Performance Enhancement of Adaptive Filters for Sparse Echo Cancellation

by

Mujahid Ado Alkassim

Submitted to the Institute of Graduate Studies in Science and Engineering

in partial fulfillment of

the requirements for the degree of

Master of Science

in

Mevlana (Rumi) University

June, 2017

ABSTRACT

PERFORMANCE ENHANCEMENT OF ADAPTIVE FILTERS FOR SPARSE ECHO CANCELLATION

Mujahid Ado ALKASSIM

M.Sc. Thesis, 2014

Thesis supervisor: Assist. Prof. Dr. Mohammad Shukri SALMAN

Keywords: Sparse Adaptive Filters, Acoustic Impulse Response, VSSLMS Algorithm, NNCLMS Algorithm, Norm Constraint, AWGN, ACGN.

Sparse adaptive filters have been studied for a long time to address the problem of acoustic echo cancelation in communication systems such as mobile phones, digital cameras, etc. The solution is shaped in estimating the acoustic room system. Up till now, the available techniques are either high computationally complex or have poor performance. However, if the impulse response of the acoustic room is sparse, this performance could be improved further.

In this thesis, we propose a new approach that enables us to identify the sparse echo path of the acoustic room using the variable step-size least-mean-square (VSSLMS) adaptive algorithm. The performance of the VSSLMS algorithm is modified by employing a p-norm penalty term in the cost function of the VSSLMS algorithm. This penalty term imposes a zero attraction of the filter coefficients according to the relative value of each filter coefficient among all the entries which, in turn, leads to an improved performance when the system is sparse.

The convergence analysis of the proposed algorithm is presented and stability condition is derived. Different experiments, in the context of echo path estimation, have been conducted to investigate the performance of the proposed algorithm compared to the performances of other algorithms such as; zero-attracting least-mean-square (ZA-LMS), reweighted zero-attracting LMS (RZA-LMS), non-uniform norm constraint LMS

(NNCLMS), proportionate normalized LMS (PNLMS) and improved PNLMS (IPNLMS) algorithms. Simulation results show that the proposed algorithm outperforms the aforementioned algorithms in terms of mean-square-deviation (MSD) and convergence rate.

ACKNOWLEDGEMENTS

I thank Almighty Allah, the merciful and all-knowing, for sparing my life in sound health and giving me the opportunity to accomplish this thesis.

I wish to express my deepest gratitude to my supervisor Assist. Prof. Dr. Mohammad Shukri Salman for his guidance, advice, criticism, encouragement and insight throughout the research.

I am highly indebted to my parents for their encouragement, support and unlimited love.

I also wish to express my appreciation to Kano State Government under the leadership of His Excellency Engr. Dr. Rabi'u Musa Kwankwaso for granting me masters scholarship.

Finally, i wish to extend a special thanks to my colleagues for their valuable support and company. They really made my life a fabulous one.

Dedicated to my parents

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS/ABBREVIATIONS

δ	Regularization parameter
ε	Reweighting parameter
γ	Control parameter
κ	Adjusting parameter
λ	Eigenvalue
μ	Step-size
$\wp$	Correlation parameter
σ^2	Variance
$\mathbf{I}$	Identity matrix
$\mathbf{Q}$	control matrix
N	Filter length
$\mathbf{R}$	Autocorrelation matrix
$\mathbf{w}$	Tap weights vector
$\mathbf{x}$	Tap input vector
AEC	Acoustic echo canceller
AWGN	Additive white Gaussian noise
ACGN	Additive correlated Gaussian noise
CS	Compressive sensing
DSP	Digital signal processing
FIR	Finite impulse response
IPNLMS	Improved proportionate normalized least-mean-square
LMS	Least-mean-square
MSE	Mean-square-error
MSD	Mean square deviation
NLMS	Normalized least-mean-square
NNCLMS	Non-uniform norm constraint LMS
PNLMS	Proportionate normalized least-mean-square
RZA-LMS	Reweighted zero-attracting LMS

SNR	Signal-to-noise ratio
VSSLMS	Variable step size LMS
ZA-LMS	Zero-attracting LMS

CHAPTER 1

INTRODUCTION

1.1. Overview

The alarming research interest in the field of digital signal processing (DSP) nowadays is due to improvements in digital circuit design and implementation [1]. Adaptive filtering techniques are one of the well-known fields of DSP that has become prevalent and widely used in digital devices such as cell phones, digital cameras, medical instruments, etc. Unlike non-adaptive filters, the advantages of adaptive filters are their self-adjusting ability and iterative solution according to a certain optimization algorithm such least-mean-square (LMS) algorithm. As a result, adaptive filters have been applied to various fields such as control, communications, signal processing, etc., [2], [3].

Nowadays, the rapid emergence and use of hands-free telephones and internet phones have attracted a considerable research interest in echo cancellation as specific application of adaptive filtering techniques. An echo occurs in a communication system when an interruption and perhaps interfered versions of a signal are reproduced back to the sending end of that signal. This interruption version is only observable if the amount of the echo is substantially large [4]. For instance, in the case of hands-free mobile communication, the acoustic coupling between headset and speaker phone may occasionally be strong enough to create echo to the extent that a user may seriously be irritated. And in internet phones such as skype or distance learning education via video-streaming, not just perfect data transmission is important, but better audio quality is also needed in these hypermedia role. For these reasons, effective echo cancellation is highly crucial to improve the audio quality and clarity of a call [5], [6]. This requires adaptive filter to identify the echo path and efficiently cancel out the echo.

However, the nature of the echo path to be identified by the adaptive filter is sparse,and requires adaptive filters with sparsity exploitation property. Example of a sparse system that usually occurs in telecommunication networks is where a telephone equipment is connected to a packet switched network. In this case, the energy of the signal is typically of length 64-128ms, out of which the active region of the energy spans duration of 8-12ms, the remaining region is dominated by zero energy of the signal, making the impulse response sparse. The dominant portion is due to the presence of bulk delay caused by network propagation, encoding and jitter buffer delays [7], [8], [9]. Even though, sparse adaptive filters have been extensively studied for a long time to address the problem of echo in communication system. Up to date, the available filters fail to provide a tolerable performance due to challenging sparse nature of the echo path. For this particular application, where the system is sparse, conventional LMS filter performs poorly due to lack of sparsity exploitation property of the system. Hence, it is highly desirable to design a more efficient and robust adaptive filter that can provide acceptable performance for identifying sparse echo path. This can be achieved by exploiting the sparsity property of the systems [10].

1.2. The Aim of Adaptive Filter

In practical applications, an adaptive filter is employed to filter out a signal with an unknown frequency response. Therefore, the main aim of an adaptive filter is to adjust its parameters to adapt with changes of system parameters [11]. It does so by setting its parameters in such away that its output tries to minimize a meaningful objective function involving reference signal. In most cases, the objective function is a function of the input, the reference and the adaptive filter output signal. An adaptive algorithm is composed of three basic items: definition of the minimization algorithm, definition of the objective function and definition of the error signal. The error signal is usually defined as the difference between the filter output and a desired response. The optimal filter parameters are found through minimization of a cost function of the error signal. A useful approach is based on minimizing the mean-square value of the error signal. The general configuration of an adaptive filtering system is shown in Fig. 1.1. The input signal is denoted as $x(n)$, the output signal as $y(n)$, the desired response as $d(n)$ and the error signal as $e(n) = d(n) - y(n)$. The error signal is used to form a performance function that is iteratively minimized by the adaptation algorithm in order to determine the appropriate updating of the filter coefficients. The minimization of the objective function implies that the adaptive filter output signal matches the desired signal in some sense [12].

Basically, adaptive filters are classified into two major categories according to their impulse response, namely the finite impulse response (FIR); whose impulse response is of a finite duration because it settles to zero in finite time, and infinite impulse response (IIR) filters [13]; which have internal feedback and may continue to respond indefinitely. The most widely used filter is the FIR filter which will also be used throughout this work. Adaptive FIR filters have many structures such as adaptive transversal filters, the lattice predictor, the systolic array, etc. [2]. This adaptive transversal filters structure is shown in Fig. 1.2.

From the figure, $\mathbf{x}(n)$ denotes the tap input vector at time n, $\hat{d}(n|X_n)$ denotes the corresponding estimate of the desired response at the filter output, and X_n denotes the space spanned by the tap inputs $x(n), x(n-1), \ldots, x(n-N+1)$. By comparing this with the actual desired response $d(n)$, then estimate error can be calculated as $e(n) = d(n) - \hat{d}(n|X_n)$.

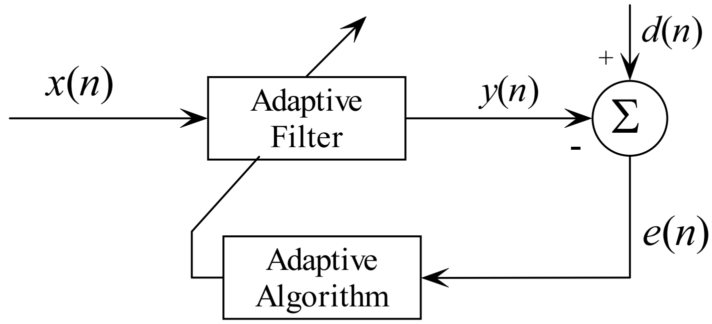

Figure 1.1. General Adaptive Filter Configuration, [2].

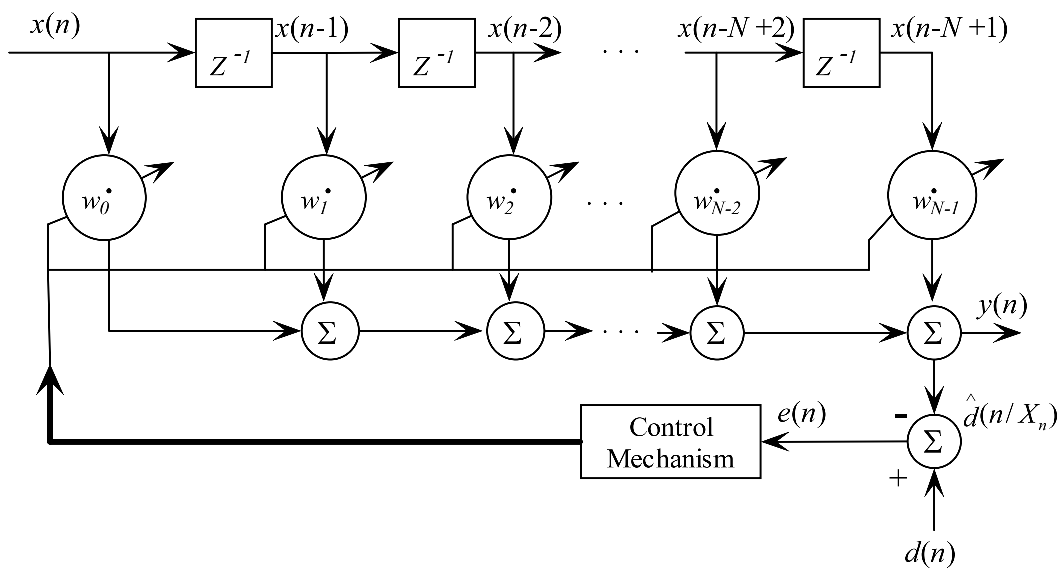

Figure 1.2. Structure of Adaptive Transversal Filter, [11].

1.3. Application of Adaptive filters

As previously discussed, an adaptive filter has the ability to self-adjust itself to operate well in an unknown environment and track variations of input signal statistics. These features make it a powerful device in many signal processing and control applications. Although, the nature of its applications differ in various aspects but have a common feature: the input signal and a desired response, which are used to estimate the error signal. Depending on the way the desired response is extracted, the applications of an adaptive filter can be classified into four categories: System identification, noise cancellation, channel equalization and linear prediction [2]. The configuration of each class will be briefly explained.

1.3.1. Adaptive System Identification

The adaptive system identification is basically used to estimate a transfer function of an unknown system such as the response of an unknown communication channel or the frequency response of an auditorium. In this case, the desired signal $d(n)$ is the output of an unknown system $u(n)$ when excited by the input signal $x(n)$. Both the the unknown system and the adaptive filter are driven by a common input $x(n)$ as shown in Fig. 1.3. The output of the adaptive filter $y(n)$ is subtracted from the output of the unknown system resulting in an error estimate $e(n)$. The error is used to control the filter coefficients of the adaptive system. When the error signal is minimized, the adaptive filter represents the model for the unknown system [14].

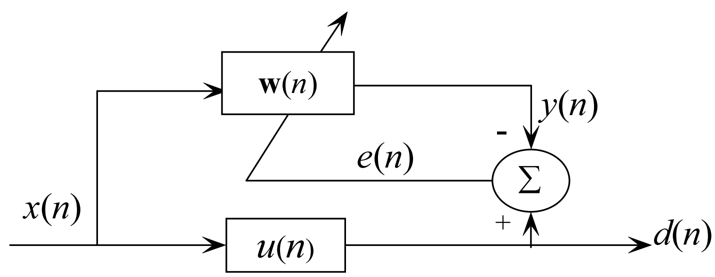

Figure 1.3. Adaptive System Identification Configuration, [11].

