

**THE SIGNED REGRESSOR LEAST MEAN FOURTH
(SRLMF) ADAPTIVE ALGORITHM**

BY

MOHAMMED MUJAHID ULLA FAIZ

A Thesis Presented to the
DEANSHIP OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

In

ELECTRICAL ENGINEERING

November 2009

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DHAHRAN 31261, SAUDI ARABIA

DEANSHIP OF GRADUATE STUDIES

This thesis, written by **MOHAMMED MUJAHID ULLA FAIZ** under the direction of his thesis adviser and approved by his thesis committee, has been presented to and accepted by the Dean of Graduate Studies, in partial fulfillment of the requirements for the degree of **MASTER OF SCIENCE IN ELECTRICAL ENGINEERING**.

Thesis Committee

Azzedine Zerguine

Dr. Azzedine Zerguine (Adviser)

Asrar Ul-Haq Sheikh

Dr. Asrar Ul-Haq Sheikh (Member)

Omar A. Al-Swailem
Dr. Omar A. Al-Swailem (Member)

Samir H. Abdul-Jauwad
Dr. Samir H. Abdul-Jauwad
Department Chairman

Salafi A. Zummo
Dr. Salafi A. Zummo
Dean of Graduate Studies

11/1/15
Date



To my beloved mother

ACKNOWLEDGMENTS

I am truly grateful to Allah, the most merciful, the most gracious for his countless blessings on me. It is He who gave me the wisdom, motivation, determination, strength, courage and patience to complete this work.

I am very grateful to King Fahd University of Petroleum & Minerals (KFUPM) for providing me with the research assistantship to pursue my M.S. program at the Department of Electrical Engineering. Studying at KFUPM has been a professionally rewarding experience for me.

I would like to express my profound gratitude and appreciation to my thesis adviser Dr. Azzedine Zerguine for providing me an opportunity to work on this research topic. This research topic has deepened my understanding of adaptive filtering and has furthered my appreciation of its impact on modern engineering applications. Dr. Zerguine has made elegant contributions to the theory and practice of adaptive filters. It was under his guidance that I developed interest in research.

I would like to thank my thesis committee members Dr. Asrar Ul-Haq Sheikh and Dr. Omar A. Al-Swailem for their assistance and valuable suggestions. It was a pleasure to take two courses with Dr. Asrar. I wish he could offer more graduate courses.

I would also like to take this opportunity to acknowledge two of my teachers at KFUPM: Dr. Tareq Y. Al-Naffouri for motivating us to carry out our research in adaptive filtering and Dr. Samir Al-Ghadban for helping me with MATLAB simulations. I was fortunate to take courses with them.

I would like to thank my friends Umair Bin Mansoor, Ahmed Abdul Quadeer, Muhammad Saqib Sohail, Babar Khan, Muhammad Omer Bin Saeed, Syed Asad, Ahmed Salim, and Syed Ali Aamir Imam for their sincere and kind responses whenever I used to approach them for help during my work on this thesis. I am very thankful to Mohammed E. Eltayeb for being my roommate. I would also like to thank all my friends at KFUPM. They made my stay at the university very pleasant and joyful.

Last but by no means least, I cannot forget to express my deepest gratitude to my mother Noorjahaan, for her patience, support, and encouragement during the many years of my education. For all the sacrifices she has endured to guarantee my education, I dedicate this thesis to her.

Nomenclature

Symbols

J_i	Cost function.
\mathcal{M}	Misadjustment.
μ	Step-size.
\mathbf{u}_i	Regressor vector (a row vector).
v_i	Additive noise.
y_i	Adaptive filter output.
d_i	Desired signal.
e_i	Estimation error signal.
\mathbf{w}_i	Weight vector (a column vector).
$\tilde{\mathbf{w}}_i$	Weight error vector (a column vector).
M	Filter length.
e_{a_i}	a priori estimation error.
e_{p_i}	a posteriori estimation error.
σ_u^2	Regressor variance.
σ_v^2	Noise variance.
\mathbf{R}	Regressor covariance matrix.
$\bar{\mu}_i$	Pseudo-inverse of the regressor vector.
ζ_v^k	k^{th} order moment of v_i .
ζ	Excess-mean-square error.

$\ x\ ^2$	Squared Euclidean norm of x .
$\ x\ _W^2$	Weighted squared Euclidean norm of x .
ρ	Eigenvalue spread.

Operators

$()^*$	Hermitian transpose operation.
$()^T$	Transpose operation.
$\text{Re}(x)$	Real part of x .
\lim	Limit operator.
$\text{Tr}(\mathbf{A})$	Trace of the matrix \mathbf{A} .
$\text{sign}[\mathbf{u}_i]$	Sign function of the regressor vector.
$\text{csgn}[\mathbf{u}_i]$	Complex sign function of the regressor vector.
$g[e_i]$	Some function of the estimation error signal.
$\mathbf{H}[\mathbf{u}_i]$	Some positive-definite Hermitian matrix-valued function of \mathbf{u}_i .
$\mathbf{E}[\]$	Expectation operation.

Abbreviations

LMS	Least mean squares algorithm.
LMF	Least mean fourth algorithm.
SA	Sign algorithm.
SRA	Signed regressor algorithm.
SSA	Sign sign algorithm.
SRLMS	Signed regressor least mean squares algorithm.
SRLMF	Signed regressor least mean fourth algorithm.
FIR	Finite impulse response.

MSE	Mean-square error.
MSD	Mean-square deviation.
EMSE	Excess-mean-square error.
AWGN	Additive white Gaussian noise.
SNR	Signal-to-noise ratio.
ISI	Inter-symbol-interference.
LPC	Linear prediction coding.
i.i.d.	Independent and identically distributed.
MIMO	Multiple-input-multiple-output.
OFDM	Orthogonal frequency division multiplexing.
SM	Spatial modulation.

TABLE OF CONTENTS

NOMENCLATURE	iv
LIST OF TABLES	x
LIST OF FIGURES	xi
ABSTRACT (ENGLISH)	xv
1 INTRODUCTION	1
1.1 Adaptive Filters	2
1.2 Applications of Adaptive Filters	3
1.2.1 Identification	3
1.2.2 Inverse Modelling	4
1.2.3 Prediction	5
1.2.4 Interference Cancellation	6
1.3 Adaptive Filtering Algorithms	7
1.3.1 The LMS Algorithm	7
1.3.2 The LMF Algorithm	8
1.3.3 The SRLMS Algorithm	9
1.4 Thesis Objectives and Organization	10
2 THE SRLMF ALGORITHM	13
2.1 Introduction	13
2.2 Motivation	13

2.3	The SRLMF Algorithm Update Recursion	15
2.4	Computational Load	17
2.5	Conclusion	17
3	STEADY-STATE ANALYSIS OF THE SRLMF ALGORITHM	19
3.1	Introduction	19
3.2	Stationary Data Model	20
3.3	Energy-Conservation Relation	21
3.4	Variance Relation	23
3.5	Mean-Square Analysis of the SRLMF algorithm	25
	3.5.1 Real-Valued Data	25
	3.5.2 Complex-Valued Data	29
3.6	Conclusion	32
4	TRACKING ANALYSIS OF THE SRLMF ALGORITHM	33
4.1	Introduction	33
4.2	Nonstationary Data Model	33
4.3	Tracking Analysis of the SRLMF algorithm	37
	4.3.1 Real-Valued Data	37
	4.3.2 Complex-Valued Data	39
4.4	Conclusion	40
5	TRANSIENT ANALYSIS OF THE SRLMF ALGORITHM	41
5.1	Introduction	41
5.2	Weighted Energy-Conservation Relation	42
5.3	Weighted Variance Relation	44
5.4	Transient Analysis of the SRLMF algorithm	46
5.5	Conclusion	56
6	PERFORMANCE ANALYSIS OF THE SRLMF ALGORITHM	58
6.1	Introduction	58
6.2	Mean-Square Performance of the SRLMF algorithm	59

6.3	Tracking Performance of the SRLMF algorithm	88
6.3.1	Random-Walk Channel	88
6.3.2	Rayleigh Fading Channel	90
6.4	Transient Performance of the SRLMF algorithm	97
6.5	Conclusion	102
7	THESIS CONTRIBUTIONS, CONCLUSIONS AND RECOM-	
	MENDATIONS FOR FUTURE WORK	103
7.1	Thesis Contributions	103
7.2	Conclusions	104
7.3	Recommendations for Future Work	104
	APPENDIX A: Price's Theorem for Complex Sign Function	106
	REFERENCES	108
	VITAE	113

LIST OF TABLES

2.1	Computational load per iteration for LMF and SRLMF algorithms when data is real.	17
2.2	Computational load per iteration for LMF and SRLMF algorithms when data is complex.	17

LIST OF FIGURES

1.1	Identification scenario.	4
1.2	Inverse Modelling scenario.	5
1.3	Prediction scenario.	6
1.4	Interference Cancellation scenario.	6
6.1	Comparison of the MSE learning curves of LMF and SRLMF algorithms in an AWGN environment with SNR=0 dB.	61
6.2	Comparison of the third-tap weight learning curves of LMF and SRLMF algorithms in an AWGN environment with SNR=0 dB.	62
6.3	Comparison of the MSE learning curves of LMF and SRLMF algorithms in an AWGN environment with SNR=10 dB.	63
6.4	Comparison of the third-tap weight learning curves of LMF and SRLMF algorithms in an AWGN environment with SNR=10 dB.	64
6.5	Comparison of the MSE learning curves of LMF and SRLMF algorithms in an AWGN environment with SNR=20 dB.	65
6.6	Comparison of the third-tap weight learning curves of LMF and SRLMF algorithms in an AWGN environment with SNR=20 dB.	66
6.7	Comparison of the MSE learning curves of LMF and SRLMF algorithms when there is a sudden burst in an AWGN environment with SNR=20 dB.	67
6.8	Comparison of the third-tap weight learning curves of LMF and SRLMF algorithms when there is a sudden burst in an AWGN environment with SNR=20 dB.	68

6.9	Comparison of the MSE learning curves of LMF and SRLMF algorithms in uniform noise environment with SNR=0 dB.	70
6.10	Comparison of the third-tap weight learning curves of LMF and SRLMF algorithms in uniform noise environment with SNR=0 dB.	71
6.11	Comparison of the MSE learning curves of LMF and SRLMF algorithms in uniform noise environment with SNR=10 dB.	72
6.12	Comparison of the third-tap weight learning curves of LMF and SRLMF algorithms in uniform noise environment with SNR=10 dB.	73
6.13	Comparison of the MSE learning curves of LMF and SRLMF algorithms in uniform noise environment with SNR=20 dB.	74
6.14	Comparison of the third-tap weight learning curves of LMF and SRLMF algorithms in uniform noise environment with SNR=20 dB.	75
6.15	Comparison of the MSE learning curves of LMF and SRLMF algorithms in Laplacian noise environment with SNR=0 dB.	76
6.16	Comparison of the third-tap weight learning curves of LMF and SRLMF algorithms in Laplacian noise environment with SNR=0 dB.	77
6.17	Comparison of the MSE learning curves of LMF and SRLMF algorithms in Laplacian noise environment with SNR=10 dB.	78
6.18	Comparison of the third-tap weight learning curves of LMF and SRLMF algorithms in Laplacian noise environment with SNR=10 dB.	79
6.19	Comparison of the MSE learning curves of LMF and SRLMF algorithms in Laplacian noise environment with SNR=20 dB.	80
6.20	Comparison of the third-tap weight learning curves of LMF and SRLMF algorithms in Laplacian noise environment with SNR=20 dB.	81
6.21	Comparison of the MSE learning curves of the SRLMF algorithm in Gaussian, uniform and Laplacian noise environments with SNR=10 dB.	82

6.22	Comparison of the third-tap weight learning curves of the SRLMF algorithm in Gaussian, uniform and Laplacian noise environments with SNR=10 dB.	83
6.23	Theoretical and simulated MSE of the SRLMF algorithm using white Gaussian regressors with shift structure with SNR=30 dB. .	85
6.24	Theoretical and simulated MSE of the SRLMF algorithm using correlated Gaussian regressors with shift structure with SNR=30 dB.	86
6.25	Theoretical and simulated MSE of the SRLMF algorithm using Gaussian regressors with an eigenvalue spread=5 without shift structure with SNR=30 dB.	87
6.26	Theoretical and simulated MSE of the SRLMF algorithm for a random-walk channel as a function of the step-size with SNR=30 dB.	89
6.27	Theoretical and simulated MSE of the SRLMF algorithm for a single-path Rayleigh fading channel as a function of the step-size with SNR=30 dB.	91
6.28	Theoretical and simulated MSE of the SRLMF algorithm for a single-path Rayleigh fading channel as a function of the Doppler frequency with SNR=30 dB and step-size=0.01.	92
6.29	Theoretical and simulated MSE of the SRLMF algorithm for a multipath Rayleigh fading channel as a function of the step-size with SNR=30 dB.	94
6.30	Theoretical and simulated MSE of the SRLMF algorithm for a multipath Rayleigh fading channel as a function of the Doppler frequency with SNR=30 dB and step-size=0.01.	95
6.31	A typical trajectory of the amplitude of the first Rayleigh fading ray and its estimate with SNR=20 dB and step-size=0.01.	96
6.32	Theoretical and simulated MSD (top) and MSE (bottom) learning curves of the SRLMF algorithm using white Gaussian regressors with SNR=50 dB and step-size=0.01.	98

6.33	Theoretical and simulated MSD (top) and MSE (bottom) learning curves of the SRLMF algorithm using white non-Gaussian regressors with SNR=50 dB and step-size=0.01.	99
6.34	Theoretical and simulated MSD (top) and MSE (bottom) learning curves of the SRLMF algorithm using Gaussian regressors with an eigenvalue spread=5, SNR=50 dB and step-size=0.01.	100
6.35	Theoretical and simulated MSD (top) and MSE (bottom) learning curves of the SRLMF algorithm using non-Gaussian regressors with an eigenvalue spread=5, SNR=50 dB and step-size=0.01.	101

THESIS ABSTRACT

NAME: Mohammed Mujahid Ulla Faiz
TITLE OF STUDY: The Signed Regressor Least Mean Fourth (SRLMF)
Adaptive Algorithm
MAJOR FIELD: Electrical Engineering
DATE OF DEGREE: November 2009

In this thesis, a novel algorithm, called the signed regressor least mean fourth (SRLMF) adaptive algorithm, that reduces the computational cost and complexity while maintaining good performance is presented. Expressions are derived for the steady-state excess-mean-square error (EMSE) of the SRLMF algorithm in a stationary environment. Also, expressions are obtained for the tracking EMSE of the SRLMF algorithm in a nonstationary environment. An optimum value of the step-size μ is also derived. Moreover, the weighted variance relation has been extended in order to derive expressions for the mean-square error (MSE) and the mean-square deviation (MSD) of the proposed algorithm during the transient phase. Computer simulations are carried out to corroborate the theoretical findings. It is shown that there is a good match between the theoretical and simulated

results. It is also shown that the SRLMF algorithm has no performance degradation when compared with the least mean fourth (LMF) algorithm. The results in this study emphasize the usefulness of this algorithm in applications requiring reduced implementation costs for which the LMF algorithm is too complex.

CHAPTER 1

INTRODUCTION

The subject of adaptive filters constitutes an important part of the statistical signal processing. When the filter is required to operate in a stationary environment, where the signal statistics (i.e., mean and correlation) are known, the use of Wiener filter provides a solution, which is optimum in the mean-square error sense. However, when the filter is required to operate in a nonstationary environment, where the signal statistics are unknown, the use of an adaptive filter offers an attractive solution to the problem. In a nonstationary environment, adaptive filters provide significant improvement in performance over fixed filters, which are designed by conventional methods. Therefore, adaptive filters have been successfully applied in many diverse fields such as biomedicine, communications, control, radar, sonar, seismology, just to name a few [1].

1.1 Adaptive Filters

Adaptive filters have the ability of adapting their characteristics in order to achieve the desired objectives. Adaptation is accomplished automatically by adjusting the filter coefficients in accordance with the input data. Thus making the adaptive filter nonlinear. The performance of an adaptive filter is evaluated in terms of its transient behavior and its steady-state behavior. The former provides information about how fast a filter learns, while the latter provides information about how well a filter learns. Such performance analysis are usually challenging since adaptive filters are, by design, time-variant, nonlinear, and stochastic systems [2]. For this reason, it has been common in the literature to study different adaptive schemes separately due to the differences that exist in their update equations. An adaptive filter is said to be linear if its input-output map obeys the principle of superposition whenever, at any particular instant of time, the filter's parameters are all fixed [1].

The adaptive filter usually relies on a recursive algorithm for its operation, which makes it possible for the filter to perform satisfactorily in an environment where complete knowledge of the relevant signal characteristics is not available. The algorithm starts from some predetermined set of initial conditions, representing complete ignorance about the environment. In a stationary environment, after successive iterations, the algorithm tries to converge to the optimum Wiener solution in some statistical sense. In a nonstationary environment, the algorithm offers a tracking ability, in that it can track time variations in the statistics of the

input data, provided that the variations are sufficiently slow [1].

1.2 Applications of Adaptive Filters

Adaptive filters have been successfully applied in many diverse fields such as biomedicine, communications, control, radar, sonar, seismology, just to name a few. Although these applications are quite different in nature, nevertheless, they have one basic common feature: An input signal and a desired response are used to compute an estimation error, which is in turn used to control the values of a set of adjustable filter coefficients. Depending on the filter structure employed, the adjustable coefficients may take the form of tap weights, reflection coefficients, or rotation parameters. However, the main difference among the various applications arises in the manner in which the desired response is extracted. On this basis, adaptive filters are classified into the following four basic classes [1].

1.2.1 Identification

In this class of applications, an adaptive filter is used to provide a linear model that represents the best fit to an unknown plant as shown in Fig. 1.1. Both the adaptive filter and the unknown plant are driven by the same input \mathbf{u}_i . v_i is the additive noise. The adaptive filter output y_i is subtracted from the unknown plant output d_i . The resulting error signal e_i is used to update the adaptive filter coefficients. The unknown plant to be identified can be either stationary or time varying. This class of adaptive filters are used in system identification and layered

earth modeling [1].

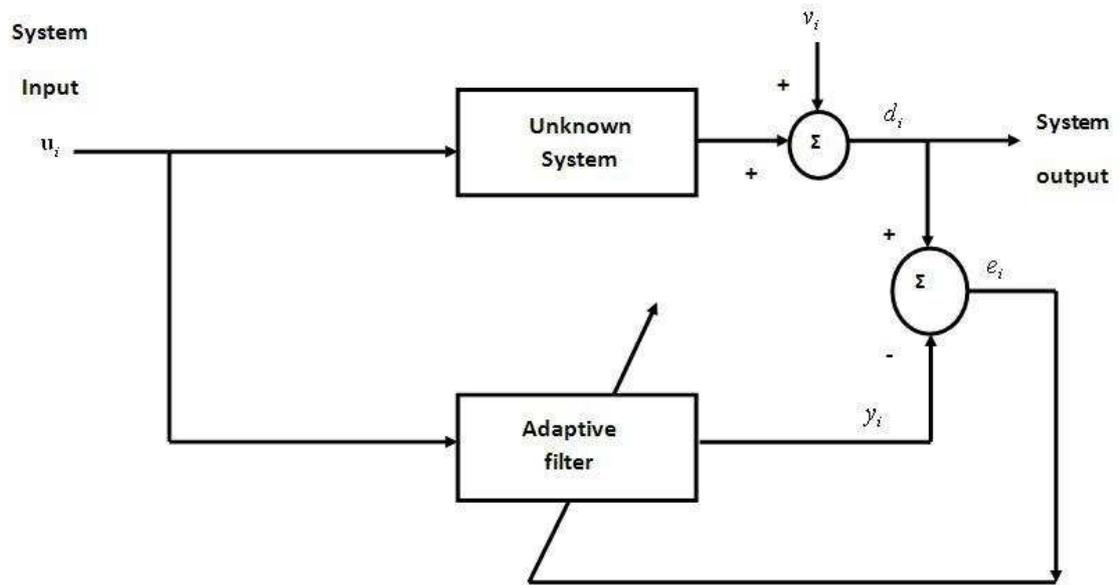


Figure 1.1: Identification scenario.

1.2.2 Inverse Modelling

In this class of applications, the adaptive filter is used to provide an inverse model that represents the best fit of an unknown plant as shown in Fig. 1.2. Thus, at convergence, the adaptive filter has a best transfer function equal to the reciprocal of the unknown plant's transfer function, such that the combination of the two constitutes an ideal transmission medium. A delayed version of the unknown plant input serves as the desired response d_i for the adaptive filter. In some applications, the unknown plant input is used without delay as the desired response. This class of adaptive filters are used in equalization to mitigate the effect of inter-symbol-interference (ISI) in digital receivers [3].

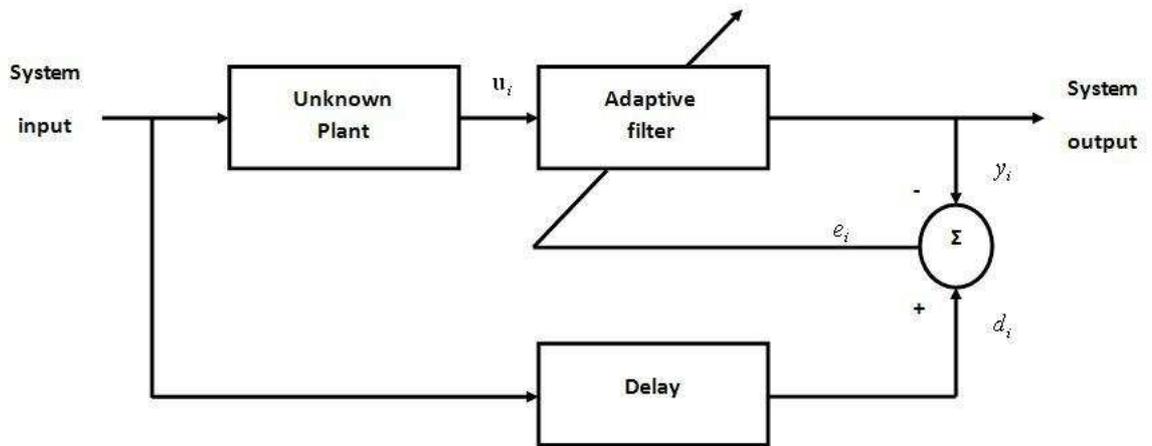


Figure 1.2: Inverse Modelling scenario.

1.2.3 Prediction

In this class of applications, the adaptive filter is used to provide the best prediction of the present value of a random signal as shown in Fig. 1.3. The present value of the random signal serves as the desired response d_i for the adaptive filter. Past values of the random signal supply the input \mathbf{u}_i to the adaptive filter. Depending on the application of interest, the adaptive filter output or the estimation (prediction) error may serve as the system output. In the former case, the system operates as a predictor, whereas in the latter case, it operates as a prediction-error filter. This class of adaptive filters are used in linear prediction coding (LPC) of speech [4] and spectrum analysis [1].

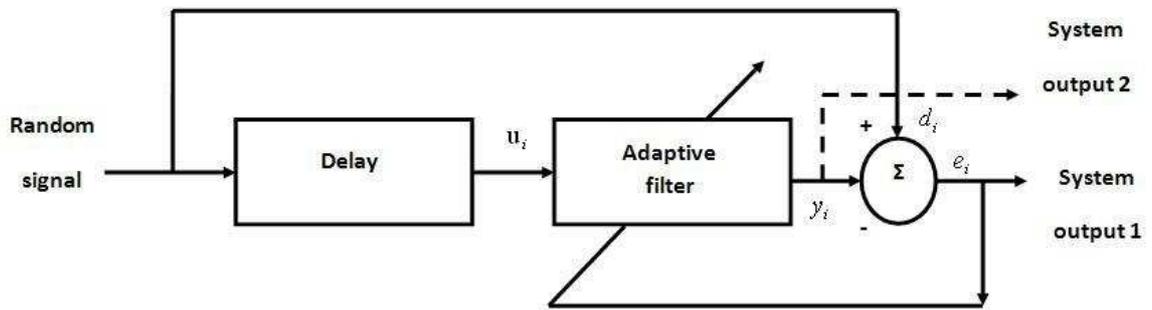


Figure 1.3: Prediction scenario.

1.2.4 Interference Cancellation

Finally, in this class of applications, the adaptive filter is used to cancel the unknown interference contained in a primary signal as shown in Fig. 1.4. The primary signal serves as the desired response d_i for the adaptive filter. A reference (auxiliary) signal derived from a sensor is applied as the input u_i to the adaptive filter. This class of adaptive filters are used in beamforming [1] and noise cancellation [5].

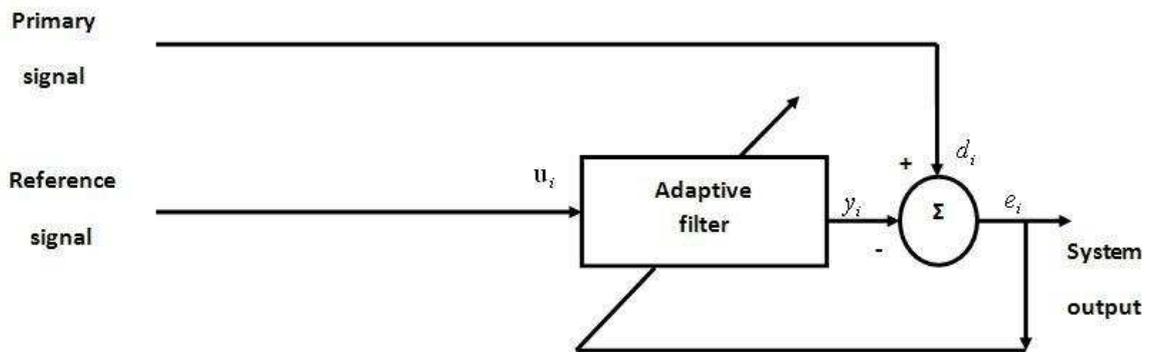


Figure 1.4: Interference Cancellation scenario.

1.3 Adaptive Filtering Algorithms

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

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \mathbf{u}_i^* g[e_i], \quad i \geq 0, \quad (1.1)$$

where μ is called the step-size parameter since it affects how small or how large the correction term is, \mathbf{u}_i is the input sequence, $g[e_i]$ denotes some function of the estimation error signal, and $e_i = d_i - \mathbf{u}_i \mathbf{w}_{i-1}$. Some of the well known algorithms are presented below.

1.3.1 The LMS Algorithm

If $g[e_i] = e_i$ in (1.1), the least mean squares (LMS) algorithm [6] is obtained

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \mathbf{u}_i^* e_i, \quad i \geq 0. \quad (1.2)$$

If we assume that the data is real-valued then (1.2) can be written as

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \mathbf{u}_i^T e_i, \quad i \geq 0. \quad (1.3)$$

The LMS algorithm allows error minimization in the mean square sense. It is one of the most commonly used algorithm in adaptive filtering as it is relatively simple to implement and requires less number of computations. It is capable of achieving satisfactory performance under the right conditions. Its major limitation is relatively slower rate of convergence for the case of highly correlated data.

In a non-stationary environment, the orientation of the error-performance surface varies continuously with time. In this case, the LMS algorithm has the added task of continually tracking the bottom of the error performance surface. Indeed, tracking will occur provided that the input data varies slowly compared to the learning rate of the LMS algorithm [1].

1.3.2 The LMF Algorithm

If $g[e_i] = e_i^3$ in (1.1), the least mean fourth (LMF) algorithm [7] is obtained

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \mathbf{u}_i^* e_i^3, \quad i \geq 0. \quad (1.4)$$

If we assume that the data is real-valued then (1.4) can be written as

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \mathbf{u}_i^T e_i^3, \quad i \geq 0. \quad (1.5)$$

Adaptive algorithms based on higher order moments of the error signal have been shown to perform better mean square estimation than the well known LMS algorithm in some important applications. The LMF is one such algorithm. It

allows error minimization in the mean fourth sense and can be viewed as an extension of the Widrow-Hoff LMS algorithm. It has been shown that the LMF algorithm can outperform the LMS algorithm for Gaussian, uniform, and sinusoidal noise distributions [7]–[8]. In such a case, the LMF algorithm can lead to considerably smaller excess-mean-square error (EMSE) for the same speed of convergence.

Strictly speaking, it has been shown in [9] that the LMF algorithm can never be mean-square stable for any step-size when the regressor sequence is not strictly bounded. For input distributions with infinite support, even for the Gaussian distribution, the LMF algorithm always has a nonzero probability of divergence, no matter how small the step-size is chosen. Since practically all actual regressor sequences are bounded, this means that the algorithm is very sensitive to larger values of the regressor sequence, even if they occur very rarely, as in the case of Gaussian regressors.

