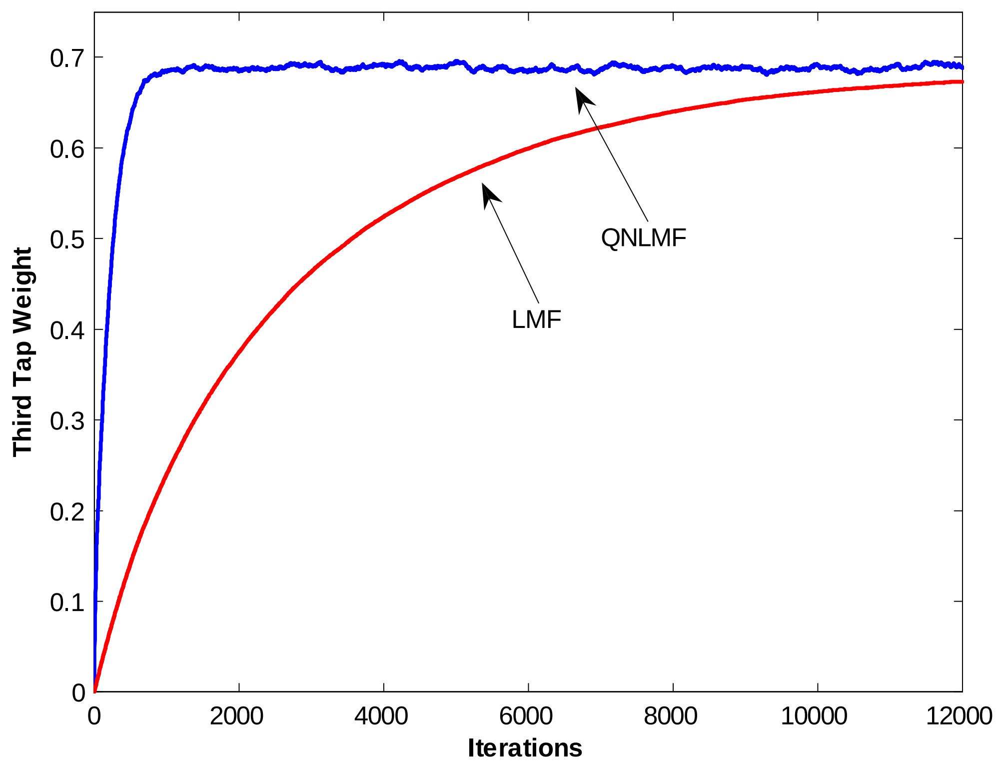

Figure 6.3: Comparison of the learning curves for the third-tap of the LMF and the proposed QNLMF in AWGN environment with SNR = 0 dB.

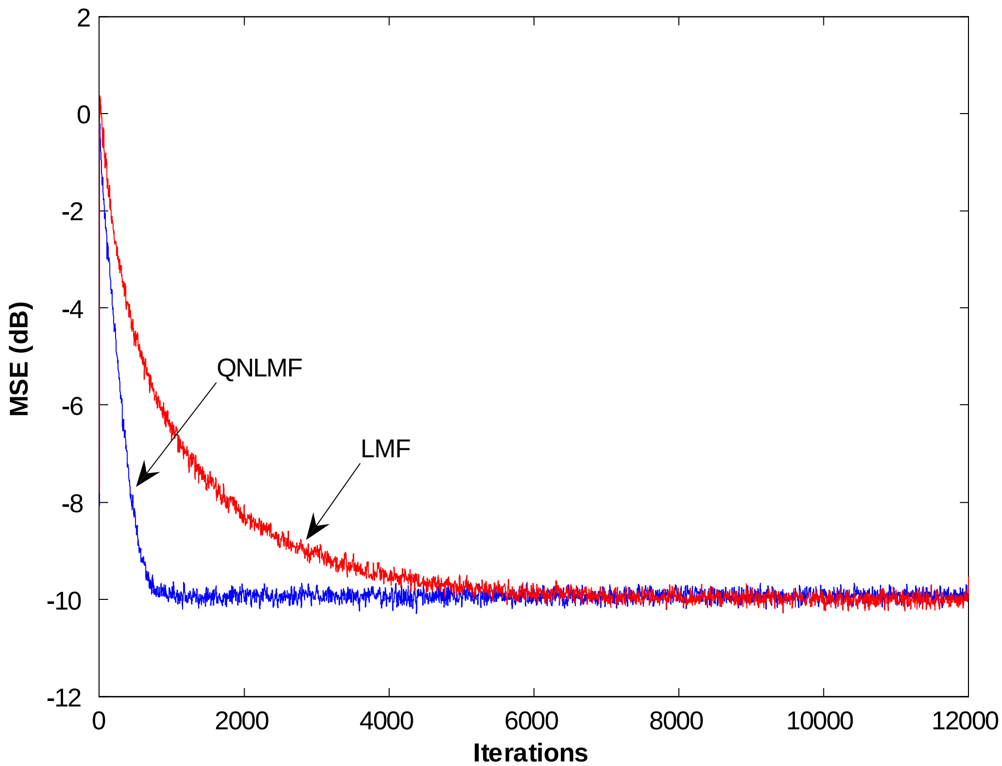

Figure 6.4: Comparison of the convergence speed of the LMF and the proposed QNLMF in AWGN environment with SNR = 10 dB.

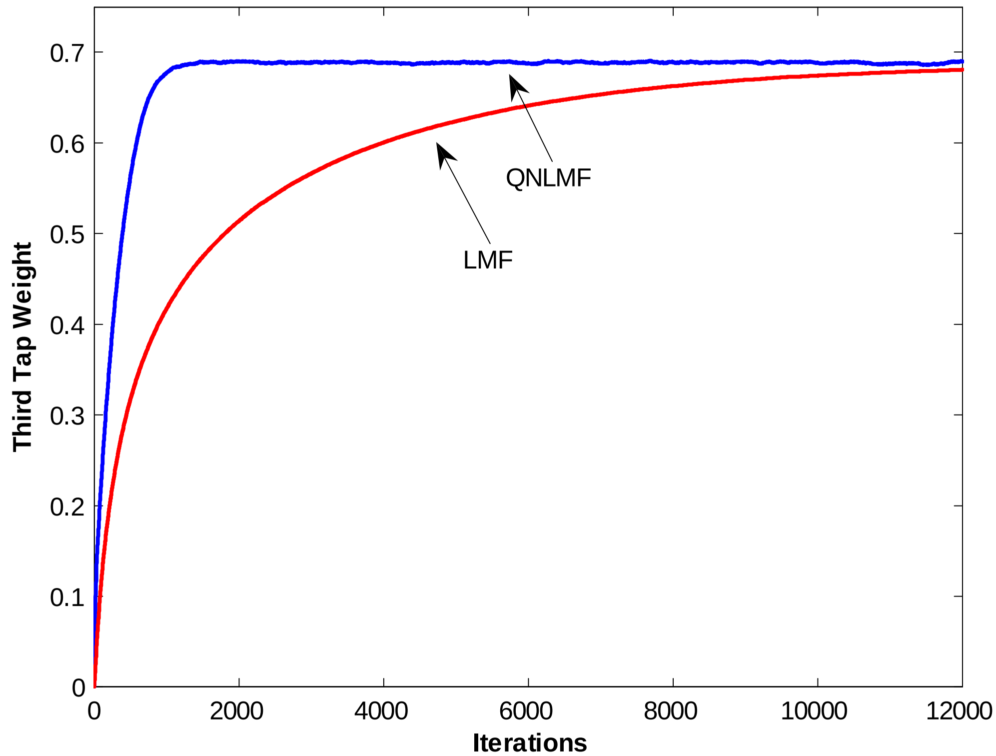

Figure 6.5: Comparison of the learning curves for the third-tap of the LMF and the proposed QNLMF in AWGN environment with SNR = 10 dB.

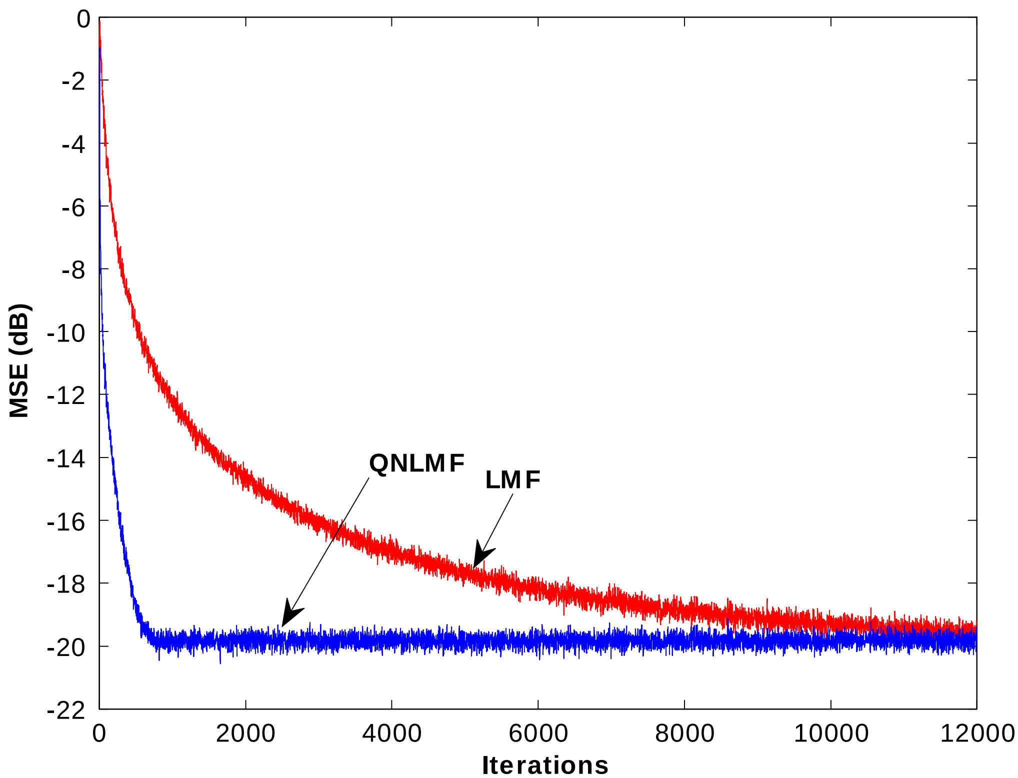

Figure 6.6: Comparison of the convergence speed of the LMF and the proposed QNLMF in AWGN environment with SNR = 20 dB.

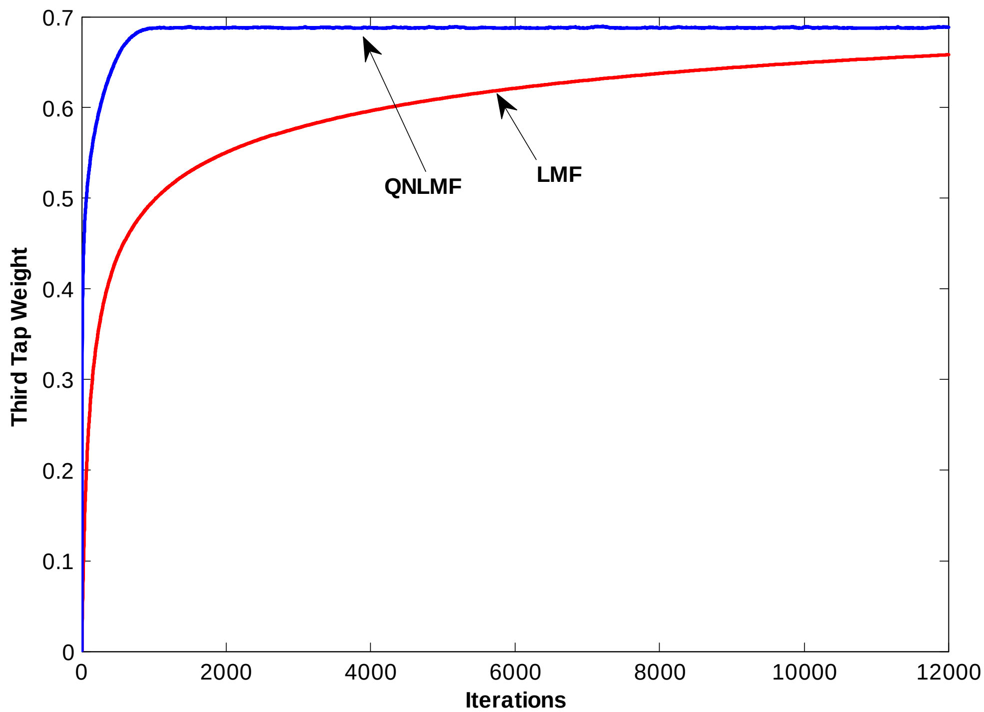

Figure 6.7: Comparison of the learning curves for the third-tap of the LMF and the proposed QNLMF in AWGN environment with SNR = 20 dB.