1.3.2. Adaptive Noise Cancellation

In noise cancellation, adaptive filters let you remove noise from a signal in real time application. The adaptive noise cancellation configuration is illustrated in Fig. 1.4. The signal of interest $s(n)$ is corrupted with uncorrelated additive noise $N_0(n)$, the combined signal is then used as the desired response $d(n)$ and is compared with the reference signal $x(n)$ that is uncorrelated with the noise signal $N_0(n)$ located at a different point. In this case, the adaptive filter provides an estimate $y(n)$ of the noise $N_0(n)$, by exploiting the correlation between $N_0(n)$ and $N_1(n)$ so that the error signal will be a noiseless version of the target signal $s(n)$, [11], [15].

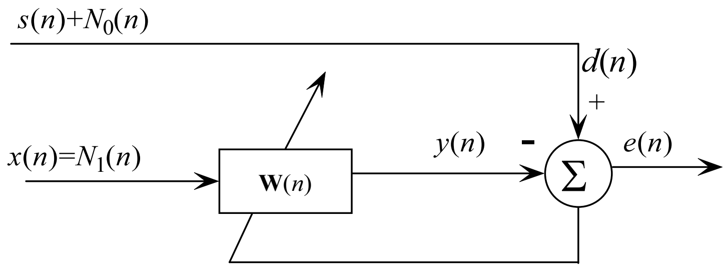

Figure 1.4. Adaptive Noise Cancellation Configuration, [11].

1.3.3. Adaptive Linear Prediction

In adaptive linear prediction application, the adaptive filter is used to estimate the values of a signal based on the past values of the signal. This application is essentially used in speech and image compression techniques. In this case, the desired signal $d(n)$ is a forward (or eventually backward) version of the adaptive filter input signal as shown in Fig. 1.5. When the system converged, the adaptive filter represents a model for the input signal and can used as a predictor model for the input signal [16].

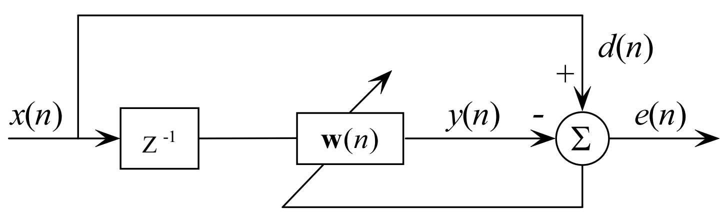

Figure 1.5. Adaptive Linear Prediction, [11].

1.3.4. Adaptive Channel Equalization

Channel equalization also known as inverse filtering consists of estimating a transfer function to compensate for the linear distortion caused by the channel due to spectral changes [17], [18]. For this case, the input signal $x(n)$ is applied through the unknown system $u(n)$ and then through the adaptive filter resulting in an output $y(n)$ as shown in Fig. 1.6. The same input is also sent through a delay to attain $d(n)$. In this class of application, the adaptive system is said to be converged when the transfer function of the adaptive system is close to the reciprocal of the unknown system's transfer function.

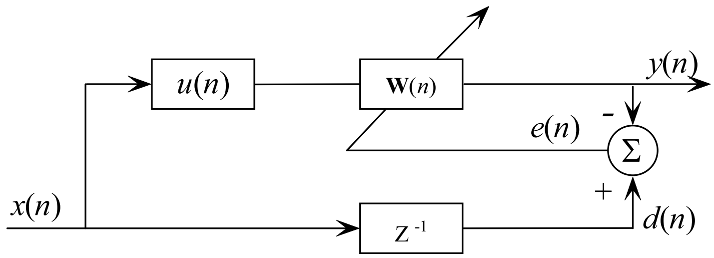

Figure 1.6. Adaptive Channel Equalization Configuration, [11].

1.4. Sparse Systems

A system is called sparse if its finite impulse response (FIR) model contains a few active taps in the presence of many negligible or zero ones [5], [19]. Sparse systems are classied into two main categories:

- General Sparse Systems: This consists a few active taps interspersed among many negligible ones along the entire response of the system [20].
- Clustering Sparse Systems: This consists a cluster or gathering of large active taps at one or more positions along the response of the system [21].

A typical example of a single clustering sparse system is the acoustic echo path, while the echo path of a satellite links is an example of multi-clustering sparse system [22]. These systems are illustrated in Fig. 1.7.

1.5. Echo Cancellation

Even though echo cancellation has been studied for a long time and outstanding performance has been achieved by existing echo cancellers, latest advances of hands-free telephony systems, internet phones such as VoIP or skype, and teleconferencing systems need additional enhancement to increase the voice quality [23]. In communication systems, an echo can be categorized into:

1. Acoustic echo: This is caused due to acoustic coupling between microphone and loudspeaker (e.g., as in speaker phone).
2. Network echo: This occurs when there is an unbalanced coupling between 2-wire and 4-wire circuits.

In both cases, the adaptive filter has to model an unknown system, i.e., the echo path. However, this thesis focuses on acoustic aspect of the echo. Fig. 1.8 illustrates a scenario of acoustic echo cancellation (AEC) configuration, where an adaptive filter is used to identify the unknown echo path by adaptively adjusting its

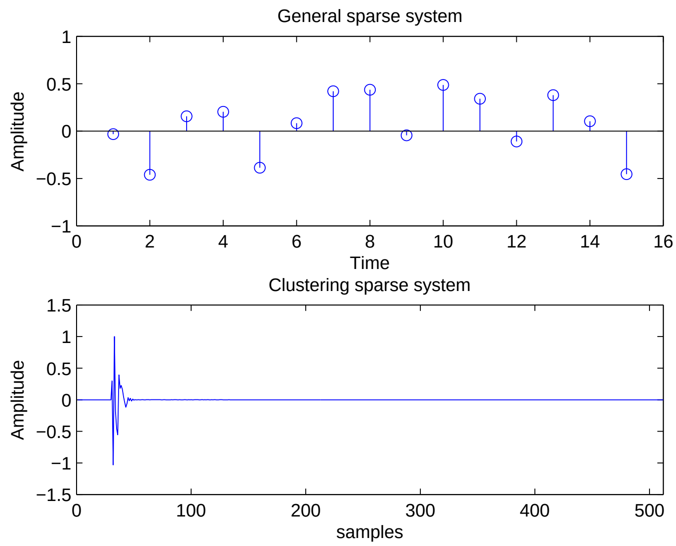

Figure 1.7. Typical sparse system.

coefficients. The estimated coefficients are used to provide a replica of the echoes which can be subtracted from the target signal to achieve cancellation.

1.5.1. Notations and Definitions

- n is the discrete time index,
- N is the length of the adaptive filter,
- $x(n)$ is the far-end signal (i.e., input signal of the adaptive filter and loudspeaker),
- $\mathbf{x}(n) = [x(n), x(n-1), \ldots, x(n-N+1)]^T$ is the input tap vector,
- $\mathbf{w}_0 = [w_0, w_1, \ldots, w_{N-1}]^T$ is the impulse response of the system (i.e., the echo path).
- $\mathbf{w}(n) = [w_0(n), w_1(n), \ldots, w_{N-1}(n)]^T$ is the estimated impulse response at time n (i.e., the adaptive filter at time n,
- $y(n) = \mathbf{w}_0^T\mathbf{x}(n)$ is the echo signal,
- $\hat{y}(n) = \mathbf{w}^T(n)\mathbf{x}(n)$ is output of the adaptive filter,
- $d(n) = y(n) + v(n)$ is the reference signal or desired response.

Consider a loudspeaker and microphone in the same acoustic room. When the loudspeaker broadcast a signal $x(n)$, the microphone in the room will record the local signal $v(n)$ (i.e. near-end speech). In this case, the microphone signal to be transmitted back to the far-end side will be:

$$\begin{aligned} d(n) &= y(n) + v(n) \\ &= \mathbf{w}_0^T\mathbf{x}(n) + v(n). \end{aligned} \tag{1.1}$$

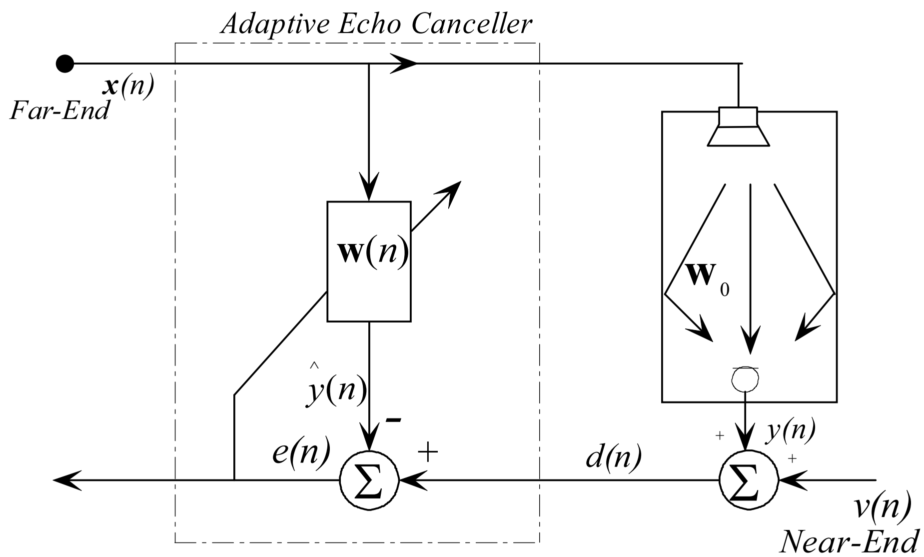

Figure 1.8. Block diagram of acoustic room system.

If the echo path transfer function is modelled as FIR filter $\mathbf{w}_0$, then the echo signal can be considered as filtered version of the loudspeaker signal i.e., $y(n) = \mathbf{w}_0^T\mathbf{x}(n)$ and the estimated output from the adaptive filter is $\hat{y}(n) = \mathbf{w}^T(n)\mathbf{x}(n)$. The function of the AEC in this case is to identify the unknown room impulse response $\mathbf{w}_0$ and hence subtract an estimate of the echo signal from the microphone signal. Hence, the error signal is,

$$\begin{aligned} e(n) &= d(n) - \mathbf{w}^T(n)\mathbf{x}(n) \\ &= [\mathbf{w}_0^T - \mathbf{w}^T(n)]x(n) + v(n), \end{aligned} \tag{1.2}$$

where $\mathbf{w}(n)$ is an estimate of $\mathbf{w}_0$. When the error signal is minimized, the adaptive filter estimate $\mathbf{w}(n)$ represents the echo path transfer function $\mathbf{w}_0$ and therefore converges [24], [25].

1.6. Performance Measures

Evaluation of performance measures influences the choice of one algorithm among many others. In echo cancellation problem, the aim is to measure how much the undesired echo is attenuated. There are many ways to measure this attenuation but the most important ones which are commonly adopted in this context would next be explained.

1.6.1. Convergence Rate

This metric indicates the speed at which the filter converges to its steady state. A faster convergence rate is an efficient property of an adaptive filter. Usually, there is a trade-off between the convergence rate and other performance measures.

1.6.2. Mean-Square-Error

The mean-square-error (MSE) is the mean square value of the difference between the desired signal and the filter output. An adaptive system has effectively converged to the true solution of the system if the MSE value is very small. It is given by:

$$MSE = \mathrm{E}\{[d(n) - \hat{y}(n))]^2\}, \tag{1.3}$$

where $\mathrm{E}\{\cdot\}$ denotes expected value [14].

1.6.3. Mean-Square-Deviation

Mean-square-deviation (MSD) is the most used performance measure in echo cancellation. It quantifies directly how well an adaptive filter converges to the impulse response of the system that need to be identified. The MSD is defined as:

$$MSD = \mathrm{E}\|\mathbf{w}_0 - \mathbf{w}(n)\|^2, \tag{1.4}$$

where $\|\cdot\|^2$ is the l_2-norm. It measures the closeness of an estimated system to that of the true system and is particulary useful to study the tracking capability of a time-varying system [26].

CHAPTER 2

PROBLEM STATEMENT AND THESIS ORGANIZATION

2.1. Introduction

This chapter states the specific challenges that are mostly encountered in acoustic echo cancellation problem and our possible contributions to overcome some of the challenges. We briefly explain the available techniques and there consequences and then present our possible solution.

2.2. Statement of the Problem

Among the various applications of adaptive filtering techniques, echo cancellation is well known to be the most tricky one. This is so because its explicit nature represent a lot of challenges for any adaptive filter. There are quite a lot of issues related to this crucial application, among which a few are as follows. First, it is well known that the echo paths can have excessive lengths in time, e.g., up to or even more than hundreds of milliseconds. For instance, in network echo cancellation, the usual lengths are in the range between 32 and 128 milliseconds, while in acoustic echo cancellation, these lengths can be even higher [7]. As a result, long length adaptive filters are readily required (with hundreds or even thousands of coefficients), affecting the convergence rate of the adaptive algorithm. Alongside, the echo paths are time-variant systems, requiring efficient tracking abilities for the echo canceller. Second, the undesired echo signal is usually combined with the near-end signal; conceptually, the function of the adaptive filter here is to segregate this mixture and offer an estimate of the echo at its output along with an estimate of the near-end from the error signal. This is quite a difficult task since the near-end signal may include either or both the background noise and the near-end speech; this noise can also be variant and powerful while the near-end speech can be like a big disturbance. Also, the input of the adaptive filter is a times speech sequence, which is a time-varying and highly correlated signal that can affect the whole performance of adaptive algorithms. In addition, the echo path is sparse in nature, requiring adaptive algorithms with good sparsity exploitation properties.