1.3.3 The SRLMS Algorithm

New algorithms that make use of the signum (polarity) of either the estimation error or the input data, or both, have been derived from the LMS algorithm for the simplicity of implementation, enabling a significant reduction in computing time, particularly the time required for “multiplications” [10]–[12]. It should be noted that clipping either the estimation error or the input data, or both, basically reduces the number of multiplications necessary at each algorithm iteration.

The algorithm based on clipping of the input data is known as the signed regressor algorithm (SRA) [13]–[16]. The SRA or the signed regressor least mean squares (SRLMS) algorithm is given by

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \text{csgn}[\mathbf{u}_i]^* e_i, \quad i \geq 0. \quad (1.6)$$

If we assume that the data is real-valued then (1.6) can be written as

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \text{sign}[\mathbf{u}_i]^T e_i, \quad i \geq 0. \quad (1.7)$$

1.4 Thesis Objectives and Organization

In this thesis, the signed regressor least mean fourth (SRLMF) algorithm is proposed. On the basis of the above discussion, the update recursion for the SRLMF algorithm can be written as follows:

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \text{csgn}[\mathbf{u}_i]^* e_i^3, \quad i \geq 0. \quad (1.8)$$

If we assume that the data is real-valued then (1.8) can be written as

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \text{sign}[\mathbf{u}_i]^T e_i^3, \quad i \geq 0. \quad (1.9)$$

The objectives of this thesis are as follows: First, to derive the SRLMF algorithm update recursion and to estimate the computational load per iteration

of the proposed algorithm. Second, to study the steady-state performance of the SRLMF algorithm and to derive expressions for the steady-state EMSE in a stationary environment. Third, to study the tracking performance of the SRLMF algorithm and to derive expressions for the tracking EMSE in a nonstationary environment. Fourth, to study the transient performance of the SRLMF algorithm and to derive expressions for the mean-square error (MSE) and the mean-square deviation (MSD) during the transient phase. Finally, to support the analytical results by computer simulations.

This thesis is organized as follows: In Chapter 2, the motivation behind this work is presented and the update recursion of the SRLMF algorithm is derived. Also in Chapter 2, a comparison between the computational load of the SRLMF algorithm and the LMF algorithm is presented.

In Chapter 3, expressions for the steady-state EMSE in a stationary environment are derived. The framework used in this study, and pursued further in Chapters 4 and 5, relies on energy-conservation arguments [2].

In Chapter 4, expressions for the tracking EMSE in a nonstationary environment are derived. An optimum value of the step-size μ is also derived. The presentation in Chapter 4 will reveal that the tracking results can be obtained by inspection from the mean-square results as there are only minor differences.

Transient analysis is more conveniently performed by relying on a weighted energy-conservation relation, as opposed to the unweighted version that was employed in Chapters 3 and 4. In Chapter 5, the weighted variance relation has been

extended in order to derive expressions for the MSE and the MSD of the SRLMF algorithm during the transient phase.

In Chapter 6, computer simulations are carried out to corroborate the theoretical findings, where it is shown that the theoretical and simulated results are in good agreement. Moreover, the results show that both the SRLMF algorithm and the LMF algorithm have a similar performance for the same steady-state EMSE.

Finally, conclusions, contributions and recommendations for future work are presented in Chapter 7.

CHAPTER 2

THE SRLMF ALGORITHM

2.1 Introduction

In this chapter, the motivation behind this work is presented and the update recursion of the SRLMF algorithm is derived. Also in Chapter 2, a comparison between the computational load of the SRLMF algorithm and the LMF algorithm is presented.

2.2 Motivation

Reduction in complexity of the least mean square (LMS) algorithm has always received attention in the area of adaptive filtering [17]–[19]. This reduction is usually done by clipping either the estimation error or the input data, or both to reduce the number of multiplications necessary at each algorithm iteration. The algorithm based on clipping of the estimation error is known as the signed error or more commonly the sign algorithm (SA) [20]–[24], the algorithm based on clipping

of the input data is known as the signed regressor algorithm (SRA) [13]–[16], and the algorithm based on clipping of both the estimation error and the input data is known as the sign sign algorithm (SSA) [25]–[26]. These algorithms result in a performance loss when compared with the conventional LMS algorithm [13]–[14]. However, significant reduction in computational cost and simplified hardware implementation can justify this poor performance in applications requiring reduced implementation costs [10]–[11].

The behavior of the SRA algorithm depends on the input data. It is shown in [15] that for some inputs the LMS algorithm is stable while the SRA algorithm is unstable. This is a drawback of the SRA algorithm when compared with the SA algorithm since the latter is more stable than the LMS algorithm [11], [20]. The SRA algorithm is always stable when the input data is Gaussian as in the case of speech processing. Also, the performance of the SRA algorithm is superior to that of the SA algorithm for Gaussian input data. It is shown in [14] that the SRA algorithm is much faster than the SA algorithm in achieving the desired steady-state mean-square error for white Gaussian data. Theoretical studies of the SRA algorithm with correlated Gaussian data in both stationary and nonstationary environments are found in [16].

The convergence rate and the steady-state mean-square error of the SRA algorithm is only slightly inferior to those of the LMS algorithm for the same parameter setting. In [14], the convergence rate of the SRA algorithm is compared with that of the LMS algorithm to show that the SRA algorithm converges slower than the

LMS algorithm by a factor of $2/\pi$ for the same steady-state mean-square error.

It is shown in [27] that the SRA algorithm exhibits significantly higher robustness against the impulse noise than the LMS algorithm.

The above mentioned advantages motivates us to analyze and design the proposed signed regressor least mean fourth (SRLMF) adaptive algorithm.

2.3 The SRLMF Algorithm Update Recursion

The SRLMF algorithm is based on clipping of the regression vector \mathbf{u}_i (row vector). Consider now the adaptive filter, which updates its coefficients according to the following recursion [2]:

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \mathbf{H}[\mathbf{u}_i] \mathbf{u}_i^* g[e_i], \quad i \geq 0, \quad (2.1)$$

where \mathbf{w}_i (column vector) is the updated weight vector at time i , μ is the step-size, $\mathbf{H}[\mathbf{u}_i]$ is some positive-definite Hermitian matrix-valued function of \mathbf{u}_i , $g[e_i]$ denotes some function of the estimation error signal given by

$$e_i = d_i - \mathbf{u}_i \mathbf{w}_{i-1}, \quad (2.2)$$

where d_i is the desired signal. When the data is real-valued, the general update form in (2.1) becomes

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \mathbf{H}[\mathbf{u}_i] \mathbf{u}_i^T g[e_i], \quad i \geq 0. \quad (2.3)$$

The error function for LMF is $g[e_i] = e_i^3$. Therefore, (2.3) becomes

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \mathbf{H}[\mathbf{u}_i] \mathbf{u}_i^T e_i^3, \quad i \geq 0. \quad (2.4)$$

Now if

$$\mathbf{H}[\mathbf{u}_i] = \text{diag} \left\{ \frac{1}{|\mathbf{u}_{i_1}|}, \frac{1}{|\mathbf{u}_{i_2}|}, \dots, \frac{1}{|\mathbf{u}_{i_M}|} \right\}, \quad (2.5)$$

then the update form in (2.4) reduces to

$$\begin{aligned} \mathbf{w}_i &= \mathbf{w}_{i-1} + \mu \text{diag} \left\{ \frac{1}{|\mathbf{u}_{i_1}|}, \frac{1}{|\mathbf{u}_{i_2}|}, \dots, \frac{1}{|\mathbf{u}_{i_M}|} \right\} \mathbf{u}_i^T e_i^3, \quad i \geq 0, \\ &= \mathbf{w}_{i-1} + \mu \text{sign}[\mathbf{u}_i]^T e_i^3, \quad i \geq 0, \end{aligned} \quad (2.6)$$

where M is the filter length. The SRLMF algorithm update recursion in (2.6) can be regarded as a special case of the general update form in (2.4) for some matrix data nonlinearity that is implicitly defined by the following relation:

$$\text{sign}[\mathbf{u}_i]^T = \mathbf{H}[\mathbf{u}_i] \mathbf{u}_i^T. \quad (2.7)$$

For complex-valued data, the update recursion in (2.6) becomes

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \text{csgn}[\mathbf{u}_i]^* e_i^3, \quad i \geq 0. \quad (2.8)$$

2.4 Computational Load

Tables 2.1 and 2.2 present the estimated computational load per iteration for both real- and complex-valued data in terms of the number of real additions, real multiplications, real divisions, and comparisons with zero (or sign evaluations). We know that one complex multiplication requires four real multiplications and two real additions, while one complex addition requires two real additions.

Table 2.1: Computational load per iteration for LMF and SRLMF algorithms when data is real.

Algorithm	+	\times	/	sign
LMF	$2M$	$2M + 3$		
SRLMF	$2M$	$2M + 2$		1

Table 2.2: Computational load per iteration for LMF and SRLMF algorithms when data is complex.

Algorithm	+	\times	/	sign
LMF	$8M + 1$	$8M + 5$		
SRLMF	$6M + 1$	$6M + 3$		2

2.5 Conclusion

In this chapter, we have presented the motivation behind this work and the derivation for the update recursion of the SRLMF algorithm. It should be noted that replacement of the regressor by its sign in the update recursion limits the range of search directions that are followed by the SRLMF algorithm and, therefore, performance degradation in terms of convergence speed (and even possibly divergence) can occur relative to a conventional LMF algorithm. Also in Chapter 2, a comparison between the estimated computational load per iteration of the SRLMF

algorithm and the LMF algorithm is presented for both real- and complex-valued data. It should be noted that clipping the input data basically reduces the number of multiplications necessary at each algorithm iteration.

CHAPTER 3

STEADY-STATE ANALYSIS OF THE SRLMF ALGORITHM

3.1 Introduction

In this chapter, expressions for the steady-state EMSE of the SRLMF algorithm in a stationary environment are derived. The framework used in this chapter, and pursued further in Chapters 4 and 5, relies on energy conservation arguments [2].

Steady-state behavior relates to determining the steady-state values of $E[||\tilde{\mathbf{w}}_i||^2]$, $E[|e_{a_i}|^2]$, and $E[|e_i|^2]$, where $\tilde{\mathbf{w}}_i$ is the weight error vector defined by $\tilde{\mathbf{w}}_i = \mathbf{w}^o - \mathbf{w}_i$, e_{a_i} is the a priori estimation error defined by $e_{a_i} = \mathbf{u}_i \tilde{\mathbf{w}}_{i-1}$, and e_i is the estimation error defined in (2.2).

An adaptive filter is said to operate in steady-state if it holds that

$$\lim_{i \rightarrow \infty} \mathbf{E}[\tilde{\mathbf{w}}_i] = \lim_{i \rightarrow \infty} \mathbf{E}[\tilde{\mathbf{w}}_{i-1}]. \quad (3.1)$$

$$\lim_{i \rightarrow \infty} \mathbf{E}[\tilde{\mathbf{w}}_i \tilde{\mathbf{w}}_i^*] = \lim_{i \rightarrow \infty} \mathbf{E}[\tilde{\mathbf{w}}_{i-1} \tilde{\mathbf{w}}_{i-1}^*] = \mathbf{C}. \quad (3.2)$$

That is, the mean and covariance matrix of the weight error vector tend to some finite constant values. In particular, it follows that the following condition holds:

$$\lim_{i \rightarrow \infty} \mathbf{E}[|\tilde{\mathbf{w}}_i|^2] = \lim_{i \rightarrow \infty} \mathbf{E}[|\tilde{\mathbf{w}}_{i-1}|^2] = \text{Tr}(\mathbf{C}). \quad (3.3)$$

3.2 Stationary Data Model

We shall assume that the data $\{d_i, \mathbf{u}_i\}$ satisfy the following conditions of the stationary data model [2]:

- A.1** There exists an optimal weight vector \mathbf{w}^o such that $d_i = \mathbf{u}_i \mathbf{w}^o + v_i$.
- A.2** The noise sequence v_i is independent and identically distributed (i.i.d.) with variance $\sigma_v^2 = \mathbf{E}[|v_i|^2]$ and is independent of \mathbf{u}_j for all i, j .
- A.3** The initial condition \mathbf{w}_{-1} is independent of the zero mean random variables $\{d_i, \mathbf{u}_i, v_i\}$.
- A.4** The regressor covariance matrix is $\mathbf{R} = \mathbf{E}[\mathbf{u}_i^* \mathbf{u}_i] > \mathbf{0}$.

3.3 Energy-Conservation Relation

Let us consider the adaptive filter updates of the generic form given below:

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \mathbf{H}[\mathbf{u}_i] \mathbf{u}_i^* g[e_i], \quad i \geq 0. \quad (3.4)$$

Subtracting both sides of (3.4) from \mathbf{w}^o we get

$$\tilde{\mathbf{w}}_i = \tilde{\mathbf{w}}_{i-1} - \mu \mathbf{H}[\mathbf{u}_i] \mathbf{u}_i^* g[e_i]. \quad (3.5)$$

If we multiply both sides of (3.5) by \mathbf{u}_i from the left we get

$$e_{p_i} = e_{a_i} - \mu \|\mathbf{u}_i\|_{\mathbf{H}}^2 g[e_i], \quad (3.6)$$

where e_{p_i} is the a posteriori estimation error defined by $e_{p_i} = \mathbf{u}_i \tilde{\mathbf{w}}_i$, and $\|\mathbf{u}_i\|_{\mathbf{H}}^2 = \mathbf{u}_i \mathbf{H}[\mathbf{u}_i] \mathbf{u}_i^*$. Two cases can be considered here, that is,

Case 1: $\|\mathbf{u}_i\|_{\mathbf{H}}^2 = 0$.

In this case, $\tilde{\mathbf{w}}_i = \tilde{\mathbf{w}}_{i-1}$ and $e_{a_i} = e_{p_i}$ so that $\|\tilde{\mathbf{w}}_i\|^2 = \|\tilde{\mathbf{w}}_{i-1}\|^2$ and $|e_{a_i}|^2 = |e_{p_i}|^2$.

Case 2: $\|\mathbf{u}_i\|_{\mathbf{H}}^2 \neq 0$.

In this case, we use (3.6) to solve for $g[e_i]$,

$$g[e_i] = \frac{1}{\mu \|\mathbf{u}_i\|_{\mathbf{H}}^2} (e_{a_i} - e_{p_i}), \quad (3.7)$$

and substitute into (3.5) to obtain

$$\tilde{\mathbf{w}}_i = \tilde{\mathbf{w}}_{i-1} - \frac{\mathbf{H}[\mathbf{u}_i]\mathbf{u}_i^*}{\|\mathbf{u}_i\|_{\mathbb{H}}^2}(e_{a_i} - e_{p_i}). \quad (3.8)$$

Expression (3.8) can be rearranged as

$$\tilde{\mathbf{w}}_i + \frac{\mathbf{H}[\mathbf{u}_i]\mathbf{u}_i^*}{\|\mathbf{u}_i\|_{\mathbb{H}}^2}e_{a_i} = \tilde{\mathbf{w}}_{i-1} + \frac{\mathbf{H}[\mathbf{u}_i]\mathbf{u}_i^*}{\|\mathbf{u}_i\|_{\mathbb{H}}^2}e_{p_i}. \quad (3.9)$$

By evaluating the energies (i.e, squared Euclidean norms) of both sides of (3.9)

we find,

$$\left\| \tilde{\mathbf{w}}_i + \frac{\mathbf{H}[\mathbf{u}_i]\mathbf{u}_i^*}{\|\mathbf{u}_i\|_{\mathbb{H}}^2}e_{a_i} \right\|^2 = \left\| \tilde{\mathbf{w}}_{i-1} + \frac{\mathbf{H}[\mathbf{u}_i]\mathbf{u}_i^*}{\|\mathbf{u}_i\|_{\mathbb{H}}^2}e_{p_i} \right\|^2. \quad (3.10)$$

After a straightforward calculation, the following energy-conservation results [2]:

$$\|\tilde{\mathbf{w}}_i\|^2 + \frac{1}{\|\mathbf{u}_i\|_{\mathbb{H}}^2}|e_{a_i}|^2 = \|\tilde{\mathbf{w}}_{i-1}\|^2 + \frac{1}{\|\mathbf{u}_i\|_{\mathbb{H}}^2}|e_{p_i}|^2. \quad (3.11)$$

The energy-conservation relation in (3.11) can be further simplified to look like

$$\|\tilde{\mathbf{w}}_i\|^2 + \bar{\mu}_i|e_{a_i}|^2 = \|\tilde{\mathbf{w}}_{i-1}\|^2 + \bar{\mu}_i|e_{p_i}|^2, \quad (3.12)$$

where

$$\bar{\mu}_i = \begin{cases} \frac{1}{\|\mathbf{u}_i\|_{\mathbb{H}}^2}, & \text{if } \mathbf{u}_i \neq 0, \\ 0, & \text{otherwise.} \end{cases} \quad (3.13)$$

3.4 Variance Relation

By taking expectation on both sides of (3.12) we get

$$\mathbb{E}[|\tilde{\mathbf{w}}_i|^2] + \mathbb{E}[\bar{\mu}_i |e_{a_i}|^2] = \mathbb{E}[|\tilde{\mathbf{w}}_{i-1}|^2] + \mathbb{E}[\bar{\mu}_i |e_{p_i}|^2]. \quad (3.14)$$

Taking the limit as $i \rightarrow \infty$ and using the steady-state condition (3.3), we obtain

$$\lim_{i \rightarrow \infty} \mathbb{E}[\bar{\mu}_i |e_{a_i}|^2] = \lim_{i \rightarrow \infty} \mathbb{E}[\bar{\mu}_i |e_{p_i}|^2]. \quad (3.15)$$

Substituting (3.6) into (3.15) we get

$$\lim_{i \rightarrow \infty} \mathbb{E}[\bar{\mu}_i |e_{a_i}|^2] = \lim_{i \rightarrow \infty} \mathbb{E}[\bar{\mu}_i |e_{a_i} - \mu \|\mathbf{u}_i\|_{\mathbb{H}}^2 \mathbf{g}[e_i]|^2]. \quad (3.16)$$

This relation can be simplified by expanding the term on the right-hand side as follows (the argument of \mathbf{g} is dropped for compactness of notation):

$$\begin{aligned} \bar{\mu}_i |e_{a_i} - \mu \|\mathbf{u}_i\|_{\mathbb{H}}^2 \mathbf{g}|^2 &= \bar{\mu}_i |e_{a_i}|^2 + \mu^2 \bar{\mu}_i \|\mathbf{u}_i\|_{\mathbb{H}}^4 |\mathbf{g}|^2 - \mu \bar{\mu}_i \|\mathbf{u}_i\|_{\mathbb{H}}^2 \mathbf{g} e_{a_i}^* \\ &\quad - \mu \bar{\mu}_i \|\mathbf{u}_i\|_{\mathbb{H}}^2 \mathbf{g}^* e_{a_i}. \end{aligned} \quad (3.17)$$

From (3.13) it is obvious that the product $\bar{\mu}_i \|\mathbf{u}_i\|_{\mathbb{H}}^2$ is unity for all \mathbf{u}_i except for the single event $\mathbf{u}_i = 0$, we find that

$$\bar{\mu}_i |e_{a_i} - \mu \|\mathbf{u}_i\|_{\mathbb{H}}^2 \mathbf{g}|^2 = \bar{\mu}_i |e_{a_i}|^2 + \mu^2 \|\mathbf{u}_i\|_{\mathbb{H}}^2 |\mathbf{g}|^2 - \mu e_{a_i} \mathbf{g}^* - \mu e_{a_i}^* \mathbf{g}. \quad (3.18)$$

Taking expectations of both sides of (3.18) we obtain

$$\begin{aligned} \mathbf{E} [\bar{\mu}_i |e_{a_i} - \mu \|\mathbf{u}_i\|_{\mathbf{H}}^2 |g|^2] &= \mathbf{E} [\bar{\mu}_i |e_{a_i}|^2] + \mu^2 \mathbf{E} [\|\mathbf{u}_i\|_{\mathbf{H}}^2 |g|^2] - \mu \mathbf{E} [e_{a_i} g^*] \\ &\quad - \mu \mathbf{E} [e_{a_i}^* g]. \end{aligned} \quad (3.19)$$

Substituting this result into (3.16) leads to

$$\lim_{i \rightarrow \infty} \mu \mathbf{E} [\|\mathbf{u}_i\|_{\mathbf{H}}^2 |g|^2] = \lim_{i \rightarrow \infty} \mathbf{E} [e_{a_i} g^* + e_{a_i}^* g], \quad (3.20)$$

which can be written as [2]:

$$\lim_{i \rightarrow \infty} \mu \mathbf{E} [\|\mathbf{u}_i\|_{\mathbf{H}}^2 |g[e_i]|^2] = \lim_{i \rightarrow \infty} 2 \operatorname{Re}(\mathbf{E} [e_{a_i}^* g[e_i]]). \quad (3.21)$$

The variance relation in (3.21) holds for any adaptive filter of the form (3.4), and for any data $\{d_i, \mathbf{u}_i\}$, assuming filter operation in steady-state.

For real-valued data, this variance relation becomes

$$\lim_{i \rightarrow \infty} \mu \mathbf{E} [\|\mathbf{u}_i\|_{\mathbf{H}}^2 g^2[e_i]] = \lim_{i \rightarrow \infty} 2 \mathbf{E} [e_{a_i} g[e_i]]. \quad (3.22)$$

3.5 Mean-Square Analysis of the SRLMF algorithm

Let us distinguish between real- and complex-valued data as the definition of the sign function is different in both cases. However, the final expression for the EMSE turns out to be identical except for a scaling factor.

3.5.1 Real-Valued Data

Since

$$e_i = e_{a_i} + v_i, \quad (3.23)$$

therefore, when the data is real-valued, $g[e_i]$ becomes

$$\begin{aligned} g[e_i] &= e_i^3, \\ &= (e_{a_i} + v_i)[e_{a_i}^2 + v_i^2 + 2e_{a_i}v_i]. \end{aligned} \quad (3.24)$$

By using the fact that e_{a_i} and v_i are independent, we reach at the following expression for the term $\mathbf{E}[e_{a_i}g[e_i]]$:

$$\mathbf{E}[e_{a_i}g[e_i]] = 3\sigma_v^2\mathbf{E}[e_{a_i}^2] + \mathbf{E}[e_{a_i}^4]. \quad (3.25)$$

If we ignore third- and higher-order terms of e_{a_i} , then (3.25) becomes

$$\begin{aligned} \mathbb{E}[e_{a_i}g[e_i]] &\approx 3\sigma_v^2\mathbb{E}[e_{a_i}^2], \\ &\approx a\mathbb{E}[e_{a_i}^2], \end{aligned} \quad (3.26)$$

where $a = 3\sigma_v^2$. Ignoring higher order terms of e_{a_i} is reasonable since the error e_{a_i} becomes small in steady-state. To get more accurate expression for the steady-state EMSE we may consider the higher-order terms of e_{a_i} .

To evaluate the term $\mathbb{E}[\|\mathbf{u}_i\|_{\mathbb{H}}^2g^2[e_i]]$, we start by noting that

$$\begin{aligned} g^2[e_i] &= e_i^6, \\ &= e_{a_i}^6 + 6e_{a_i}^5v_i + 6e_{a_i}v_i^5 + 15e_{a_i}^4v_i^2 + 15e_{a_i}^2v_i^4 + 20e_{a_i}^3v_i^3 + v_i^6. \end{aligned} \quad (3.27)$$

If we multiply (3.27) by $\|\mathbf{u}_i\|_{\mathbb{H}}^2$ from the left, use the fact that v_i is independent of both \mathbf{u}_i and e_{a_i} , and if we again ignore third- and higher-order terms of e_{a_i} , we obtain

$$\begin{aligned} \mathbb{E}[\|\mathbf{u}_i\|_{\mathbb{H}}^2g^2[e_i]] &\approx 6\mathbb{E}[\|\mathbf{u}_i\|_{\mathbb{H}}^2e_{a_i}v_i^5] + 15\mathbb{E}[\|\mathbf{u}_i\|_{\mathbb{H}}^2e_{a_i}^2v_i^4] + \mathbb{E}[\|\mathbf{u}_i\|_{\mathbb{H}}^2v_i^6], \\ &\approx 6\mathbb{E}[\|\mathbf{u}_i\|_{\mathbb{H}}^2e_{a_i}]E[v_i^5] + 15\mathbb{E}[\|\mathbf{u}_i\|_{\mathbb{H}}^2e_{a_i}^2]E[v_i^4] + \mathbb{E}[\|\mathbf{u}_i\|_{\mathbb{H}}^2]E[v_i^6], \\ &\approx 6\mathbb{E}[\|\mathbf{u}_i\|_{\mathbb{H}}^2e_{a_i}]E[v_i^5] + 15\mathbb{E}[\|\mathbf{u}_i\|_{\mathbb{H}}^2e_{a_i}^2]\xi_v^4 + \mathbb{E}[\|\mathbf{u}_i\|_{\mathbb{H}}^2]\xi_v^6, \end{aligned} \quad (3.28)$$

where $\xi_v^4 = \mathbb{E}[|v_i|^4]$, and $\xi_v^6 = \mathbb{E}[|v_i|^6]$ denote the forth- and sixth-order moments of v_i , respectively.

From Price's theorem [28] we have

$$\mathbb{E}[x \operatorname{sign}(y)] = \sqrt{\frac{2}{\pi}} \frac{1}{\sigma_y} \mathbb{E}[xy], \quad (3.29)$$

where x and y denote two real-valued zero-mean jointly-Gaussian random variables with variances σ_x^2 and σ_y^2 , respectively. Therefore, using (3.29), in the case of the SRLMF algorithm the evaluation of $\mathbb{E}[||\mathbf{u}_i||_{\mathbb{H}}^2]$ becomes straight forward and is given by

$$\begin{aligned} \mathbb{E}[||\mathbf{u}_i||_{\mathbb{H}}^2] &= \mathbb{E}[\mathbf{u}_i \mathbf{H}[\mathbf{u}_i] \mathbf{u}_i^{\mathbb{T}}], \\ &= \mathbb{E}[\mathbf{u}_i \operatorname{sign}[\mathbf{u}_i] \mathbf{u}_i^{\mathbb{T}}], \\ &= \sqrt{\frac{2}{\pi \sigma_u^2}} \operatorname{Tr}(\mathbf{R}). \end{aligned} \quad (3.30)$$

Substituting (3.30) into (3.28) we get

$$\begin{aligned} \mathbb{E}[||\mathbf{u}_i||_{\mathbb{H}}^2 \mathbf{g}^2[e_i]] &\approx 6\mathbb{E}[||\mathbf{u}_i||_{\mathbb{H}}^2 e_{a_i}] \mathbb{E}[v_i^5] + 15\mathbb{E}[||\mathbf{u}_i||_{\mathbb{H}}^2 e_{a_i}^2] \xi_v^4 + \sqrt{\frac{2}{\pi \sigma_u^2}} \operatorname{Tr}(\mathbf{R}) \xi_v^6, \\ &\approx b \sqrt{\frac{2}{\pi \sigma_u^2}} \operatorname{Tr}(\mathbf{R}) + c \mathbb{E}[||\mathbf{u}_i||_{\mathbb{H}}^2 e_{a_i}^2] + 6\mathbb{E}[||\mathbf{u}_i||_{\mathbb{H}}^2 e_{a_i}] \mathbb{E}[v_i^5], \end{aligned} \quad (3.31)$$

where $b = \xi_v^6$ and $c = 15\xi_v^4$.