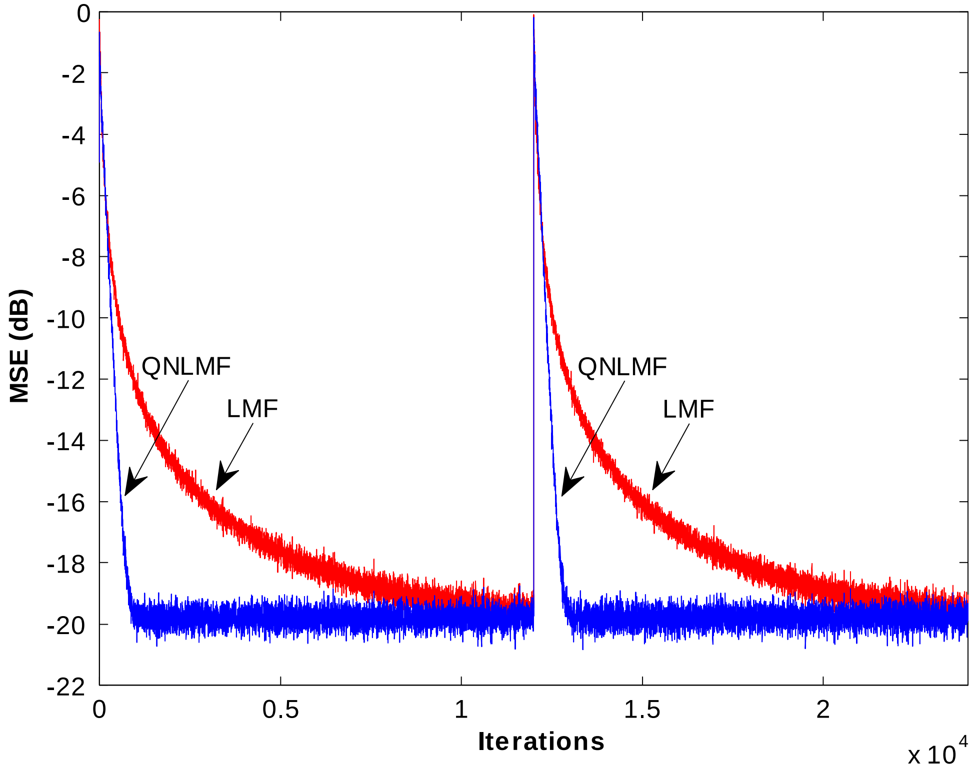

Figure 6.8: Comparison of the convergence speed of the LMF and the proposed QNLMF when there is a sudden burst in AWGN environment with SNR = 20 dB.

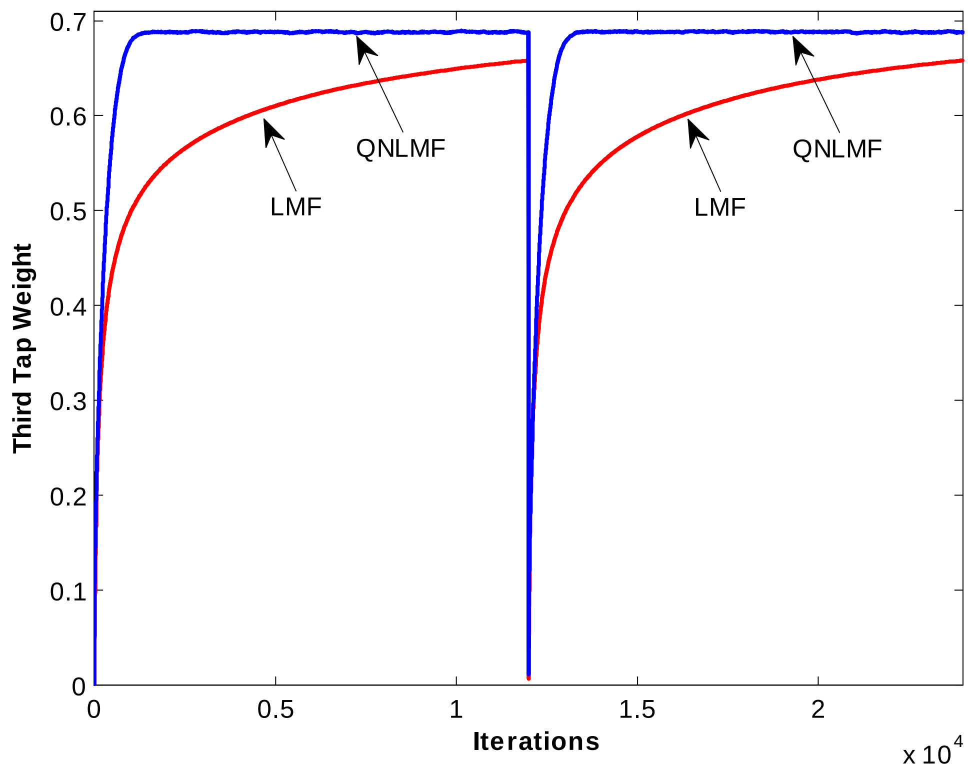

Figure 6.9: Comparison of the learning curves for the third-tap of the LMF and the proposed QNLMF when there is a sudden burst in AWGN environment with SNR = 20 dB.

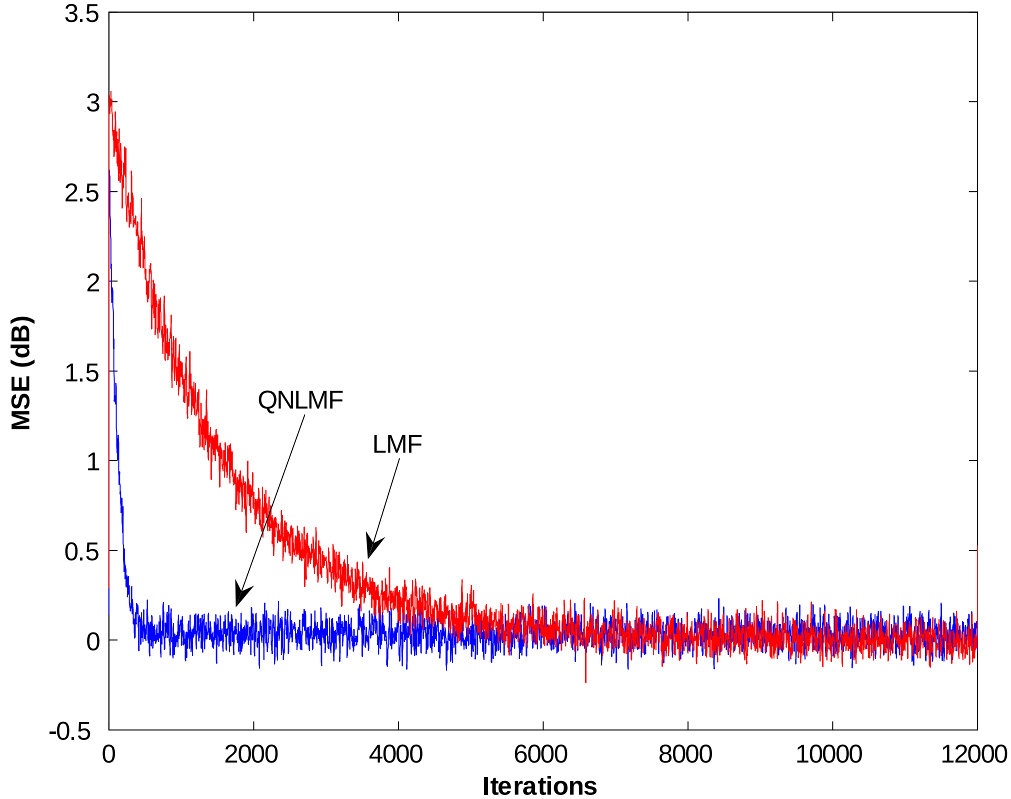

Figure 6.10: Comparison of the convergence speed of the LMF and the proposed QNLMF in Uniform noise environment with SNR = 0 dB.

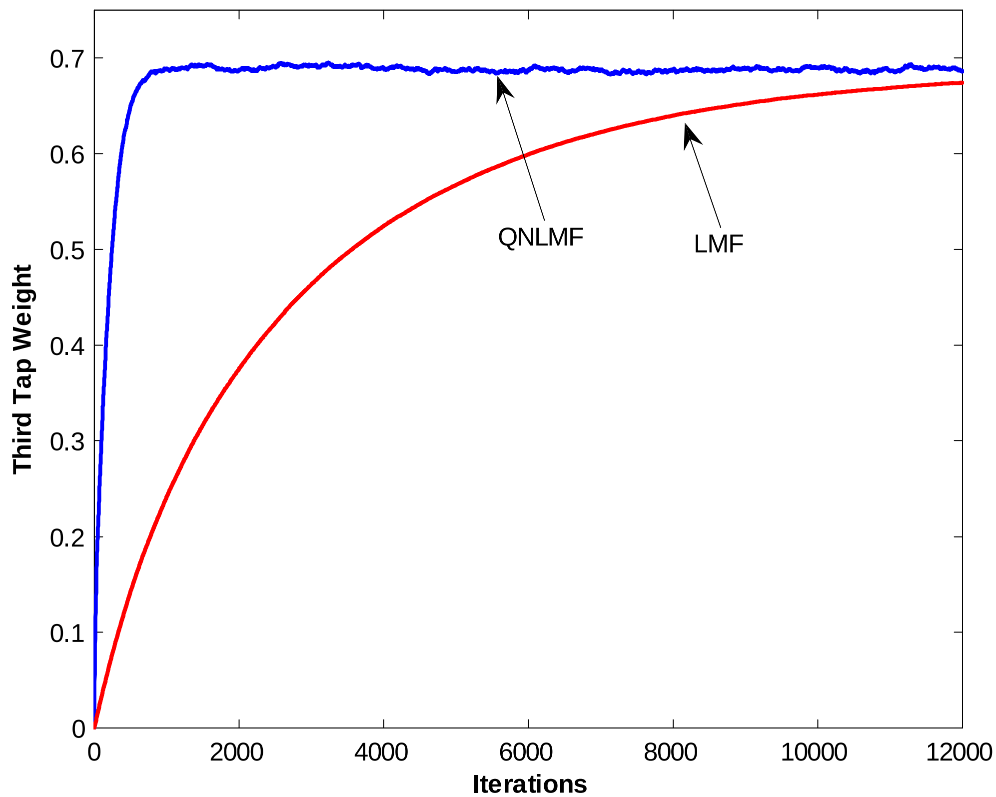

Figure 6.11: Comparison of the learning curves for the third-tap of the LMF and the proposed QNLMF in Uniform noise environment with SNR = 0 dB.