Over the paste decades, numerous types of adaptive filters have been used for echo cancellation. The normalized least-mean-square (NLMS) algorithm is one of the most popular among them, due to its numerical stability and moderate computational complexity. However, its use of a uniform step-size across all filter coefficients limits its convergence speed when estimating a sparse signal [16]. To overcome this problem, Duttweiller in [7] proposed a proportionate updating technique by assigning different step-sizes across filter taps independently to promote sparsity exploitation. Other approaches for sparsity exploitation apply subset selection scheme during the filtering process through statistical detection of active taps or sequential partial updating [34], [35]. However, both of these approaches are somewhat tricky and computationally complex whose performances degrade with the variation of sparseness level of the echo path. In addition, the aforementioned techniques fail to provide a satisfactory performance in a high correlated environment. The problem of identifying sparse echo paths has gained increasing interest due to the recently introduced framework of Compressive Sensing (CS) [26], [30], [32]. As a result, the LMS algorithm was modified to exploit sparsity property of a signal by employing l_0-norm or l_1-norm constraint into the cost function of the standard LMS [22], [29], [36], [37], [38]. The norm constraints accelerate the convergence of small active taps for identification of sparse echo path. Unfortunately, the resulting modified LMS filters suffer from the norm constraint adaptation during filtering process and produce estimation bias for identifying systems with a variety of sparseness levels due to lack of adjustable factor. To limit the estimation bias and enable the quantitative adjustment of the norm constraint adaptation, a non-uniform norm constraint (NNCLMS) was proposed in [40] which employs a p-norm like constraint to modify the cost function of LMS filter. The main challenge of this approach is its inability to maintain its performance when the input signal is highly correlated such as speech signal [41]. The variable step-size LMS (VSSLMS) was proposed by Harris et. al. [42] to stabilize the performance of the conventional LMS, but still has limited ability to exploit sparsity of the system due to its no use of sparsity characteristics [43], [56].

2.3. Our Contributions

In this thesis, we propose a new approach of identifying a sparse echo path. The proposed approach will be shown to overcome some of the above mentioned limitations. The approach combines a VSSLMS and a p-norm constraint. The variable step-size portion stabilizes the sparse system when the input signal is correlated where as the p-norm constraint exploits the system's sparsity by imposing a zero attraction of the filter coefficients according to the relative value of each filter coefficient among all the entries which, in turn, leads to an improved performance when the system is sparse. It would be shown to have a superior performance compared to the conventional approaches. We also carry out the convergence analysis and establish a stability condition of the proposed algorithm. The performance of the proposed algorithm is compared with diverse l_1-norm and p-norm based sparse adaptive filters in AEC settings using two noise types; Additive White Gaussian Noise (AWGN) and Additive Correlated Gaussian Noise (ACGN) and using acoustic echo paths of length $N = 256$ and $N = 512$ respectively. Also, the performance of the proposed algorithm has been extensively investigated in other sparse systems with a variety of sparseness degree. Simulation results demonstrate that the proposed algorithm outperforms different l_1-norm and p-norm based sparse filters in a sparse system identification.

2.4. Thesis Organization

The structure of the thesis is organized as follows:

- In Chapter 3, a general review of the most important adaptive filters used for echo cancellation application is presented.
- In Chapter 4, the proposed algorithm is presented. A review of the VSSLMS algorithm and a broad concept of the p-norm constraint are provided. The mean square convergence analysis and a stability criterion of the proposed algorithm are also carried out and presented.
- In Chapter 5, an experimental study is provided in order to compare the performance of the proposed filter with other l_1-norm and p-norm based sparse adaptive filters in the context of AEC.
- In Chapter 6, conclusions and a discussion on possibilities for future work are provided.

2.5. List of Publications

Some of the research presented in this thesis have been published. These publications are as follows.

Journal Paper:

- M. L. Aliyu, M. A. Alkassim and Mohammad Shukri Salman, “A p-Norm Variable Step-Size LMS Algorithm for Sparse System Identification,” Signal Image and Video Processing, Springer, 2013, DOI: 10.1007/s11760-013-0610-7.

Conference paper:

- T. R. Gwadabe, M. L. Aliyu, M.A. Alkassim, Mohammad Shukri Salman and H. Haddad, “A New Sparse Leaky LMS Type Algorithm,” IEEE 22nd Signal Processing and Communications Applications Conference (SIU2014), Trabzon, Turkey, April 2014.

CHAPTER 3

REVIEW OF SPARSE ADAPTIVE ALGORITHMS

3.1. Introduction

This chapter provides a brief review of the well known sparse adaptive filters used for echo cancellation. Firstly, we present the proportionate-based adaptive algorithms as background for estimating a sparse impulse response, we then subsequently discussed the zero attracting sparse adaptive filters used in the field due to their robustness and efficiency in performance. These filters operate based on l_1-norm optimization such as used in CS techniques [26] rather than proportionate updating based approach [7].

3.2. Least Mean Square Algorithm

The LMS algorithm is basically the most popular algorithm commonly found in the literature of adaptive filtering due to the fact that it is easy to implement, easy to understand, and robust in various aspects. It is a stochastic gradient algorithm in which evaluation of a gradient vector is made possible by iteratively modifying a cost function. The objective of the LMS algorithm is to sequentially estimate the coefficients of the filter using the input-output relationship of the signal. With reference to Fig. 1.8, the cost function $J(n)$ of the conventional LMS is defined as [2], [58]:

$$J(n) = \frac{1}{2}e^2(n), \tag{3.1}$$

where $e(n)$ is the instantaneous error given by:

$$e(n) = d(n) - \mathbf{w}^T(n)\mathbf{x}(n). \tag{3.2}$$

The minimum of the cost function can be obtained recursively using the gradient method [2]:

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \mu\frac{\partial J(n)}{\partial \mathbf{w}(n)}.$$

Hence the filter coefficient update for the LMS is then:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu e(n)\mathbf{x}(n), \tag{3.3}$$

where μ is the step-size parameter controlling the convergence and steady-state behavior of the LMS algorithm and is given by,

$$0 < \mu < \frac{1}{\lambda_{max}(\mathbf{R})}, \tag{3.4}$$

where λ_{max} is the maximum eigenvalue of **R**, and **R** is the autocorrelation matrix of the input tap vector. Normally, there exist a compromise between the convergence rate and MSE value in LMS. Table 3.1 gives the summary of the LMS algorithm.

Table 3.1. Summary of the LMS algorithm.

Filter Output	$y(n) = \mathbf{w}(n)^T\mathbf{x}(n)$
Estimate Error	$e(n) = d(n) - y(n)$
Tap-Weight Adaptation	$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu e(n)\mathbf{x}(n)$

3.3. Normalized LMS Algorithm

Despite the rampant usage of the LMS algorithm in many applications, selection of the step-size parameter μ posses a lot of challenges, especially, when the input signal $x(n)$ is highly correlated. As a result, the LMS algorithm suffers from a gradient noise amplification problem [11]. Normalized LMS (NLMS) was proposed to address these challenges. In NLMS algorithm, the step-size μ is normalized by the energy of input-tap vector as:

$$\mu_{NLMS} = \frac{\mu}{\mathbf{x}^T(n)\mathbf{x}(n)}, \tag{3.5}$$

Substituting (3.5) in (3.3), then the update scheme for NLMS is [7]:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu e(n)\mathbf{x}(n)}{\mathbf{x}^T(n)\mathbf{x}(n) + \delta}, \tag{3.6}$$

where δ is a very small positive constant to avoid dividing by zero.

However, for identification of long, sparse and dynamic acoustic echo path, NLMS algorithm has the disadvantages of slow convergence and poor tracking ability. However, the filter taps update for many of the adaptive algorithms can be generalized by the following set of equations [28]:

$$\begin{aligned} e(n) &= d(n) - \mathbf{w}^T(n)\mathbf{x}(n) \\ &= [\mathbf{w}_0^T - \mathbf{w}^T(n)]\mathbf{x}(n) + v(n), \end{aligned} \tag{3.7}$$

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu \mathbf{Q}(n) e(n) \mathbf{x}(n)}{\mathbf{x}^T(n) \mathbf{Q}(n) \mathbf{x}(n) + \delta}, \tag{3.8}$$

$$\mathbf{Q}(n) = diag \left\{ q_0(n), q_1(n), \ldots, q_{N-1}(n) \right\}, \tag{3.9}$$

where $\mathbf{w}_0$ is the optimum solution of the filter and $v(n)$ is the observation noise. $\mathbf{Q}(n)$ is the diagonal step-size control matrix and is introduced here to determine the step-size of each filter tap and is dependent on the particular algorithm. Since the step-size for NLMS algorithm is uniform for all filter taps, then the diagonal matrix is identity in this case, which is given by:

$$\mathbf{Q}(n) = \mathbf{I}_{N \times \ N}, \tag{3.10}$$

where $\mathbf{I}_{N \times \ N}$ is an $N \times \ N$ identity matrix. Table 3.2 gives the summary of the NLMS algorithm.

Table 3.2. Summary of the NLMS algorithm.

Filter Output	$y(n) = \mathbf{w}(n)^T \mathbf{x}(n)$
Estimate Error	$e(n) = d(n) - y(n)$
Tap-Weight Adaptation	$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu \mathbf{Q}(n) e(n) \mathbf{x}(n)}{\mathbf{x}^T(n) \mathbf{Q}(n) \mathbf{x}(n) + \delta}$

3.4. Proportionate NLMS Algorithm

In order to achieve a fast tracking of the sparse echo path, the proportionate NLMS (PNLMS) algorithm was proposed in [7], [33]. In the PNLMS algorithm, each filter coefficient is updated independently of others by regulating the adaptation step-size in proportion to the value of the filter tap estimates. The PNLMS algorithm allocates higher step-sizes to taps with larger magnitudes using a control matrix $\mathbf{Q}(n)$. For this case, the entries of the control matrix can be expressed as:

$$q_i(n) = \frac{k_i(n)}{\sum_{i=0}^{N-1} k_i(n)}, \ 0 \ \leq i < \ N-1, \tag{3.11}$$

and,

$$k_i(n) = \max \left\{ \rho \times \max \left\{ \gamma, |\ w_0(n)|, \ldots, |\ w_{N-1}(n)| \right\}, |\ w_i(n)| \right\}, \ 0 \ \leq i < \ N-1, \tag{3.12}$$

It has been indicated in [7] that the best choices for γ and ρ values are 0.01 and $\frac{5}{N}$, respectively. The function of γ is to prevent the taps from being inactive when they are very much smaller than the largest taps and the function of ρ is to prevent $w_i(n)$ from being inactive during initialization stage. Also, the regularization parameter δ for PNLMS is taken as:

$$\delta_{PNLMS} = \delta_{NLMS}, \tag{3.13}$$

where $\delta_{PNLMS} = \sigma_x^2$ is the variance of the input sinal. And for $q_i = 1, \forall i$, the PNLMS algorithm becomes equivalent to NLMS algorithm. A summary of the PNLMS algorithm is given in Table 3.3.

Table 3.3. Summary of the PNLMS algorithm.

Filter Output	$y(n) = \mathbf{w}(n)^T \mathbf{x}(n)$
Estimate Error	$e(n) = d(n) - y(n)$
Tap-Weight Adaptation	$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu \mathbf{Q}(n) e(n) \mathbf{x}(n)}{\mathbf{x}^T(n) \mathbf{Q}(n) \mathbf{x}(n) + \delta}$
	$\mathbf{Q}(n) = diag\left\{q_0(n), q_1(n), \ldots, q_{N-1}(n)\right\}$
	$q_i(n) = \frac{k_i(n)}{\sum_{i=0}^{N-1} k_i(n)}$
	$k_i(n) = \max\left\{\rho \times \max\left\{\gamma, \mid w_0(n) \mid, \ldots, \mid w_{N-1}(n) \mid\right\}, \mid w_i(n) \mid\right\}$
	$0 \le i < N-1$

3.5. The Improved PNLMS Algorithm

The PNLMS algorithm provides good performances such as fast initial convergence and well tracking ability when the echo path is sparse. And when the echo path is densely sparse, it unluckily converges slowly even much more slower than the NLMS algorithm. This makes the proportionality rule very aggresive. To overcome this aggresiveness, the improved PNLMS (IPNLMS) algorithm combines the NLMS update and proportionate term in the same rule each controlled by a factor α. Entries of the control matrix $\mathbf{Q}(n)$ for this case are given by [44],[45]:

$$q_i(n) = \frac{1-\alpha}{2N} + \frac{(1+\alpha) \mid w_i(n) \mid}{2 \| \mathbf{w}(n) \|_1 + \varepsilon},\ 0 \le i < N-1. \tag{3.14}$$

where ε is a small positive number to avoid dividing by zero, and $\| \cdot \|_1$ is the l_1-norm operator. It can be seen that the IPNLMS algorithm is the same as the NLMS algorithm when $\alpha = -1$ and PNLMS when $\alpha = 1$. Results from [46] shows that the good choice for $\alpha = 0, 0.5, -0.7$, is a favorable choice for most AEC applications. Table 3.4 is the summary of the IPNLMS.

Table 3.4. Summary of the IPNLMS algorithm.