Substituting (3.26) and (3.31) into (3.22) we get

$$2a\mathbb{E}[e_{a_i}^2] = \mu b \sqrt{\frac{2}{\pi \sigma_u^2}} \operatorname{Tr}(\mathbf{R}) + \mu c \mathbb{E}[||\mathbf{u}_i||_{\mathbb{H}}^2 e_{a_i}^2] + 6\mu \mathbb{E}[||\mathbf{u}_i||_{\mathbb{H}}^2 e_{a_i}] \mathbb{E}[v_i^5]. \quad (3.32)$$

In order to simplify (3.32) and arrive at an expression for the steady-state EMSE $\zeta = \mathbb{E}[e_{a_i}^2]$, we consider two cases:

1. Sufficiently small step-sizes:

Small step-sizes lead to small values of $\mathbb{E}[e_{a_i}^2]$ and e_{a_i} at steady-state. Therefore, for smaller values of μ , the last two terms in (3.32) can be ignored, and therefore we get

$$\zeta = \frac{\mu b}{2a} \sqrt{\frac{2}{\pi \sigma_u^2}} \text{Tr}(\mathbf{R}). \quad (3.33)$$

Substituting the values of a and b in (3.33) results in

$$\zeta = \frac{\mu \xi_v^6}{6\sigma_v^2} \sqrt{\frac{2}{\pi \sigma_u^2}} \text{Tr}(\mathbf{R}). \quad (3.34)$$

2. Separation principle:

For larger values of μ , we resort to the separation assumption, namely, that at steady-state, $\|\mathbf{u}_i\|_{\mathbb{H}}^2$ is independent of e_{a_i} . In this case, the last term in (3.32) will be zero since e_{a_i} has zero mean, we then obtain

$$2a\mathbb{E}[e_{a_i}^2] = \mu b \sqrt{\frac{2}{\pi \sigma_u^2}} \text{Tr}(\mathbf{R}) + \mu c \sqrt{\frac{2}{\pi \sigma_u^2}} \text{Tr}(\mathbf{R}) \mathbb{E}[e_{a_i}^2]. \quad (3.35)$$

Ultimately, the EMSE of the SRLMF algorithm after substituting the values of a , b and c in (3.35) looks like the following:

$$\zeta = \frac{\mu \xi_v^6 \sqrt{\frac{2}{\pi \sigma_u^2}} \text{Tr}(\mathbf{R})}{\left(6\sigma_v^2 - 15\mu \xi_v^4 \sqrt{\frac{2}{\pi \sigma_u^2}} \text{Tr}(\mathbf{R})\right)}. \quad (3.36)$$

3.5.2 Complex-Valued Data

When the data is complex-valued, $g[e_i]$ becomes

$$\begin{aligned}
g[e_i] &= |e_i|^3, \\
&= |e_{a_i}|^3 + e_{a_i}|v_i|^2 + e_{a_i}[e_{a_i}^*v_i + e_{a_i}v_i^*] + v_i|e_{a_i}|^2 + |v_i|^3 \\
&\quad + v_i[e_{a_i}^*v_i + e_{a_i}v_i^*].
\end{aligned} \tag{3.37}$$

For complex-valued data, we assume further that the noise sequence v_i is circular i.e., $\mathbb{E}[v_i^2] = 0$. This assumption leads to the following expression for the term $\mathbb{E}[e_{a_i}^*g[e_i]]$:

$$\begin{aligned}
\mathbb{E}[e_{a_i}^*g[e_i]] &= \mathbb{E}[|e_{a_i}|^2|v_i|^2] + \mathbb{E}[|e_{a_i}|^2[e_{a_i}^*v_i + e_{a_i}v_i^*]] \\
&\quad + \mathbb{E}[e_{a_i}^*v_i[e_{a_i}^*v_i + e_{a_i}v_i^*]].
\end{aligned} \tag{3.38}$$

If we ignore third- and higher-order terms of e_{a_i} , then (3.38) becomes

$$\begin{aligned}
\mathbb{E}[e_{a_i}^*g[e_i]] &\approx 2\sigma_v^2\mathbb{E}[|e_{a_i}|^2], \\
&\approx a'\mathbb{E}[|e_{a_i}|^2],
\end{aligned} \tag{3.39}$$

where $a' = 2\sigma_v^2$.

To evaluate the term $\mathbb{E}[||\mathbf{u}_i||_{\mathbb{H}}^2|g[e_i]|^2]$, we start by noting that (time index i is

omitted for compactness of notation):

$$\begin{aligned}
|g[e]|^2 &= |e|^6, \\
&= |e_a|^6 + |v|^6 + 3|e_a|^4[e_a^*v + e_av^*] + 3|v|^4[e_a^*v + e_av^*] \\
&\quad + |e_a|^2e_a^*v[3e_a^*v + 2e_av^*] + |e_a|^2e_av^*[3e_av^* + 2e_a^*v] + 5|v|^2|e_a|^4 \\
&\quad + 5|e_a|^2|v|^4 + |v|^2e_a^*v[3e_a^*v + 2e_av^*] + |v|^2e_av^*[3e_av^* + 2e_a^*v] \\
&\quad + 9|v|^2|e_a|^2[e_a^*v + e_av^*] + e_a^{*3}v^3 + e_a^3v^{*3}. \tag{3.40}
\end{aligned}$$

If we multiply (3.40) by $\|\mathbf{u}_i\|_{\mathbb{H}}^2$ from the left, use the fact that v_i is independent of both \mathbf{u}_i and e_{a_i} , and if we again ignore third- and higher-order terms of e_{a_i} , we obtain

$$\begin{aligned}
\mathbb{E} [\|\mathbf{u}_i\|_{\mathbb{H}}^2 |g[e_i]|^2] &\approx \xi_v^6 \mathbb{E} [\|\mathbf{u}_i\|_{\mathbb{H}}^2] + 9\xi_v^4 \mathbb{E} [\|\mathbf{u}_i\|_{\mathbb{H}}^2 |e_{a_i}|^2] \\
&\quad + 3\mathbb{E} [\|\mathbf{u}_i\|_{\mathbb{H}}^2 |v_i|^4 [e_{a_i}^* v_i + e_{a_i} v_i^*]]. \tag{3.41}
\end{aligned}$$

From Price's theorem [28] we have

$$\mathbb{E} [\operatorname{Re}[x^* \operatorname{csgn}(y)]] = \sqrt{\frac{2}{\pi}} \frac{\sqrt{2}}{\sigma_y} \mathbb{E} [\operatorname{Re}[x^* y]], \tag{3.42}$$

where $x = x_r + jx_i$ and $y = y_r + jy_i$ denote two complex-valued jointly-Gaussian random variables. Therefore, $\mathbb{E}[\|\mathbf{u}_i\|_{\mathbb{H}}^2]$ as evaluated in Appendix A results in

$$\mathbb{E}[\|\mathbf{u}_i\|_{\mathbb{H}}^2] = \frac{4\operatorname{Tr}(\mathbf{R})}{\sqrt{\pi\sigma_u^2}}. \tag{3.43}$$

Substituting (3.43) into (3.41) we get

$$\begin{aligned} \mathbb{E} [|\|\mathbf{u}_i\|_{\mathbb{H}}^2 |g[e_i]|^2] &\approx b' \left[\frac{4\text{Tr}(\mathbf{R})}{\sqrt{\pi\sigma_u^2}} \right] + c' \mathbb{E} [|\|\mathbf{u}_i\|_{\mathbb{H}}^2 |e_{a_i}|^2] \\ &+ 3\mathbb{E} [|\|\mathbf{u}_i\|_{\mathbb{H}}^2 |v_i|^4 [e_{a_i}^* v_i + e_{a_i} v_i^*]], \end{aligned} \quad (3.44)$$

where $b' = \xi_v^6$ and $c' = 9\xi_v^4$.

Substituting (3.39) and (3.44) into (3.21) we get

$$\begin{aligned} 2a' \mathbb{E} [e_{a_i}^2] &= \mu b' \left[\frac{4\text{Tr}(\mathbf{R})}{\sqrt{\pi\sigma_u^2}} \right] + \mu c' \mathbb{E} [|\|\mathbf{u}_i\|_{\mathbb{H}}^2 |e_{a_i}|^2] \\ &+ 3\mu \mathbb{E} [|\|\mathbf{u}_i\|_{\mathbb{H}}^2 |v_i|^4 [e_{a_i}^* v_i + e_{a_i} v_i^*]]. \end{aligned} \quad (3.45)$$

Here too, as was in the case of real-valued data, in order to simplify (3.45) and arrive at an expression for the steady-state EMSE of the SRLMF algorithm, we consider two cases:

1. Sufficiently small step-sizes:

For smaller values of μ , the last two terms in (3.45) can be ignored, we get

$$\zeta = \frac{\mu b'}{2a'} \left[\frac{4}{\sqrt{\pi\sigma_u^2}} \right] \text{Tr}(\mathbf{R}). \quad (3.46)$$

Substituting the values of a' and b' in (3.46) results in

$$\zeta = \frac{\mu \xi_v^6 \text{Tr}(\mathbf{R})}{\sigma_v^2 \sqrt{\pi\sigma_u^2}}. \quad (3.47)$$

2. Separation principle:

For larger values of μ , the last term in (3.45) will be zero, we then obtain

$$\zeta = \frac{\mu b' \left[\frac{4}{\sqrt{\pi\sigma_u^2}} \right] \text{Tr}(\mathbf{R})}{\left(2a' - \mu c' \left[\frac{4}{\sqrt{\pi\sigma_u^2}} \right] \text{Tr}(\mathbf{R}) \right)}. \quad (3.48)$$

Ultimately, the EMSE of the SRLMF algorithm after substituting the values of a' , b' and c' in (3.48) looks like the following:

$$\zeta = \frac{\mu \xi_v^6 \left[\frac{4}{\sqrt{\pi\sigma_u^2}} \right] \text{Tr}(\mathbf{R})}{\left(4\sigma_v^2 - 9\mu \xi_v^4 \left[\frac{4}{\sqrt{\pi\sigma_u^2}} \right] \text{Tr}(\mathbf{R}) \right)}. \quad (3.49)$$

3.6 Conclusion

In this chapter, expressions for the steady-state EMSE of the SRLMF algorithm are evaluated by relying on energy conservation arguments. In the process of this evaluation, we distinguished between real- and complex-valued data as the definition of the sign function is different in both cases. However, the final expression for the EMSE of the SRLMF algorithm turned out to be identical except for a scaling factor. A conclusion that stands out from the expressions for the steady-state EMSE is that the performance of the SRLMF algorithm is dependent on the step-size μ and the input covariance matrix \mathbf{R} .

CHAPTER 4

TRACKING ANALYSIS OF THE SRLMF ALGORITHM

4.1 Introduction

In this chapter, expressions for the tracking EMSE of the SRLMF algorithm in a nonstationary environment are derived. As in the case of Chapter 3, this chapter proceeds by defining some important assumptions that will be used in the study of the tracking of the SRLMF algorithm. Finally, an optimum value of the step-size μ of the SRLMF algorithm is obtained.

4.2 Nonstationary Data Model

We shall assume that the data $\{d_i, \mathbf{u}_i\}$ satisfy the following conditions of the nonstationary data model [2]:

A.5 There exists an optimal weight vector \mathbf{w}_i^o such that $d_i = \mathbf{u}_i \mathbf{w}_i^o + v_i$.

A.6 The weight vector varies according to the random-walk model $\mathbf{w}_i^o = \mathbf{w}_{i-1}^o + \mathbf{q}_i$, and the sequence \mathbf{q}_i is i.i.d. with covariance matrix \mathbf{Q} . Moreover, \mathbf{q}_i is independent of $\{v_j, \mathbf{u}_j\}$ for all i, j .

A.7 The initial conditions $\{\mathbf{w}_{-1}, \mathbf{w}_{-1}^o\}$ are independent of the zero mean random variables $\{d_i, \mathbf{u}_i, v_i, \mathbf{q}_i\}$.

In this regard, the energy-conservation relation in (3.12) can be set up accordingly to look like

$$\|\mathbf{w}_i^o - \mathbf{w}_i\|^2 + \bar{\mu}_i |e_{a_i}|^2 = \|\mathbf{w}_i^o - \mathbf{w}_{i-1}\|^2 + \bar{\mu}_i |e_{p_i}|^2. \quad (4.1)$$

By taking expectation on both sides of (4.1) we get

$$\mathbb{E}[\|\tilde{\mathbf{w}}_i\|^2] + \mathbb{E}[\bar{\mu}_i |e_{a_i}|^2] = \mathbb{E}[\|\mathbf{w}_i^o - \mathbf{w}_{i-1}\|^2] + \mathbb{E}[\bar{\mu}_i |e_{p_i}|^2]. \quad (4.2)$$

The random-walk model used in this study is governed by the following recursion:

$$\mathbf{w}_i^o = \mathbf{w}_{i-1}^o + \mathbf{q}_i. \quad (4.3)$$

Eventually, this random-walk model allows us to relate $\mathbb{E} [\|\mathbf{w}_i^o - \mathbf{w}_{i-1}\|^2]$ to $\mathbb{E} [\|\tilde{\mathbf{w}}_{i-1}\|^2]$ as follows:

$$\begin{aligned}
\mathbb{E} [\|\mathbf{w}_i^o - \mathbf{w}_{i-1}\|^2] &= \mathbb{E} [\|\mathbf{w}_{i-1}^o + \mathbf{q}_i - \mathbf{w}_{i-1}\|^2], \\
&= \mathbb{E} [\|\tilde{\mathbf{w}}_{i-1} + \mathbf{q}_i\|^2], \\
&= \mathbb{E} [(\tilde{\mathbf{w}}_{i-1} + \mathbf{q}_i)^* (\tilde{\mathbf{w}}_{i-1} + \mathbf{q}_i)], \\
&= \mathbb{E} [\|\tilde{\mathbf{w}}_{i-1}\|^2] + \mathbb{E} [\|\mathbf{q}_i\|^2] + \mathbb{E} [\tilde{\mathbf{w}}_{i-1}^* \mathbf{q}_i] + \mathbb{E} [\mathbf{q}_i^* \tilde{\mathbf{w}}_{i-1}], \\
&= \mathbb{E} [\|\tilde{\mathbf{w}}_{i-1}\|^2] + \text{Tr}(\mathbf{Q}), \tag{4.4}
\end{aligned}$$

where $\text{Tr}(\mathbf{Q}) = \mathbb{E} [\|\mathbf{q}_i\|^2]$.

Substituting (4.4) into (4.2) we get

$$\mathbb{E} [\|\tilde{\mathbf{w}}_i\|^2] + \mathbb{E} [\bar{\mu}_i |e_{a_i}|^2] = \mathbb{E} [\|\tilde{\mathbf{w}}_{i-1}\|^2] + \text{Tr}(\mathbf{Q}) + \mathbb{E} [\bar{\mu}_i |e_{p_i}|^2]. \tag{4.5}$$

Taking the limit as $i \rightarrow \infty$ and using the steady-state condition (3.3), we obtain

$$\lim_{i \rightarrow \infty} \mathbb{E} [\bar{\mu}_i |e_{a_i}|^2] = \lim_{i \rightarrow \infty} \text{Tr}(\mathbf{Q}) + \mathbb{E} [\bar{\mu}_i |e_{p_i}|^2]. \tag{4.6}$$

Also, by substituting (3.6) into (4.6) we get

$$\lim_{i \rightarrow \infty} \mathbb{E} [\bar{\mu}_i |e_{a_i}|^2] = \lim_{i \rightarrow \infty} \text{Tr}(\mathbf{Q}) + \mathbb{E} [\bar{\mu}_i |e_{a_i} - \mu| \|\mathbf{u}_i\|_{\text{Hfg}}^2 [e_i]|^2]. \tag{4.7}$$

This relation can be simplified by expanding the term on the right-hand side as follows (the argument of g is dropped for compactness of notation):

$$\begin{aligned} \bar{\mu}_i |e_{a_i} - \mu \|\mathbf{u}_i\|_{\mathbb{H}}^2 g|^2 &= \bar{\mu}_i |e_{a_i}|^2 + \mu^2 \bar{\mu}_i \|\mathbf{u}_i\|_{\mathbb{H}}^4 |g|^2 - \mu \bar{\mu}_i \|\mathbf{u}_i\|_{\mathbb{H}}^2 g e_{a_i}^* \\ &\quad - \mu \bar{\mu}_i \|\mathbf{u}_i\|_{\mathbb{H}}^2 g^* e_{a_i}. \end{aligned} \quad (4.8)$$

But since the product $\bar{\mu}_i \|\mathbf{u}_i\|_{\mathbb{H}}^2$ is unity for all \mathbf{u}_i except for the single event $\mathbf{u}_i = \mathbf{0}$, we find that

$$\bar{\mu}_i |e_{a_i} - \mu \|\mathbf{u}_i\|_{\mathbb{H}}^2 g|^2 = \bar{\mu}_i |e_{a_i}|^2 + \mu^2 \|\mathbf{u}_i\|_{\mathbb{H}}^2 |g|^2 - \mu e_{a_i} g^* - \mu e_{a_i}^* g. \quad (4.9)$$

Taking expectations of both sides of (4.9) we obtain

$$\begin{aligned} \mathbb{E} [\bar{\mu}_i |e_{a_i} - \mu \|\mathbf{u}_i\|_{\mathbb{H}}^2 g|^2] &= \mathbb{E} [\bar{\mu}_i |e_{a_i}|^2] + \mu^2 \mathbb{E} [\|\mathbf{u}_i\|_{\mathbb{H}}^2 |g|^2] - \mu \mathbb{E} [e_{a_i} g^*] \\ &\quad - \mu \mathbb{E} [e_{a_i}^* g]. \end{aligned} \quad (4.10)$$

Substituting this result into (4.7) leads to

$$\mu \mathbb{E} [\|\mathbf{u}_i\|_{\mathbb{H}}^2 |g|^2] + \mu^{-1} \text{Tr}(\mathbf{Q}) = \mathbb{E} [e_{a_i} g^* + e_{a_i}^* g],$$

which can be written as [2]:

$$\lim_{i \rightarrow \infty} \mu \mathbb{E} [\|\mathbf{u}_i\|_{\mathbb{H}}^2 |g[e_i]|^2] + \mu^{-1} \text{Tr}(\mathbf{Q}) = \lim_{i \rightarrow \infty} 2\text{Re}(\mathbb{E} [e_{a_i}^* g[e_i]]). \quad (4.11)$$

The variance relation in (4.11) holds for any adaptive filter of the form (3.4), and for any data $\{d_i, \mathbf{u}_i\}$, assuming filter operation in steady-state.

For real-valued data, this variance relation becomes

$$\lim_{i \rightarrow \infty} \mu \mathbb{E} [\|\mathbf{u}_i\|_{\text{H}}^2 g^2[e_i]] + \mu^{-1} \text{Tr}(\mathbf{Q}) = \lim_{i \rightarrow \infty} 2\mathbb{E} [e_{a_i} g[e_i]]. \quad (4.12)$$

4.3 Tracking Analysis of the SRLMF algorithm

In this section, the tracking performance of the SRLMF algorithm is carried out for two scenarios, the real- and complex-valued data. Each of these scenarios will be dealt alone.

4.3.1 Real-Valued Data

Tracking results can be obtained by inspection from the mean-square results as there are only minor differences. Therefore, by substituting (3.26) and (3.31) into (4.12) we get

$$\begin{aligned} 2a\mathbb{E}[e_{a_i}^2] &= \mu^{-1} \text{Tr}(\mathbf{Q}) + \mu b \sqrt{\frac{2}{\pi \sigma_u^2}} \text{Tr}(\mathbf{R}) + \mu c \mathbb{E}[\|\mathbf{u}_i\|_{\text{H}}^2 e_{a_i}^2] \\ &\quad + 6\mu \mathbb{E}[\|\mathbf{u}_i\|_{\text{H}}^2 e_{a_i}] \mathbb{E}[v_i^5], \end{aligned} \quad (4.13)$$

where $a = 3\sigma_v^2$, $b = \xi_v^6$ and $c = 15\xi_v^4$.

In order to simplify (4.13) and arrive at an expression for the tracking EMSE of the SRLMF algorithm, we consider two cases:

1. Sufficiently small step-sizes:

Small step-sizes lead to small values of $E[e_{a_i}^2]$ and e_{a_i} at steady-state. Therefore, for smaller values of μ , the last two terms in (4.13) can be ignored, and therefore we get

$$\zeta = \frac{\mu^{-1}\text{Tr}(\mathbf{Q}) + \mu b \sqrt{\frac{2}{\pi\sigma_u^2}} \text{Tr}(\mathbf{R})}{2a}. \quad (4.14)$$

An optimum value of the step-size of the SRLMF algorithm is obtained by minimizing (4.14) with respect to μ . Setting the derivative of ζ with respect to μ equal to zero gives

$$\mu_{\text{opt}} = \sqrt{\frac{\text{Tr}(\mathbf{Q})}{b \sqrt{\frac{2}{\pi\sigma_u^2}} \text{Tr}(\mathbf{R})}}. \quad (4.15)$$

2. Separation principle:

For larger values of μ , the last term in (4.13) will be zero, we then obtain

$$\zeta = \frac{\mu^{-1}\text{Tr}(\mathbf{Q}) + \mu b \sqrt{\frac{2}{\pi\sigma_u^2}} \text{Tr}(\mathbf{R})}{\left(2a - \mu c \sqrt{\frac{2}{\pi\sigma_u^2}} \text{Tr}(\mathbf{R})\right)}, \quad (4.16)$$

and eventually the optimum step-size of the SRLMF algorithm is given by

$$\mu_{\text{opt}} = \sqrt{\text{Tr}(\mathbf{Q}) \left[\frac{c^2 \text{Tr}(\mathbf{Q})}{4a^2 b^2} + \frac{1}{b \sqrt{\frac{2}{\pi\sigma_u^2}} \text{Tr}(\mathbf{R})} \right]} - \frac{c}{2ab} \text{Tr}(\mathbf{Q}). \quad (4.17)$$

4.3.2 Complex-Valued Data

Similarly, by substituting (3.39) and (3.44) into (4.11) we get

$$\begin{aligned}
2a'E[|e_{a_i}|^2] &= \mu^{-1}\text{Tr}(\mathbf{Q}) + \mu b' \left[\frac{4\text{Tr}(\mathbf{R})}{\sqrt{\pi\sigma_u^2}} \right] + \mu c' E[\|\mathbf{u}_i\|_{\mathbb{H}}^2 |e_{a_i}|^2] \\
&\quad + 3\mu E[\|\mathbf{u}_i\|_{\mathbb{H}}^2 |v_i|^4 [e_{a_i}^* v_i + e_{a_i} v_i^*]],
\end{aligned} \tag{4.18}$$

where $a' = 2\sigma_v^2$, $b' = \xi_v^6$ and $c' = 9\xi_v^4$.

In order to simplify (4.18) and arrive at an expression for the tracking EMSE of the SRLMF algorithm, we consider two cases:

1. Sufficiently small step-sizes:

For smaller values of μ , the last two terms in (4.18) can be ignored, we get

$$\zeta = \frac{\mu^{-1}\text{Tr}(\mathbf{Q}) + \mu b' \left[\frac{4}{\sqrt{\pi\sigma_u^2}} \right] \text{Tr}(\mathbf{R})}{2a'}, \tag{4.19}$$

and the associated optimum step-size value is equal to

$$\mu_{\text{opt}} = \sqrt{\frac{\sqrt{\pi\sigma_u^2}\text{Tr}(\mathbf{Q})}{4b'\text{Tr}(\mathbf{R})}}. \tag{4.20}$$

2. Separation principle:

For larger values of μ , the last term in (4.18) will be zero, we then obtain

$$\zeta = \frac{\mu^{-1}\text{Tr}(\mathbf{Q}) + \mu b' \left[\frac{4}{\sqrt{\pi\sigma_u^2}} \right] \text{Tr}(\mathbf{R})}{\left(2a' - \mu c' \left[\frac{4}{\sqrt{\pi\sigma_u^2}} \right] \text{Tr}(\mathbf{R}) \right)}, \quad (4.21)$$

and finally the optimum step-size corresponding to this scenario is given by

$$\mu_{\text{opt}} = \sqrt{\text{Tr}(\mathbf{Q}) \left[\frac{c'^2 \text{Tr}(\mathbf{Q})}{4a'^2 b'^2} + \frac{\sqrt{\pi\sigma_u^2}}{4b' \text{Tr}(\mathbf{R})} \right]} - \frac{c'}{2a'b'} \text{Tr}(\mathbf{Q}). \quad (4.22)$$

4.4 Conclusion

In this chapter, expressions for the tracking EMSE of the SRLMF algorithm are evaluated by relying on energy conservation arguments. The term $\mu^{-1}\text{Tr}(\mathbf{Q})$ in the tracking EMSE of the SRLMF algorithm reflects the effect of the nonstationarity on filter performance. We observe that $\text{Tr}(\mathbf{Q})$ appears multiplied by μ^{-1} so that the larger the step-size the smaller the effect of the nonstationarity on the EMSE. This behavior is intuitive since a smaller step-size signifies faster adaptation, in which case the SRLMF algorithm will have a better chance at learning and at following the data statistics. A small step-size, on the other hand, leads to smaller EMSE under stationary conditions, but it may also lead to poor tracking performance. This discussion suggests that there exists a compromise choice for the step-size μ . Therefore, an optimum value of the step-size is also evaluated in Chapter 4.

CHAPTER 5

TRANSIENT ANALYSIS OF THE SRLMF ALGORITHM

5.1 Introduction

Transient analysis is more conveniently performed by relying on a weighted energy-conservation relation, as opposed to the unweighted version that was employed in Chapters 3 and 4. In this chapter, the weighted variance relation presented in [2] has been extended in order to derive expressions for the MSE and the MSD of the SRLMF algorithm during the transient phase.

Here, we shall assume that the data $\{d_i, \mathbf{u}_i\}$ satisfy the conditions of the stationary data model described in Chapter 3.

5.2 Weighted Energy-Conservation Relation

To be able to analyze the transient behavior of the SRLMF algorithm, the energy-conservation relation developed earlier in the previous chapters needs to be revisited for this purpose. In the ensuing analysis the derivation of the weighted energy-conservation relation is developed.

Theorem 1: For any adaptive filter of the form (3.4), any positive-definite Hermitian matrix Σ , and for any data $\{d_i, \mathbf{u}_i\}$, it holds that [2]:

$$\|\mathbf{u}_i\|_{\mathbf{H}\Sigma\mathbf{H}}^2 \|\tilde{\mathbf{w}}_i\|_{\Sigma}^2 + |e_{a_i}^{\mathbf{H}\Sigma}|^2 = \|\mathbf{u}_i\|_{\mathbf{H}\Sigma\mathbf{H}}^2 \|\tilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2 + |e_{p_i}^{\mathbf{H}\Sigma}|^2, \quad (5.1)$$

where $e_{a_i}^{\mathbf{H}\Sigma} = \mathbf{u}_i \mathbf{H}[\mathbf{u}_i] \Sigma \tilde{\mathbf{w}}_{i-1}$, $e_{p_i}^{\mathbf{H}\Sigma} = \mathbf{u}_i \mathbf{H}[\mathbf{u}_i] \Sigma \tilde{\mathbf{w}}_i$, $\tilde{\mathbf{w}}_i = \mathbf{w}^o - \mathbf{w}_i$, and $\|\mathbf{u}_i\|_{\mathbf{H}\Sigma\mathbf{H}}^2 = \mathbf{u}_i (\mathbf{H}[\mathbf{u}_i] \Sigma \mathbf{H}[\mathbf{u}_i]) \mathbf{u}_i^*$.

Proof: Let us consider the adaptive filter updates of the generic form given below:

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \mathbf{H}[\mathbf{u}_i] \mathbf{u}_i^* g[e_i], \quad i \geq 0. \quad (5.2)$$

Subtracting both sides of (5.2) from \mathbf{w}^o we get

$$\tilde{\mathbf{w}}_i = \tilde{\mathbf{w}}_{i-1} - \mu \mathbf{H}[\mathbf{u}_i] \mathbf{u}_i^* g[e_i]. \quad (5.3)$$

If we multiply both sides of (5.3) by $\mathbf{u}_i \mathbf{H}[\mathbf{u}_i] \Sigma$ from the left we get

$$e_{p_i}^{\text{H}\Sigma} = e_{a_i}^{\text{H}\Sigma} - \mu \|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2 g[e_i]. \quad (5.4)$$

Two cases can be considered here.

Case 1: $\|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2 = 0$.

In this case, $\tilde{\mathbf{w}}_i = \tilde{\mathbf{w}}_{i-1}$ and $e_{a_i}^{\text{H}\Sigma} = e_{p_i}^{\text{H}\Sigma}$ so that $\|\tilde{\mathbf{w}}_i\|_{\Sigma}^2 = \|\tilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2$ and $|e_{a_i}^{\text{H}\Sigma}|^2 = |e_{p_i}^{\text{H}\Sigma}|^2$.

Case 2: $\|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2 \neq 0$.

In this case, we use (5.4) to solve for $g[e_i]$,

$$g[e_i] = \frac{1}{\mu \|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2} (e_{a_i}^{\text{H}\Sigma} - e_{p_i}^{\text{H}\Sigma}). \quad (5.5)$$

Substituting (5.5) into (5.3) we get

$$\tilde{\mathbf{w}}_i = \tilde{\mathbf{w}}_{i-1} - \frac{\mathbf{H}[\mathbf{u}_i] \mathbf{u}_i^*}{\|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2} (e_{a_i}^{\text{H}\Sigma} - e_{p_i}^{\text{H}\Sigma}). \quad (5.6)$$

Expression (5.6) can be rearranged as

$$\tilde{\mathbf{w}}_i + \frac{\mathbf{H}[\mathbf{u}_i] \mathbf{u}_i^*}{\|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2} e_{a_i}^{\text{H}\Sigma} = \tilde{\mathbf{w}}_{i-1} + \frac{\mathbf{H}[\mathbf{u}_i] \mathbf{u}_i^*}{\|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2} e_{p_i}^{\text{H}\Sigma}. \quad (5.7)$$

Evaluating the energies of both sides of (5.7) results in

$$\left\| \tilde{\mathbf{w}}_i + \frac{\mathbf{H}[\mathbf{u}_i] \mathbf{u}_i^*}{\|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2} e_{a_i}^{\text{H}\Sigma} \right\|_{\Sigma}^2 = \left\| \tilde{\mathbf{w}}_{i-1} + \frac{\mathbf{H}[\mathbf{u}_i] \mathbf{u}_i^*}{\|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2} e_{p_i}^{\text{H}\Sigma} \right\|_{\Sigma}^2. \quad (5.8)$$

After a straightforward calculation, the following weighted energy-conservation results:

$$\|\tilde{\mathbf{w}}_i\|_{\Sigma}^2 + \frac{1}{\|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2} |e_{a_i}^{\text{H}\Sigma}|^2 = \|\tilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2 + \frac{1}{\|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2} |e_{p_i}^{\text{H}\Sigma}|^2. \quad (5.9)$$

The weighted energy-conservation relation in (5.9) can also be written as

$$\|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2 \|\tilde{\mathbf{w}}_i\|_{\Sigma}^2 + |e_{a_i}^{\text{H}\Sigma}|^2 = \|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2 \|\tilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2 + |e_{p_i}^{\text{H}\Sigma}|^2. \quad (5.10)$$

5.3 Weighted Variance Relation

Similarly here, in this section, the weighted variance relation is developed.