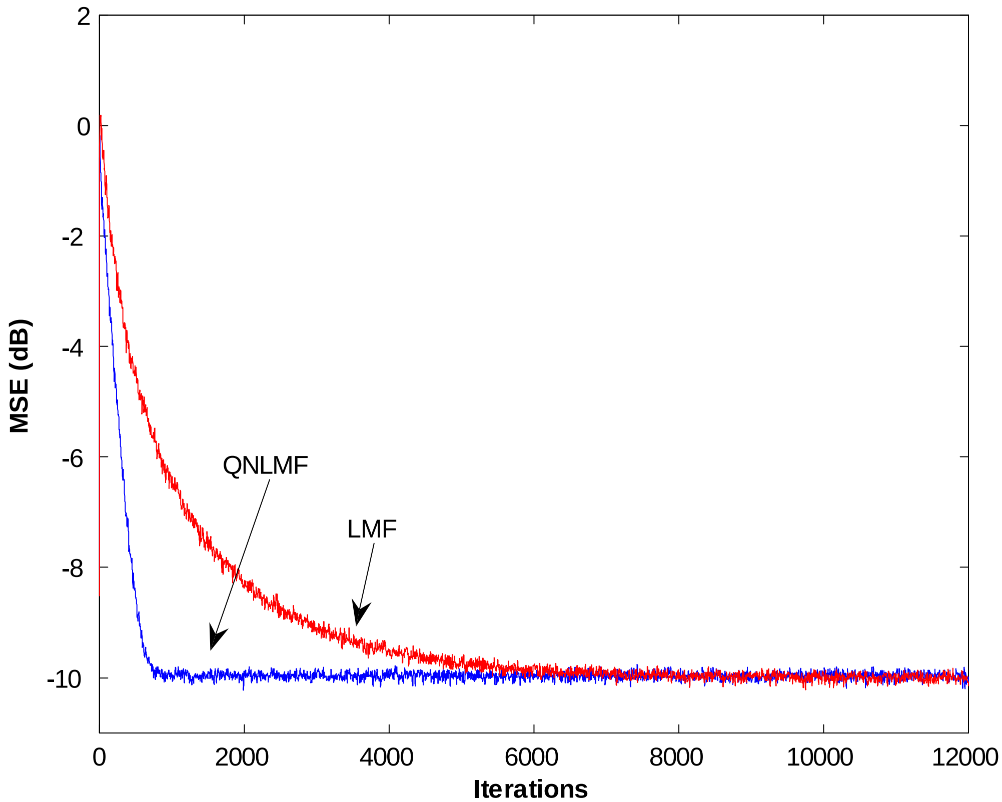

Figure 6.12: Comparison of the convergence speed of the LMF and the proposed QNLMF in Uniform noise environment with SNR = 10 dB.

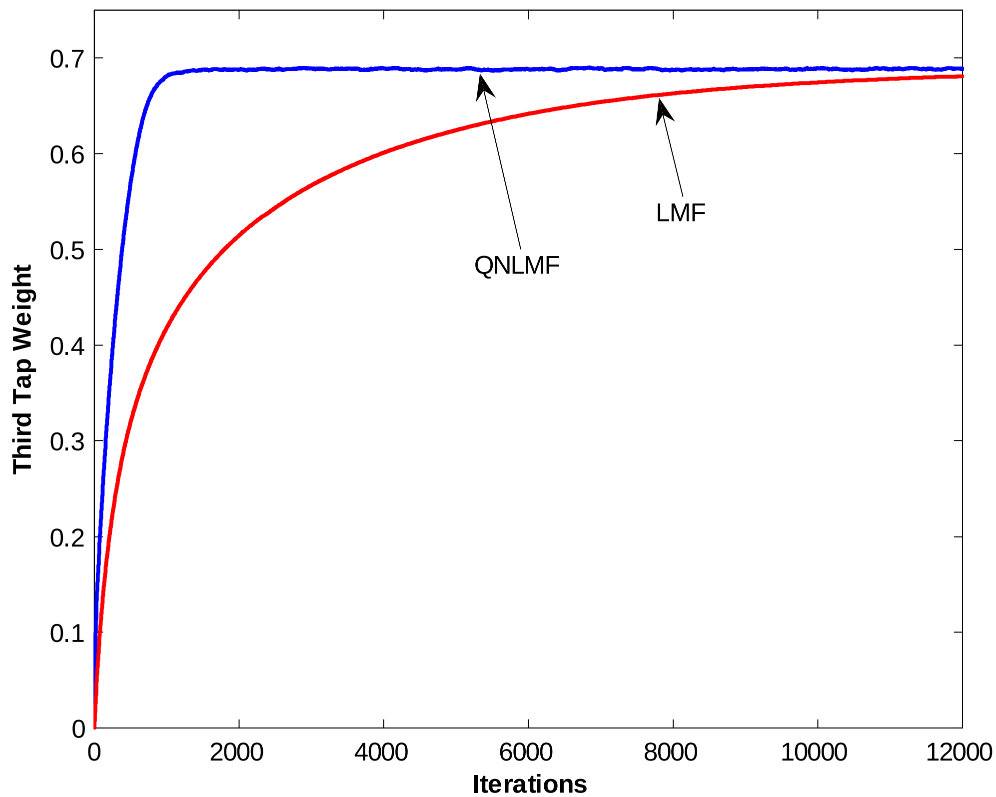

Figure 6.13: Comparison of the learning curves for the third-tap of the LMF and the proposed QNLMF in Uniform noise environment with SNR = 10 dB.

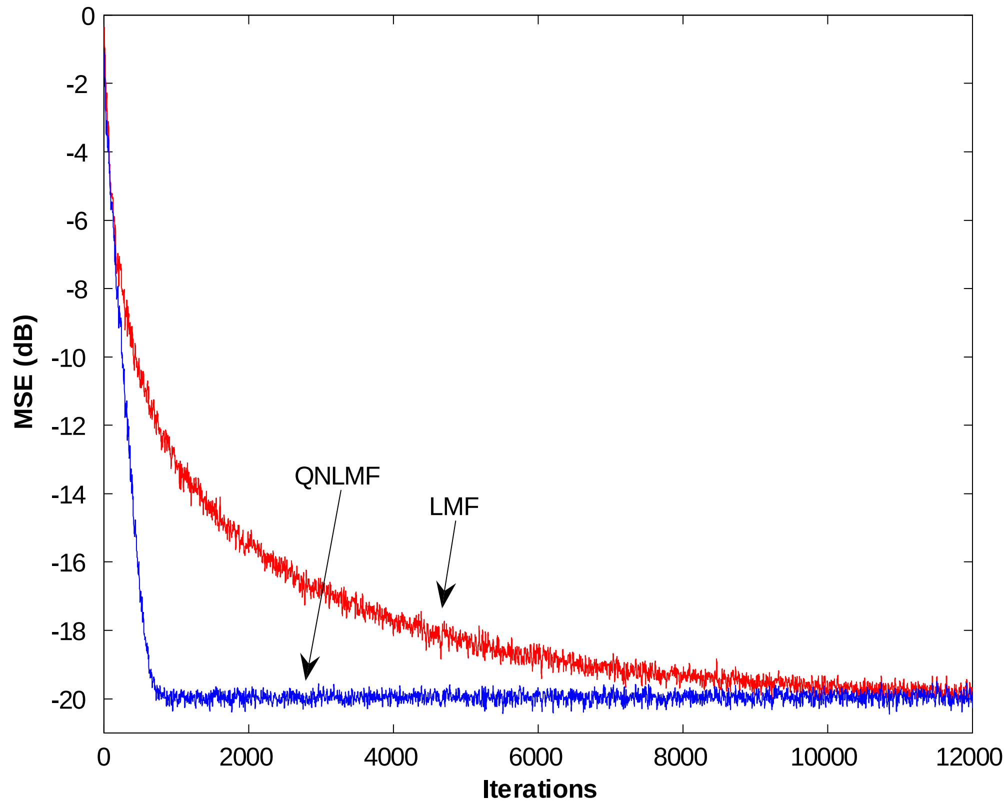

Figure 6.14: Comparison of the convergence speed of the LMF and the proposed QNLMF in Uniform noise environment with SNR = 20 dB.

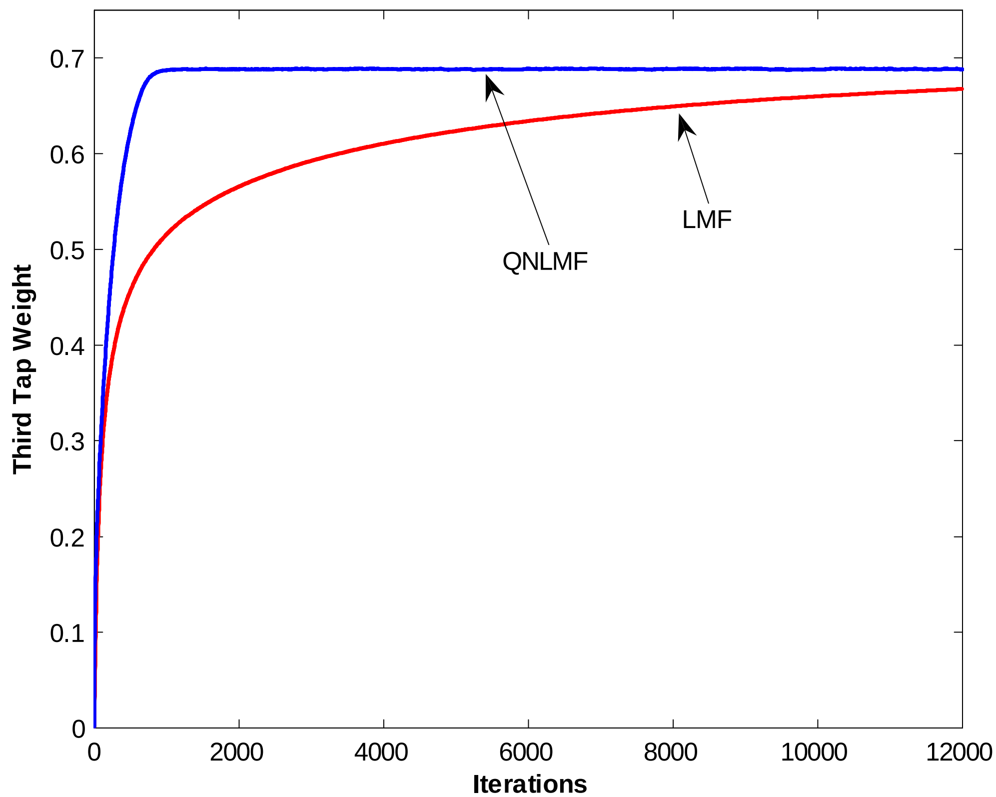

Figure 6.15: Comparison of the learning curves for the third-tap of the LMF and the proposed QNLMF in Uniform noise environment with SNR = 20 dB.

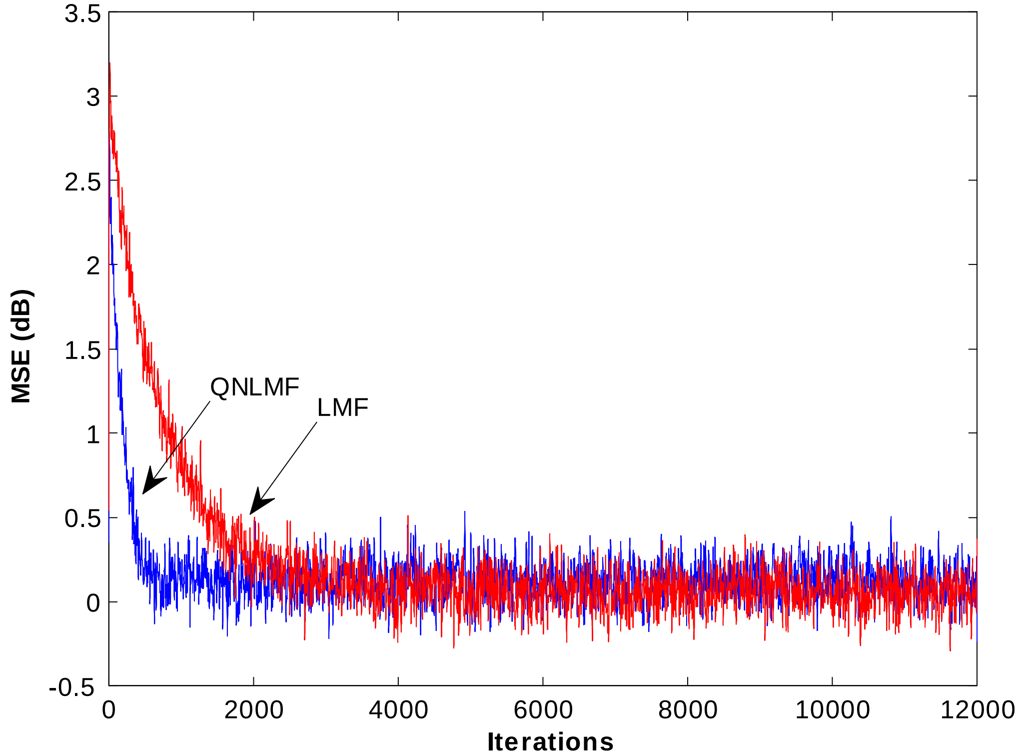

Figure 6.16: Comparison of the convergence speed of the LMF and the proposed QNLMF in Laplacian noise environment with SNR = 0 dB.