Filter Output	$y(n) = \mathbf{w}(n)^T\mathbf{x}(n)$
Estimate Error	$e(n) = d(n) - y(n)$
Tap-Weight Adaptation	$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{\mu\mathbf{Q}(n)e(n)\mathbf{x}(n)}{\mathbf{x}^T(n)\mathbf{Q}(n)\mathbf{x}(n)+\delta}$
	$\mathbf{Q}(n) = diag\left\{q_0(n), q_1(n), \ldots, q_{N-1}(n)\right\}$
	$q_i(n) = \frac{1-\alpha}{2N} + \frac{(1+\alpha)\mid w_i(n)\mid}{2\Vert\ \mathbf{w}(n)\Vert_1+\varepsilon}$
	$0 \leq i < N-1$

3.6. Zero Attracting Sparse Adaptive filters

In many scenarios of acoustic echo cancellation, the sparse echo path exhibits a few number of active coefficients having large magnitudes among many negligible ones. This requires adaptive filters with ability to track those few active taps. In other situations, the impulse response may contain only a few nonzero values requiring adaptive filters with zero attracting properties, where the conventional LMS and NLMS algorithms can not provide a satisfactory performance. This problem was recently addressed by CS famework [26], [30] where a modified LMS algorithm was derived by inserting either l_0-norm [22], [36] or l_1-norm [37], [38] into the cost function of the standard LMS algorithm. This enables a zero attraction for all filter coefficients during the filtering process. This technique lead to the derivation of zero attracting LMS (ZA-LMS) algorithm [21], [20], [47], [50], [52]. Also a reweighted ZA-LMS (RZA-LMS) was subsequently derived similar to the ZA-LMS. With different reweighted zero attractors for different filter taps, RZA-LMS selectively induces filter taps to zero. This leads to a superior performance compared with the ZA-LMS algorithm.

3.6.1. Zero Attracting LMS Algorithm

The Zero-Attracting LMS (ZA-LMS) algorithm exploits the sparsity of a system by introducing the l_1-norm of filter taps in the quadratic cost function of the LMS [20]. ZA-LMS algorithm can be derived by inserting an l_1-norm constraint into (3.1) as:

$$J_1(n) = \frac{1}{2}e^2(n) + \gamma\|\mathbf{w}(n)\|_1, \tag{3.15}$$

where γ is the control parameter that determines the degree of attractor of the l_1-norm for $\mathbf{w}(n)$. The minimum of $J_1(n)$ can be sought recursively using the gradient method [2]:

$$
\begin{aligned}
\mathbf{w}(n+1) &= \mathbf{w}(n) - \mu\frac{\partial J_1(n)}{\partial \mathbf{w}(n)} \\
&= \mathbf{w}(n) - \rho \mathrm{sgn}(\mathbf{w}(n)) + \mu e(n)\mathbf{x}(n),
\end{aligned} \tag{3.16}
$$

where $\rho = \mu\gamma$ and $\mathrm{sgn}(\cdot)$ is a component-wise sign function defined as:

$$
sgn(x) = \begin{cases} \frac{x}{|x|} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0, \end{cases} \tag{3.17}
$$

Comparing (3.16) with the tap weight update of the conventional LMS (3.3), it can be seen that the ZA-LMS algorithm has the additional term $-\rho\mathrm{sgn}(\mathbf{w}(n))$ which constantly attracts the filter coefficients to zero. This term is known as the zero-attractor [20], whose strength is controlled by ρ. It accelerates convergence when most of the system coefficients are zero, i.e., the system is sparse. Table 3.5 gives the summary of the ZA-LMS algorithm.

Table 3.5. Summary of the ZA-LMS algorithm.

Filter Output	$y(n) = \mathbf{w}(n)^T\mathbf{x}(n)$
Estimate Error	$e(n) = d(n) - y(n)$
Tap-Weight Adaptation	$\mathbf{w}(n+1) = \mathbf{w}(n) - \rho\mathrm{sgn}(\mathbf{w}(n)) + \mu e(n)\mathbf{x}(n)$
	$sgn(x) = \begin{cases} \frac{x}{\lvert x\rvert} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$

3.6.2. Reweighted Zero Attracting LMS Algorithm

While the ZA-LMS algorithm uses the same zero-attractor to uniformly update all filter coefficients, the reweighted zero-attracting LMS (RZA-LMS) uses a distinct zero attractor for different filter coefficients. The reweighted zero attractors are obtained by adding a log-sum penalty of the coefficients vector to the cost function of the LMS algorithm. This is just an approximation of the l_0-norm penalty [60]. So the new cost function is defined by:

$$
J_2(n) = \frac{1}{2}e^2(n) + \gamma' \sum_{i=0}^{N-1} \log\left(1 + \frac{|\mathbf{w}_i|}{\zeta'}\right), \tag{3.18}
$$

where ζ' and γ' are positive constants. Using the gradient method, the update equation of the RZA-LMS algorithm becomes:

$$\begin{aligned}\mathbf{w}(n+1) &= \mathbf{w}(n) - \mu\frac{\partial J_2(n)}{\partial \mathbf{w}(n)} \\ &= \mathbf{w}(n) - \frac{\rho \mathrm{sgn}(\mathbf{w}(n))}{1+\zeta|\mathbf{w}(n)|} + \mu e(n)\mathbf{x}(n). \end{aligned} \tag{3.19}$$

where $\rho = \mu\gamma'/\zeta'$ and $\zeta = 1/\zeta'$. A summary of the RZA-LMS algorithm is given in Table 3.6.

Table 3.6. Summary of the RZA-LMS algorithm.

Filter Output	$y(n) = \mathbf{w}(n)^T\mathbf{x}(n)$
Estimate Error	$e(n) = d(n) - y(n)$
Tap-Weight Adaptation	$\mathbf{w}(n+1) = \mathbf{w}(n) - \frac{\rho \mathrm{sgn}(\mathbf{w}(n))}{1+\zeta|\mathbf{w}(n)|} + \mu e(n)\mathbf{x}(n)$
	$sgn(x) = \begin{cases} \frac{x}{|x|} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$

Despite the improvement brought by zero attracting LMS based algorithms, such approaches face some difficulties for identifying unknown systems associated with a variety of sparseness levels. They therefore have the effect of producing estimation bias while achieving sparsity exploitation because of the difficulties of norm constraint adaption [48]. This is due to lack of adjustable factor that can adapt the norm constraint to the unknown system associated with different sparsity. As a result, the effectiveness of both the ZA-LMS and RZA-LMS algorithms significantly decreases with the variation of sparsity of the unknown system. To overcome this problem, a non-uniform norm constraint LMS (NNCLMS) algorithm was proposed in [40], which enables the norm constraint to seek a trade-off between the sparsity exploitation effect and the estimation bias it produces.

3.6.3. Non-Uniform Norm Constraint LMS Algorithm

To limit the effect of estimation bias caused by norm constraint adaptation due to different sparsity of the unknown system, a classic p-norm like is employed to modify the cost function of the LMS filter. The classic p-norm like is split to form a non-uniform norm constraint that can be made more flexible for quantitative adjustment of the norm constraint adaptation. The NNCLMS algorithm is derived by first inserting the p-norm constraint into the cost function of the LMS filter.

$$J_3(n) = \frac{1}{2}e^2(n) + \gamma\|\mathbf{w}(n)\|_p^p \tag{3.20}$$

$$= \frac{1}{2}|d(n) - \mathbf{x}^T(n)\mathbf{w}(n)|^2 + \gamma\|\mathbf{w}(n)\|_p^p, \tag{3.21}$$

where $\gamma\|\mathbf{w}(n)\|_p^p$ is the p-norm like constraint term, and $\gamma > 0$ is the factor to balance the constraint term and estimation error [48], [49]. The term $\|\mathbf{w}(n)\|_p^p$ is called "p-norm like" which is different from euclidean norm and defined as:

$$|\mathbf{w}(n)\|_p^p = \sum |w_i|^p, \quad 0 \le p \le 1, \tag{3.22}$$

that is,

$$\lim_{p \to 0} \|\mathbf{w}(n)\|_p^p = \|\mathbf{w}(n)\|_0, \tag{3.23}$$

$$\lim_{p \to 1} \|\mathbf{w}(n)\|_p^p = \|\mathbf{w}(n)\|_1 = \sum_{i=1}^{n} |w_i|. \tag{3.24}$$

Hence, (3.23) and (3.24) which means counting the number of non-zero coefficients. The optimization problem represented in (3.18) could be solved by applying a gradient technique, the gradient of the cost function can be obtained with respect to $\mathbf{w}(n)$ as:

$$\widehat{\boldsymbol{\nabla}} J_3(n) = \frac{1}{2}\frac{\partial |e(n)|^2}{\partial \mathbf{w}} + \lambda \frac{\partial \|\mathbf{w}(n)\|_p^p}{\partial \mathbf{w}}, \tag{3.25}$$

Therefore, the gradient recursion of filter coefficient vector is:

$$\begin{aligned} w_i(n+1) &= w_i(n) - \mu \widehat{\boldsymbol{\nabla}} J_3(n) \\ &= w_i(n) + \mu e(n)x(n-i) - \frac{\kappa f \operatorname{sgn}[w_i(n)]}{1 + \varepsilon |w_i(n)|^{1-p}}, \ \forall 0 \le i < N, \end{aligned} \tag{3.26}$$

where f can be obtained as:

$$f = \frac{\operatorname{sgn}[\mathrm{E}\left[|w_i(n)|\right] - |w_i(n)|] + 1}{2}, \ \forall 0 \le i < N. \tag{3.27}$$

The last term in (3.26) imposes a non-uniform norm related to zero attraction on filter coefficients. Unlike l_1-norm zero attraction term, the impact of the non-uniform norm term is dependents on the value of each coefficient with respect to the expectation. The non-uniform zero-attractor is exerted to enhance the convergence of small coefficient, and disappear to remove the estimation bias caused by large coefficients, thus seeks a trade-off between these two types of impact on the performance of the algorithm. More details of the non-uniform p-norm vector will be explained in the next chapter. The main challenge of this filter is its inability to stabilize the sparse system when the input signal is highly correlated. The summary of the NNCLMS is provided in Table 3.7.

Table 3.7. Summary of the NNCLMS algorithm.

Initialize	$\mathbf{w} = zeros(N, 1)$
initial values of	$\kappa, \varepsilon, \mu, N$
for	$n = 1, 2, ...$ do
	$e(n) = d(n) - \mathbf{w}^T(n)\mathbf{x}(n)$
	$f = \frac{\text{sgn}[\text{E}[\lvert w_i(n)\rvert] - \lvert w_i(n)\rvert] + 1}{2}, \ \forall\, 0 \leq i < N,$
	$w_i(n+1) = w_i(n) + \mu e(n) x(n-i) - \kappa f \frac{\text{sgn}[w_i(n)]}{1+\varepsilon\lvert w_i(n)\rvert},$

CHAPTER 4

PROPOSED METHOD

4.1. Introduction

As has been previously shown that the NNCLMS algorithm alleviates the difficulties caused by norm constraint adaptation and offers a superior performance compared to zero attracting algorithms due to its non-uniform p-norm like constraint adjustment. This is because the norm constraint adaptation is adjusted to cope up with non-linearity of the norm constraint during the filtering process and result in efficient tracking of a sparsity changing system. On the other hand, the NNCLMS algorithm suffers from stability challenges in a correlated environment due to the step-size of the LMS algorithm. Hence, we propose an alternative approach to identify such a sparse system by using a VSSLMS algorithm employing the p-norm like constraint. The variable step-size update would stabilize the system and the non-uniform p-norm constraint would overcome the difficulties of the norm constraint adaptation. The proposed approach would be shown to perform more efficient than the existing approaches.

This chapter begins with a review of the VSSLMS algorithm followed by the derivation of the proposed algorithm. we, firstly, broadly demonstrate how the concept p-norm vector is optimized to a non-uniform p-norm like constraint. We then carry out the mean square convergence analysis and establish the stability condition of the proposed algorithm.

4.2. Review of Variable Step-Size LMS Algorithm

The VSSLMS algorithm was derived from the conventional LMS algorithm with a variable step-size [42], [56]. Recalling the cost function of the conventional LMS:

$$J(n) = \frac{1}{2}e^2(n), \tag{4.1}$$

where $e(n)$ is the instantaneous error and is given by,

$$e(n) = y(n) - \mathbf{w}^T(n)\mathbf{x}(n). \tag{4.2}$$

The filter coefficient vector is then updated by [12]:

$$\begin{aligned}\mathbf{w}(n+1) &= \mathbf{w}(n) - \mu\frac{\partial J(n)}{\partial \mathbf{w}(n)} \\ &= \mathbf{w}(n) + \mu e(n)\mathbf{x}(n),\end{aligned} \tag{4.3}$$

where μ is the step-size which serves as the condition of the LMS algorithm given by,

$$0 < \mu < \frac{1}{\lambda_{max}(\mathbf{R})}, \tag{4.4}$$

where λ_{max} is the maximum eigenvalue of **R**, and **R** is the autocorrelation matrix of the input tap vector. Usually, there is a trade off between the convergence rate and MSE value in LMS algorithm. The VSSLMS algorithm uses a variable step-size, as proposed in [42], in order to avoid such a trade off, and is given by,

$$\mu'(n+1) = \alpha\mu'(n) + \gamma e^2(n), \tag{4.5}$$

with $0 < \alpha < 1$ and $\gamma > 0$, then

$$\mu(n) = \begin{cases} \mu_{max} & \text{if } \mu'(n+1) > \mu_{max} \\ \mu_{min} & \text{if } \mu'(n+1) < \mu_{min} \\ \mu'(n+1) & \text{otherwise,} \end{cases} \tag{4.6}$$

where $0 < \mu_{min} < \mu_{max}$. Therefore, the step size is a positive value whose strength can be controlled by α, γ and the instantaneous error $e(n)$ as in (4.5). In other words, if the instantaneous error is large, the step size will increase to provide faster tracking. If the prediction error decreases, the step size will be decreased to reduce the misadjustment and, hence, provides a small steady-state error. The perfection of the algorithm could be attained at μ_{max}, μ_{max} is chosen in such a way to assure constrained MSE [43], [47], [57].