Theorem 2: For any adaptive filter of the form (5.2), any positive-definite Hermitian matrix Σ , and for any data $\{d_i, \mathbf{u}_i\}$, it holds that

$$\begin{aligned} \mathbf{E}[\|\tilde{\mathbf{w}}_i\|_{\Sigma}^2] &= \mathbf{E}[\|\tilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2] + \mu^2 \mathbf{E} [\|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2 |g[e_i]|^2] \\ &\quad - 2\mu \text{Re} (\mathbf{E} [e_{a_i}^{\text{H}\Sigma*} g[e_i]]), \quad \text{as } i \rightarrow \infty. \end{aligned} \quad (5.11)$$

Similarly, for real-valued data, the above weighted variance relation becomes

$$\begin{aligned} \mathbf{E}[\|\tilde{\mathbf{w}}_i\|_{\Sigma}^2] &= \mathbf{E}[\|\tilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2] + \mu^2 \mathbf{E} [\|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2 g^2[e_i]] \\ &\quad - 2\mu \mathbf{E} [e_{a_i}^{\text{H}\Sigma} g[e_i]], \quad \text{as } i \rightarrow \infty. \end{aligned} \quad (5.12)$$

Proof: Squaring both sides of (5.4) we get

$$|e_{p_i}^{\text{H}\Sigma}|^2 = |e_{a_i}^{\text{H}\Sigma} - \mu \|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2 \mathfrak{g}[e_i]|^2. \quad (5.13)$$

For compactness of notation let us omit the argument of \mathfrak{g} so that (5.13) looks like

$$|e_{p_i}^{\text{H}\Sigma}|^2 = |e_{a_i}^{\text{H}\Sigma}|^2 + \mu^2 \|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^4 |\mathfrak{g}|^2 - \mu e_{a_i}^{\text{H}\Sigma} \|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2 \mathfrak{g}^* - \mu e_{a_i}^{\text{H}\Sigma*} \|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2 \mathfrak{g}. \quad (5.14)$$

Substituting (5.14) into (5.10) we get

$$\begin{aligned} \|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2 \|\tilde{\mathbf{w}}_i\|_{\Sigma}^2 &= \|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2 \|\tilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2 + \mu^2 \|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^4 |\mathfrak{g}|^2 - \mu e_{a_i}^{\text{H}\Sigma} \|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2 \mathfrak{g}^* \\ &\quad - \mu e_{a_i}^{\text{H}\Sigma*} \|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2 \mathfrak{g}. \end{aligned} \quad (5.15)$$

Dividing both sides of (5.15) by $\|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2$ (of course here $\|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2 \neq 0$) we get

$$\|\tilde{\mathbf{w}}_i\|_{\Sigma}^2 = \|\tilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2 + \mu^2 \|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2 |\mathfrak{g}|^2 - \mu e_{a_i}^{\text{H}\Sigma} \mathfrak{g}^* - \mu e_{a_i}^{\text{H}\Sigma*} \mathfrak{g}. \quad (5.16)$$

Taking expectations of both sides of (5.16) we obtain

$$\begin{aligned} \mathbb{E}[\|\tilde{\mathbf{w}}_i\|_{\Sigma}^2] &= \mathbb{E}[\|\tilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2] + \mu^2 \mathbb{E}[\|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2 |\mathfrak{g}[e_i]|^2] \\ &\quad - \mu \mathbb{E}[e_{a_i}^{\text{H}\Sigma} \mathfrak{g}[e_i]^* + e_{a_i}^{\text{H}\Sigma*} \mathfrak{g}[e_i]]. \end{aligned} \quad (5.17)$$

or in the following format:

$$\begin{aligned} \mathbb{E}[\|\tilde{\mathbf{w}}_i\|_{\Sigma}^2] &= \mathbb{E}[\|\tilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2] + \mu^2 \mathbb{E}[\|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2 \mathbb{g}[e_i]^2] \\ &\quad - 2\mu \text{Re}(\mathbb{E}[e_{a_i}^{\text{H}\Sigma} \mathbb{g}[e_i]]), \quad \text{as } i \rightarrow \infty. \end{aligned} \quad (5.18)$$

For real-valued data, the weighted variance relation in (5.18) becomes

$$\begin{aligned} \mathbb{E}[\|\tilde{\mathbf{w}}_i\|_{\Sigma}^2] &= \mathbb{E}[\|\tilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2] + \mu^2 \mathbb{E}[\|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2 \mathbb{g}^2[e_i]] \\ &\quad - 2\mu \mathbb{E}[e_{a_i}^{\text{H}\Sigma} \mathbb{g}[e_i]], \quad \text{as } i \rightarrow \infty. \end{aligned} \quad (5.19)$$

5.4 Transient Analysis of the SRLMF algorithm

The transient analysis of the class of filters in (5.2) is more challenging due to the presence of the error nonlinearity. Nevertheless, by using some approximations, the analysis can be carried out to provide some useful insights about the performance of the SRLMF algorithm.

To start, the expectations $\mathbb{E}[\|\mathbf{u}_i\|_{\text{H}\Sigma\text{H}}^2 \mathbb{g}^2[e_i]]$ and $\mathbb{E}[e_{a_i}^{\text{H}\Sigma} \mathbb{g}[e_i]]$ are evaluated in the ensuing analysis in terms of the weighted norm of $\tilde{\mathbf{w}}_{i-1}$. Since, these expectations are involved mathematically; therefore, we shall rely on the following assumption in order to facilitate their evaluation [2]:

A.8 The a priori estimation errors $\{e_{a_i}, e_{a_i}^{\text{H}\Sigma}\}$ are jointly circular Gaussian.

Evaluation of $\mathbb{E}[e_{a_i}^{\text{H}\Sigma} \mathbb{g}[e_i]]$:

From Price's theorem, if x and y are jointly Gaussian random variables that

are independent from a third random variable z , then it holds that [28]:

$$\mathbf{E}[xg(y+z)] = \frac{\mathbf{E}[xy]}{\mathbf{E}[y^2]} \mathbf{E}[yg(y+z)]. \quad (5.20)$$

Applying this result to the term $\mathbf{E}[e_{a_i}^{\text{H}\Sigma} g[e_i]]$, and using (3.23), we get

$$\begin{aligned} \mathbf{E}[e_{a_i}^{\text{H}\Sigma} g[e_i]] &= \mathbf{E}[e_{a_i}^{\text{H}\Sigma} g[e_{a_i} + v_i]], \\ &= \mathbf{E}[e_{a_i}^{\text{H}\Sigma} e_{a_i}] \left[\frac{\mathbf{E}[e_{a_i} g[e_i]]}{\mathbf{E}[e_{a_i}^2]} \right]. \end{aligned} \quad (5.21)$$

In view of the assumption (A.8), the expectation $\mathbf{E}[e_{a_i} g[e_i]]$ depends on e_{a_i} only through its second moment, $\mathbf{E}[e_{a_i}^2]$. Therefore, we can define the following function of $\mathbf{E}[e_{a_i}^2]$:

$$\mathcal{Z}_1 = \frac{\mathbf{E}[e_{a_i} g[e_i]]}{\mathbf{E}[e_{a_i}^2]}. \quad (5.22)$$

For the SRLMF algorithm, $g[e_i] = e_i^3$, therefore

$$\begin{aligned} \mathbf{E}[e_{a_i} g[e_i]] &= \mathbf{E}[e_{a_i} (e_{a_i} + v_i)^3], \\ &= \mathbf{E}[e_{a_i}^4 + 3e_{a_i}^3 v_i + 3e_{a_i}^2 v_i^2 + v_i^3 e_{a_i}]. \end{aligned} \quad (5.23)$$

Now since e_{a_i} and v_i are zero mean Gaussian and independent random variables with variances $\mathbf{E}[e_{a_i}^2]$ and σ_v^2 , respectively, we obtain

$$\mathbf{E}[e_{a_i} g[e_i]] = \mathbf{E}[e_{a_i}^4] + 3\sigma_v^2 \mathbf{E}[e_{a_i}^2]. \quad (5.24)$$

By using the fact that for circular Gaussian e_{a_i} it holds that $\mathbb{E}[e_{a_i}^4] = 3\mathbb{E}[e_{a_i}^2]^2$, we get

$$\begin{aligned} \mathbb{E}[e_{a_i}g[e_i]] &= 3\mathbb{E}[e_{a_i}^2]^2 + 3\sigma_v^2\mathbb{E}[e_{a_i}^2], \\ &= 3\mathbb{E}[e_{a_i}^2] [\mathbb{E}[e_{a_i}^2] + \sigma_v^2]. \end{aligned} \quad (5.25)$$

Substituting (5.25) into (5.22) we get

$$\mathcal{Z}_1 = 3 [\mathbb{E}[e_{a_i}^2] + \sigma_v^2]. \quad (5.26)$$

The expression for \mathcal{Z}_1 is related to the desired term $\mathbb{E}[e_{a_i}^{\text{HS}}g[e_i]]$ through the equality

$$\mathbb{E}[e_{a_i}^{\text{HS}}g[e_i]] = \mathcal{Z}_1\mathbb{E}[e_{a_i}^{\text{HS}}e_{a_i}]. \quad (5.27)$$

Evaluation of $\mathbb{E}[\|\mathbf{u}_i\|_{\text{HS}}^2g^2[e_i]]$:

In order to facilitate the evaluation of the term $\mathbb{E}[\|\mathbf{u}_i\|_{\text{HS}}^2g^2[e_i]]$ we use the separation principle, namely, we assume that the filter is long enough (say filter length of five, $M = 5$) so that the following assumption holds [2]:

A.9 $\|\mathbf{u}_i\|_{\text{HS}}^2$ is independent of e_i .

Therefore,

$$\mathbb{E}[\|\mathbf{u}_i\|_{\text{HS}}^2g^2[e_i]] = (\mathbb{E}[\|\mathbf{u}_i\|_{\text{HS}}^2]) (\mathbb{E}[g^2[e_i]]). \quad (5.28)$$

Since e_{a_i} is Gaussian and independent of the noise, the expectation $\mathbb{E}[g^2[e_i]]$ depends on e_{a_i} through its second moment only. Therefore, we can define the fol-

lowing function of $E[e_{a_i}^2]$:

$$\mathcal{Z}_2 = E[g^2[e_i]]. \quad (5.29)$$

For the SRLMF algorithm, $g[e_i] = e_i^3$. Since e_{a_i} and v_i are zero mean Gaussian and independent random variables with variances $E[e_{a_i}^2]$ and σ_v^2 , we have $\sigma_e^2 = E[e_i^2] = E[e_{a_i}^2] + \sigma_v^2$. Moreover from [2], $E[e_i^6] = 15\sigma_e^6$. Thus

$$\begin{aligned} \mathcal{Z}_2 &= E[e_i^6], \\ &= 15\sigma_e^6, \\ &= 15(\sigma_e^2)^3, \\ &= 15(E[e_{a_i}^2] + \sigma_v^2)^3, \\ &= 15(E[e_{a_i}^2])^3 + 45\sigma_v^2(E[e_{a_i}^2])^2 + 45\xi_v^4 E[e_{a_i}^2] + 15\xi_v^6. \end{aligned} \quad (5.30)$$

The expression for \mathcal{Z}_2 is related to the desired term $E[||\mathbf{u}_i||_{\text{H}\Sigma\text{H}}^2 g^2[e_i]]$ through the equality

$$\begin{aligned} E[||\mathbf{u}_i||_{\text{H}\Sigma\text{H}}^2 g^2[e_i]] &= \mathcal{Z}_2 E[||\mathbf{u}_i||_{\text{H}\Sigma\text{H}}^2], \\ &= \mathcal{Z}_2 E[||\text{sign}[\mathbf{u}_i]||_{\Sigma}^2]. \end{aligned} \quad (5.31)$$

Since

$$\begin{aligned}
\mathbb{E} [\|\mathbf{u}_i\|_{\mathbf{H}\Sigma\mathbf{H}}^2] &= \mathbb{E}[\mathbf{u}_i\mathbf{H}[\mathbf{u}_i]\Sigma\mathbf{H}[\mathbf{u}_i]\mathbf{u}_i^T], \\
&= \mathbb{E}[\text{sign}[\mathbf{u}_i]\Sigma\text{sign}[\mathbf{u}_i]^T], \\
&= \mathbb{E} [\|\text{sign}[\mathbf{u}_i]\|_{\Sigma}^2].
\end{aligned} \tag{5.32}$$

Substituting (5.27) and (5.31) into (5.19) we get

$$\mathbb{E} [\|\tilde{\mathbf{w}}_i\|_{\Sigma}^2] = \mathbb{E} [\|\tilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2] + \mu^2 \mathcal{Z}_2 \mathbb{E} [\|\text{sign}[\mathbf{u}_i]\|_{\Sigma}^2] - 2\mu \mathcal{Z}_1 \mathbb{E} [e_{a_i}^{\text{H}\Sigma} e_{a_i}]. \tag{5.33}$$

Independence Assumption

If we assume that the regressor sequence $\{\mathbf{u}_i\}$ is i.i.d. then

$$\begin{aligned}
\mathbb{E} [e_{a_i}^{\text{H}\Sigma} e_{a_i}] &= \mathbb{E} [\tilde{\mathbf{w}}_{i-1}^T \Sigma \mathbf{H}[\mathbf{u}_i] \mathbf{u}_i^T \mathbf{u}_i \tilde{\mathbf{w}}_{i-1}], \\
&= \mathbb{E} \left[\|\tilde{\mathbf{w}}_{i-1}\|_{\Sigma \mathbf{H} \mathbf{u}_i^T \mathbf{u}_i}^2 \right].
\end{aligned} \tag{5.34}$$

In this way, the terms $\{\mathbb{E} [e_{a_i}^{\text{H}\Sigma} e_{a_i}], \mathcal{Z}_1, \mathcal{Z}_2\}$ become all functions of $\tilde{\mathbf{w}}_{i-1}$. Therefore, (5.33) becomes

$$\begin{aligned}
\mathbb{E} [\|\tilde{\mathbf{w}}_i\|_{\Sigma}^2] &= \mathbb{E} [\|\tilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2] + \mu^2 \mathcal{Z}_2 \mathbb{E} [\|\text{sign}[\mathbf{u}_i]\|_{\Sigma}^2] - 2\mu \mathcal{Z}_1 \mathbb{E} \left[\|\tilde{\mathbf{w}}_{i-1}\|_{\Sigma \mathbf{H} \mathbf{u}_i^T \mathbf{u}_i}^2 \right], \\
&= \mathbb{E} [\|\tilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2] + \mu^2 \mathcal{Z}_2 \mathbb{E} [\|\text{sign}[\mathbf{u}_i]\|_{\Sigma}^2] - 2\mu \mathcal{Z}_1 \mathbb{E} \left[\|\tilde{\mathbf{w}}_{i-1}\|_{\Sigma \text{sign}[\mathbf{u}_i]^T \mathbf{u}_i}^2 \right], \\
&= \mathbb{E} [\|\tilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2] + \mu^2 \mathcal{Z}_2 \mathbb{E} [\|\text{sign}[\mathbf{u}_i]\|_{\Sigma}^2] - \sqrt{\frac{8}{\pi \sigma_u^2}} \mu \mathcal{Z}_1 \mathbb{E} [\|\tilde{\mathbf{w}}_{i-1}\|_{\Sigma \mathbf{R}}^2].
\end{aligned} \tag{5.35}$$

We thus find that studying the transient behavior of the SRLMF algorithm in effect has reduced to evaluating the functions \mathcal{Z}_1 and \mathcal{Z}_2 and studying the resulting variance relation (5.35). Let us now illustrate the application of the above results for white and correlated input data:

White Input Data

For white input data \mathbf{R} is diagonal, say $\mathbf{R} = \sigma_u^2 \mathbf{I}$. Therefore, if we select $\Sigma = \mathbf{I}$, the variance relation (5.35) becomes

$$\mathbb{E} [\|\tilde{\mathbf{w}}_i\|^2] = \mathbb{E} [\|\tilde{\mathbf{w}}_{i-1}\|^2] + \mu^2 \mathcal{Z}_2 \mathbb{E} [\|\text{sign}[\mathbf{u}_i]\|^2] - \sqrt{\frac{8\sigma_u^2}{\pi}} \mu \mathcal{Z}_1 \mathbb{E} [\|\tilde{\mathbf{w}}_{i-1}\|^2]. \quad (5.36)$$

Now since

$$\begin{aligned} e_{a_i}^2 &= \tilde{\mathbf{w}}_{i-1}^T \mathbf{u}_i^T \mathbf{u}_i \tilde{\mathbf{w}}_{i-1}, \\ &= \|\tilde{\mathbf{w}}_{i-1}\|_{\mathbf{u}_i^T \mathbf{u}_i}^2. \end{aligned} \quad (5.37)$$

Substituting (5.37) into (5.30) we get

$$\begin{aligned} \mathcal{Z}_2 &= 15 \left(\mathbb{E} \left[\|\tilde{\mathbf{w}}_{i-1}\|_{\mathbf{u}_i^T \mathbf{u}_i}^2 \right] \right)^3 + 45\sigma_v^2 \left(\mathbb{E} \left[\|\tilde{\mathbf{w}}_{i-1}\|_{\mathbf{u}_i^T \mathbf{u}_i}^2 \right] \right)^2 + 45\xi_v^4 \mathbb{E} \left[\|\tilde{\mathbf{w}}_{i-1}\|_{\mathbf{u}_i^T \mathbf{u}_i}^2 \right] \\ &\quad + 15\xi_v^6, \\ &= 15 \left(\mathbb{E} \left[\|\tilde{\mathbf{w}}_{i-1}\|_{\mathbf{R}}^2 \right] \right)^3 + 45\sigma_v^2 \left(\mathbb{E} \left[\|\tilde{\mathbf{w}}_{i-1}\|_{\mathbf{R}}^2 \right] \right)^2 + 45\xi_v^4 \mathbb{E} \left[\|\tilde{\mathbf{w}}_{i-1}\|_{\mathbf{R}}^2 \right] + 15\xi_v^6, \\ &= 15 \left(\sigma_u^2 \mathbb{E} \left[\|\tilde{\mathbf{w}}_{i-1}\|^2 \right] \right)^3 + 45\sigma_v^2 \left(\sigma_u^2 \mathbb{E} \left[\|\tilde{\mathbf{w}}_{i-1}\|^2 \right] \right)^2 + 45\xi_v^4 \sigma_u^2 \mathbb{E} \left[\|\tilde{\mathbf{w}}_{i-1}\|^2 \right] \\ &\quad + 15\xi_v^6. \end{aligned} \quad (5.38)$$

Similarly by substituting (5.37) into (5.26) we get

$$\mathcal{Z}_1 = 3 (\sigma_u^2 \mathbb{E} [|\tilde{\mathbf{w}}_{i-1}|^2] + \sigma_v^2). \quad (5.39)$$

Substituting (5.38) and (5.39) into (5.36) we get

$$\begin{aligned} \mathbb{E} [|\tilde{\mathbf{w}}_i|^2] &= \mathbb{E} [|\tilde{\mathbf{w}}_{i-1}|^2] + \mu^2 [15 (\sigma_u^2 \mathbb{E} [|\tilde{\mathbf{w}}_{i-1}|^2])^3 + 45 \sigma_v^2 (\sigma_u^2 \mathbb{E} [|\tilde{\mathbf{w}}_{i-1}|^2])^2 \\ &\quad + 45 \xi_v^4 \sigma_u^2 \mathbb{E} [|\tilde{\mathbf{w}}_{i-1}|^2] + 15 \xi_v^6] \mathbb{E} [|\text{sign}[\mathbf{u}_i]|^2] \\ &\quad - 3 \sqrt{\frac{8 \sigma_u^2}{\pi}} \mu (\sigma_u^2 \mathbb{E} [|\tilde{\mathbf{w}}_{i-1}|^2] + \sigma_v^2) \mathbb{E} [|\tilde{\mathbf{w}}_{i-1}|^2]. \end{aligned} \quad (5.40)$$

Since $\mathbb{E} [|\text{sign}[\mathbf{u}_i]|^2] = M$, the recursion in (5.40) becomes

$$\begin{aligned} \mathbb{E} [|\tilde{\mathbf{w}}_i|^2] &= \mathbb{E} [|\tilde{\mathbf{w}}_{i-1}|^2] + 15 \mu^2 M \sigma_u^6 (\mathbb{E} [|\tilde{\mathbf{w}}_{i-1}|^2])^3 + 45 \mu^2 M \sigma_v^2 \sigma_u^4 (\mathbb{E} [|\tilde{\mathbf{w}}_{i-1}|^2])^2 \\ &\quad + 45 \mu^2 M \xi_v^4 \sigma_u^2 \mathbb{E} [|\tilde{\mathbf{w}}_{i-1}|^2] + 15 \mu^2 M \xi_v^6 - 6 \sqrt{\frac{2 \sigma_u^2}{\pi}} \mu \sigma_u^2 (\mathbb{E} [|\tilde{\mathbf{w}}_{i-1}|^2])^2 \\ &\quad - 6 \sqrt{\frac{2 \sigma_u^2}{\pi}} \mu \sigma_v^2 \mathbb{E} [|\tilde{\mathbf{w}}_{i-1}|^2], \\ &= f \mathbb{E} [|\tilde{\mathbf{w}}_{i-1}|^2] + 15 \mu^2 M \xi_v^6, \end{aligned} \quad (5.41)$$

where

$$\begin{aligned} f &= 1 + 3 \mu \left(15 \mu M \sigma_u^2 \xi_v^4 - 2 \sqrt{\frac{2 \sigma_u^2}{\pi}} \sigma_v^2 \right) + 3 \mu \sigma_u^2 \left(15 \mu M \sigma_u^2 \sigma_v^2 - 2 \sqrt{\frac{2 \sigma_u^2}{\pi}} \right) \mathbb{E} [|\tilde{\mathbf{w}}_{i-1}|^2] \\ &\quad + 15 \mu^2 M \sigma_u^6 (\mathbb{E} [|\tilde{\mathbf{w}}_{i-1}|^2])^2. \end{aligned} \quad (5.42)$$

We see that the transient behavior of the SRLMF algorithm is described by a

nonlinear recursion in $E [||\tilde{\mathbf{w}}_i||^2]$ due to the presence of the factor $E [||\tilde{\mathbf{w}}_{i-1}||^2]$ inside f .

Correlated Input Data

For uncorrelated data, the variance relation (5.36) shows that only unweighted norms of $\tilde{\mathbf{w}}_i$ and $\tilde{\mathbf{w}}_{i-1}$ appear on both sides of the equation. However, for correlated data, different weighing matrices will appear on both sides of (5.36).

If $\Sigma = \mathbf{I}$ in (5.35) we get

$$\begin{aligned} E [||\tilde{\mathbf{w}}_i||^2] &= E [||\tilde{\mathbf{w}}_{i-1}||^2] + \mu^2 \mathcal{Z}_2 E [||\text{sign}[\mathbf{u}_i]||^2] \\ &\quad - \sqrt{\frac{8}{\pi\sigma_u^2}} \mu \mathcal{Z}_1 E [||\tilde{\mathbf{w}}_{i-1}||_{\mathbf{R}}^2]. \end{aligned} \quad (5.43)$$

If $\Sigma = \mathbf{R}$ in (5.35) we get

$$\begin{aligned} E [||\tilde{\mathbf{w}}_i||_{\mathbf{R}}^2] &= E [||\tilde{\mathbf{w}}_{i-1}||_{\mathbf{R}}^2] + \mu^2 \mathcal{Z}_2 E [||\text{sign}[\mathbf{u}_i]||_{\mathbf{R}}^2] \\ &\quad - \sqrt{\frac{8}{\pi\sigma_u^2}} \mu \mathcal{Z}_1 E [||\tilde{\mathbf{w}}_{i-1}||_{\mathbf{R}^2}^2]. \end{aligned} \quad (5.44)$$

Similarly if $\Sigma = \mathbf{R}^{M-1}$ in (5.35) we get

$$\begin{aligned} E [||\tilde{\mathbf{w}}_i||_{\mathbf{R}^{M-1}}^2] &= E [||\tilde{\mathbf{w}}_{i-1}||_{\mathbf{R}^{M-1}}^2] + \mu^2 \mathcal{Z}_2 E [||\text{sign}[\mathbf{u}_i]||_{\mathbf{R}^{M-1}}^2] \\ &\quad - \sqrt{\frac{8}{\pi\sigma_u^2}} \mu \mathcal{Z}_1 E [||\tilde{\mathbf{w}}_{i-1}||_{\mathbf{R}^M}^2]. \end{aligned} \quad (5.45)$$

The term $\mathbb{E} [\|\tilde{\mathbf{w}}_i\|_{\mathbf{R}^M}^2]$ can be inferred from the prior weighting factors

$$\{\mathbb{E} [\|\tilde{\mathbf{w}}_i\|^2], \mathbb{E} [\|\tilde{\mathbf{w}}_i\|_{\mathbf{R}}^2], \mathbb{E} [\|\tilde{\mathbf{w}}_i\|_{\mathbf{R}^2}^2], \dots, \mathbb{E} [\|\tilde{\mathbf{w}}_i\|_{\mathbf{R}^{M-1}}^2]\}, \quad (5.46)$$

by expressing \mathbf{R}^M as a linear combination of its lower-order powers using the Cayley-Hamilton theorem. Thus let $p(x) = \det(x\mathbf{I} - \mathbf{R})$ denote the characteristic polynomial of \mathbf{R} , say

$$p(x) = x^M + p_{M-1}x^{M-1} + p_{M-2}x^{M-2} + \dots + p_1x + p_0. \quad (5.47)$$

Then we know that [2]:

$$\mathbf{R}^M = -p_{M-1}\mathbf{R}^{M-1} - p_{M-2}\mathbf{R}^{M-2} - \dots - p_1\mathbf{R} - p_0\mathbf{I}. \quad (5.48)$$

Using this fact we have

$$\mathbb{E} [\|\tilde{\mathbf{w}}_i\|_{\mathbf{R}^M}^2] = -p_0\mathbb{E} [\|\tilde{\mathbf{w}}_i\|^2] - p_1\mathbb{E} [\|\tilde{\mathbf{w}}_i\|_{\mathbf{R}}^2] - \dots - p_{M-1}\mathbb{E} [\|\tilde{\mathbf{w}}_i\|_{\mathbf{R}^{M-1}}^2]. \quad (5.49)$$

We can collect the above results into a compact vector notation by writing (5.43)–(5.45) as

$$\mathcal{W}_i = \mathcal{F}\mathcal{W}_{i-1} + \mu^2\mathcal{Z}_2\mathcal{Y}, \quad (5.50)$$

and the excess mean-square error is defined as

$$\text{EMSE} \triangleq \lim_{i \rightarrow \infty} \text{E} [|e_{a_i}|^2], \quad (5.54)$$

where

$$\text{E} [|e_{a_i}|^2] = \text{E} [|\tilde{\mathbf{w}}_{i-1}|_{\mathbf{R}}^2]. \quad (5.55)$$

The evolution of $\text{E} [|e_{a_i}|^2]$ is described by the second entry of the state vector \mathcal{W}_i in (5.50). The resulting learning curve of the filter is $\text{E} [|e_i|^2] = \sigma_v^2 + \text{E} [|e_{a_i}|^2]$.

We know that, the mean-square deviation is defined as

$$\text{MSD} \triangleq \lim_{i \rightarrow \infty} \text{E} [|\tilde{\mathbf{w}}_i|^2]. \quad (5.56)$$

The evolution of $\text{E} [|\tilde{\mathbf{w}}_i|^2]$ is described by the first entry of the state vector \mathcal{W}_i in (5.50).

5.5 Conclusion

The transient analysis is challenging due to the presence of the error nonlinearity in the update recursion of the SRLMF algorithm. Nevertheless, by using some simplifying assumptions, the analysis has been carried out in order to provide useful insights about the performance of the proposed algorithm. Also, transient analysis can be more conveniently performed by relying on a weighted energy-conservation relation, as opposed to the unweighted version that was employed

in Chapters 3 and 4. In this chapter, we have extended the weighted variance relation presented in [2] in order to derive expressions for the MSE and the MSD of the SRLMF algorithm during the transient phase for the case of white and correlated input data.