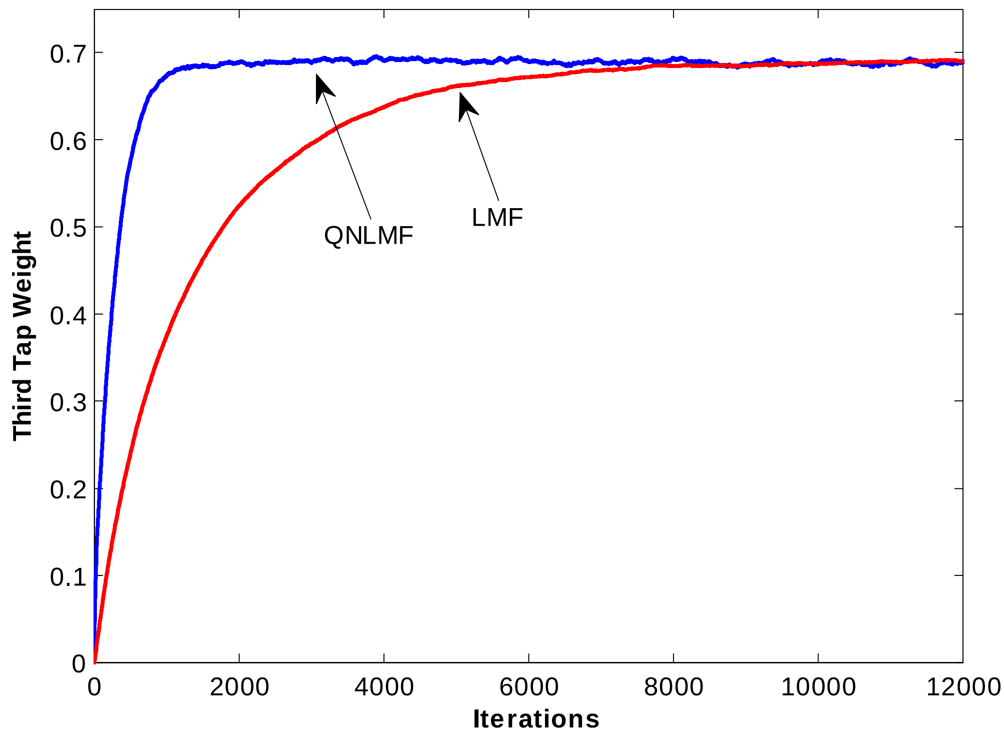

Figure 6.17: Comparison of the learning curves for the third-tap of the LMF and the proposed QNLMF in Laplacian noise environment with SNR = 0 dB.

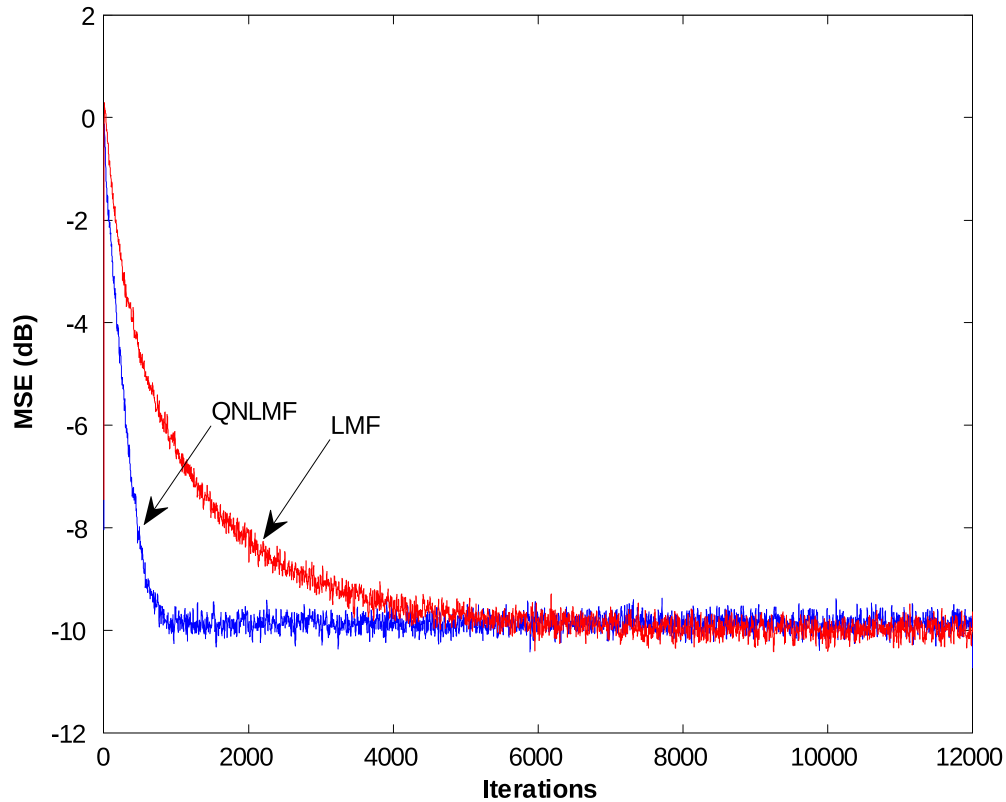

Figure 6.18: Comparison of the convergence speed of the LMF and the proposed QNLMF in Laplacian noise environment with SNR = 10 dB.

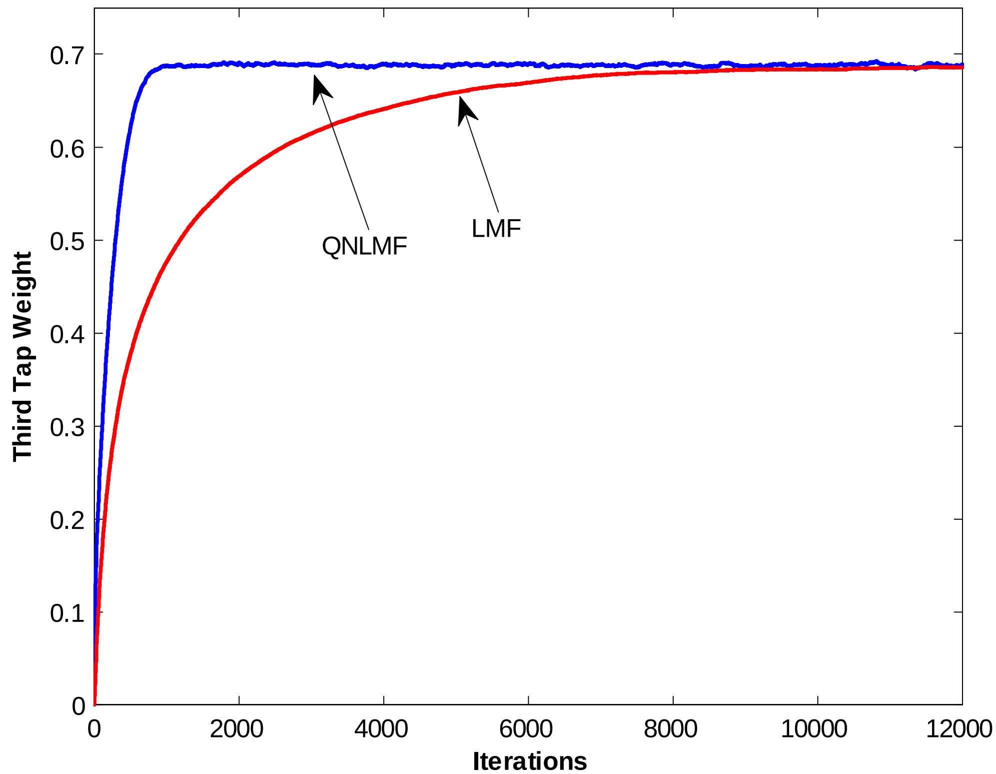

Figure 6.19: Comparison of the convergence speed of the LMF and the proposed QNLMF in Laplacian noise environment with SNR = 10 dB.

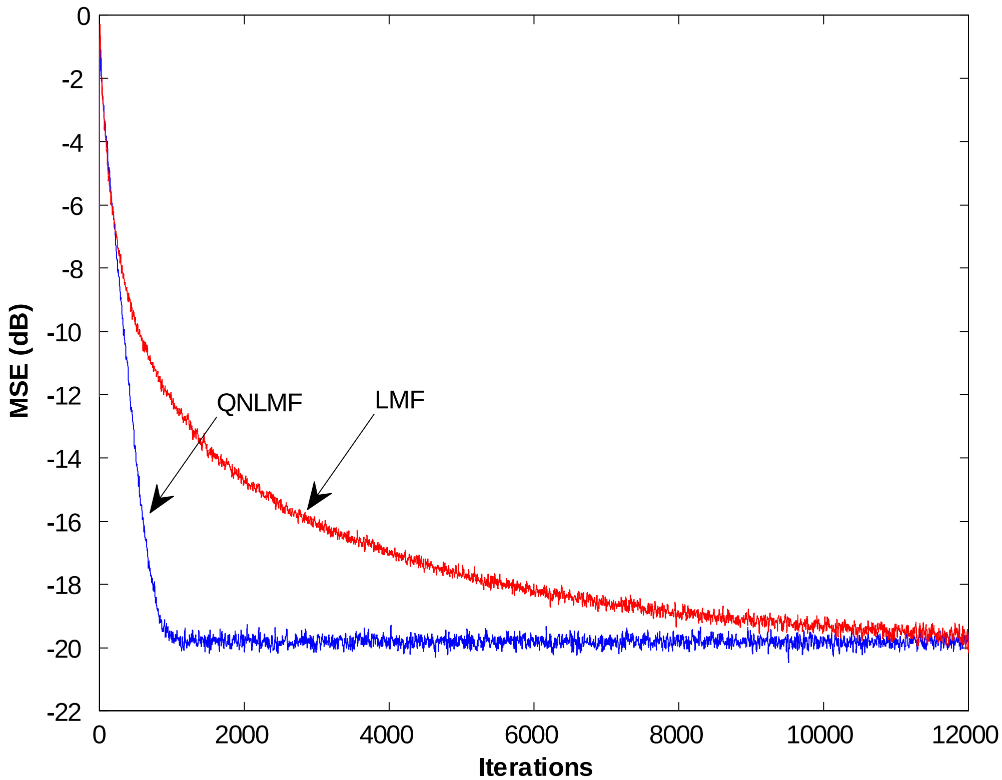

Figure 6.20: Comparison of the convergence speed of the LMF and the proposed QNLMF in Laplacian noise environment with SNR = 20 dB.