4.3. Proposed Algorithm

Despite the fact that the VSSLMS algorithm provides a good performance, its performance can be improved further by imposing the sparsity condition of the system. The proposed algorithm is derived by solving the optimization problem:

$$\mathbf{w}(n) = \arg\min_{\mathbf{w}} J_{1,n}(\mathbf{w}), \tag{4.7}$$

where $J_{1,n}(\mathbf{w})$ is the modified cost function in (4.1) which is achieved by incorporating a p-norm penalty function as:

$$J_{1,n}(\mathbf{w}) = \frac{1}{2}|e(n)|^2 + \lambda\|\mathbf{w}(n)\|_p^p. \tag{4.8}$$

The last term in (4.8) is a p-norm like constraint penalty, where λ is a positive constant whose value is used to adjust the penalty function. This function has a definition which is different from the classic Euclidean norm [40], defined as:

$$\|\mathbf{w}(n)\|_p^p = \sum_{i=1}^{n} |w_i|^p, \tag{4.9}$$

where $0 \leq p \leq 1$ and we can deduce that:

$$\lim_{p\to 0} \|\mathbf{w}(n)\|_p^p = \|\mathbf{w}(n)\|_0, \tag{4.10}$$

which counts the number of non-zero coefficients, and

$$\lim_{p\to 1} \|\mathbf{w}(n)\|_p^p = \|\mathbf{w}(n)\|_1 = \sum_{i=1}^{n} |w_i|. \tag{4.11}$$

Both (4.10) and (4.11) are utilized for proper solution and analysis of sparse system derived by the l_0 and l_1-norm algorithms as stated in [20]. Since the p-norm has been analyzed, next is to find a solution for the optimization problem in (4.7) by using a gradient minimization techniques, the gradient of the cost function with respect to the filter coefficient vector $\mathbf{w}(n)$ is:

$$\widehat{\boldsymbol{\nabla}} J_{1,n}(\mathbf{w}) = \frac{1}{2}\frac{\partial |e(n)|^2}{\partial \mathbf{w}} + \lambda\ \frac{\partial \|\mathbf{w}(n)\|_p^p}{\partial \mathbf{w}}, \tag{4.12}$$

whose solution is found to be

$$\widehat{\boldsymbol{\nabla}} J_{1,n}(\mathbf{w}) = \mathbf{w}(n) - \mathbf{x}(n)e(n) + \lambda p \frac{\text{sgn}[\mathbf{w}(n)]}{\|\mathbf{w}(n)\|^{1-p}}. \tag{4.13}$$

Thus, from the gradient descent recursion,

$$\begin{aligned} w_i(n+1) &= w_i(n) - \mu(n)\widehat{\nabla} J_{1,n}(w_i) \\ &= w_i(n) + \mu(n)e(n)x(n-i) - \kappa(n)p\frac{\text{sgn}[w_i(n)]}{\sigma + |w_i(n)|^{1-p}},\ \ \forall\, 0\ \leq i <\ N, \end{aligned} \tag{4.14}$$

where $\forall\, 0 \leq i < N$, $\mu(n)$ is the variable step-size of the algorithm given by (4.1), $\kappa = \mu\lambda > 0$ is an adjustable parameter controlling the stability of the system, σ is a very small positive constant to avoid dividing by zero and $\mathrm{sgn}[w_i(n)]$ is a component-wise sign function.

The introduction of the p-norm facilitates the optimization of the norm constraint, this can be achieved by adjusting the parameter p as in (4.9). This parameter has effect on both the estimation bias and the intensity of the sparsity measure, hence the trade-off makes it difficult to achieve the best solution for the optimization problem.

To address these problems, the classic p-norm as in (4.7) is riven into a non-uniform p-norm like [40]. In this method, a different value of p is assigned for each entry of $\mathbf{w}(n)$ as:

$$\|\mathbf{w}(n)\|_{p,N}^{p} = \sum_{i=1}^{N} |w_i|^{p_i}, \tag{4.15}$$

where $0 \leq p_i \leq 1$ and the new cost function, which is subjected to (4.14), can be deduced from the gradient descent recursion equation as,

$$w_i(n+1) = w_i(n) + \mu(n)e(n)x(n-i) - \kappa(n)p_i \frac{\mathrm{sgn}[w_i(n)]}{\sigma + |w_i(n)|^{1-p_i}}. \tag{4.16}$$

where $\forall\, 0 \leq i < N$, and the introduction of $\mathbf{p}_i$ vector makes it feasible to control the effect of estimation bias and sparsity correction measure by assigning a different value of p for every tap weight vector. The last part of (4.16) suggests that a metric of the absolute value of $w_i(n)$ can be introduced to quantitatively classify the filter coefficients into small and large categories.

By considering the range of the expected value of the tap weight vector, the absolute value expectation can be defined as:

$$h_i(n) = \mathrm{E}\left[|w_i(n)|\right], \ \forall\, 0 \leq i < N, \tag{4.17}$$

Since we are interested in the minimum possible value of **p** rather than its index, and minimizing the term $\frac{p_i}{|w_i(n)|^{1-p_i}}$ in (4.16) is equivalent to minimizing $p_i|w_i(n)|^{1-p_i}$, then the optimization of the large category for each entry of **p** can be expressed as:

$$\min_{p_i}[p_i\,|w_i(n)|^{p_i-1}] = 0, \ \ sub.\,to: \ w_i(n) > h_i(n)], \tag{4.18}$$

and for the small category, p_i is set to be unity so as to avoid imbalance between the extremely great or slight intensity caused by various values of small $w_i(n)$. Therefore, the comprehensive optimization of the non-uniform norm constraint will cause p_i to take a value of either a 0 or 1 when $w_i(n) > h_i(n)$ or $w_i(n) < h_i(n)$, respectively. With these, the update equation becomes:

$$w_i(n+1) = w_i(n) + \mu(n)e(n)x(n-i) - \kappa(n) f \,\mathrm{sgn}[w_i(n)], \;\; \forall\, 0 \leq i < N, \tag{4.19}$$

where f_i can be obtained by:

$$f_i = \frac{\mathrm{sgn}[\mathrm{E}\left[|w_i(n)|\right] - |w_i(n)|] + 1}{2}, \;\; \forall\, 0 \leq i < N. \tag{4.20}$$

The second term in (4.19) provides a variable step size update while the last term imposes a non-uniform norm constraint whose function attracts small filter coefficients to zero. Unlike other norm constraint algorithms, the norm exertion here depends on the value of individual coefficients with respect to the expectation of the tap weight. Also, the zero-attractor of the non-uniform norm increases the convergence rate of small coefficients and eliminates the estimation bias caused by large coefficients, hence improves the performance of the algorithm [51]. A summary of the proposed algorithm is given in Table 4.1.

Table 4.1. Summary of the proposed algorithm.

Initialize	$\mathbf{w} = zeros(N,1)$
initial values of	$\kappa(n), \varepsilon, \mu_{max}, \mu_{min}, \gamma, \alpha$
for	$n = 1, 2, ...$ do
	$e(n) = d(n) - \mathbf{w}^T(n)\mathbf{x}(n)$
	$w_i(n+1) = w_i(n) + \mu(n)e(n)x(n-i) - \kappa(n) f_i \frac{\mathrm{sgn}[w_i(n)]}{1+\varepsilon\|w_i(n)\|}$,
	$f_i = \frac{\mathrm{sgn}[\mathrm{E}[\|w_i(n)\|]-\|w_i(n)\|]+1}{2}, \;\; \forall\, 0 \leq i < N$,
where	$\mu'(n+1) = \alpha\mu'(n) + \gamma e^2(n)$,
	$\mu(n) = \begin{cases} \mu_{max} & \text{if } \mu'(n+1) > \mu_{max} \\ \mu_{min} & \text{if } \mu'(n+1) < \mu_{min} \\ \mu'(n+1) & \text{otherwise.} \end{cases}$

Moreover, the performance of the proposed algorithm can be improved further by reweighting the term introduced in (4.20) which makes it completely behaves as a reweighted zero attractor [20]. By this, the update equation of the proposed algorithm can be expressed as:

$$w_i(n+1) = w_i(n) + \mu(n)e(n)x(n-i) - \kappa f_i \frac{\text{sgn}[w_i(n)]}{1+\varepsilon|w_i(n)|}, \ \forall 0 \leq i < N, \tag{4.21}$$

where $\varepsilon > 0$ is an adjustable parameter whose value controls the reweighing factor. Table 4.1 provides a summary of the proposed algorithm.

4.4. Convergence Analysis of the Proposed Algorithm

This section presents the convergence analysis of the proposed algorithm and derivation of a stability criteria. The convergence analysis is aimed to ensure that the algorithm satisfies the conditions needed for the application requirements. Now, lets begin by substituting (4.20) in (4.21),

$$\begin{aligned} w_i(n+1) &= w_i(n) + \mu(n)e(n)x(n-i) \\ &- \kappa\left(\frac{\text{sgn}(\text{E}\left[|w_i(n)|\right] - |w_i(n)|) + 1}{2(1+\varepsilon|w_i(n)|)}\right)\text{sgn}\left[w_i(n)\right]. \end{aligned} \tag{4.22}$$

Assuming that an i.i.d zero mean Gaussian input signal $x(n)$ corrupted by a zero-mean white noise $v(n)$, the filter misalignment vector can be defined as:

$$\delta\mathbf{w}(n) = \mathbf{w}(n) - \mathbf{w}_0 \tag{4.23}$$

where $\mathbf{w}_0$ represents the unknown system coefficients vector.
The mean and the auto-covariance matrix of $\delta\mathbf{w}(n)$ can be written as,

$$\boldsymbol{\epsilon}(n) = \text{E}[\delta\mathbf{w}(n)], \tag{4.24}$$

$$\mathbf{S}(n) = \text{E}\left[\mathbf{q}(n)\mathbf{q}^T(n)\right], \tag{4.25}$$

where $\mathbf{q}(n)$ is the zero-mean misalignment vector defined as:

$$\mathbf{q}(n) = \delta\mathbf{w}(n) - \text{E}[\delta\mathbf{w}(n)], \tag{4.26}$$

The instantaneous mean-square-deviation (MSD) can be defined as:

$$\begin{aligned} J(n) &= \text{E}[\|\delta\mathbf{w}(n)\|_p^p] \\ &= \sum_{i=0}^{N-1}\Lambda_i(n), \end{aligned} \tag{4.27}$$

where $\Lambda_i(n)$ denotes the i^{th}-tap MSD and is defined with respect to the i^{th} element of $\delta\mathbf{w}(n)$

$$\begin{aligned}\Lambda_i(n) &= \mathrm{E}[\delta_i^2(n)] \\ &= S_{ii}(n) + \epsilon_i^2(n);\;\; i = 0, \ldots, N-1\end{aligned} \tag{4.28}$$

where $S_{ii}(n)$ and $\epsilon_i(n)$ are the i^{th} diagonal element and the i^{th} element of the $S(n)$ and $\epsilon(n)$, respectively [31].

Substituting $d(n) = \mathbf{w}_0^T\mathbf{x}(n) + v(n)$ and (4.2) in (4.22) gives

$$\begin{aligned}w_i(n+1) &= w_i(n) + \mu(n)[\mathbf{x}^T(n)\mathbf{w}_0 + v(n) - \mathbf{x}^T(n)\mathbf{w}(n)]x(n-i) \\ &- \kappa\left(\frac{\mathrm{sgn}(\mathrm{E}\left[|w_i(n)|\right] - |w_i(n)|) + 1}{2(1+\varepsilon|w_i(n)|)}\right)\mathrm{sgn}\left[w_i(n)\right],\end{aligned} \tag{4.29}$$

equation (4.29) can be rewritten in vector form as,

$$\begin{aligned}\mathbf{w}(n+1) &= \mathbf{w}(n) + \mu(n)\mathbf{x}(n)\left[\mathbf{x}^T(n)\mathbf{w}_0 + v(n) - \mathbf{x}^T(n)\mathbf{w}(n)\right] \\ &- \frac{\kappa}{2}\left(\{\mathrm{sgn}(\mathrm{E}\left[|\mathbf{w}(n)|\right] - |\mathbf{w}(n)|) + \mathbf{1}\} \odot \mathrm{sgn}\left[\mathbf{w}(n)\right]\right) \\ &\oslash \left(\mathbf{1} + \varepsilon|\mathbf{w}(n)|\right),\end{aligned} \tag{4.30}$$

where $|\mathbf{w}|$ is the element wise absolute of $\mathbf{w}$, $\mathbf{1}$ denotes a vector of ones of the same size of $\mathbf{w}$, and $\odot$ and $\oslash$ denote element-by-element vector multiplication and division, respectively. Subtracting $\mathbf{w}_0$ from both sides of (4.30) and substituting (4.23) yields

$$\begin{aligned}\boldsymbol{\delta}(n+1) &= \mathbf{A}(n)\boldsymbol{\delta}(n) + \mu(n)\mathbf{x}(n)v(n) - \frac{\kappa(n)}{2}(\{\mathrm{sgn}(\mathrm{E}\left[|\mathbf{w}(n)|\right] - |\mathbf{w}(n)|) + \mathbf{1}\} \\ &\odot \;\mathrm{sgn}\left[\mathbf{w}(n)\right]) \oslash \left(\mathbf{1} + \varepsilon|\mathbf{w}(n)|\right),\end{aligned} \tag{4.31}$$

$$\mathbf{A}(n) = \mathbf{I}_N - \mu(n)\mathbf{x}(n)\mathbf{x}^T(n), \tag{4.32}$$

where $\mathbf{I}_N$ denotes the $N \times N$ identity matrix.