CHAPTER 6

PERFORMANCE ANALYSIS OF THE SRLMF ALGORITHM

6.1 Introduction

In this chapter, simulations are carried out to corroborate the theoretical findings, where it is shown that the theoretical and simulated results are in good agreement. Moreover, the results show that both the SRLMF algorithm and the LMF algorithm have a similar performance for the same misadjustment. The simulations reported in this chapter are based on unknown system identification setup shown in Fig. 1.1 with filter length of five ($M = 5$).

The different performance indexes studied in the previous chapters are tested to find out the closeness of the simulations to the theoretical findings. Mean-square performance, tracking performance, and transient performance of the SRLMF algorithm are all investigated in different scenarios to prove their effectiveness.

6.2 Mean-Square Performance of the SRLMF algorithm

In this section, the SRLMF algorithm is compared with the LMF algorithm in terms of convergence rate. The comparison is made in the presence of three different noise environments namely Gaussian, uniform, and Laplacian and for three different signal-to-noise ratios equal to 0 dB, 10 dB, and 20 dB. It has been shown that both the SRLMF algorithm and the LMF algorithm take approximately the same number of iterations to converge to the same steady-state value in all the studied noise environments and signal-to-noise ratios.

In order to compare the performance of adaptive filters, it is customary to adopt a common performance measure across filters. The practice that is most widely used in the literature of adaptive filtering is to have the same misadjustment, which is defined as

$$\mathcal{M} \triangleq \frac{\text{EMSE}}{J_{\min}}, \quad (6.1)$$

where J_{\min} is the minimum value of the cost function which is equal to the variance of the noise v_i , i.e., $J_{\min} = \sigma_v^2$. Therefore, setting the misadjustment of the SRLMF algorithm equal to that of the LMF algorithm gives:

$$\frac{\mu_{\text{SRLMF}} \xi_v^6}{6\sigma_v^2} \sqrt{\frac{2}{\pi\sigma_u^2}} \text{Tr}(\mathbf{R}) = \frac{\mu_{\text{LMF}} \xi_v^6}{6\sigma_v^2} \text{Tr}(\mathbf{R}), \quad (6.2)$$

which can be used to solve for the step-size of the SRLMF algorithm in terms of

the LMF algorithm

$$\mu_{\text{SRLMF}} = \sqrt{\frac{\pi\sigma_u^2}{2}}\mu_{\text{LMF}}. \quad (6.3)$$

Assume regressor variance is $\sigma_u^2 = 1$. Therefore, the step-size of the SRLMF algorithm used in this part of the simulations is given by:

$$\mu_{\text{SRLMF}} = \sqrt{\frac{\pi}{2}}\mu_{\text{LMF}}. \quad (6.4)$$

Hence, whatever value of μ_{LMF} used, the step-size corresponding to the SRLMF algorithm is set by (6.4).

Figures 6.1-6.22 use the following specifications: Let us consider a real-valued white Gaussian regression sequence. Run the filter for 12000 iterations and the MSE learning curve is the average over 1000 independent runs. The unknown system is characterized by the following channel:

$$\mathbf{w}^o = [0.227 \quad 0.460 \quad 0.688 \quad 0.460 \quad 0.227]^T. \quad (6.5)$$

The results in Figs. 6.1, 6.3 and 6.5 show that both the SRLMF algorithm and the LMF algorithm converge to the same steady-state value in approximately 1000, 6000, and 9000 iterations in an additive white Gaussian noise (AWGN) environment with signal-to-noise ratios equal to 0 dB, 10 dB, and 20 dB, respectively. The result in Fig. 6.2 shows that the behavior of the third-tap weight learning curves is same for both the algorithms in an AWGN environment with SNR = 0 dB. However, this behavior of the third-tap weight learning curve gets slightly bet-

ter for the LMF algorithm than the SRLMF algorithm for higher signal-to-noise ratios as seen from Figs. 6.4 and 6.6.

The result in Fig. 6.7 shows that both the SRLMF algorithm and the LMF algorithm converge to the same steady-state value in approximately 8000 iterations when there is a sudden burst in an AWGN environment with $\text{SNR} = 20$ dB. Also, the behavior of the third-tap weight learning curve gets slightly better for the LMF algorithm than the SRLMF algorithm as seen from Fig. 6.8.

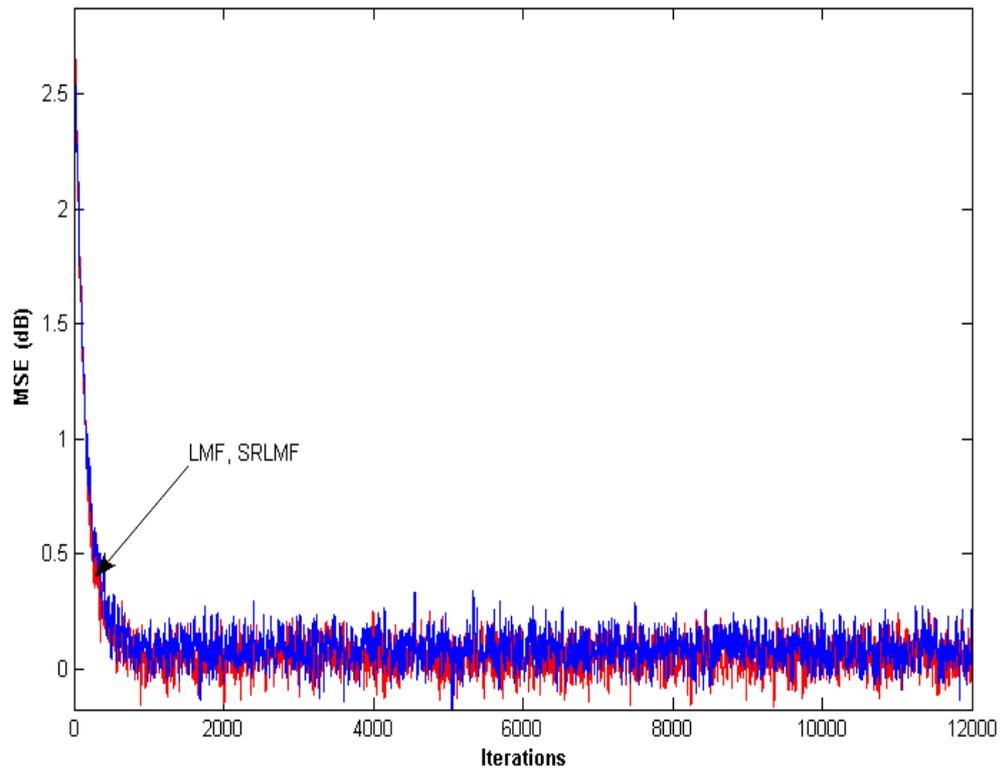


Figure 6.1: Comparison of the MSE learning curves of LMF and SRLMF algorithms in an AWGN environment with $\text{SNR}=0$ dB.

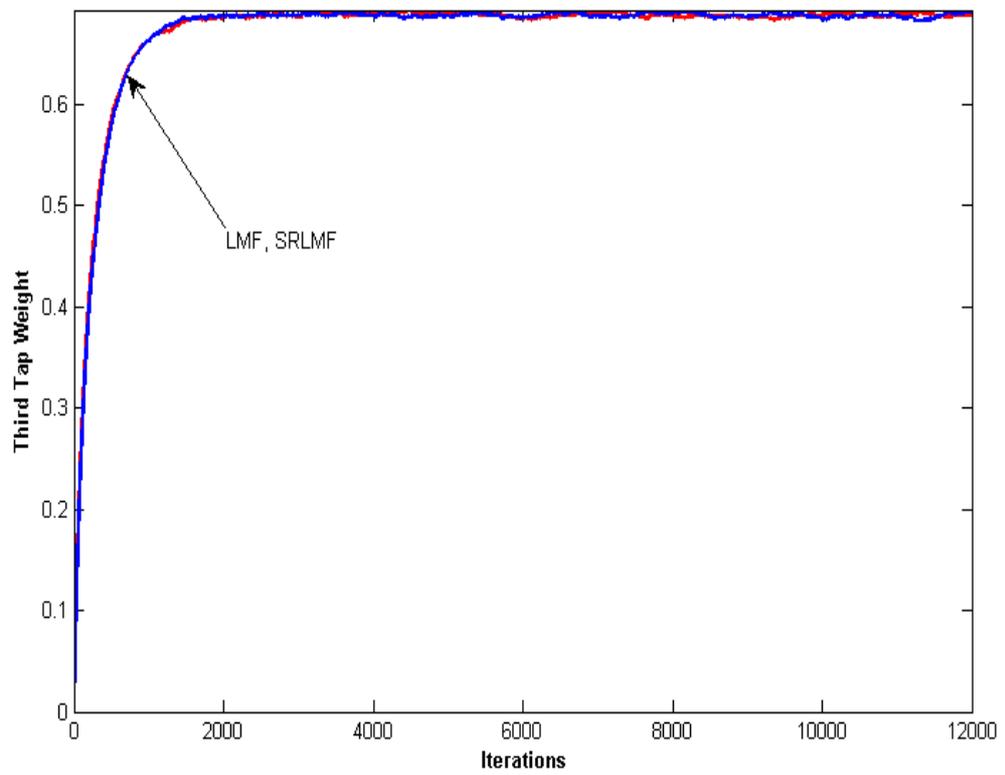


Figure 6.2: Comparison of the third-tap weight learning curves of LMF and SRLMF algorithms in an AWGN environment with SNR=0 dB.

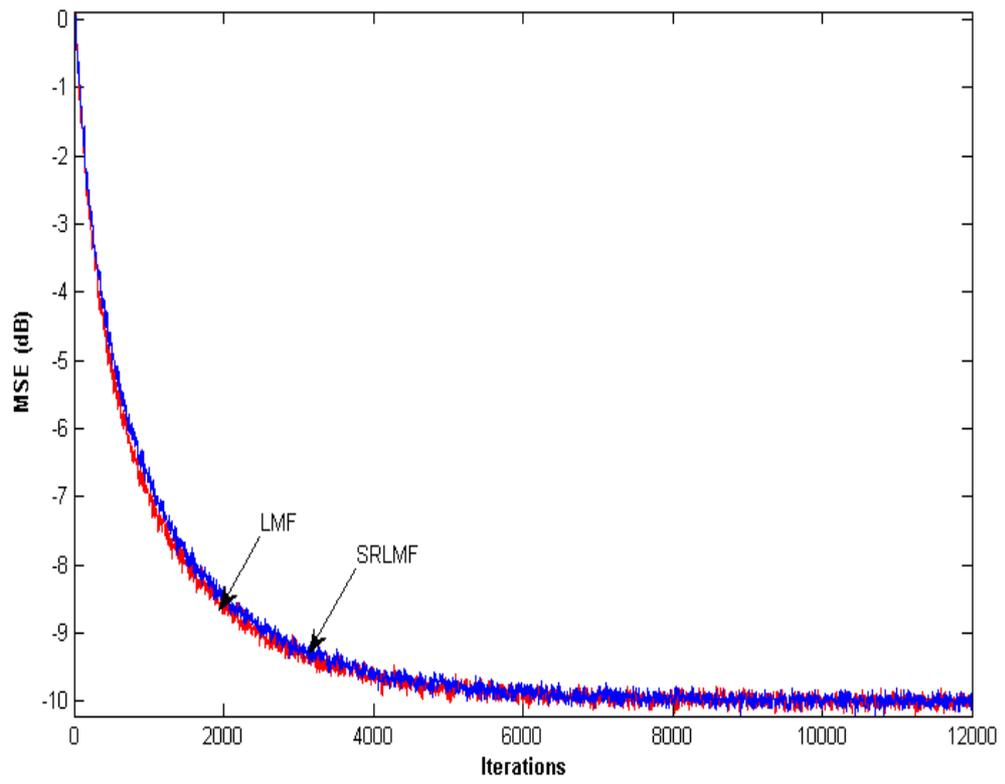


Figure 6.3: Comparison of the MSE learning curves of LMF and SRLMF algorithms in an AWGN environment with SNR=10 dB.

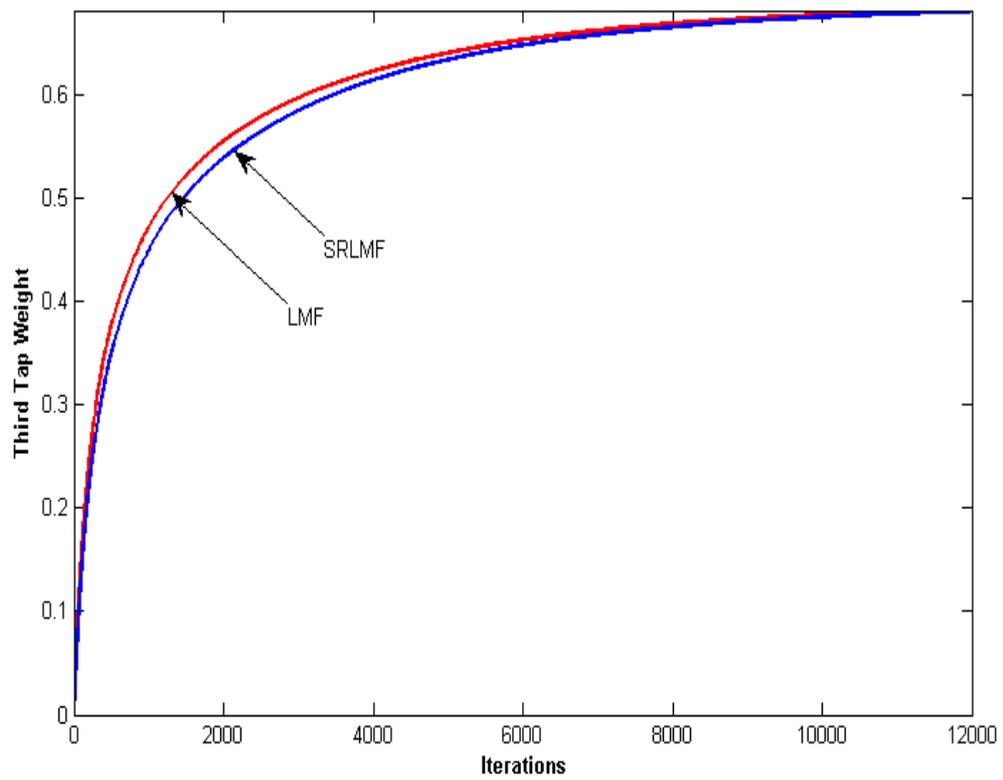


Figure 6.4: Comparison of the third-tap weight learning curves of LMF and SRLMF algorithms in an AWGN environment with SNR=10 dB.

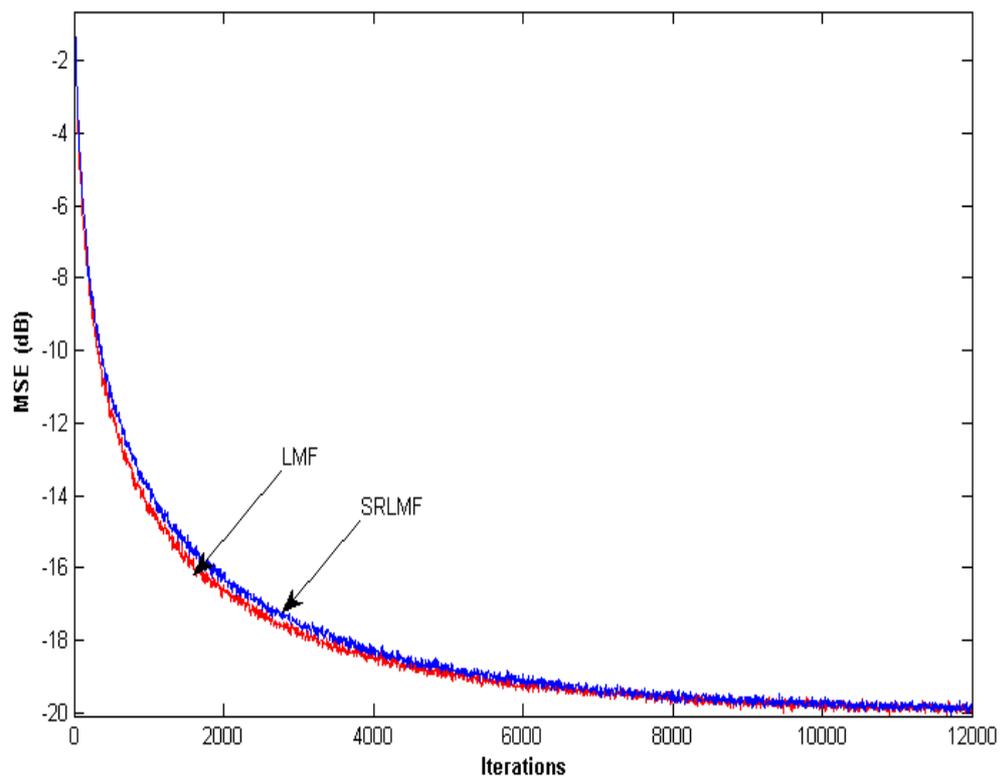


Figure 6.5: Comparison of the MSE learning curves of LMF and SRLMF algorithms in an AWGN environment with SNR=20 dB.

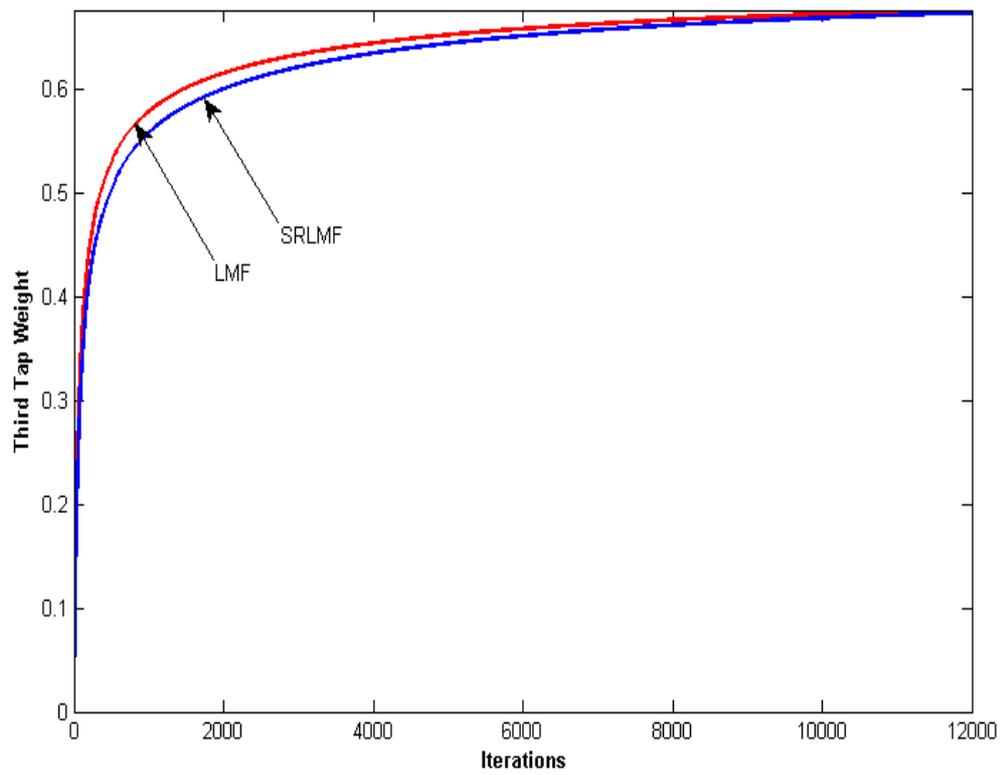


Figure 6.6: Comparison of the third-tap weight learning curves of LMF and SRLMF algorithms in an AWGN environment with SNR=20 dB.

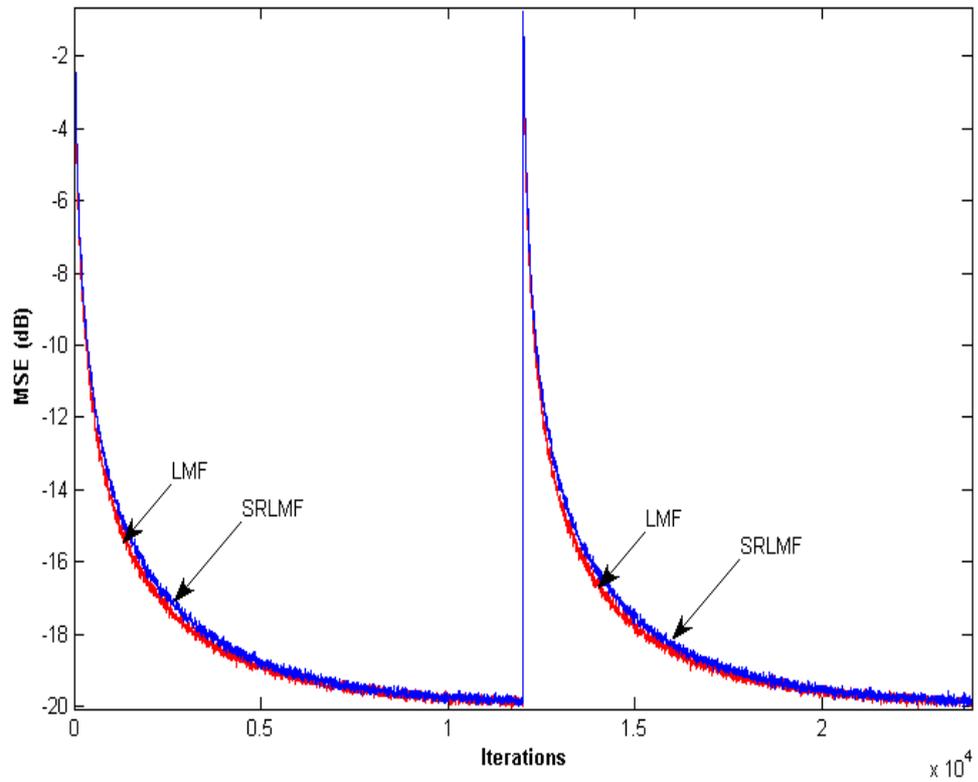


Figure 6.7: Comparison of the MSE learning curves of LMF and SRLMF algorithms when there is a sudden burst in an AWGN environment with SNR=20 dB.

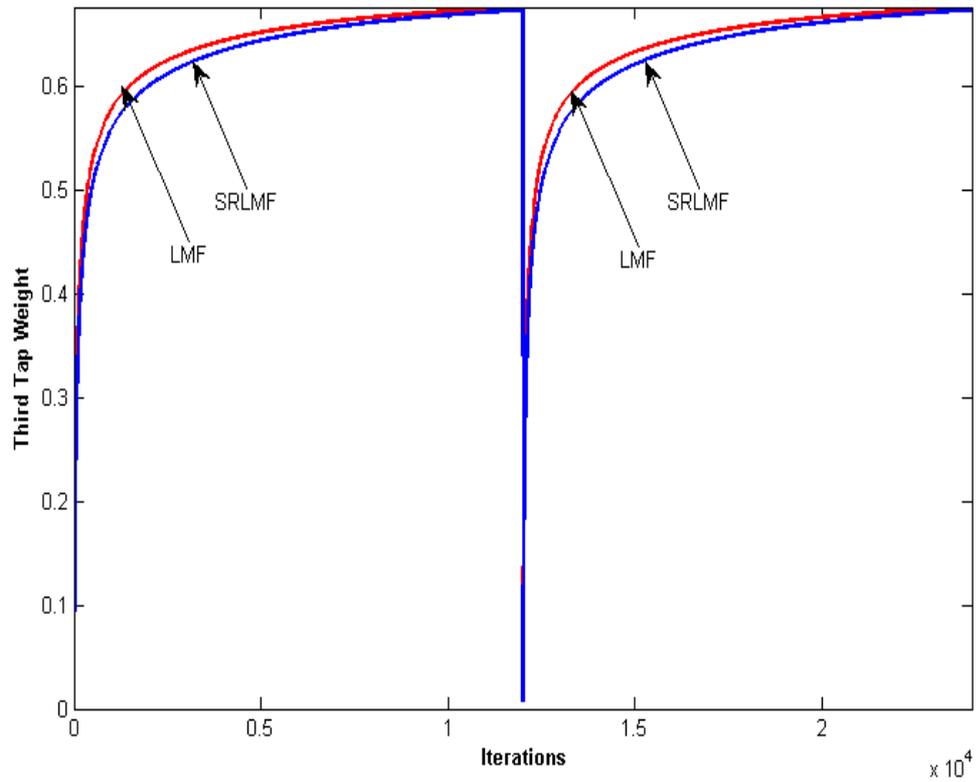


Figure 6.8: Comparison of the third-tap weight learning curves of LMF and SRLMF algorithms when there is a sudden burst in an AWGN environment with SNR=20 dB.

The results in Figs. 6.9, 6.11 and 6.13 show that both the SRLMF algorithm and the LMF algorithm converge to the same steady-state value in approximately 1000, 6000, and 10000 iterations in uniform noise environment with signal-to-noise ratios equal to 0 dB, 10 dB, and 20 dB, respectively. The result in Fig. 6.10 shows that the behavior of the third-tap weight learning curves is same for both the algorithms in uniform noise environment with $\text{SNR} = 0$ dB. However, this behavior of the third-tap weight learning curve gets slightly better for the LMF algorithm than the SRLMF algorithm for higher signal-to-noise ratios as seen from Figs. 6.12 and 6.14.

The results in Figs. 6.15, 6.17 and 6.19 show that both the SRLMF algorithm and the LMF algorithm converge to the same steady-state value in approximately 1500, 7000, and 11000 iterations in Laplacian noise environment with signal-to-noise ratios equal to 0 dB, 10 dB, and 20 dB, respectively. The result in Fig. 6.16 shows that the behavior of the third-tap weight learning curves is same for both the algorithms in Laplacian noise environment with $\text{SNR} = 0$ dB. However, this behavior of the third-tap weight learning curve gets slightly better for the LMF algorithm than the SRLMF algorithm for higher signal-to-noise ratios as seen from Figs. 6.18 and 6.20.

The result in Fig. 6.21 shows that SRLMF converges to the same steady-state value in approximately 1500, 3000, and 8000 iterations in uniform, Gaussian, and Laplacian noise environments, respectively, with $\text{SNR} = 10$ dB. Also, the behavior of the third-tap weight learning curve is better for uniform noise environment as

compared to Gaussian and Laplacian noise environments as seen from Fig. 6.22.

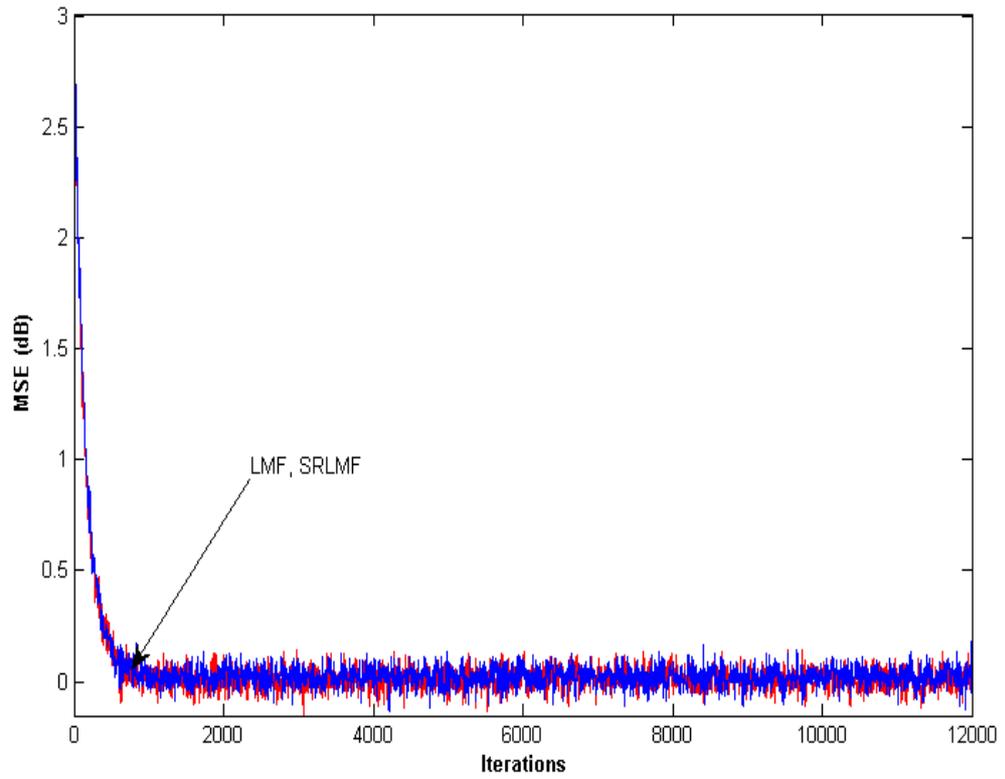


Figure 6.9: Comparison of the MSE learning curves of LMF and SRLMF algorithms in uniform noise environment with SNR=0 dB.