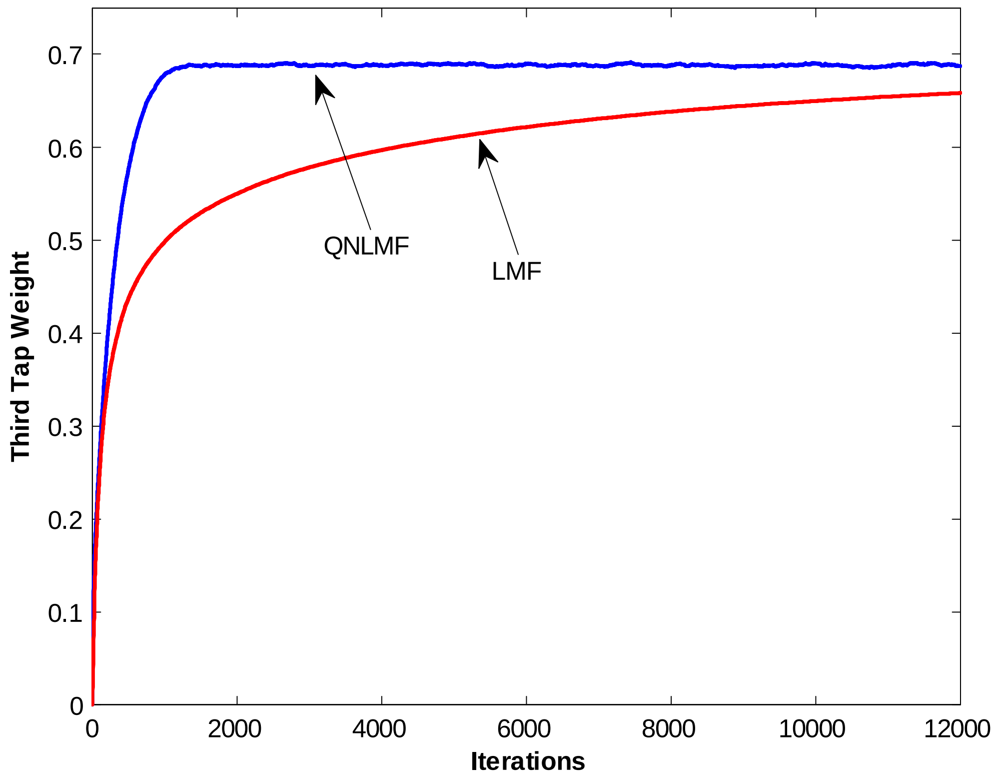

Figure 6.21: Comparison of the learning curves for the third-tap of the LMF and the proposed QNLMF in Laplacian noise environment with SNR = 20 dB.

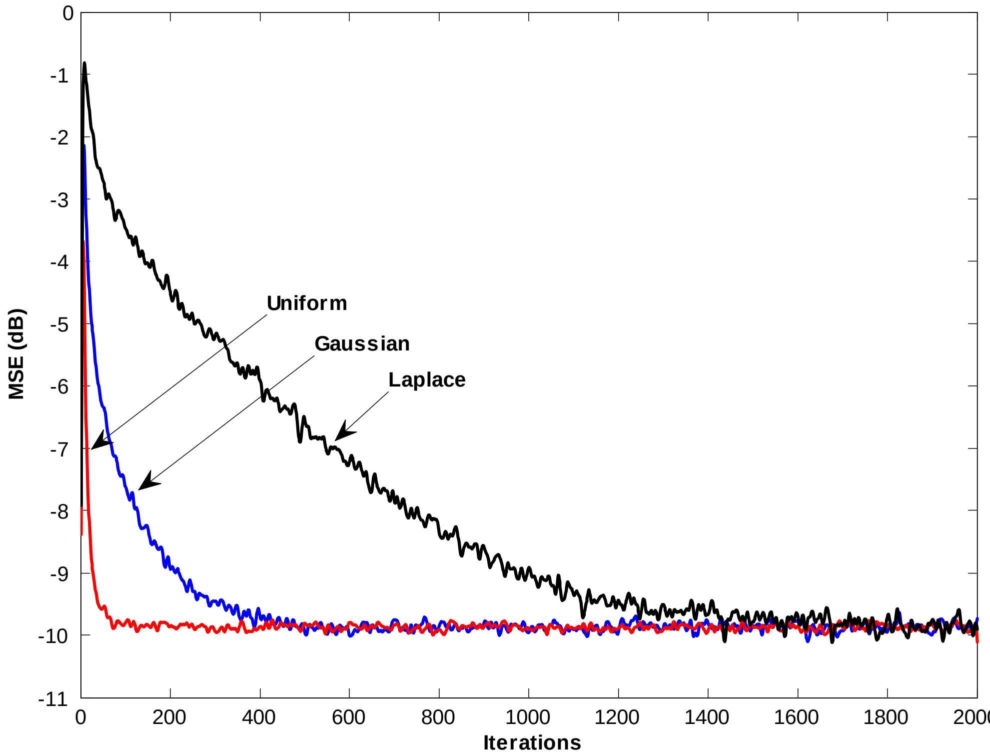

Figure 6.22: Zoomed image of the Convergence behavior of the QNLMF algorithm in presence of Gaussian, Uniform and Laplacian environment with SNR = 10 dB.

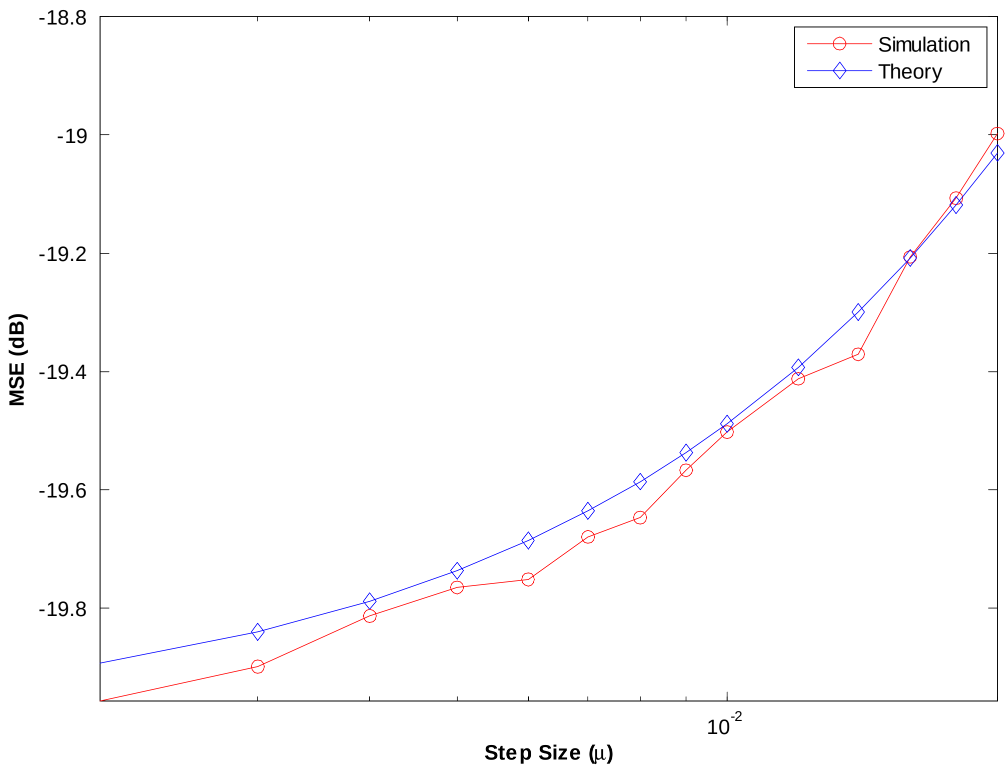

Figure 6.23: Comparison of Analytical and Experimental MSE of QNLMF in stationary environment for different step-sizes.

6.2 Tracking Performance Analysis of the Proposed Algorithms

In this section we investigate the QNLMF algorithm in tracking a constantly varying channel. Two models for the channel coefficient variation are considered here: namely Random Walk and Rayleigh Fading (both single-path and multi-path). The input vector $\{\mathbf{u}_i\}$ is a BPSK $\{\pm 1\}$ signal, and the variance of the additive Gaussian noise is set to achieve an SNR of 20 dB.

6.2.1 Random Walk Model

The random walk model for the channel coefficients is,

$$\mathbf{w}_i^o = \mathbf{w}_{i-1}^o + \mathbf{q}_i \tag{6.4}$$

where $\mathbf{q}_i$ is a white Gaussian vector sequence, whose components are uncorrelated and have zero mean and variance $\sigma_{\mathbf{q}}^2 = 10^{-7}$. The channel used for the mean-square analysis is used here as the initial channel coefficients. The values for different variables have been selected in the same way as was done for mean-square analysis. The results from Figure 6.25 and 6.26 shows the experimental and analytical behavior of EMSE for QNLMF algorithm for varying step-sizes. As it is clear from the results that initially there is some variation between experimental and analytical behavior but as the stepsize is increased a close agreement between theory and simulation results is obtained. Moreover, unlike in the stationary case, the tracking steady-state MSE is a less increasing function of the step-size μ.

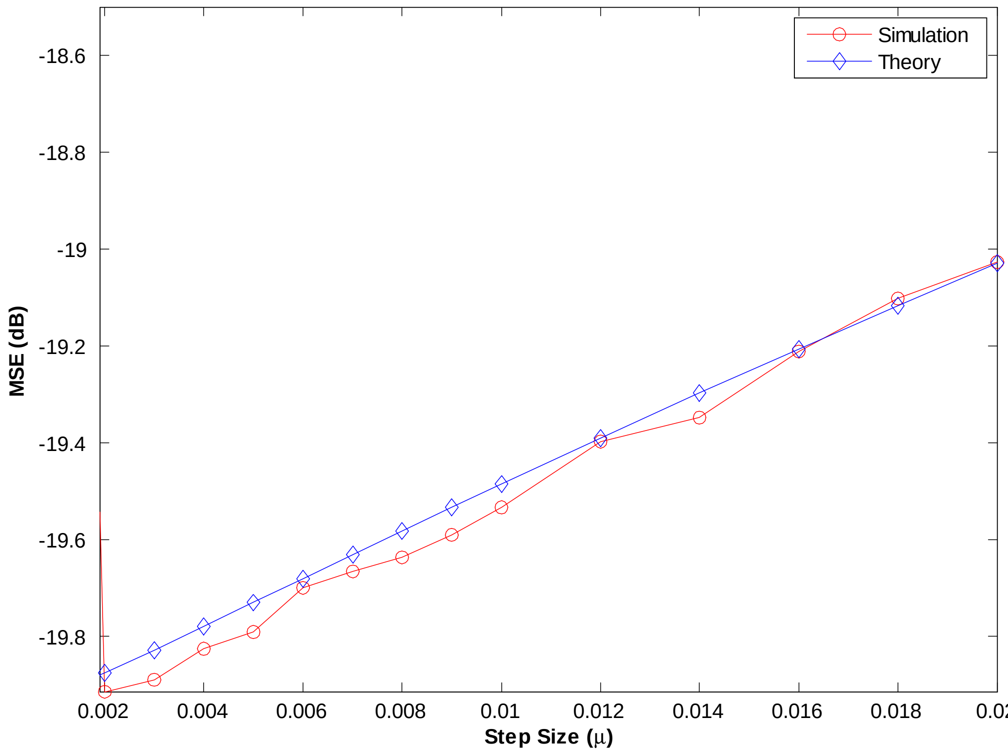

Figure 6.24: Comparison of Analytical and Experimental MSE of QNLMF for Random Walk Channel for different step-sizes.

6.2.2 Rayleigh Fading Model

For the case of single-path and multi-path , the weight-vectors that we wish to estimate looks like:

$$\begin{bmatrix} 0 & 0 & x_2(i) & 0 & 0 \end{bmatrix} \tag{6.5}$$

$$\begin{bmatrix} 0 & 0 & x_2(i) & 0 & x_4(i) \end{bmatrix} \tag{6.6}$$

where $x_2(i)$ and $x_4(i)$ represents absolute values of the Rayleigh fading coefficient. The carrier frequency and the doppler frequency was chosen to be 900 MHz and 66.67 Hz which corresponds to a vehicle moving at a speed of 80 kmph. The sampling period is Ts = 1 ms. The values for different variables have been selected in the same way as was done for Random Walk model in the previous section. The results from Figure 6.27 and 6.28 shows the experimental and analytical behavior of MSE for QNLMF algorithm for varying step-sizes under single and multi-path scenario. As can be seen from these figures, a close agreement between theory and simulation results are obtained and like in Random Walk case, tracking steady-state MSE is a lesser increasing function of the step-size μ.