Taking the expectation of (4.31) and using the independence assumption [59], [62], [61] results in

$$\begin{aligned}\boldsymbol{\epsilon}(n+1) &= [1 - \mu(n)\sigma_x^2]\boldsymbol{\epsilon}(n) - \frac{\kappa(n)}{2}\mathrm{E}[(\{\mathrm{sgn}(\mathrm{E}\left[|\mathbf{w}(n)|\right] - |\mathbf{w}(n)|) + \mathbf{1}\}) \\ &\odot \;\mathrm{sgn}\left[\mathbf{w}(n)\right])] \oslash \left(\mathbf{1} + \varepsilon|\mathbf{w}(n)|\right),\end{aligned} \tag{4.33}$$

where σ_x^2 is the variance of $x(n)$. Now, subtracting (4.33) from (4.31) and substituting (4.24) and (4.26) gives,

$$\mathbf{q}(n+1) = \mathbf{A}(n)\mathbf{q}(n) + \mu(n)\mathbf{x}(n)v(n) + \frac{\kappa(n)}{2}\mathbf{z}(n) \tag{4.34}$$

where

$$\begin{aligned} \mathbf{z}(n) &= \mathrm{E}\left[(\mathrm{sgn}(\mathrm{E}\left[|\mathbf{w}(n)|\right] - |\mathbf{w}(n)|) + \mathbf{1}) \oslash ([\mathbf{1} + \varepsilon|\mathbf{w}(n)|])\right] \odot \mathrm{sgn}\left[\mathbf{w}(n)\right] \\ &- (\mathrm{sgn}(\mathrm{E}\left[|\mathbf{w}(n)|\right] - |\mathbf{w}(n)|) + \mathbf{1}) \oslash ([\mathbf{1} + \varepsilon|\mathbf{w}(n)|]) \odot \mathrm{sgn}\left[\mathbf{w}(n)\right]. \end{aligned} \tag{4.35}$$

By (4.25),

$$\begin{aligned} \mathbf{S}(n+1) &= \mathrm{E}\left\{\mathbf{q}(n+1)\mathbf{q}^T(n+1)\right\} \\ &= \mathrm{E}\left\{ \left[\mathbf{A}(n)\mathbf{q}(n) + \mu(n)\mathbf{x}(n)v(n) + \frac{\kappa(n)}{2}\mathbf{z}(n)\right]\right. \\ &\times \left.\left[\mathbf{A}(n)\mathbf{q}(n) + \mu(n)\mathbf{x}(n)v(n) + \frac{\kappa(n)}{2}\mathbf{z}(n))\right]^T \right\}, \end{aligned} \tag{4.36}$$

$$\begin{aligned} \mathbf{S}(n+1) &= \mathrm{E}\left\{\mathbf{A}(n)\mathbf{q}(n)\mathbf{q}^T(n)\mathbf{A}^T(n)\right\} + \mathrm{E}\left\{\mathbf{A}(n)\mathbf{q}(n)\mu(n)v(n)\mathbf{x}^T(n)\right\} \\ &+ \frac{\kappa(n)}{2}\mathrm{E}\left\{\mathbf{A}(n)\mathbf{q}(n)\mathbf{z}^T(n)\right\} + \mathrm{E}\left\{\mu(n)v(n)\mathbf{x}(n)\mathbf{q}^T(n)\mathbf{A}^T(n)\right\} \\ &+ \mathrm{E}\left\{\mu^2 v^2(n)\mathbf{x}(n)\mathbf{x}^T(n)\right\} + \frac{(n)}{2}\mathrm{E}\left\{\mu(n)v(n)\mathbf{x}(n)\mathbf{z}^T(n)\right\} \\ &+ \frac{\kappa(n)}{2}\mathrm{E}\left\{\mathbf{z}(n)\mathbf{q}^T(n)\mathbf{A}^T(n)\right\} + \frac{\kappa(n)}{2}\mathrm{E}\left\{\mu(n)v^T(n)\mathbf{z}(n)\mathbf{x}^T(n)\right\} \\ &+ \left(\frac{\kappa(n)}{2}\right)^2 \mathrm{E}\left\{\mathbf{z}(n)\mathbf{z}^T(n)\right\}, \end{aligned} \tag{4.37}$$

To evaluate (4.37), we use the fact that the fourth-order moment of a Gaussian variable is three times its variance squared [55], and by the independence assumption [2] and symmetric behavior of the covariance matrix $\mathbf{S}(n)$,

$$\begin{aligned} \mathrm{E}\left\{\mathbf{A}(n)\mathbf{q}(n)\mathbf{q}^T(n)\mathbf{A}^T(n)\right\} &= \left(1 - 2\mu(n)\sigma_x^2 + 2\mu^2(n)\sigma_x^4\right)\mathbf{S}(n) \\ &+ \mu^2(n)\sigma_x^4 \mathrm{tr}[\mathbf{S}(n)] I_N, \end{aligned} \tag{4.38}$$

and

$$\begin{aligned}\mathrm{E}\left\{\mathbf{A}(n)\mathbf{q}(n)\mathbf{z}^T(n)\right\} &= \mathrm{E}^T\left\{\mathbf{z}(n)\mathbf{q}^T(n)\mathbf{A}^T(n)\right\} \\ &= (1-\mu(n)\sigma_x^2)\mathrm{E}[\mathbf{w}^T(n)\mathbf{z}(n)]. \end{aligned} \tag{4.39}$$

where $\mathrm{tr}[.]$ denotes the trace of a matrix. Now finding the trace of (4.37) and by using the results of (4.38) and (4.39), we obtain

$$\begin{aligned}\mathrm{tr}[\mathbf{S}(n+1)] &= \left(1-2\mathrm{E}\left\{\mu(n)\right\}\sigma_x^2+2\mathrm{E}\left\{\mu^2(n)\right\}\sigma_x^4\right)\mathrm{tr}[\mathbf{S}(n)] \\ &+ N\mathrm{E}\left\{\mu^2(n)\right\}\sigma_x^4\mathrm{tr}[\mathbf{S}(n)] \qquad (4.40) \\ &+ N\mathrm{E}\left\{\mu^2(n)\right\}\sigma_v^2\sigma_x^2+\kappa(1-\mathrm{E}\left\{\mu(n)\right\}\sigma_x^2)\mathrm{E}[\mathbf{w}^T(n)\mathbf{z}(n)], \qquad (4.41)\end{aligned}$$

where σ_v denotes the variance of $v(n)$.

From (4.35) it is obvious that $\mathbf{z}(n)$ is bounded, and hence, the term (4.41) $\boldsymbol{\epsilon}(n)$ converges. Thus, the adaptive filter is stable if the following holds:

$$|1-2E\{\mu(n)\}\sigma_x^2+(N+2)E\{\mu^2(n)\}\sigma_x^4|<1. \tag{4.42}$$

As the algorithm converges (n is sufficiently large), the error $e(n)\to 0$, and hence by (4.5), $\mu(n)$ becomes constant. In this case, $E\{\mu^2(n)\}=E\{\mu(n)\}^2=\mu^2(n)$. Hence the above equation simplifies to

$$0<\mu(\infty)<\frac{2}{(N+2)\sigma_x^2}. \tag{4.43}$$

This result shows that if μ satisfies (4.43), the convergence of the proposed algorithm is guaranteed.

CHAPTER 5

SIMULATION RESULTS

5.1. Introduction

This chapter presents numerical simulations that investigate the performance of the proposed algorithm in terms of convergence rate and steady-state behavior of the MSD estimate between the adaptive filter taps and acoustic impulse response taps (defined as $MSD = \mathrm{E}\|\mathbf{w}_0 - \mathbf{w}(n)\|^2$). The experiments are carried out in two main sections. In the first section, we investigate the performance of the proposed algorithm in acoustic echo cancellation setting with acoustic echo path of fix sparsity level. In the second section, the experiment investigate the performance of the proposed algorithm in other systems with a varying sparsity. All experiment are tested in different noise environments. The choice of parameters influences the performance of one algorithm over another. These parameters must be chosen in line with the convergence condition and stability criteria of the algorithm. Hence, it is important to carefully select these parameters. In our own case, extensive simulations have been conducted in selecting the parameters in order to provide optimal performances in terms of convergence rate and MSD estimate. The obtained parameters were used for the experiments, both, in AWGN and ACGN. In Table 5.1, we provide an example of showing how the parameters of the proposed algorithm are chosen for the experiment given in Section 5.2.2.

Table 5.1. Optimal paramters, $N = 256$

	μ_{max}	μ_{min}	κ	ε
Trial 1	0.009	0.001	0.0032	10
Trial 2	0.008	0.002	0.003	10
Trial 3	0.007	0.003	0.002	10
Trial 4	0.005	0.0035	0.001	10

5.2. Fixed Sparsity

All simulations of this section were done with the echo paths represented in Fig. 5.1 and Fig. 5.2 with length $N = 256$ and $N = 512$ respectively. The spareness measure of the impulse response is given

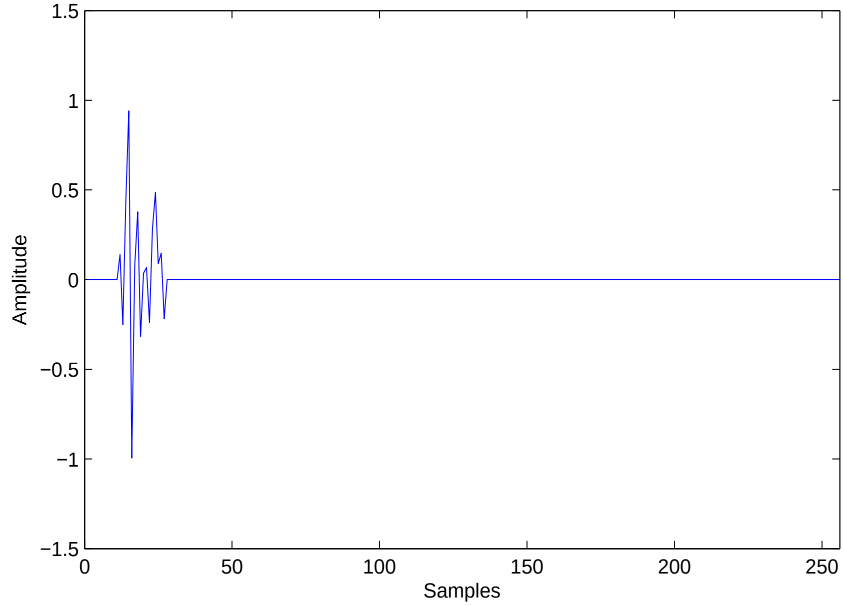

Figure 5.1. Acoustic echo path with $N = 256$.

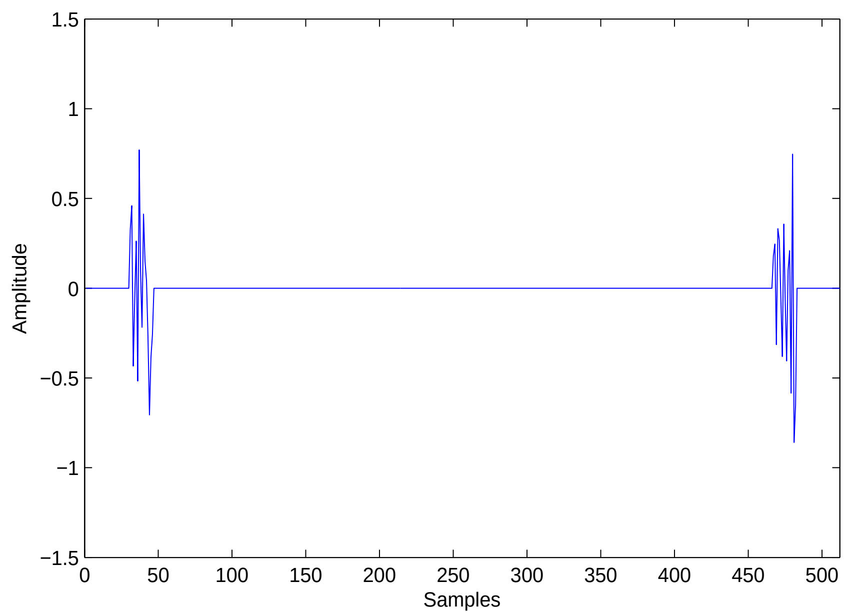

Figure 5.2. Acoustic echo path with $N = 512$.

by $\xi_{12}(\mathbf{w}_0) = \frac{N}{N-\sqrt{N}}\left(1 - \frac{\|\mathbf{w}_0\|_1}{\sqrt{N}\|\mathbf{w}_0\|_2}\right)$ [7]. The sparseness of the impulse response has been confirmed to be 0.6131, equivalent to G168 recommendation for standard acoustic impulse response [46]. For each experiment, the far-end signal was either corrupted by an additive white Gaussian noise (AWGN) or additive correlated Gaussian noise (ACGN). In all the experiments, the results were averaged over 200 independent trials and all parameters were selected by many trials to obtain the best performances.