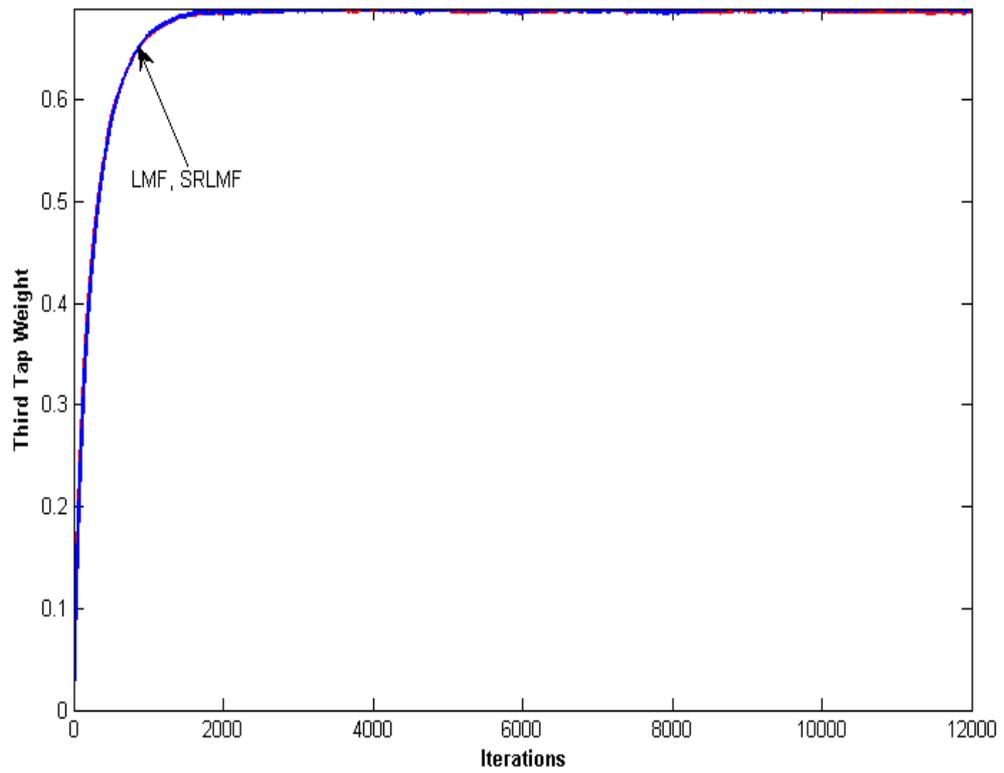


Figure 6.10: Comparison of the third-tap weight learning curves of LMF and SRLMF algorithms in uniform noise environment with SNR=0 dB.

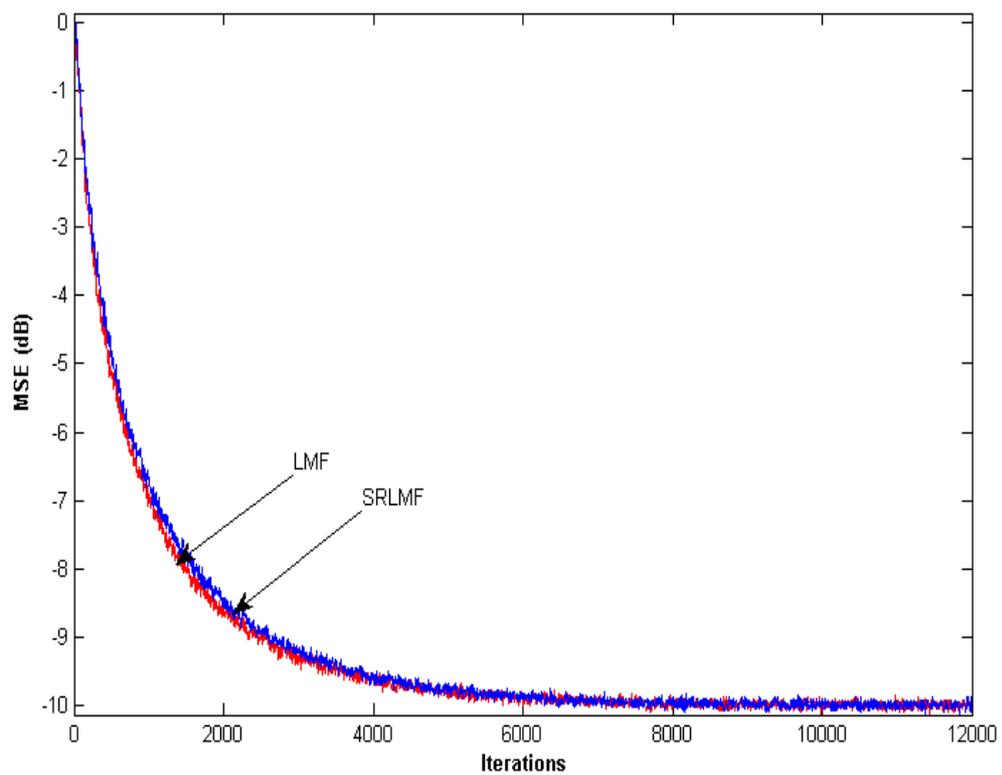


Figure 6.11: Comparison of the MSE learning curves of LMF and SRLMF algorithms in uniform noise environment with SNR=10 dB.

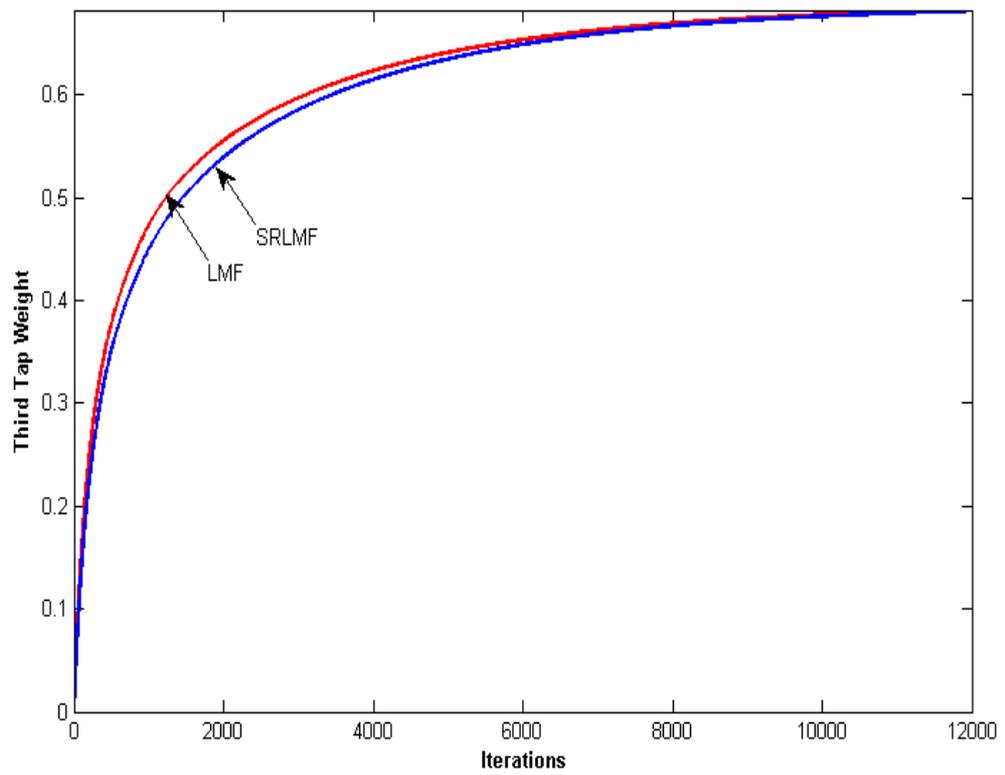


Figure 6.12: Comparison of the third-tap weight learning curves of LMF and SRLMF algorithms in uniform noise environment with SNR=10 dB.

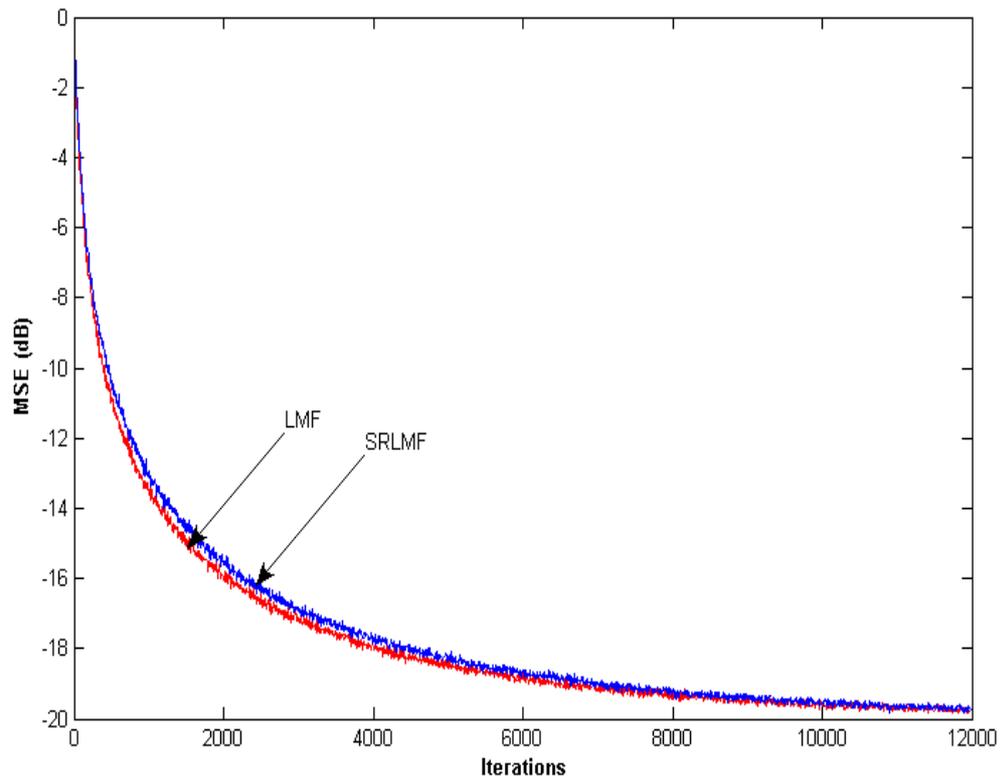


Figure 6.13: Comparison of the MSE learning curves of LMF and SRLMF algorithms in uniform noise environment with SNR=20 dB.

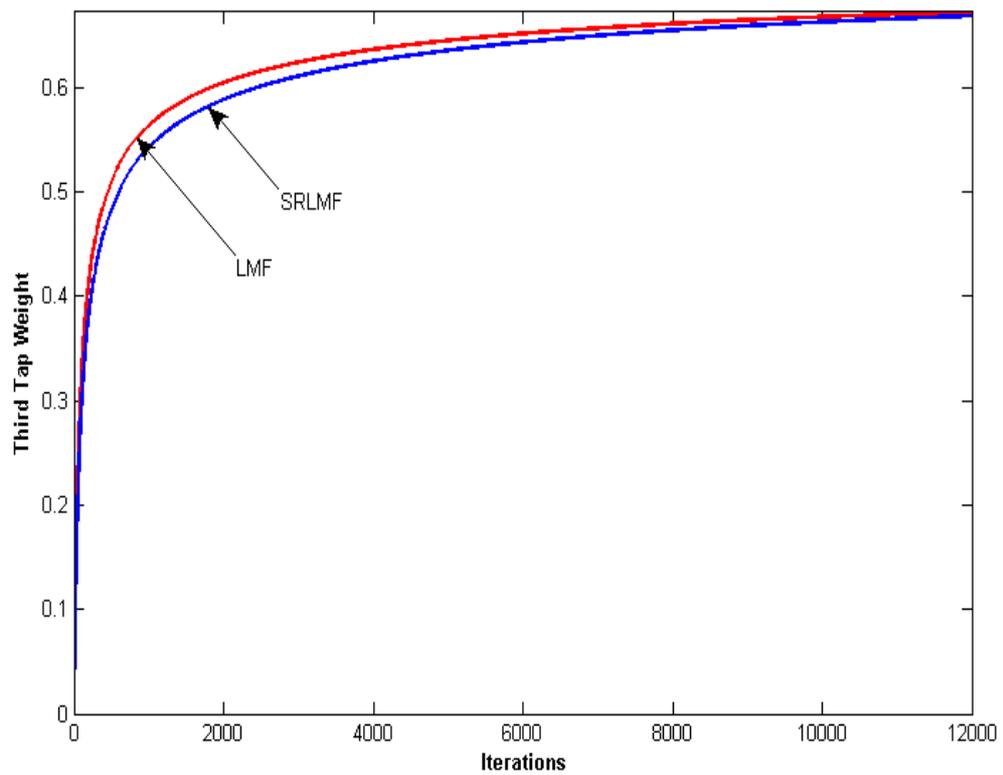


Figure 6.14: Comparison of the third-tap weight learning curves of LMF and SRLMF algorithms in uniform noise environment with SNR=20 dB.

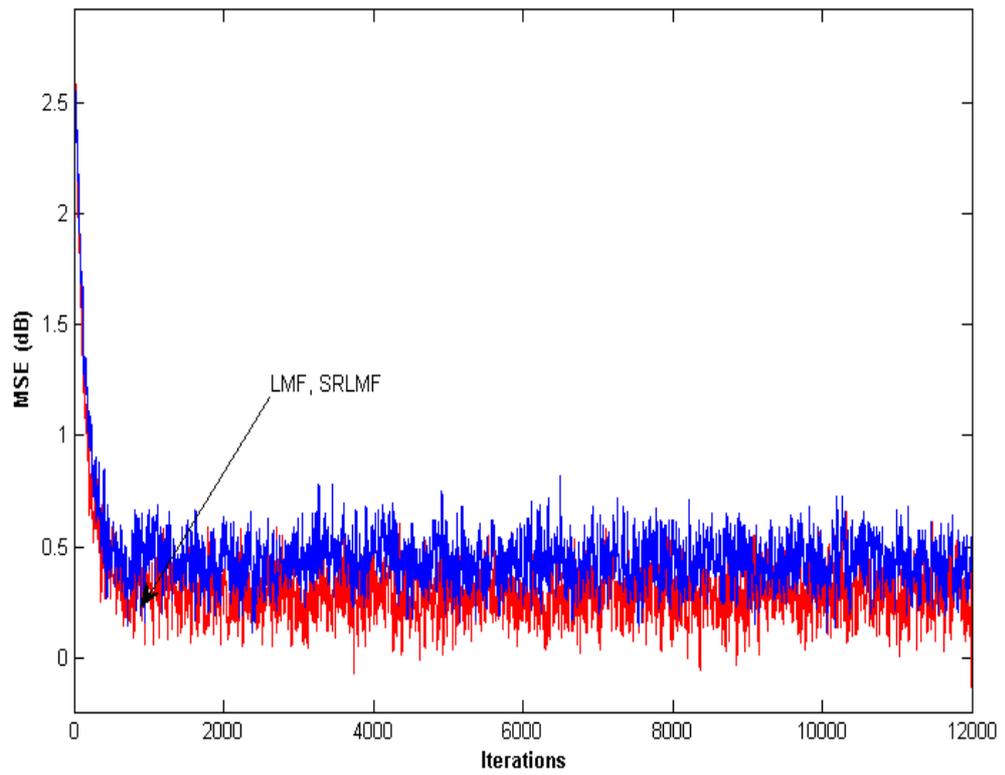


Figure 6.15: Comparison of the MSE learning curves of LMF and SRLMF algorithms in Laplacian noise environment with SNR=0 dB.

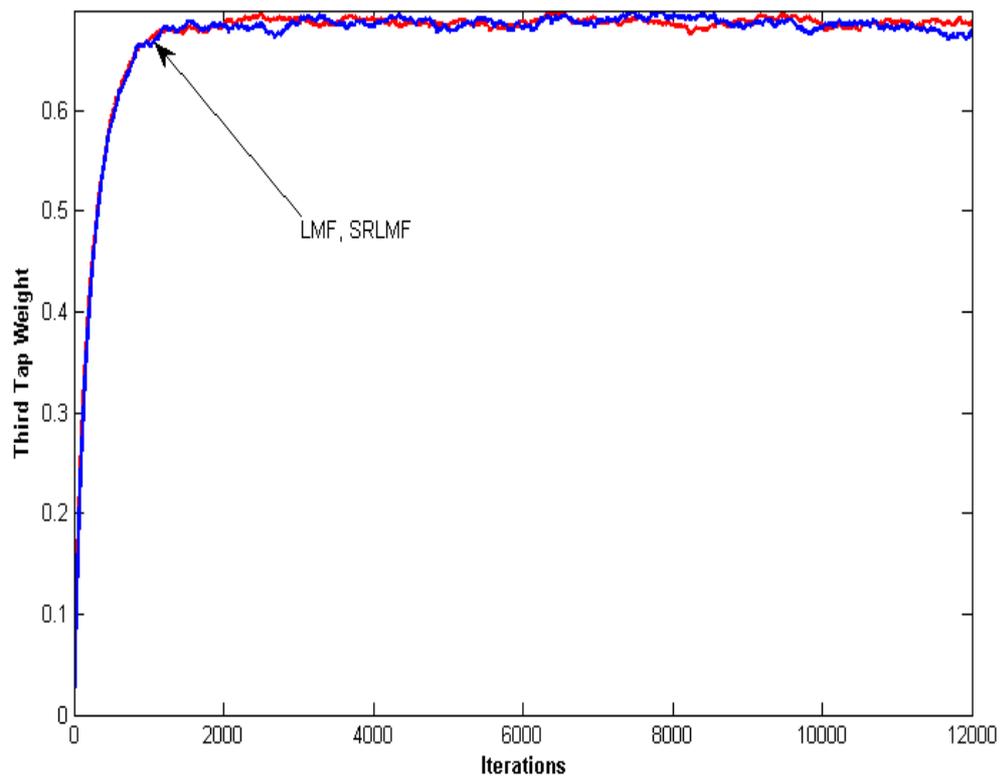


Figure 6.16: Comparison of the third-tap weight learning curves of LMF and SRLMF algorithms in Laplacian noise environment with SNR=0 dB.

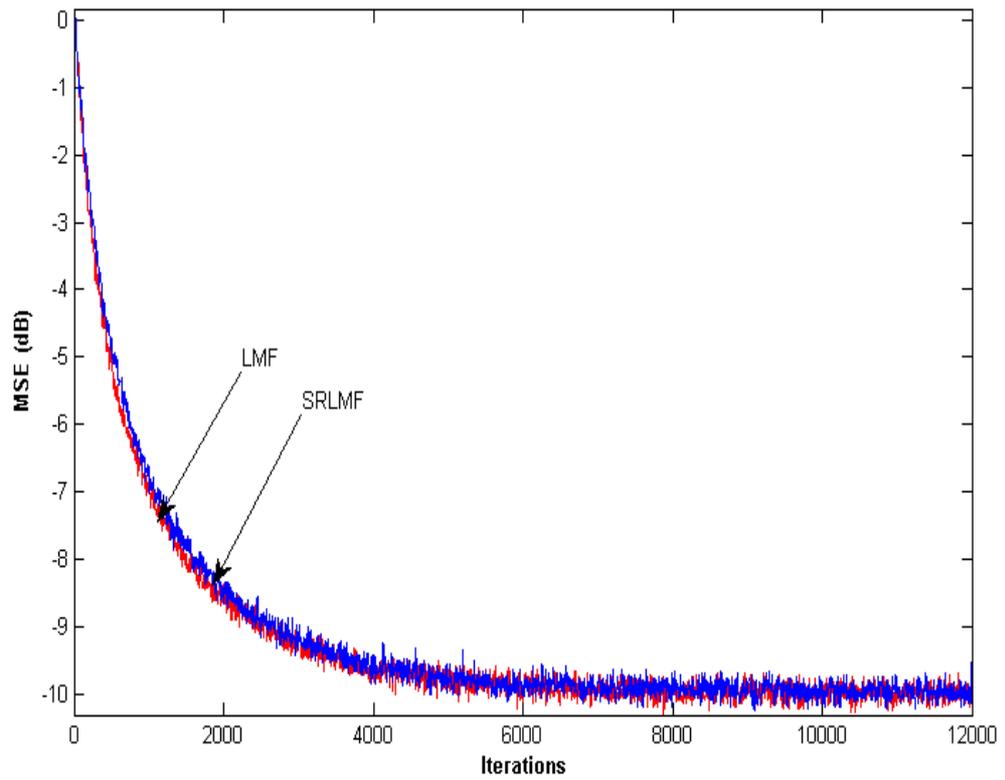


Figure 6.17: Comparison of the MSE learning curves of LMF and SRLMF algorithms in Laplacian noise environment with SNR=10 dB.

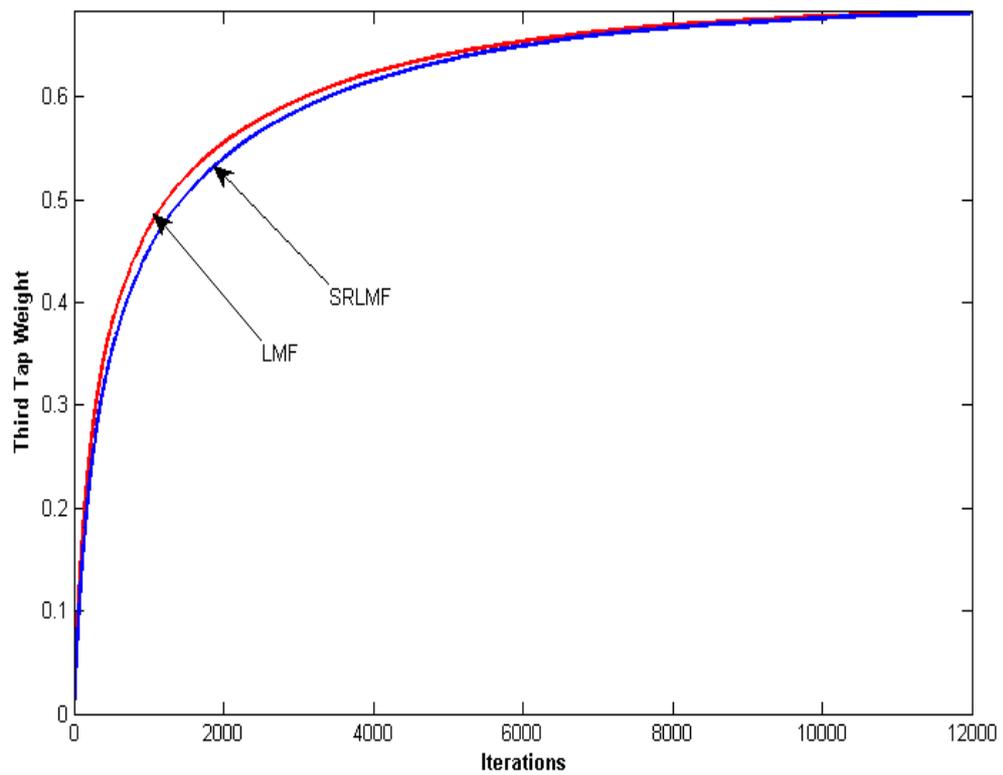


Figure 6.18: Comparison of the third-tap weight learning curves of LMF and SRLMF algorithms in Laplacian noise environment with SNR=10 dB.

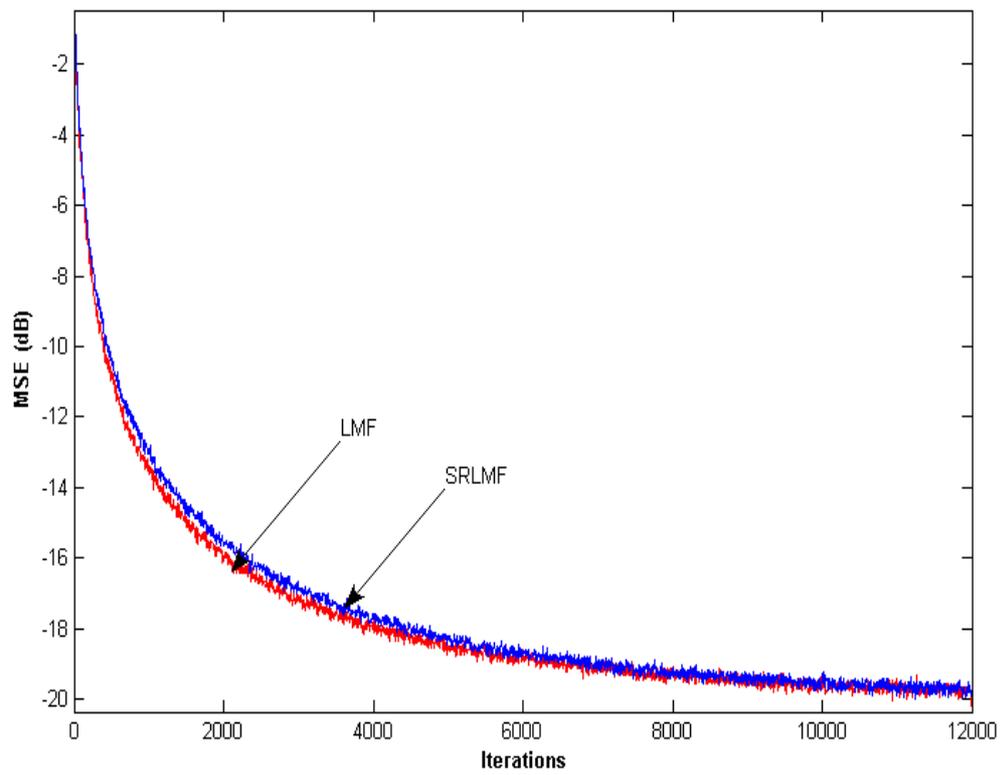


Figure 6.19: Comparison of the MSE learning curves of LMF and SRLMF algorithms in Laplacian noise environment with SNR=20 dB.

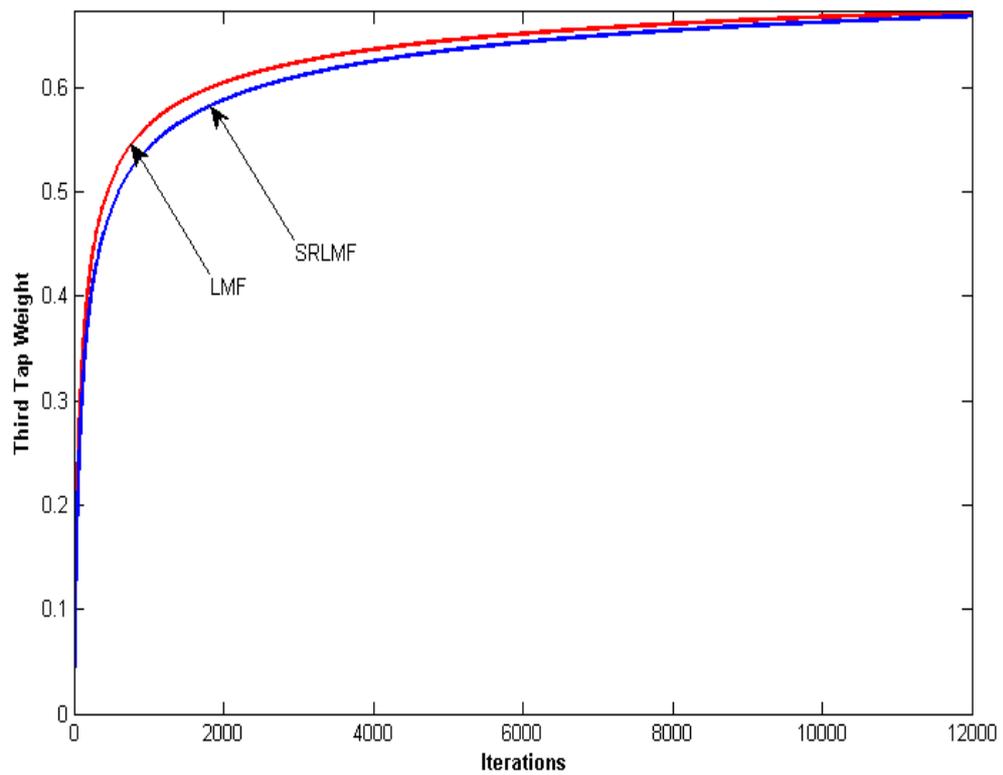


Figure 6.20: Comparison of the third-tap weight learning curves of LMF and SRLMF algorithms in Laplacian noise environment with SNR=20 dB.