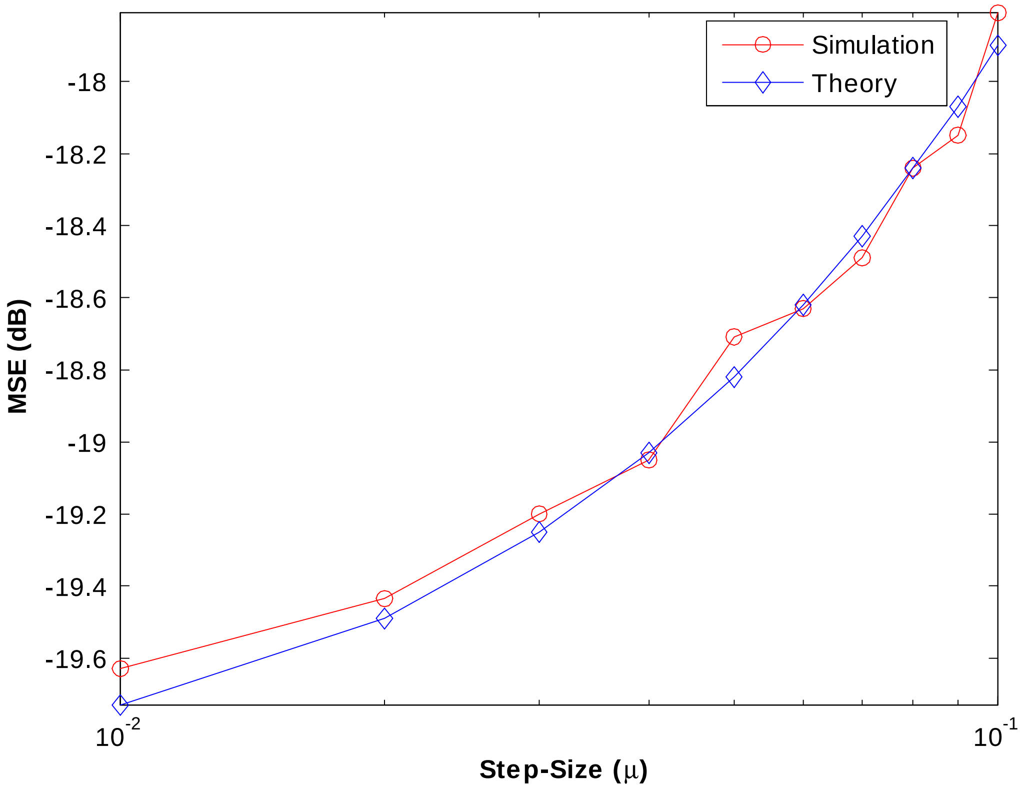

Figure 6.25: Comparison of Analytical and Experimental MSE of QNLMF for single-path Rayleigh fading Channel for different step-sizes.

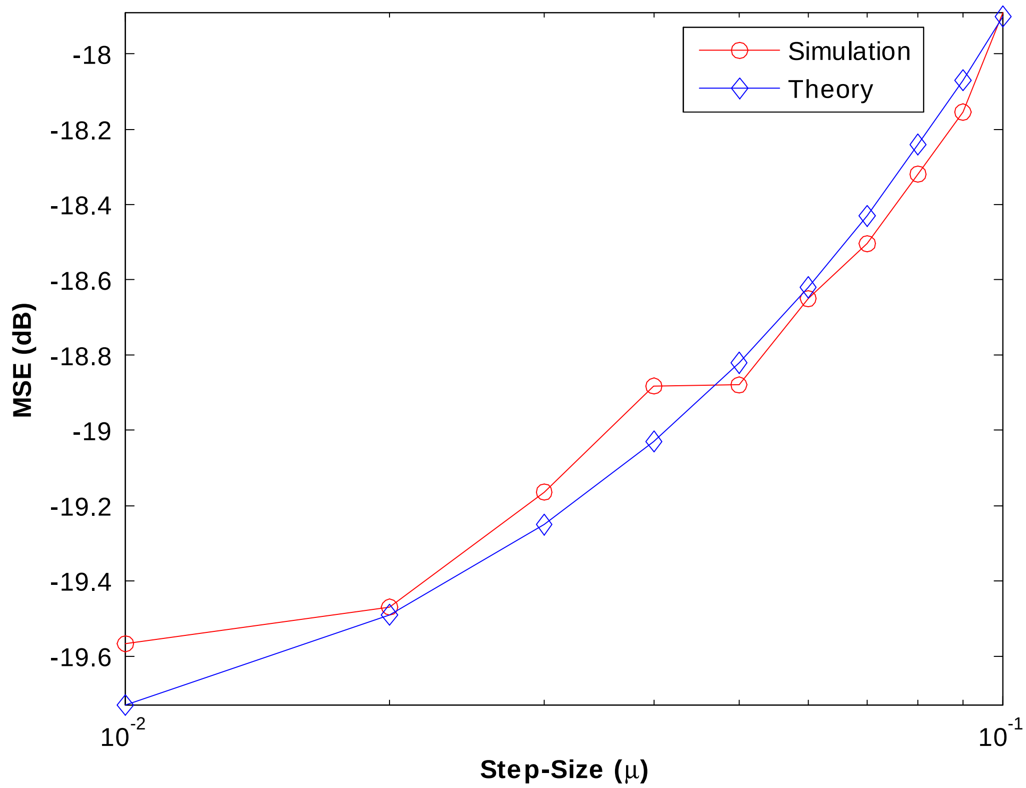

Figure 6.26: Comparison of Analytical and Experimental MSE of QNLMF for multi-path Rayleigh fading Channel for different step-sizes.

6.3 Transient Performance Analysis of the Proposed Algorithm

The transient analysis have been carried out for the uncorrelated input signal where the eigen-value spread of the covariance matrix $\mathbf{R_u}$ was chosen to be 5 and the variance of the additive Gaussian noise was kept at 10^{-5}. The experiment was run over 10^5 iterations and furthermore the results were averaged over 30 iterations. Plots concerning the MSD (mean-square deviation) vs time and MSE (mean-square error) vs time were obtained for LMF, proposed QNLMF algorithm; the results obtained were compared with the theory. As can be seen from the Figures 6.29, 6.30, 6.31 and 6.32, a close agreement between theory and simulation results are obtained for white input as well as for the correlated input.

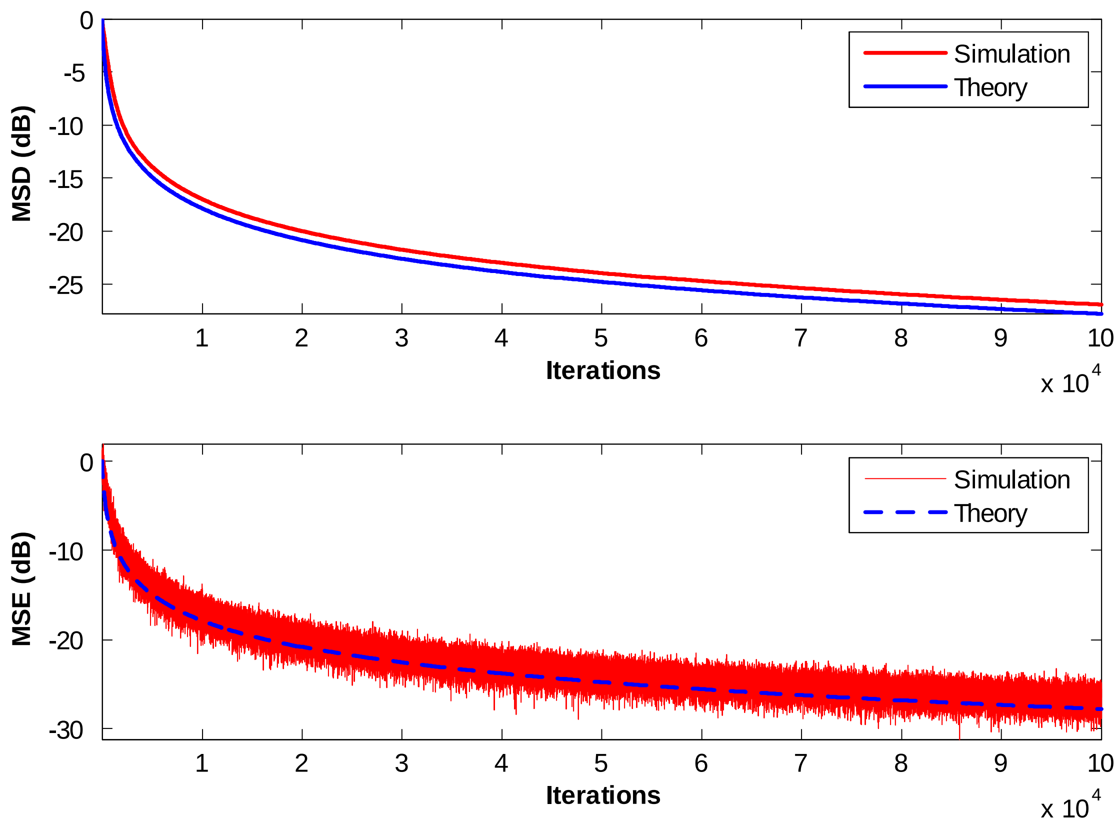

Figure 6.27: Transient Analysis of LMF adaptive algorithm MSD and MSE for White Input Data.

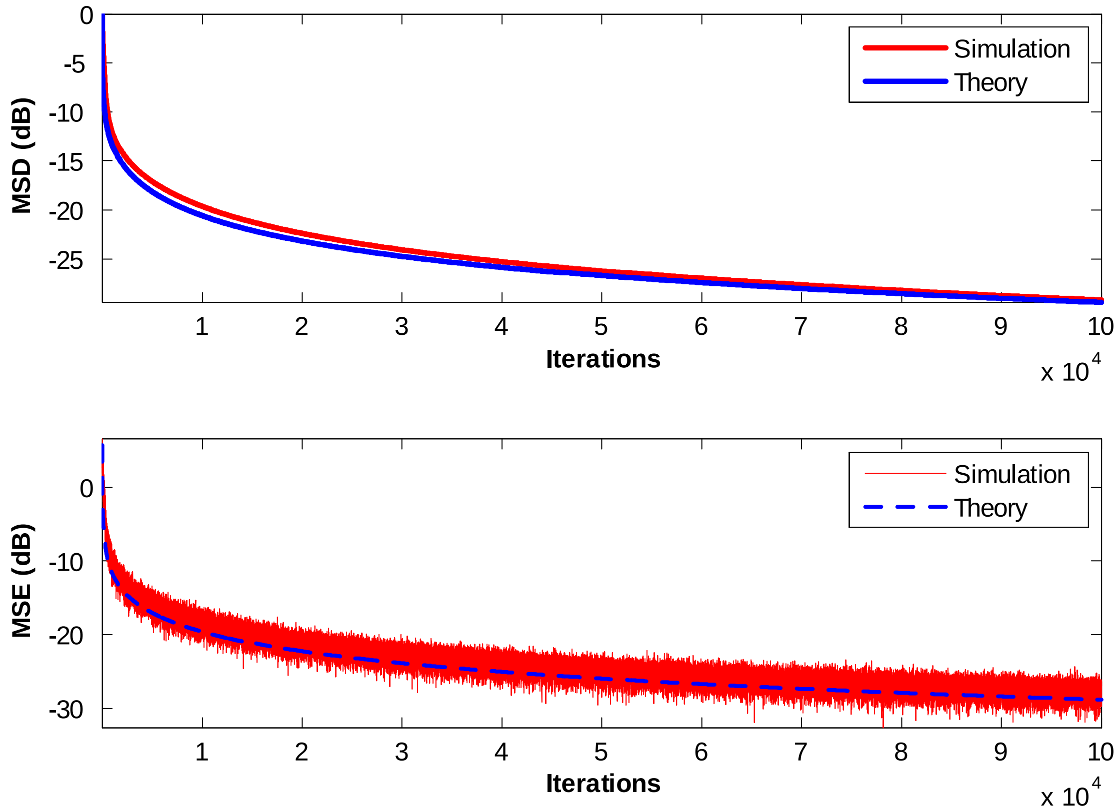

Figure 6.28: Transient Analysis of LMF adaptive algorithm MSD and MSE for Input Data with eigenvalue spread = 5.