5.2.1. Additive White Gaussian Noise

5.2.2. Filter length, $N = 256$

In this experiment, the performance of the proposed algorithm is compared to those of the NLMS, PNLMS, IPNLMS, ZA-LMS, RZA-LMS and NNCLMS algorithms in AWGN environment. The echo path is represented by the impulse response shown in Fig. 5.1. A zero mean white Gaussian sequence is used as the input signal (i.e., the far-end signal). The output of the echo path is corrupted by another zero mean sequence (i.e. the near-end signal), with a signal-to-noise ratio (SNR) of 30dB. For all the algorithms, simulations were done with the parameters given in Table 5.2. Fig. 5.3 shows that the NLMS, PNLMS and IPNLMS algorithms exhibit a slower convergence behavior than ZA-LMS algorithm with almost the same MSD estimate. The RZA-LMS algorithm converges faster than the previous algorithms with 1.5 dB lower MSD. The NNCLMS algorithm converges at the same rate with the ZA-LMS algorithm but achieves 1 dB lower MSD. However, the proposed algorithm converges at the same rate with the NNCLMS algorithm but gives 2 dB lower MSD. This will be on the account of some extra computations. However, the basic proportionate type adaptive filters displayed a slower convergence and poor tracking ability compared to the l_1-norm and p-norm based sparse adaptive filters. This is because these algorithms are designed for relatively low order filter types.

Table 5.2. Parameters used for the experiment in Section 5.2.2.

Algorithms	$N = 256$, AWGN								
	μ	ρ	ε	κ	γ	α	μ_{max}	μ_{min}	δ
Proposed	—	—	10	0.0001	0.001	0.97	0.005	0.0035	—
NNCLMS	0.005	—	10	0.0019	—	—	—	—	—
RZA-LMS	0.0056	0.000018	10	—	—	—	—	—	—
ZA-LMS	0.005	0.00018	—	—	—	—	—	—	—
IPNLMS	0.2	—	0.6	—	—	-0.75	—	—	0.000001
PNLMS	0.2	0.02	—	—	0.01	—	—	—	0.00001
NLMS	0.2	—	—	—	—	—	—	—	0.00001

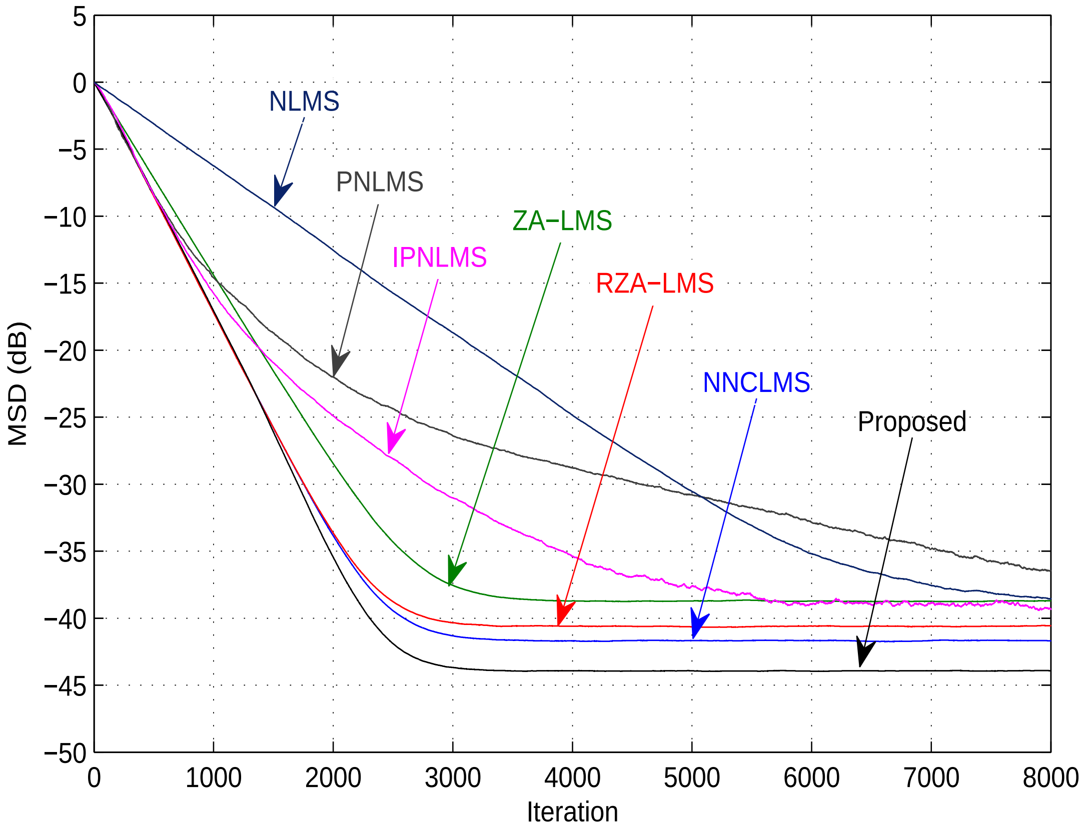

Figure 5.3. Ensemble learning curves of all algorithms in AWGN, SNR=30 dB and $N = 256$.

5.2.3. Filter length, $N = 512$

Due to the excessive length of the echo path, higher order adaptive filters are readily required for echo cancellation application. This affects the performance many of the available adaptive filters. To test the performance of the proposed algorithm due to long acoustic impulse response, we used a filter order of $N = 512$ to simulate the acoustic path represented in Fig. 5.2 with length $N = 512$. The standard sparseness degree (0.6131) is also maintained. The experiment was implemented with the same input signal and noise of the experiment in Section 5.2.1. Simulations were done with the parameters given in Table 5.3. From the table it can be observed that the step-size parameters of the proposed algorithm decreased as the filter length is increased as expected. We notice from Fig. 5.4 that, even though the filter length is increased, the proposed algorithm performance remains superior compared to the other algorithms.

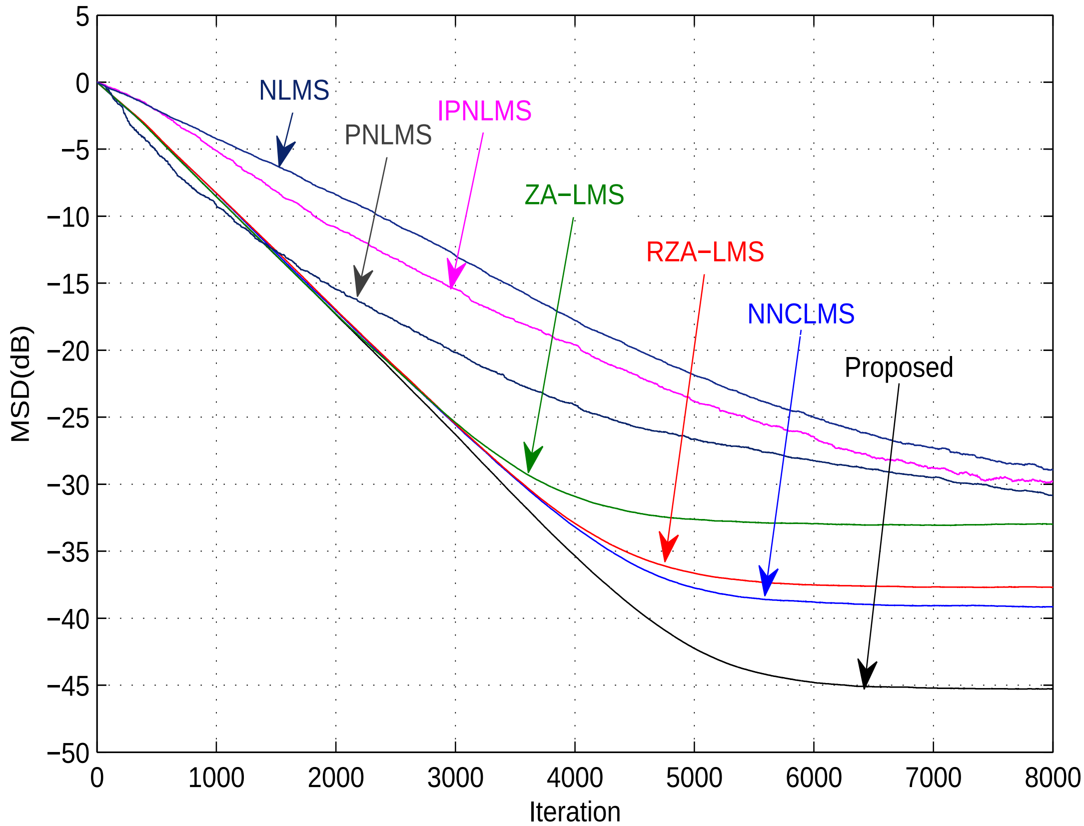

Figure 5.4. Ensemble learning curves of all algorithms in AWGN, SNR=30 dB and $N = 512$.

Table 5.3. Parameters used for the experiment in Section 5.2.3.

Algorithms	$N = 512$, AWGN								
	μ	ρ	ε	κ	γ	α	μ_{max}	μ_{min}	δ
Proposed	—	—	10	0.002	0.002	0.91	0.0025	0.0015	—
NNCLMS	0.0025	—	10	0.000016	—	—	—	—	—
RZA-LMS	0.0025	0.000018	10	—	—	—	—	—	—
ZA-LMS	0.0025	0.00011	—	—	—	—	—	—	—
IPNLMS	0.32	—	0.6	—	—	-0.75	—	—	0.000001
PNLMS	0.3	0.097	—	—	0.01	—	—	—	0.000001
NLMS	0.3	—	—	—	—	—	—	—	0.000001

5.2.4. Additive Correlated Gaussian Noise

In order to see how the proposed algorithm is robust to the noise type, an ACGN is added to the input signal. The ACGN process is created using the first-order autoregressive (AR(1)) model defined as:

$$\eta(n) = \wp\eta(n-1) + v(n), \tag{5.1}$$

where $\wp$ is a correlation parameter and $v(n)$ is a white Gaussian process with zero mean and a variance that provides a 30 dB SNR.

5.2.5. Low Correlated Gaussian Noise

In this experiment, we study the effect of low correlated Gaussian noise environment on the performance of the proposed algorithm. The low correlated Gaussian noise process was obtained from the AR(1) model generated in (5.1) by setting the correlation parameter to be $\wp = 0.45$ while maintaining the SNR value. In this case, we used a filter order of $N = 256$, which is equivalent to the length of the echo path depicted in Fig. 5.1. All the algorithms were implemented using the parameters given in Table 5.4. From Fig. 5.5, we observe that the proposed algorithm still provides the best performance with 1.5 dB lower MSD than the best of other algorithms. In the other hand, the proposed algorithm is faster than the best performer among the other algorithms by almost 200 iterations. This is by the virtue of the variable step-size which shows the ability of the proposed algorithm in suppressing correlated noise.

Table 5.4. Parameters used for the experiment in Section 5.2.5.

Algorithms	$N = 256$, ACGN								
	μ	ρ	ε	κ	γ	α	μ_{max}	μ_{min}	δ
Proposed	—	—	10	0.001	0.001	0.97	0.0056	0.0045	—
NNCLMS	0.0035	—	10	0.0019	—	—	—	—	—
RZA-LMS	0.0045	0.000018	10	—	—	—	—	—	—
ZA-LMS	0.005	0.000018	—	—	—	—	—	—	—
IPNLMS	0.2	—	0.6	—	—	-0.75	—	—	0.000001
PNLMS	0.2	0.02	—	—	0.01	—	—	—	0.00001
NLMS	0.2	—	—	—	—	—	—	—	0.00001

5.2.6. Highly Correlated Gaussian Noise

In this experiment, we investigate the performance of the proposed algorithm in relatively high correlated noise environment. The input signal is assumed to be corrupted by ACGN process generated by the AR(1)