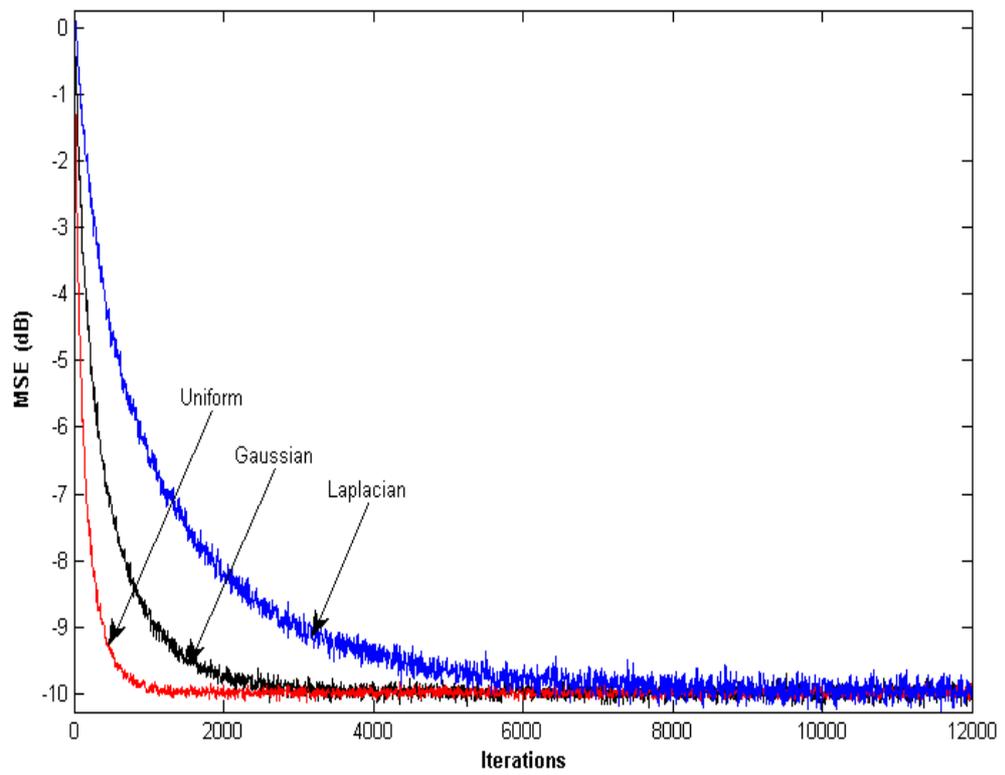


Figure 6.21: Comparison of the MSE learning curves of the SRLMF algorithm in Gaussian, uniform and Laplacian noise environments with SNR=10 dB.

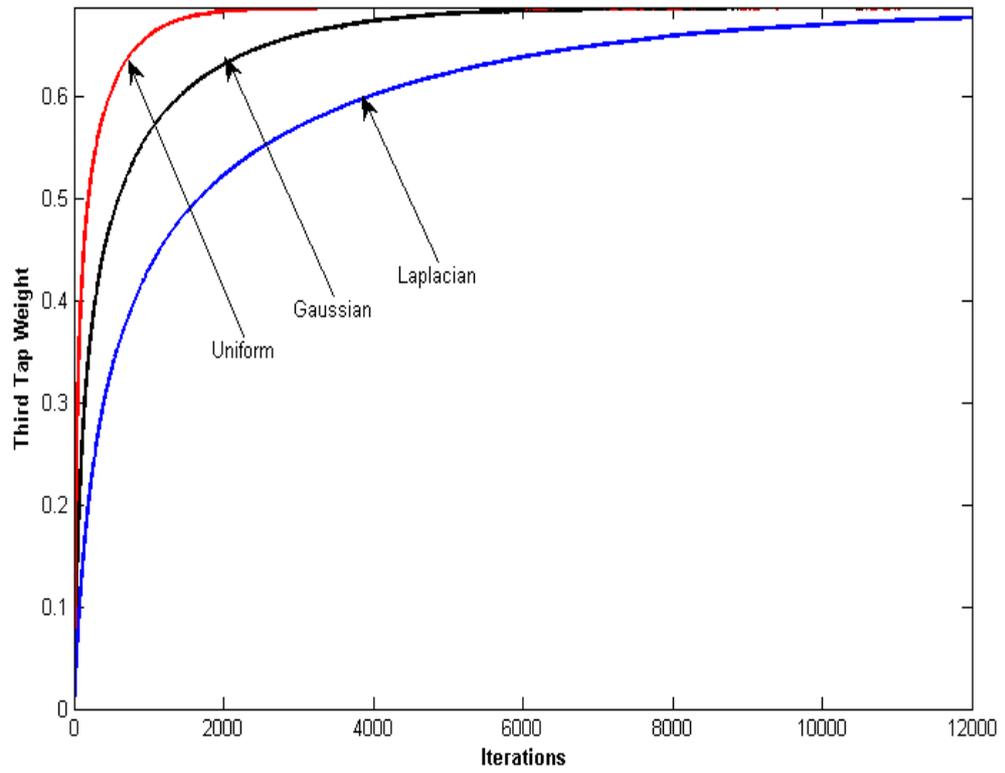


Figure 6.22: Comparison of the third-tap weight learning curves of the SRLMF algorithm in Gaussian, uniform and Laplacian noise environments with SNR=10 dB.

In order to validate the theoretical findings extensive simulations are carried out for different scenarios. Figs. 6.23-6.25 use the following specifications: Run the filter for 1×10^6 iterations, the MSE learning curve is the average over 100 independent runs and the SNR is set to 30 dB. The average of the last 1×10^5 entries of the ensemble-average curve is used as the experimental value for the MSE.

In the case of Fig. 6.23, the regressors, with shift structure, are generated by feeding a unit-variance white process into a tapped delay line. However, in Fig. 6.24, the regressors, with shift structure, are generated by passing correlated data into a tapped delay line. Here, the correlated data are obtained by passing a unit-variance i.i.d. Gaussian data through a first-order auto-regressive model with transfer function $\frac{\sqrt{1-a^2}}{(1-az^{-1})}$ and $a = 0.8$. To further test the validity of the results, Gaussian regressors with an eigenvalue spread of five without a shift structure are used, this is depicted in Fig. 6.25. As it can be seen from these figures, the simulation results match very well the theoretical results ((3.34) and (3.36)).

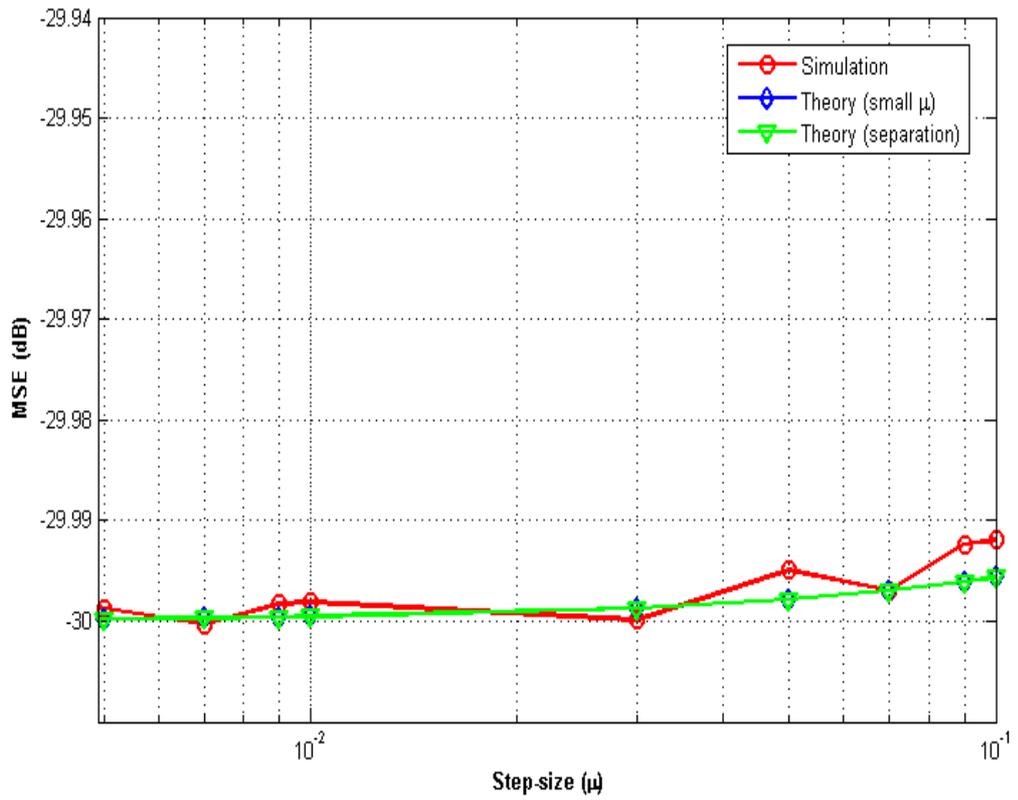


Figure 6.23: Theoretical and simulated MSE of the SRLMF algorithm using white Gaussian regressors with shift structure with SNR=30 dB.

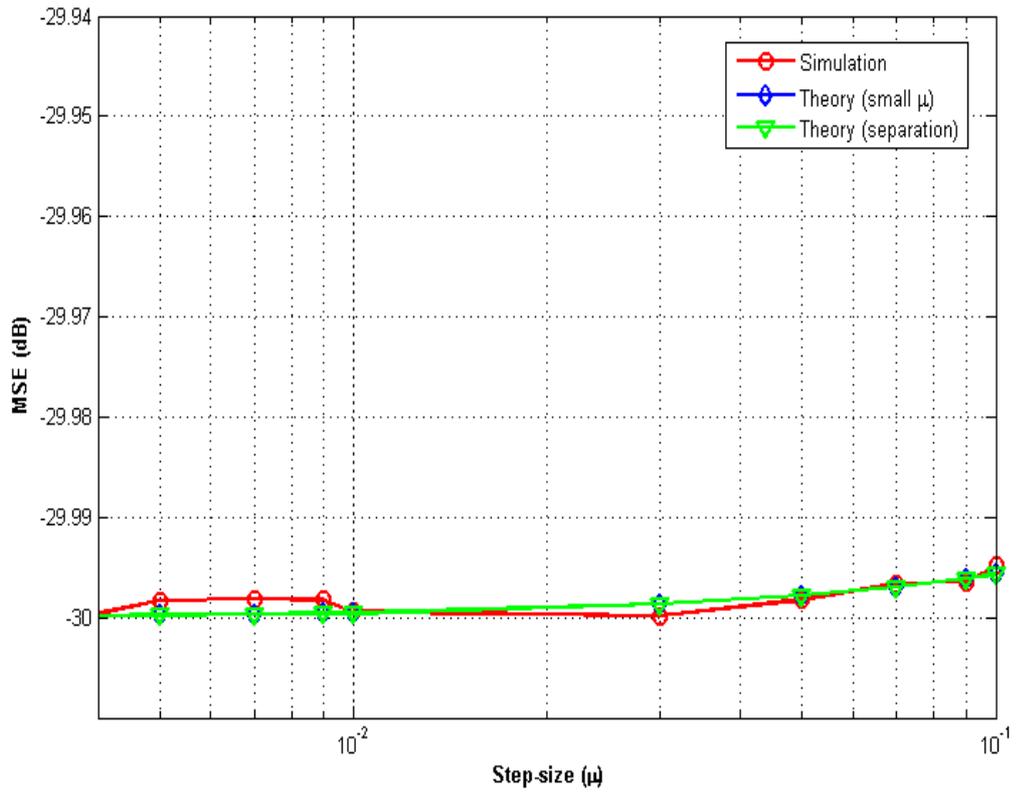


Figure 6.24: Theoretical and simulated MSE of the SRLMF algorithm using correlated Gaussian regressors with shift structure with SNR=30 dB.

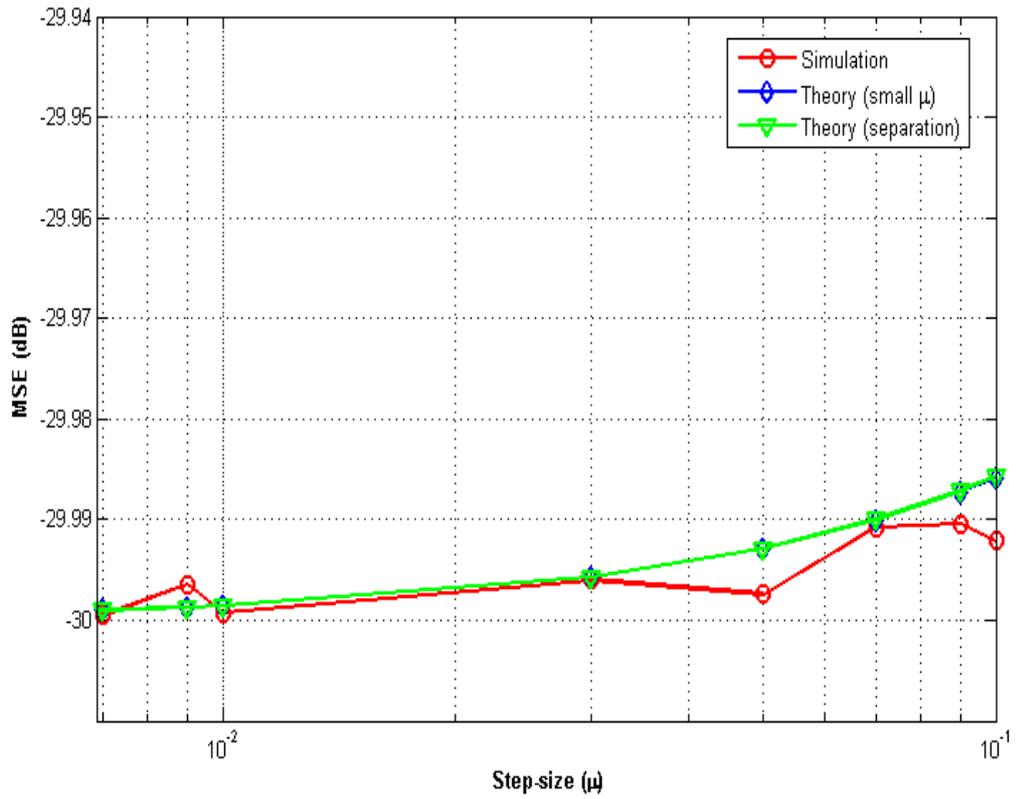


Figure 6.25: Theoretical and simulated MSE of the SRLMF algorithm using Gaussian regressors with an eigenvalue spread=5 without shift structure with SNR=30 dB.

6.3 Tracking Performance of the SRLMF algorithm

In this section, the tracking performance of the SRLMF algorithm is investigated for a random-walk and Rayleigh fading channels. All the simulations in this section use the following common specifications: Let us consider a real-valued white Gaussian regression sequence. Run the filter for 50000 iterations and the MSE learning curve is the average over 50 independent runs.

6.3.1 Random-Walk Channel

Here, the random-walk channel behaves according to

$$\mathbf{w}_i^o = \mathbf{w}_{i-1}^o + \mathbf{q}_i, \quad (6.6)$$

where \mathbf{q}_i is a Gaussian sequence with zero mean and variance $\sigma_q^2 = 10^{-9}$ and $\mathbf{w}_{-1}^o = [0.227 \ 0.460 \ 0.688 \ 0.460 \ 0.227]^T$.

The result in Fig. 6.26 shows the MSE as a function of the step-size when the SRLMF algorithm is used to track a random-walk channel (6.6) with an SNR = 30 dB. As observed from Fig. 6.26, the simulation results corroborate closely the theoretical results ((4.14) and (4.16)).

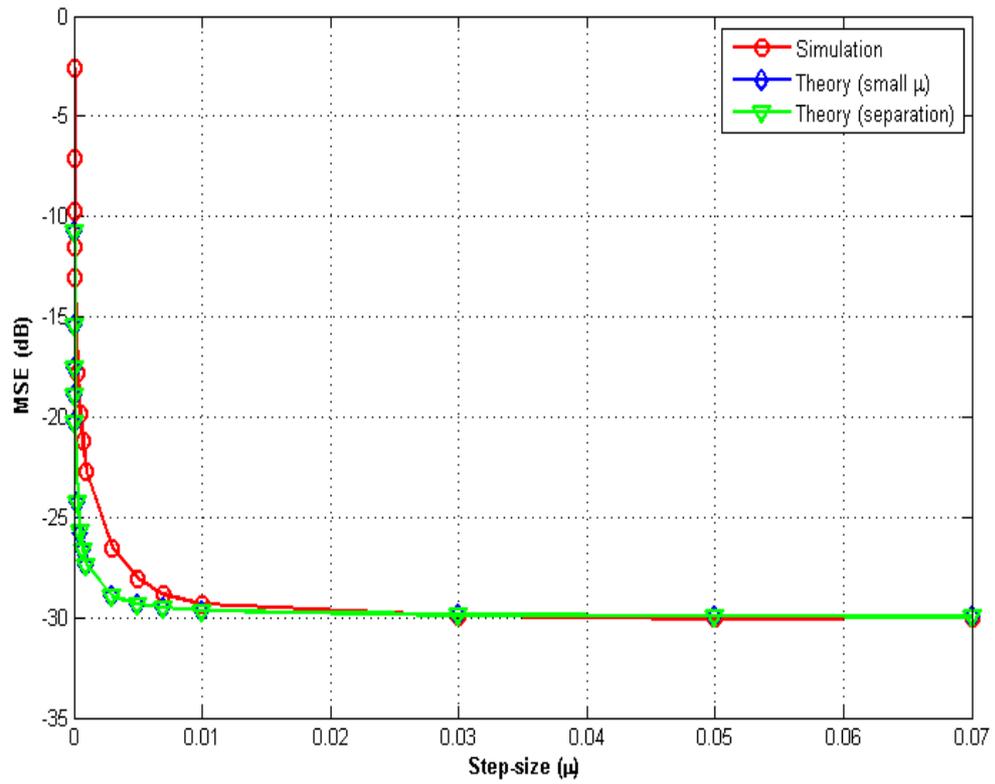


Figure 6.26: Theoretical and simulated MSE of the SRLMF algorithm for a random-walk channel as a function of the step-size with SNR=30 dB.

6.3.2 Rayleigh Fading Channel

Single-path

Let us consider a wireless channel with one Rayleigh fading ray, which is assumed to fade at a Doppler frequency of $f_D = 10\text{Hz}$. Let us fix the sampling period at $T_s = 0.025\mu\text{s}$. Let the SNR be 30 dB. The weight vector we wish to estimate has the form:

$$[0 \quad 0 \quad x_1(n) \quad 0 \quad 0], \quad (6.7)$$

where $x_1(n)$ represents the second Rayleigh fading ray.

The result in Fig. 6.27 shows the MSE as a function of the step-size when the SRLMF algorithm is used to track a single-path Rayleigh fading channel (6.7). The theoretical values are obtained by using the expressions ((4.14) and (4.16)). It is seen that there is a good match between between the theoretical and simulated results.

The result in Fig. 6.28 shows the MSE as a function of the Doppler frequency over the range 10Hz to 20Hz. Here the step-size is fixed at $\mu = 0.01$. The theoretical values are obtained by using the expressions ((4.14) and (4.16)). It is seen that as the Doppler frequency increases, the tracking performance of the SRLMF algorithm deteriorates. This behavior is expected since higher Doppler frequencies correspond to faster variations in the channel.

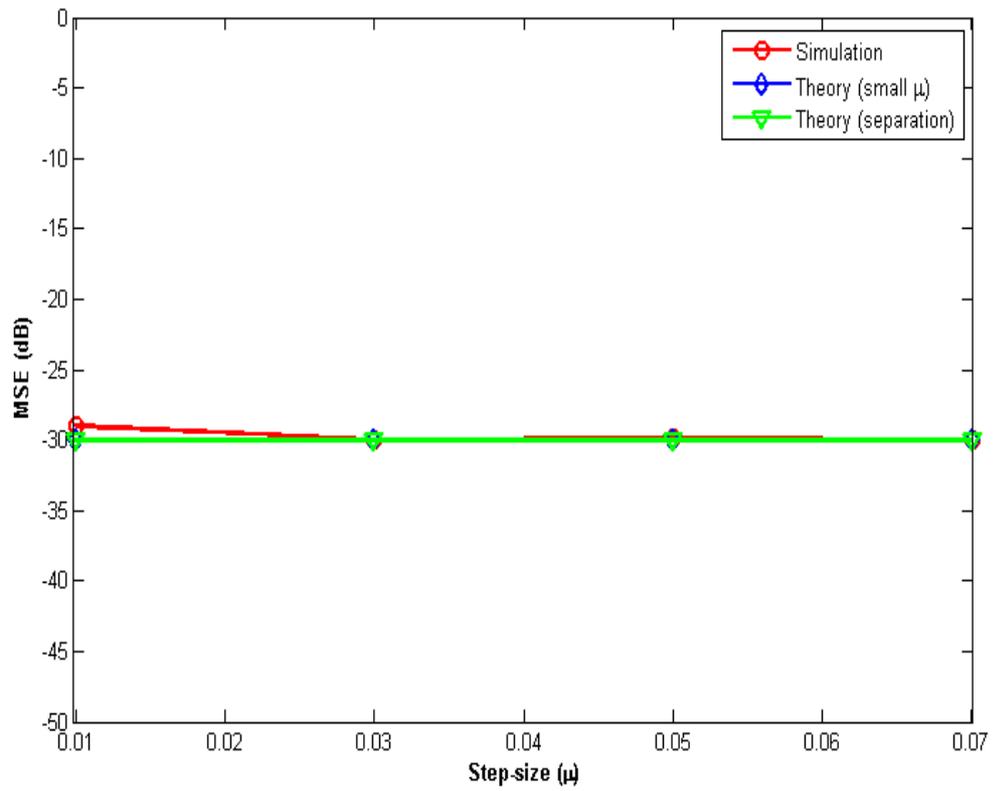


Figure 6.27: Theoretical and simulated MSE of the SRLMF algorithm for a single-path Rayleigh fading channel as a function of the step-size with SNR=30 dB.

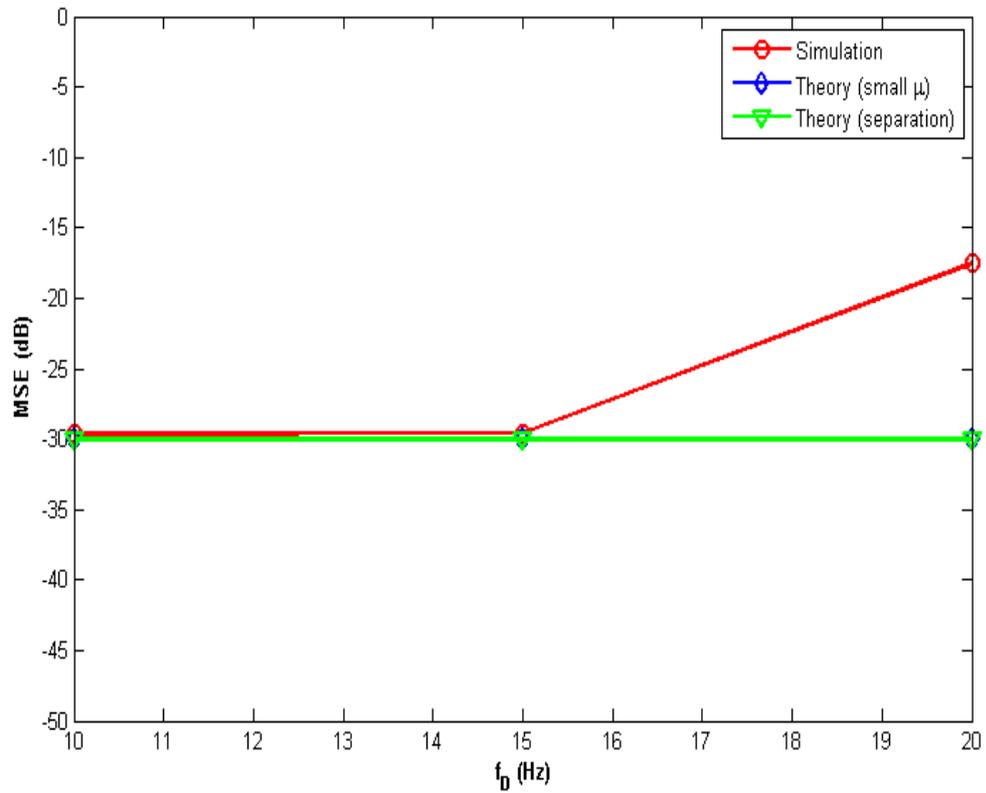


Figure 6.28: Theoretical and simulated MSE of the SRLMF algorithm for a single-path Rayleigh fading channel as a function of the Doppler frequency with SNR=30 dB and step-size=0.01.

Multipath

Let us consider a wireless channel with two Rayleigh fading rays; both rays are assumed to fade at the same Doppler frequency. The weight vector we wish to estimate has the form:

$$[0 \quad 0 \quad x_1(n) \quad 0 \quad x_2(n)], \quad (6.8)$$

where $x_1(n)$ and $x_2(n)$ represent the second and fourth Rayleigh fading rays respectively.

The result in Fig. 6.29 shows the MSE as a function of the step-size when the SRLMF algorithm is used to track a multipath Rayleigh fading channel (6.8). It is seen that the tracking performance of the SRLMF algorithm deteriorates for the multipath case.

The result in Fig. 6.30 shows the MSE as a function of the Doppler frequency over the range 15Hz to 20Hz. It is seen that as the Doppler frequency increases, the tracking performance of the SRLMF algorithm deteriorates.

The result in Fig. 6.31 shows a typical trajectory of the amplitude of the first Rayleigh fading ray and its estimate by the SRLMF algorithm operating with a step-size $\mu = 0.01$ and SNR = 20 dB. Here, the channel rays are assumed to fade at $f_D = 10\text{Hz}$ and the sampling period is fixed at $T_s = 0.8\mu\text{s}$.

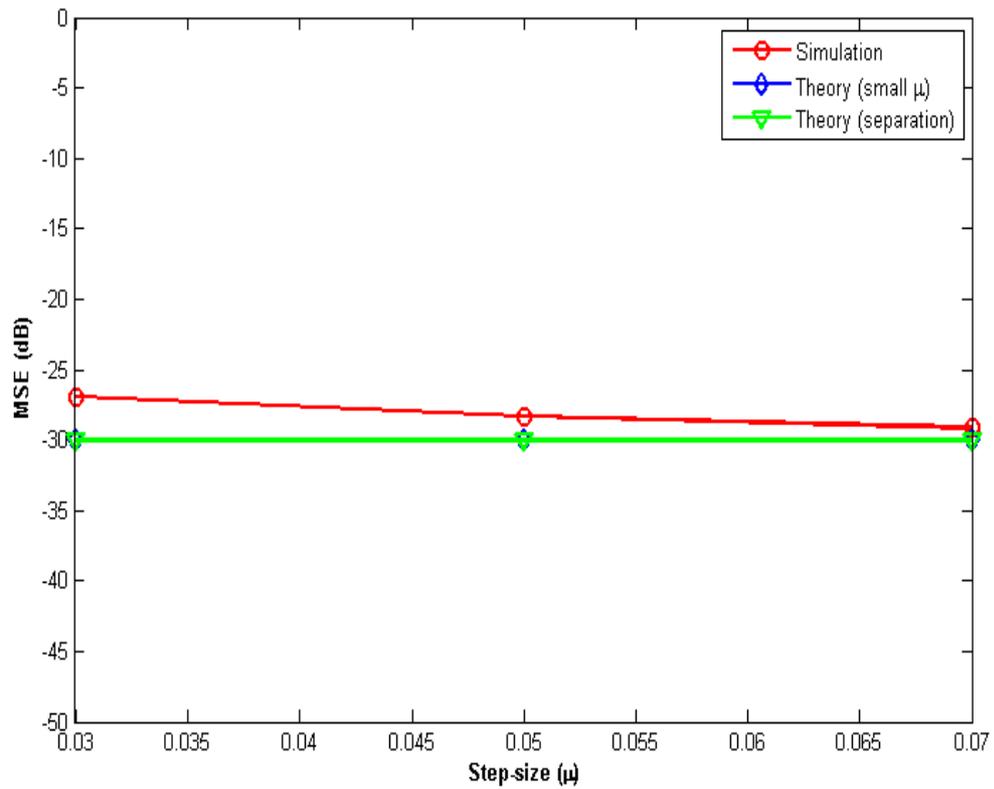


Figure 6.29: Theoretical and simulated MSE of the SRLMF algorithm for a multipath Rayleigh fading channel as a function of the step-size with SNR=30 dB.