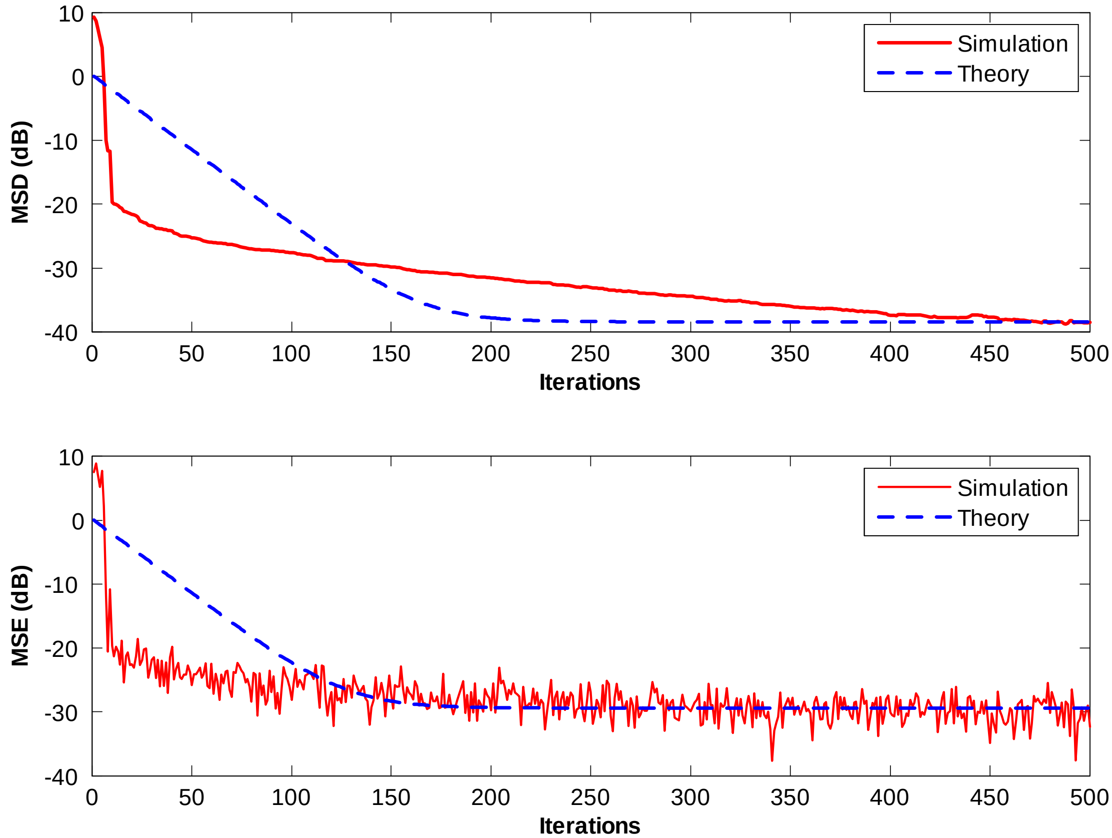

Figure 6.29: Transient Analysis of QNLMF adaptive algorithm MSD and MSE for White Input Data.

6.4 Comparison of QNLMF and RLS

When steady-state EMSE relations for both the RLS and QNLMF were evaluated, we found out that RLS performed better for all the noise enviornments. Moreover, for the same steady-state value of both the algorithms, RLS converges slightly faster then the QNLMF. So, it can be stated that the RLS algorithm performs better then the proposed QNLMF algorithm. Following is the table comparing EMSE of RLS and QNLMF over different SNR for Uniform noise environment,

EMSE in Uniform noise environment			
SNR	0dB	10dB	20dB
RLS	0.0125	0.00125	0.000125
QNLMF	0.0375	0.0032143	0.000321

Table 6.1: Theoretical steady-state EMSE of QNLMF and RLS

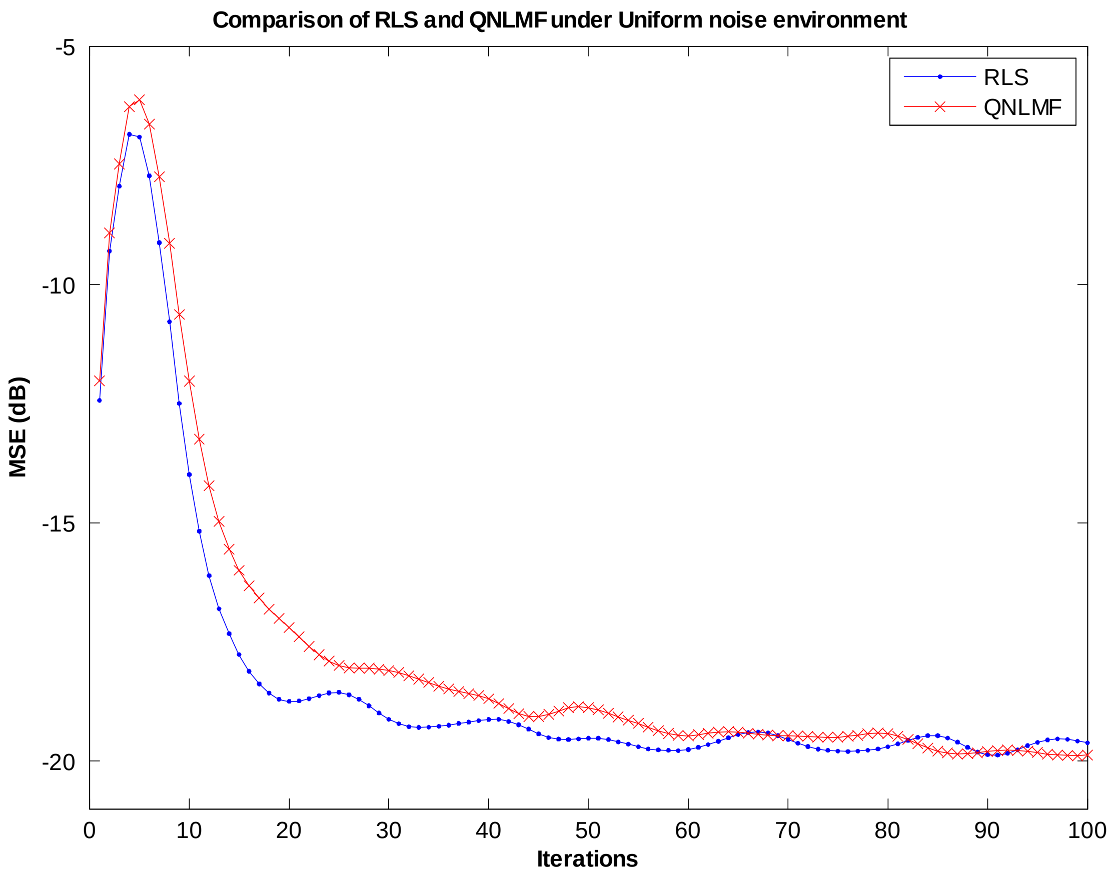

Figure 6.30: Comparison between the proposed QNLMF algorithm and the RLS algorithm.

Chapter 7

Thesis Contributions, Conclusions and Recommendations for Future Work

7.1 Thesis Contributions

This work has successfully presented a Newton's method based LMF adaptive algorithm, namely QNLMF (Quasi-Newton LMF) algorithm. This algorithm is analyzed in terms of convergence properties, steady-state performances, tracking performances and the transient behavior. The performance of the proposed algorithm have been supported by presenting the simulation scenarios. The major contribution of this thesis work are the following:

1. A novel Newton's method based LMF adaptive algorithm is proposed.

2. The convergence analysis of the proposed (QNLMF) algorithm is carried out in terms of mean square sense and the expression for the excess mean-square error is also derived using the fundamental energy relation.

3. Tracking ability of (QNLMF) algorithm is analyzed. The expression for tracking excess mean-square error is also derived.

4. The transient-state behavior of (QNLMF) is also analyzed. A mathematical model was developed to investigate the performance of the proposed algorithm.

5. Finally, the analytical results were then compared with the experimental results which supports the analysis.

7.2 Conclusions

In this thesis, we have proposed a novel Newton's method based LMF algorithm namely QNLMF for wireless environments and studies their performance both analytically and by simulations. Our study included a thorough comparison of the proposed algorithm with the well-established LMF algorithm and showed that, overall, the QNLMF enjoys a much faster convergence performance in the steady-state regime for different noise environments. Since, Newton's method based algorithms are a benchmark, the superior performance was achieved with computational complexity of the scale of RLS.

One important aspect is the stability issue of the proposed algorithm, it has been mentioned in [2] that since the recursions of LMF and LMF-related algorithms employ power of the error signal $\{e(i)\}$, and sometimes this error signal tends to assume relatively larger values during the initial stages of adaptation; therefore for implementation perspective it is sometimes advisable to implement LMF and LMF-related algorithms as follows:

If at a particular iteration it holds that $|e(i)| > 1$;

use the LMS update

else

use the LMF and LMF-related algorithm's update.

Usually, as time progresses, the error signal becomes smaller and, therefore, the steady-state performance of such an implementation would be ultimately dictated by the LMF and LMF-related algorithms and not by the LMS update. Although the above explanation seems resonable but it has not been used for any of the simulations presented in this thesis.

Comparing the computational complexity with the time it takes for an algorithm to reach the steady-state, it can be safely concluded that if the computational complexity is of paramount importance, then the method of steepest descent is the prefered iterative method for computing the tap-weight vector of adaptive traversal filter operating in a wide-sense stationary environment. If, on the other hand, the rate of convergence is the issue of interest, then Newton's method is the prefered approach. Hence, the computational complexity and the rate of convergence for an adaptive filter are usually two conflicting parameters, only one of them is usually satisfied and therefore the choice of them depends on which is of paramount importance to the application in hand.

7.3 Future Work

There are few suggestions regarding future work. In this thesis, convergence, tracking and transient analysis of the proposed algorithm is carried out under Gaussian, Uniform and Laplace distributed noise environments. It can be extended to more types of disturbances as well. As QNLMF works with a constant step-size, it can be extended to accomodate the analysis under variable step-size. These suggestions can be incorporated to augment the performane of QNLMF adaptive algorithm.

Appendix I.

Derivation of the cost function $E\left[e(i)^4\right]$

$$e(i) = d(i) - \mathbf{u}_i^T\mathbf{w}_i$$

$$\begin{aligned}
E\left(e(i)^4\right) &= E\left(d_i - \mathbf{u}_i^T\mathbf{w}_i\right)^4 \\
E\left(e(i)^4\right) &= E(d_i^4) - 4E(d_i^3\mathbf{u}_i^T\mathbf{w}_i) + 6E(d_i^2\mathbf{u}_i^T\mathbf{w}_i\mathbf{w}_i^T\mathbf{u}_i) \\
&\quad -4E(d_i\mathbf{u}_i^T\mathbf{w}_i\mathbf{w}_i^T\mathbf{u}_i\mathbf{u}_i^T\mathbf{w}_i)+E(\mathbf{u}_i^T\mathbf{w}_i\mathbf{w}_i^T\mathbf{u}_i\mathbf{u}_i^T\mathbf{w}_i\mathbf{w}_i^T\mathbf{u}_i) \\
E\left(e(i)^4\right) &= E(d_i^4) - 4E(d_i^3\mathbf{u}_i^T)\mathbf{w}_i + 6\mathbf{w}_i^T E(d_i^2\mathbf{u}_i\mathbf{u}_i^T)\mathbf{w}_i \\
&\quad -4E(d_i\mathbf{u}_i^T\mathbf{w}_i\mathbf{w}_i^T\mathbf{u}_i\mathbf{u}_i^T\mathbf{w}_i)+E(\mathbf{u}_i^T\mathbf{w}_i\mathbf{w}_i^T\mathbf{u}_i\mathbf{u}_i^T\mathbf{w}_i\mathbf{w}_i^T\mathbf{u}_i)
\end{aligned}$$

Gradient vector

$$
\begin{aligned}
\frac{\partial E\left(e(i)^4\right)}{\partial \mathbf{w}_i} &= 4E\left\{e(i)^3 \frac{\partial e(i)}{\partial \mathbf{w}_i}\right\} \\
\frac{\partial E\left(e(i)^4\right)}{\partial \mathbf{w}_i} &= 4E\left\{e(i)^3 \frac{\partial \left[d(i) - \mathbf{u}_i^T \mathbf{w}_i\right]}{\partial \mathbf{w}_i}\right\} \\
\frac{\partial E\left(e(i)^4\right)}{\partial \mathbf{w}_i} &= -4E\left\{e(i)^3 \mathbf{u}_i^T\right\}
\end{aligned}
$$

Hessian matrix

$$
\begin{aligned}
\frac{\partial^2 E\left(e(i)^4\right)}{\partial \mathbf{w}_i^2} &= -12E\left\{e(i)^2 \frac{\partial e(i)}{\partial \mathbf{w}_i} \mathbf{u}_i^T\right\} \\
\frac{\partial^2 E\left(e(i)^4\right)}{\partial \mathbf{w}_i^2} &= 12E\left\{e(i)^2 \mathbf{u}_i \mathbf{u}_i^T\right\}
\end{aligned}
$$

Appendix II.

Expansion of the term $E[\mathbf{\Phi}_i]$

$$
\begin{aligned}
E[\mathbf{\Phi}_i] &= E\left[e^2(i)\mathbf{u}_i\mathbf{u}_i^T\right] \\
&= E\left[e_a^2(i)\mathbf{u}_i\mathbf{u}_i^T\right] + 2E\left[e_a(i)n(i)\mathbf{u}_i\mathbf{u}_i^T\right] + E\left[n^2(i)\mathbf{u}_i\mathbf{u}_i^T\right] \\
&= E\left[\mathbf{u}_i e_a^2(i)\mathbf{u}_i^T\right] + E\left[n^2(i)\right]E\left[\mathbf{u}_i\mathbf{u}_i^T\right] \\
&= E\left[\mathbf{u}_i\left(\tilde{\mathbf{w}}_{i-1}^T\mathbf{u}_i\mathbf{u}_i^T\tilde{\mathbf{w}}_{i-1}\right)\mathbf{u}_i^T\right] + \sigma_n^2\mathbf{R}_\mathbf{u}.
\end{aligned}
$$

Now, assuming that the regressors $\mathbf{u}_i$ are IID, we can proceed as follows:

$$
\begin{aligned}
E[\mathbf{\Phi}_i] &= E\left[\mathbf{u}_i\tilde{\mathbf{w}}_{i-1}^T E\left(\mathbf{u}_i\mathbf{u}_i^T|\tilde{\mathbf{w}}_{i-1}\right)\tilde{\mathbf{w}}_{i-1}\mathbf{u}_i^T\right] + \sigma_n^2\mathbf{R}_\mathbf{u} \\
&= E\left[\mathbf{u}_i\left(\tilde{\mathbf{w}}_{i-1}^T\mathbf{R}_\mathbf{u}\tilde{\mathbf{w}}_{i-1}\right)\mathbf{u}_i^T\right] + \sigma_n^2\mathbf{R}_\mathbf{u} \\
&= E\left[\mathbf{u}_i\|\tilde{\mathbf{w}}_{i-1}\|_{\mathbf{R}_\mathbf{u}}^2\mathbf{u}_i^T\right] + \sigma_n^2\mathbf{R}_\mathbf{u} \\
&= E\left[\|\tilde{\mathbf{w}}_{i-1}\|_{\mathbf{R}_\mathbf{u}}^2\mathbf{u}_i\mathbf{u}_i^T\right] + \sigma_n^2\mathbf{R}_\mathbf{u} \\
&= E\left[\|\tilde{\mathbf{w}}_{i-1}\|_{\mathbf{R}_\mathbf{u}}^2\right]E\left[\mathbf{u}_i\mathbf{u}_i^T\right] + \sigma_n^2\mathbf{R}_\mathbf{u},
\end{aligned}
$$

since, $\tilde{\mathbf{w}}_{i-1}$ is independent of $\mathbf{u}_i$ as a consequence of the previous assumption, so,

$$\begin{aligned} E\left[\mathbf{\Phi}_i\right] &= E\left\|\tilde{\mathbf{w}}_{i-1}\right\|_{\mathbf{R_u}}^2 \mathbf{R_u} + \sigma_n^2 \mathbf{R_u} \\ &= \left(E\left\|\tilde{\mathbf{w}}_{i-1}\right\|_{\mathbf{R_u}}^2 + \sigma_n^2\right) \mathbf{R_u}. \end{aligned}$$

Bibliography

[1] S. Haykin, Adaptive Filter Theory, 4th edition. NJ: Prentice-Hall, 1996.

[2] A. H. Sayed, Fundamentals of Adaptive Filtering. Wiley-IEEE Press, 2003.

[3] S. A. A. Imam, “A Statistical Noise Constrained Least Mean Fourth Adaptive Algorithm,” Thesis Dessertation, Oct 2007, King Fahd University of Petroleum and Minerals.

[4] S.U.H.Qureshi, "Adaptive Equalization," IEEE Proceedings, Vol. 53, pp. 1349-1387, 1985.

[5] B. Widrow, J. Glover, J. McCool, J. Kaunitz, C. Williams, R. H. Hearn, J. R. Zeidler, E. Dong, and R. Goodin, "Adaptive Noise Cancellation: Principles and applications," Proceedings of the IEEE, Vol. 63, pp. 1692-1717, 1975.

[6] J. Makhoul, "Linear Prediction: A Tutorial," Proceedings of the IEEE, vol. 63, no. 4, pp. 561-580, 1975.

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

[8] E. Walach and B. Widrow, "The Least Mean Fourth (LMF) Adaptive Algorithm and Its Family," IEEE Transaction on Information Theory., vol. IT-30, no. 2, Mar. 1984

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

[10] S. A. A. Imam, A. Zerguine and M. Moinuddin, "A Statistical Noise Constrained Least Mean Fourth Adaptive Algorithm," IEEE International Conference on. ICASSP 2008., Mar. 31 2008-Apr. 4 2008, pp. 3817 - 3820

[11] P. I. Hübscher, J. C. M. Bermudez and V. H. Nascimento, "A Mean-Square Stability Analysis of the Least Mean Fourth Adaptive Algorithm," IEEE Trans. Signal Proc., Vol. 55, Aug. 2007, pp. 4018 - 4028

[12] S. J. Huang, S. N. Koh and H. K. Tang, "Training Algorithm Based on Newton's Method with Dynamic Error Control," IEEE International Joint Conference on Neural Networks, 1992. IJCNN., Vol. 3, 7-11 Jun. 1992, pp. 899 - 90

[13] S. H. Cho, S. D. Kim and K. Y. Jeon, "Statistical Convergence of the Adaptive Least Mean Fourth Algorithm," IEEE3rd International Conference on Signal Process, 1996., Vol 1, 14-18 Oct. 1996, pp. 610 - 613

[14] B. Widrow and M. Kamenetsky, "Statistical efficiency of adaptive algorithms," Advances in neural networks research — pp. 735-744, Vol. 16 , Jun. 2003

[15] P. S. R. Diniz, M. L. R. de Campos and A. Antoniou, "Analysis of LMS-Newton adaptive filtering algorithms with variable convergence factor," IEEE Transactions on Signal Process, Vol. 43, Mar. 1995, pp. 617 - 627

[16] R. H. Kwong and E. W. Johnston, "A variable step size LMS algorithm," IEEE Transactions on Signal Process, Vol 40, July 1992, pp. 1633 - 1642

[17] S. H. Cho and S. D. Kim, "Adaptive filters based on the high order error statistics," IEEE Asia Pacific Conference on Circuits and Systems, 18-21 Nov. 1996, pp. 109 - 112

[18] N. R. Yousef and A. H. Sayed, "Tracking analysis of the LMF and LMMN adaptive algorithms," Proceedings of 33rd Asilomar Conference on Signals, Systems, and Computers, 1999. Vol. 1, 24-27 Oct. 1999, pp. 786 - 790

[19] T. Y. Al-Naffouri and A. H. Sayed, "Transient analysis of data-normalized adaptive filters," IEEE Transactions on Signal Processing, Vol. 51, Mar. 2003, pp. 639 - 652

[20] T. Y. Al-Naffouri and A. H. Sayed, "Transient analysis of adaptive filters with error nonlinearities," IEEE Transactions on Signal Processing, Vol 51, Mar. 2003, pp. 653 - 663

[21] M. Rupp and A. H. Sayed, "Robustness of Gauss-Newton Recursive Methods: A Deterministic Feedback Analysis," IEEE Transactions on Signal Processing, vol. 50, no. 3, pp. 165-188, Jun. 1996

[22] T. Y. Al-Naffouri, A. Zerguine and M. Bettayeb, "A unifying view of error nonlinearities in LMS adaptation," IEEE International Conference on Acoustics, Speech and Signal Process. Vol. 3, 12-15 May 1998, pp. 1697 - 1700

[23] R. Heijmans, "When does the expectation of a ratio equal the ratio of expectations?," Statistical Papers 40, 107-115 (1999)

[24] J. L. Fleiss, "A note on the expectation of the reciprocal of a random variable," The American Statistician, Vol. 20, no. 1 (Feb., 1966), p. 25

[25] V. H. Nascimento and J. C. M. Bermudez, "Probability of divergence for the least-mean fourth algorithm," IEEE transactions on Signal Processing, Vol. 54, Issue 4, pp. 1376-1385, April 2006.

[26] T. Y. Al-Naffouri, A. Zerguine and M. Bettayeb, "A Unifying View of Error Nonlinearities in LMS Adaptation," in Proc. ICASSP, Seattle, USA, May 12-15, 1998.

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

[28] T. Y. Al-Naffouri and A, H, Sayed, "Optimum Error Nonlinearity for Long Adaptive Filters," IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP '02), PP. 1373-1376, vol. 2, 2002.

[29] T. Y. Al-Naffouri and A, H, Sayed, "Adaptive Filters with Error Nonlinearities: Mean-Square Analysis and Optimum Design," EURASIP Journal on Applied Signal Processing, pp. 192-205, No. 4, March, 2001.

[30] T. Y. Al-Naffouri and A.H.Sayed, "Transient Analysis of Adaptive Filters with Error Non-linearities," IEEE Transactions on Signal Processing, pp. 653-663, vol 51, March, 2003.

[31] L. Garcia, "Probability and Random Processes for Electrical Engineering," 2nd ed., Addison-Wesley, July, 1993.

[32] A. Papoulis, "Probability, Random Variables and Stochastic Processes", 3rd ed. New York McGraw-Hill, 1991.

[33] R. Price, "A useful theorem for nonlinear devices having Gaussian inputs," IRE Trans. Inform. Theory, vol. IT-4, pp. 6972, June 1958.

[34] T. Y. Al-Naffouri and A.H.Sayed, "Adaptive Filters with Error Nonlinearities: Mean-Square Analysis and Optimum Design," EURASIP Journal on Applied Signal Processing 2001:4, 192-205

[35] Azzedine Zerguine, et al., "Convergence and tracking analysis of a variable normalised LMF (XE-NLMF) algorithm," Journal of Signal Processing Volume 89, Issue 5, May 2009, pp. 778-790

Vitae

- Umair bin Mansoor.
- Born in Karachi, Pakistan.
- Received Bachelor of Engineering (B.E.) degree in Computer Systems Engineering from S.S.U.E.T., Karachi, Pakistan in 2004.
- Worked as a Network Engineer in Cybersoft Tech. World Online (WOL) currently LINKdotNET, Karachi, Pakistan.
- Joined King Fahd University of Petroleum and Minerals in September 2006.
- Email: {*umairbinmansoor@hotmail.com*}