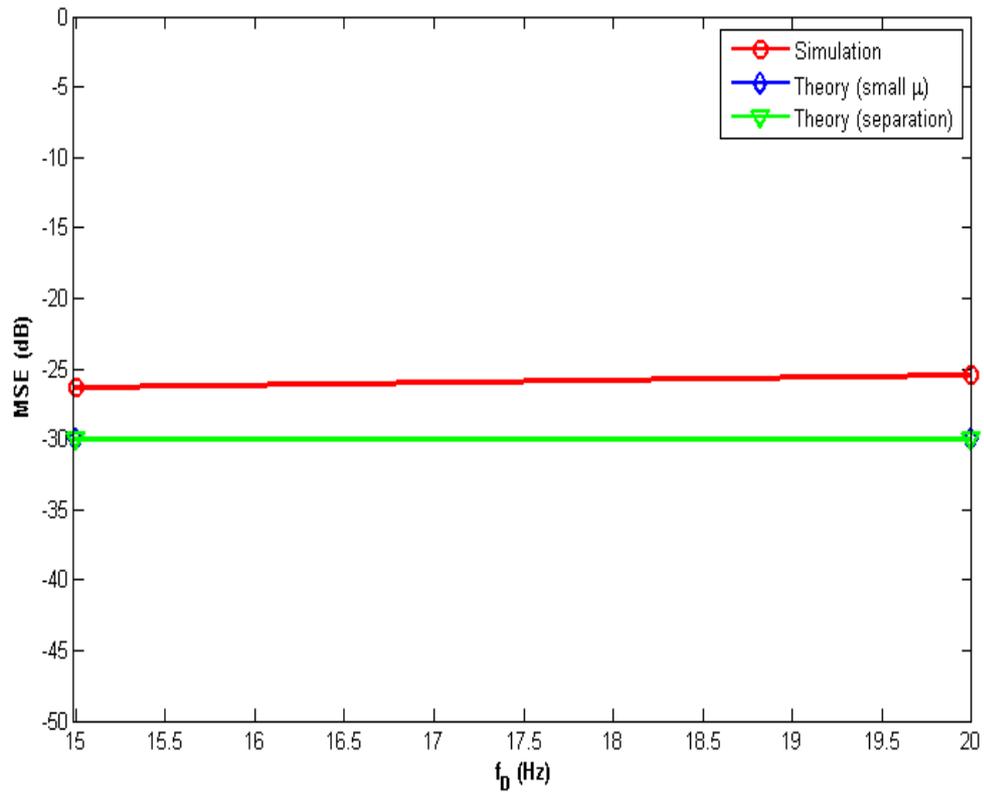


Figure 6.30: Theoretical and simulated MSE of the SRLMF algorithm for a multipath Rayleigh fading channel as a function of the Doppler frequency with SNR=30 dB and step-size=0.01.

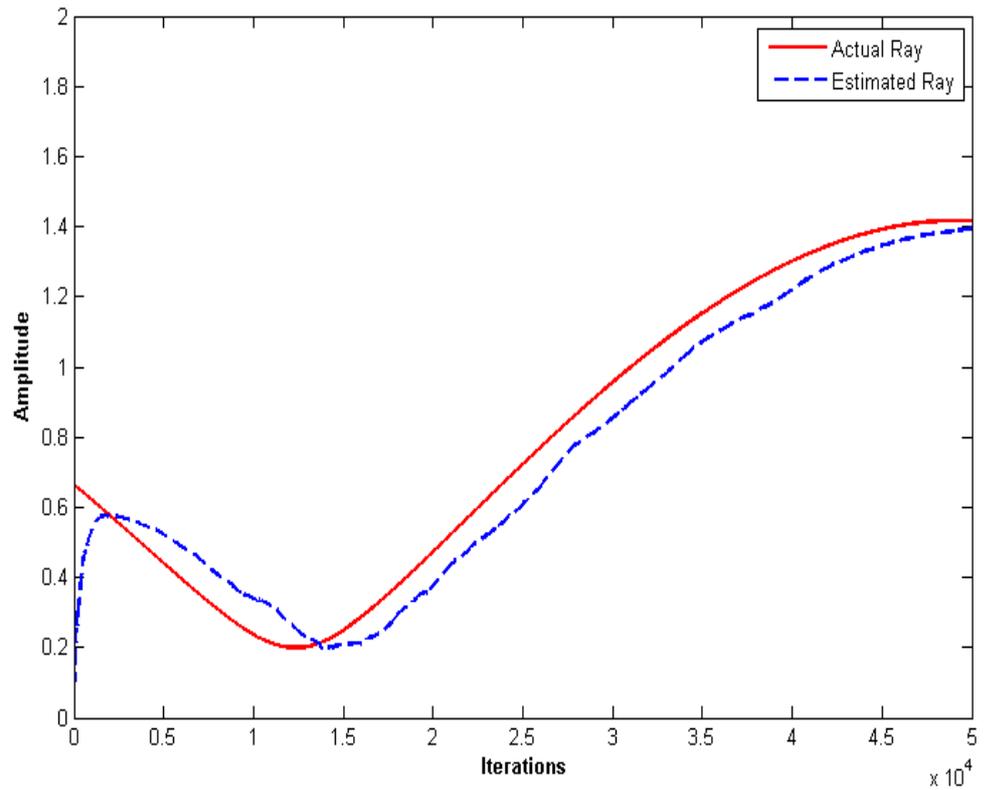


Figure 6.31: A typical trajectory of the amplitude of the first Rayleigh fading ray and its estimate with SNR=20 dB and step-size=0.01.

6.4 Transient Performance of the SRLMF algorithm

In this section, we examine the transient behavior of the SRLMF algorithm for both cases of Gaussian and non-Gaussian data. All the simulations in this section use the following specifications: Let us consider a real-valued regression sequence $\{\mathbf{u}_i\}$ with covariance matrix \mathbf{R} whose eigenvalue spread we set at $\rho = 5$. Run the filter for 100000 iterations and the MSE learning curve is the average over 30 independent runs. Let the SNR be 50 dB and the step-size is fixed at $\mu = 0.01$. The stationary channel in (6.5) is considered here.

The result in Fig. 6.32 shows the theoretical and simulated MSD and MSE learning curves of the SRLMF algorithm using white Gaussian regressors. The theoretical values are obtained by using the expression (5.50). As can be seen here, an excellent match between the theoretical and simulated results.

The result in Fig. 6.33 shows the theoretical and simulated MSD and MSE learning curves of the SRLMF algorithm using white non-Gaussian regressors. It is seen that the theoretical MSD and MSE learning curves are converging to a steady-state value that is approximately 1 dB lower than the simulated.

The results in Figs. 6.34 and 6.35 show the theoretical and simulated MSD and MSE learning curves of the SRLMF algorithm using Gaussian and non-Gaussian regressors with an eigenvalue spread equal to 5, respectively. It is seen that there is a good match between the theoretical and simulated results.

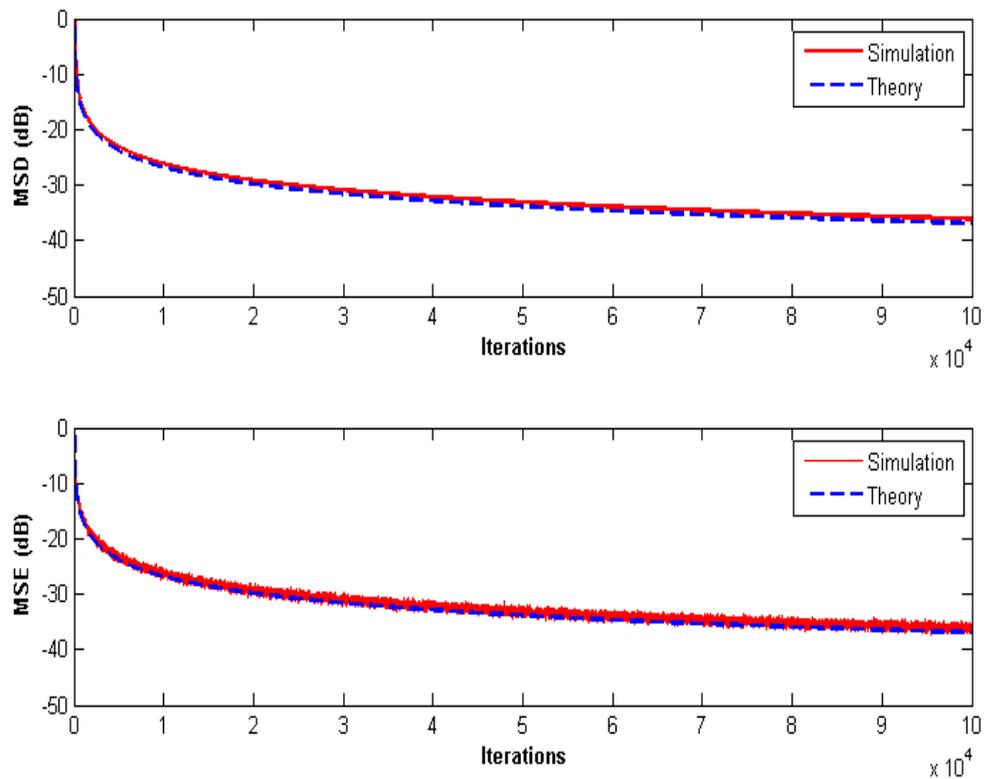


Figure 6.32: Theoretical and simulated MSD (top) and MSE (bottom) learning curves of the SRLMF algorithm using white Gaussian regressors with SNR=50 dB and step-size=0.01.

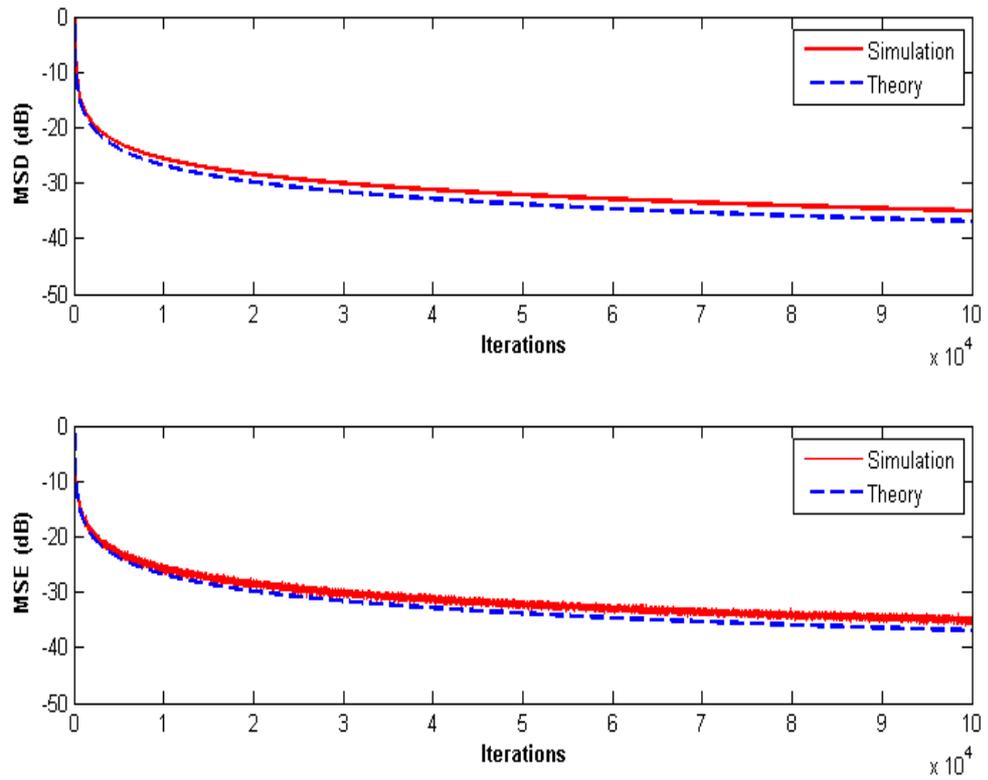


Figure 6.33: Theoretical and simulated MSD (top) and MSE (bottom) learning curves of the SRLMF algorithm using white non-Gaussian regressors with SNR=50 dB and step-size=0.01.

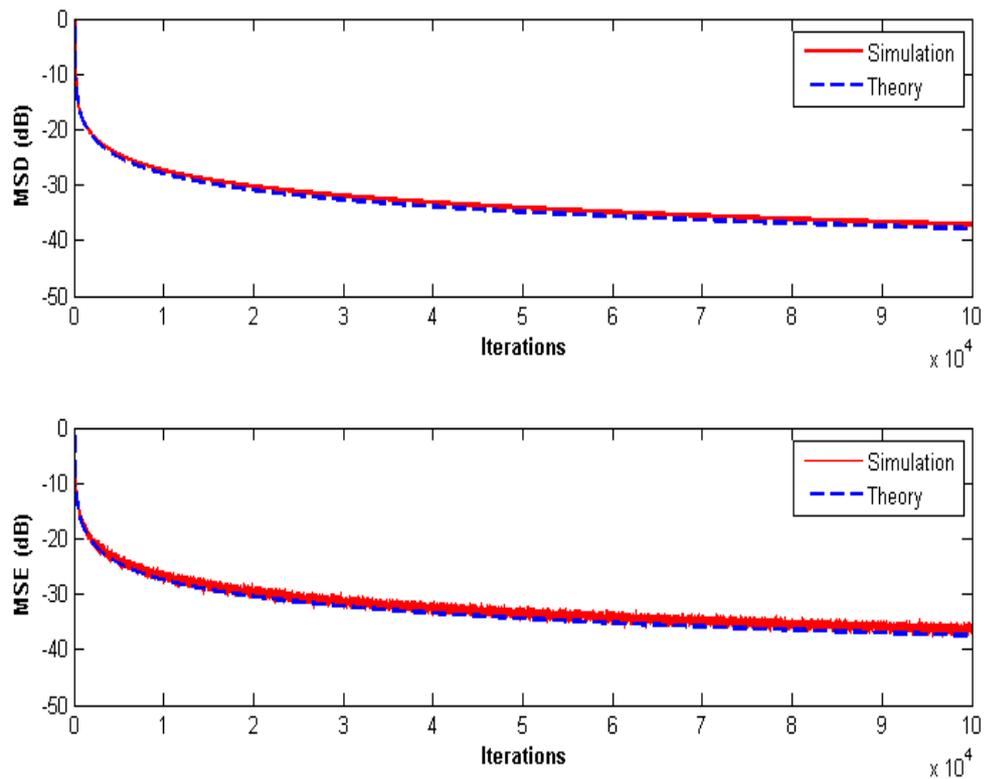


Figure 6.34: Theoretical and simulated MSD (top) and MSE (bottom) learning curves of the SRLMF algorithm using Gaussian regressors with an eigenvalue spread=5, SNR=50 dB and step-size=0.01.

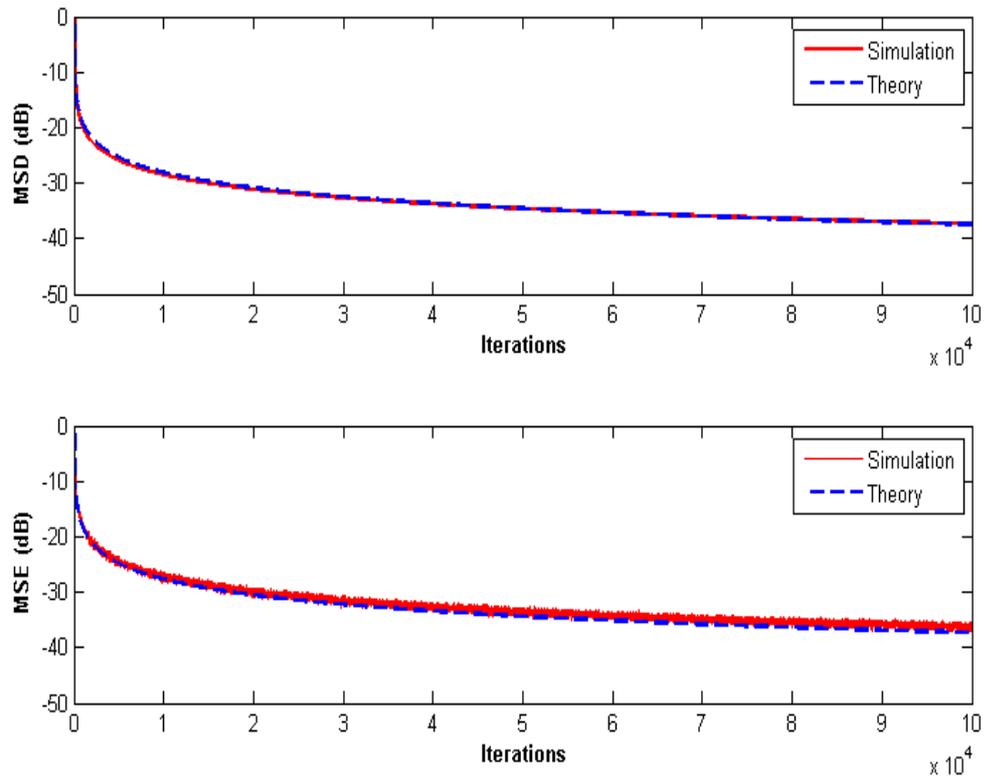


Figure 6.35: Theoretical and simulated MSD (top) and MSE (bottom) learning curves of the SRLMF algorithm using non-Gaussian regressors with an eigenvalue spread=5, SNR=50 dB and step-size=0.01.

6.5 Conclusion

In this chapter, computer simulations are carried out to corroborate the theoretical findings, where it is shown that the theoretical and simulated results are in good agreement. Moreover, the results show that both the SRLMF algorithm and the LMF algorithm have a similar performance for the same steady-state EMSE.

CHAPTER 7

THESIS CONTRIBUTIONS, CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK

7.1 Thesis Contributions

The thesis has four main contributions:

1. The first contribution is the derivation of the expressions for the steady-state EMSE of the SRLMF algorithm in a stationary environment.
2. The second contribution is the derivation of the expressions for the tracking EMSE of the SRLMF algorithm in a nonstationary environment. An optimum value of the step-size μ is also evaluated.

3. The third contribution is the extension of the weighted variance relation.
4. The fourth contribution is the derivation of the expressions for the MSE and the MSD of the SRLMF algorithm during the transient phase.

7.2 Conclusions

A new adaptive algorithm, called the SRLMF algorithm, has been presented in this work. Monte Carlo simulations have shown that there is a good agreement between the theoretical and simulated results. The simulation results indicate that both the SRLMF algorithm and the LMF algorithm converge at the same rate resulting in no performance loss. The analysis developed in this paper is believed to make practical contributions to the design of adaptive filters using the SRLMF algorithm instead of the LMF algorithm in pursuit of the reduction in computational cost and complexity while still maintaining good performance.

7.3 Recommendations for Future Work

In this thesis, the steady-state, the tracking, and the transient performance of the SRLMF algorithm have been studied with constant step-size. Many recommendations for future work can be made here and among them are the following:

1. The performance of the SRLMF algorithm can be examined with time-variant step-size. The update recursion of the SRLMF algorithm with a

time-variant step-size for real-valued data can be suggested as follows:

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu_i \text{sign}[\mathbf{u}_i]^T e_i^3, \quad i \geq 0, \quad (7.1)$$

where μ_i is the step-size at time i .

2. The performance of the SRLMF algorithm can also be examined by including the normalization factor in the update recursion. The correction term that is added to \mathbf{w}_{i-1} in the update recursion is normalized with respect to the squared-norm of the regressor \mathbf{u}_i . Moreover, the positive constant ϵ in the denominator avoids division by zero or by a small number when the regressor is zero or close to zero. The update recursion of the ϵ -NSRLMF algorithm for real-valued data can be suggested as follows:

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \frac{\mu}{\epsilon + \|\mathbf{u}_i\|^2} \text{sign}[\mathbf{u}_i]^T e_i^3, \quad i \geq 0. \quad (7.2)$$

APPENDIX A

Price's Theorem for Complex Sign Function

From Price's theorem [28] we have

$$\mathbb{E} [\operatorname{Re}[x^* \operatorname{csgn}(y)]] = \sqrt{\frac{2}{\pi}} \frac{\sqrt{2}}{\sigma_y} \mathbb{E} [\operatorname{Re}[x^* y]], \quad (\text{A.1})$$

where $x = x_r + jx_i$ and $y = y_r + jy_i$ denote two complex-valued jointly-Gaussian random variables. Therefore,

$$\begin{aligned} \mathbb{E} [|\mathbf{u}_i|_{\mathbb{H}}^2] &= \mathbb{E} [\mathbf{u}_i \mathbf{H}[\mathbf{u}_i] \mathbf{u}_i^*], \\ &= \mathbb{E} [\operatorname{Re}[\mathbf{u}_i \operatorname{csgn}[\mathbf{u}_i]^*]], \\ &= \mathbb{E} [\mathbf{u}_r \operatorname{sign}(\mathbf{u}_r)^T + \mathbf{u}_i \operatorname{sign}(\mathbf{u}_i)^T], \\ &= \sqrt{\frac{2}{\pi}} \frac{1}{\sigma_{u_r}} \operatorname{Tr}(\mathbf{R}) + \sqrt{\frac{2}{\pi}} \frac{1}{\sigma_{u_i}} \operatorname{Tr}(\mathbf{R}). \end{aligned} \quad (\text{A.2})$$

Assuming now that the real and imaginary parts of \mathbf{u}_i have identical variances, i.e., $\sigma_{u_r}^2 = \sigma_{u_i}^2$ so that $\sigma_u^2 = 2\sigma_{u_r}^2$. Therefore,

$$\sigma_{u_r} = \sigma_{u_i} = \frac{\sigma_u}{\sqrt{2}}. \quad (\text{A.3})$$

Substituting (A.3) into (A.2) we get

$$\mathbb{E}[\|\mathbf{u}_i\|_{\mathbb{H}}^2] = \frac{4\text{Tr}(\mathbf{R})}{\sqrt{\pi\sigma_u^2}}. \quad (\text{A.4})$$

REFERENCES

- [1] S. Haykin, “Adaptive Filter Theory,” 4th edition, New Jersey: Prentice Hall, 2002.
- [2] A. H. Sayed, “Fundamentals of Adaptive Filtering,” New York: Wiley Interscience, 2003.
- [3] S. U. H Qureshi, “Adaptive Equalization,” *IEEE Proceedings*, vol. 73, no. 9, pp. 1349–1387, Sep. 1985.
- [4] J. Makhoul, “Linear Prediction: A Tutorial Review,” *IEEE Proceedings*, vol. 63, no. 4, pp. 561–580, Apr. 1975.
- [5] B. Widrow, J. R. Glover, Jr., J. M. McCool, J. Kaunitz, C. S. Williams, R. H. Hearn, J. R. Zeidler, E. Dong, Jr., and R. C. Goodlin, “Adaptive Noise Cancelling: Principles and Applications,” *IEEE Proceedings*, vol. 63, no. 12, pp. 1692–1716, Dec. 1975.
- [6] B. Widrow, J. M. McCool, M. G. Larimore, and C. R. Johnson, Jr., “Stationary and Nonstationary Learning Characteristics of the LMS Adaptive Filter,” *IEEE Proceedings*, vol. 64, no. 8, pp. 1151–1162, Aug. 1976.

- [7] E. Walach and B. Widrow, “The Least Mean Fourth (LMF) Adaptive Algorithm and Its Family,” *IEEE Trans. Inf. Theory*, vol. 30, no. 2, pp. 275–283, Mar. 1984.
- [8] P. I. Hübcher and J. C. M. Bermudez, “An Improved Statistical Analysis of the Least Mean Fourth (LMF) Adaptive Algorithm,” *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 664–671, Mar. 2003.
- [9] V. H. Nascimento and J. C. M. Bermudez, “Probability of Divergence for the Least-Mean Fourth Algorithm,” *IEEE Trans. Signal Process.*, vol. 54, no. 4, pp. 1376–1385, Apr. 2006.
- [10] D. L. Duttweiler, “Adaptive Filter Performance with Nonlinearities in the Correlation Multiplier,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 30, no. 4, pp. 578–586, Aug. 1982.
- [11] A. Gersho, “Adaptive Filtering with Binary Reinforcement,” *IEEE Trans. Inform. Theory*, vol. 30, no. 2, pp. 191–199, Mar. 1984.
- [12] W. A. Sethares, “Adaptive Algorithms with Nonlinear Data and Error Functions,” *IEEE Trans. Signal Process.*, vol. 40, no. 9, pp. 2199–2206, Sep. 1992.
- [13] T. A. C. M. Claasen and W. F. G. Mecklenbrauker, “Comparison of the Convergence of Two Algorithms for Adaptive FIR Digital Filters,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 29, no. 3, pp. 670–678, June 1981.

- [14] N. J. Bershad, “Comments on ‘Comparison of the Convergence of Two Algorithms for Adaptive FIR Digital Filters,’” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 33, no. 6, pp. 1604–1606, Dec. 1985.
- [15] W. A. Sethares, I. M. Y. Mareels, B. D. O. Anderson, C. R. Johnson, Jr., and R. R. Bitmead, “Excitation Conditions for Signed Regressor Least Mean Squares Adaptation,” *IEEE Trans. Circuits Syst.*, vol. 35, no. 6, pp. 613–624, June 1988.
- [16] E. Eweda, “Analysis and Design of a Signed Regressor LMS Algorithm for Stationary and Nonstationary Adaptive Filtering with Correlated Gaussian Data,” *IEEE Trans. Circuits Syst.*, vol. 37, no. 11, pp. 1367–1374, Nov. 1990.
- [17] H. Sari, “Performance evaluation of three adaptive equalization algorithms,” *in Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing*, vol. 7, pp. 1385–1389, May 1982.
- [18] N. J. Bershad, “On the optimum data nonlinearity in LMS adaptation,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, no. 1, pp. 69–76, Feb. 1986.
- [19] C. P. Kwong, “Dual sign algorithm for adaptive filtering,” *IEEE Trans. Commun.*, vol. 34, no. 12, pp. 1272–1275, Dec. 1986.
- [20] O. Macchi, “Advances in adaptive filtering,” *in Digital Communications*, E. Biglieri and G. Prati, Eds. Amsterdam, The Netherlands: North-Holland, pp. 41–57, 1986.

- [21] V. J. Mathews and S. H. Cho, “Improved Convergence Analysis of Stochastic Gradient Adaptive Filters Using the Sign Algorithm,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 35, no. 4, pp. 450–454, Apr. 1987.
- [22] N. A. M. Verhoeckx and T. A. C. M. Claasen, “Some Considerations on the Design of Adaptive Digital Filters Equipped with the Sign Algorithm” *IEEE Trans. Commun.*, vol. 32, no. 3, pp. 258–266, Mar. 1984.
- [23] E. Eweda, “Almost Sure Convergence of a Decreasing Gain Sign Algorithm for Adaptive Filtering,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, no. 10, pp. 1669–1671, Oct. 1988.
- [24] E. Eweda, “A Tight Upper Bound of the Average Absolute Error in a Constant Step-Size Sign Algorithm,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 11, pp. 1774–1776, Nov. 1989.
- [25] S. Dasgupta and C. R. Johnson, Jr., “Some comments on the behavior of sign-sign adaptive identifiers,” *Syst. Contr. Lett.*, vol. 7, pp. 75–82, Apr. 1986.
- [26] S. Koike, “Analysis of the sign-sign algorithm based on Gaussian distributed tap weights,” in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing*, vol. 3, pp. 1673–1676, May 1998.
- [27] S. Koike, “Effects of Impulse Noise at Filter Input on Performance of Adaptive Filters Using the LMS and Signed Regressor LMS Algorithms,” in *Proc. Int. Symp. on Intelligent Signal Processing and Commun.*, pp. 829–832, Dec. 2006.

- [28] R. Price, “A Useful Theorem for Nonlinear Devices Having Gaussian Inputs,”
IRE Trans. Inform. Theory, vol. 4, no. 2, pp. 69–72, June 1958.

Vitae

- **Name:** Mohammed Mujahid Ulla Faiz
- **Nationality:** Indian
- Born in Bangalore, India, on April 4, 1980.
- Received Diploma in Electronics and Communication Engineering from MN Technical Institute, Bangalore, India, in July 1999.
- Received Bachelor's degree in Electronics and Communication Engineering from Visveswaraiah Technological University, Bangalore, India, in July 2003.
- Worked as a Lecturer, Department of Electronics and Communication, KNS Institute of Technology, Bangalore, India, from September 2004 to December 2005.
- Worked as a Software Trainee, Embedded Competency Centre, Mistral Solutions Private Limited, Bangalore, India, from December 2005 to May 2006.
- Worked as a Software Engineer, GDA Technologies Limited, Chennai, India, from June 2006 to August 2006.
- Joined as a Research Assistant, Department of Electrical Engineering, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia, in September 2006.

- Research interests include adaptive filtering analysis and design, application of adaptive filtering in communications, channel estimation for multiple-input–multiple-output (MIMO), orthogonal frequency division multiplexing (OFDM), and spatial modulation (SM) systems.
- **Present Address:** P.O. Box 8234, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia.
- **Permanent Address:** Door Number 7/1, 2nd Cross, LN Colony, Yeshwanthpur, Bangalore 560 022, Karnataka, India.
- **Telephone Number:** +966-554889175 (cell)
- **Email:** *faiz.upm@gmail.com*