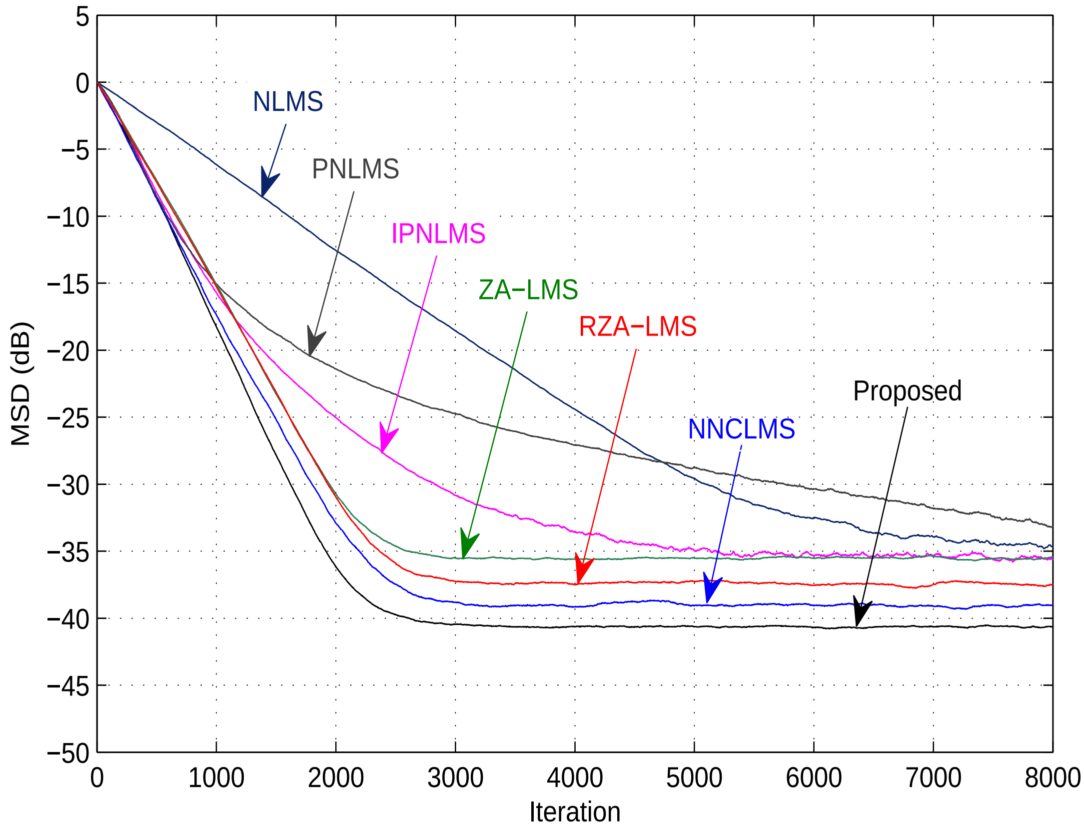

Figure 5.5. Ensemble learning curves of all algorithms in low correlated ACGN, SNR=30 dB and $N = 256$.

model created in (5.1), with the correlation parameter set to $\wp = 0.7$ (i.e., relatively high correlated Gaussian noise) with the same SNR value. The experiment was done with the parameters provided in Table 5.4. Fig. 5.6 shows that with the relatively high ACGN, the NLMS, PNLMS and IPNLMS algorithms have shown poor performances in terms of MSD and convergence rate values. However, the proposed algorithm provides the best MSD estimate among all algorithms (2 dB, 4 dB and 7 dB lower MSD than the NNCLMS, RZA-LMS and ZA-LMS algorithms, respectively). This shows that, despite the high MSD estimate provided by all the algorithms in a relatively high correlated noise environment, the proposed algorithm outperforms all other algorithms in terms of the MSD estimate.

Since the performance of all algorithms degrade a little in high correlated noise environments, we will proceed to investigate the performance of the proposed algorithm in only low correlated noise environments using a higher order filter. Therefore, in this experiment, the performance of the proposed algorithm is tested due to excessive length of the echo path in a low correlated Gaussian noise environment by repeating

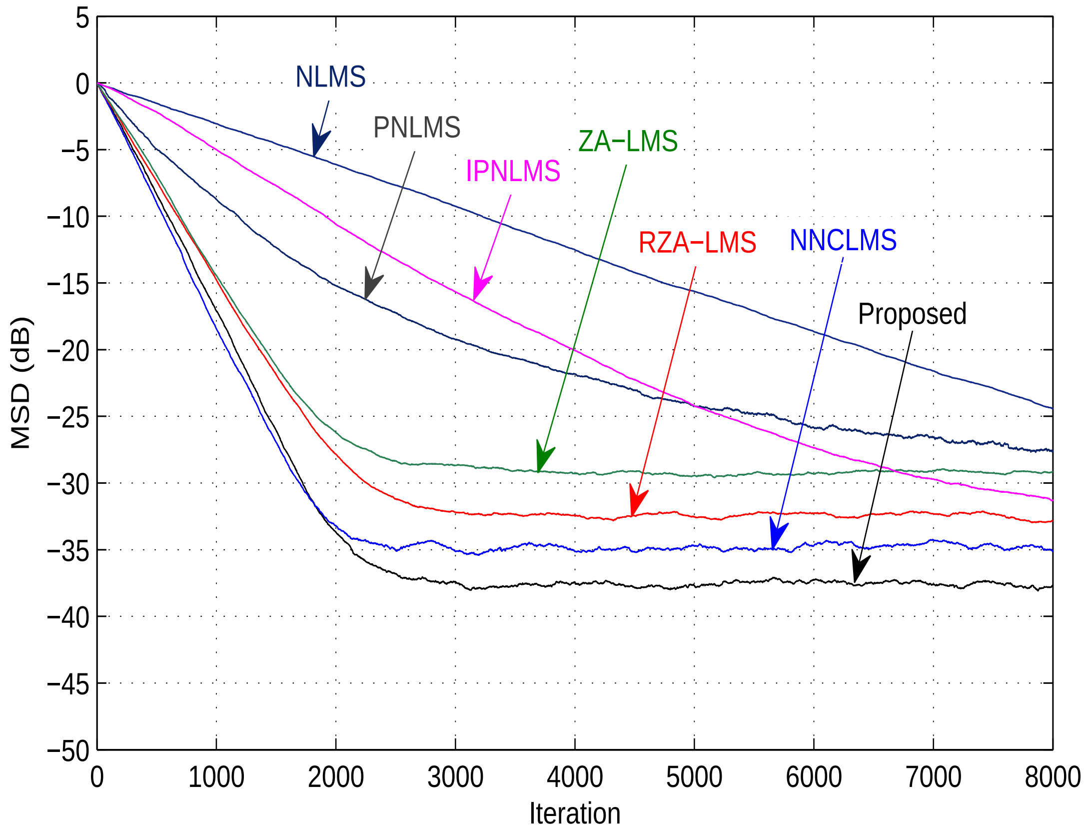

Figure 5.6. Ensemble learning curves of all the algorithms in highly correlated ACGN, SNR=30 dB and $N = 256$.

the experiment in Section 5.2.5, but with a filter length $N = 512$ and the echo path represented in Fig. 5.2. The correlated signal used in this case, which is usually encountered in acoustic echo cancellation problems is created exactly as in the same Section 5.2.5. The parameters used for the simulations are provided in Table 5.5. It can be seen from Fig. 5.7 that the NLMS, PNLMS and IPNLMS algorithms fail to reach the steady-state MSD within the time extent shown. However, even though all algorithms display slow convergence behavior, the proposed algorithm is a little bit faster than the others. In terms of MSD estimate, the proposed algorithm gives an improvement of 3 dB, 6 dB and 7 dB compared to NNCLMS, RZA-LMS and ZA-LMS algorithms, respectively. Hence, despite the length of the echo path and the effect of the correlation parameter, the proposed algorithm performs robustly and better than the rest of algorithms both in convergence rate and MSD estimate. The NLMS algorithm is clearly seen to provide poor performance among others. The proportionate updating based optimization approach utilized by PNLMS and IPNLMS algorithms is observed to be inefficient for this experimental settings and, therefore, they

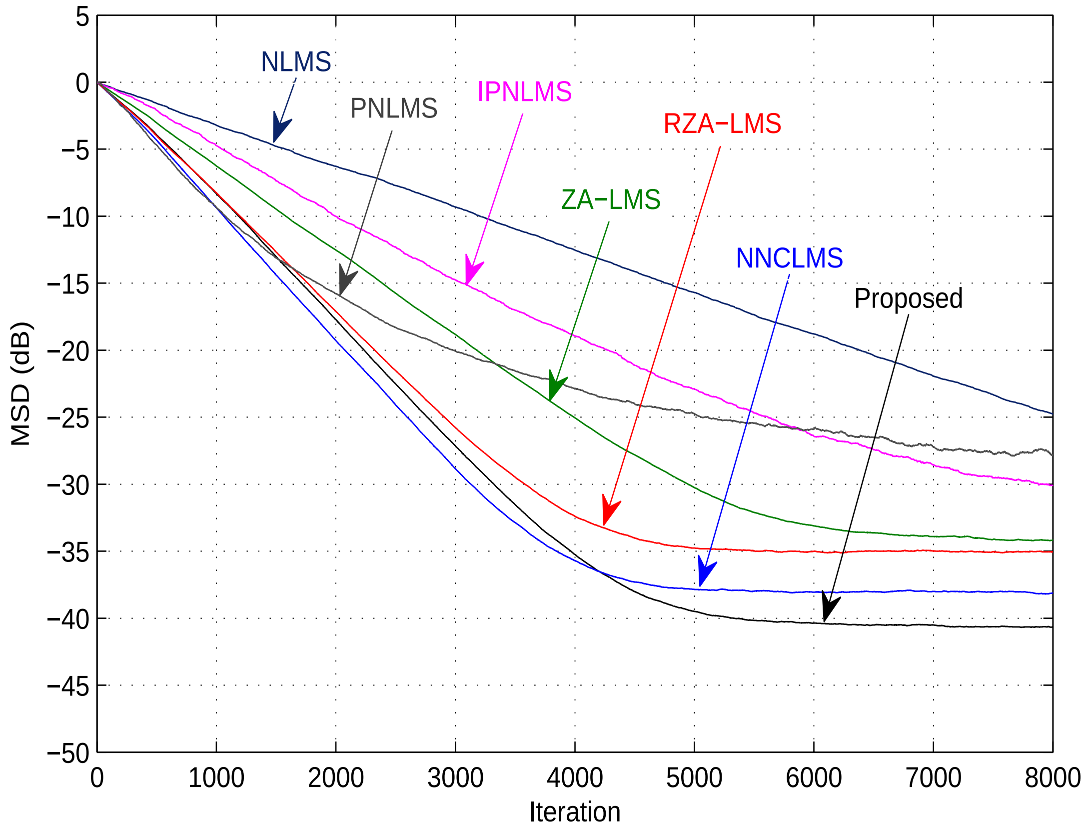

Figure 5.7. MSD curves of all algorithms in ACGN, $N = 512$.

provide unsatisfactory performances compared to NNCLMS, RZA-LMS and ZA-LMS algorithms.

Table 5.5. Parameters for the experiment in Section 5.2.6.

Algorithms	$N = 512$, ACGN								
	μ	ρ	ε	κ	γ	α	μ_{max}	μ_{min}	δ
Proposed	—	—	10	0.00001	0.0001	0.91	0.0025	0.0015	—
NNCLMS	0.0027	—	10	0.0017	—	—	—	—	—
RZA-LMS	0.0027	0.000018	10	—	—	—	—	—	—
ZA-LMS	0.0027	0.00011	—	—	—	—	—	—	—
IPNLMS	0.3	—	0.5	—	—	-0.75	—	—	0.000001
PNLMS	0.3	0.01	—	—	0.01	—	—	—	0.000001
NLMS	0.38	—	—	—	—	—	—	—	0.000001

Up to now, it has been noticed that the performances of the NLMS, PNLMS and IPNLMS algorithms are very poor in terms of MSD and convergence rate. Hence, these algorithms will be excluded from the experiments that appear in the rest of this chapter.

5.3. Performance of the Proposed Algorithm under Different Sparsity Ratios

In this section, two experiments are performed to find out the effect of the sparsity on the performance of the proposed by changing the sparsity ratio of the unknown system. One experiment is performed in AWGN and the other one is designed for ACGN environment. For ease of implementation, a filter order $N = 256$ taps was used in all the experiments. Because of the poor performances demonstrated by the NLMS, PNLMS and IPNLM algorithms in all the previous experiments under similar experimental set up, we would in this section focus on testing the performance of the proposed algorithm and compare with only those of the l_1-norm and p-norm based sparse algorithms (i.e., ZA-LMS, RZA-LMS and NNCLMS algorithms).

5.3.1. Additive White Gaussian Noise

This experiment is performed to test and compare the convergence and tracking ability of the proposed algorithm compare with those of the NNCLMS, RZA-LMS and ZA-LMS algorithms in identifying unknown systems with different sparsity conditions. The used acoustic echo path is represented in Fig. 5.1. The SNR is assumed to be 30 dB. In this scenario, the same system is made to be 75%, 50%, and 25% sparse (i.e. at 75% sparsity, 192 taps are made inactive while the rest are made active, etc.). The metric used for measuring the sparseness level is the same as the one used in Section 5.2. Simulations were done with parameters given in Table 5.2. Figs. 5.8, 5.9 and 5.10 show the results of the MSD between the echo path coefficients and the adaptive filter coefficients at 75%, 50%, and 25% sparsity ratio, respectively. Table 5.6 shows a summary of the obtained results. It can be observed that, at 75% sparse system, the proposed algorithm converges slower than the rest by 200 iterations but achieves a better MSD estimate of 5 dB, 9 dB and 11 dB than the NNCLMS, RZA-LMS and ZA-LMS algorithms, respectively. At 50% sparsity ratio, all algorithms have the same convergence rate but the proposed algorithm still outperforms the rest in terms of MSD estimate. But when the unknown system is 25% sparse, both the proposed algorithm and NNCLMS algorithm converge at the same rate but slower than the RZA-LMS and ZA-LMS algorithms with lower MSD of the proposed algorithm. The convergence behavior of RZA-LMS and ZA-LMS algorithms with such high MSD is due to their lack of adjustable factors that can adapt the norm constraint during the filtering process. In addition, we notice that the proposed algorithm shows a similar convergence in all the three sparsity ratios and shows an outstanding performance in MSD estimate compared to other algorithms in every sparsity condition of the unknown system. This experiment shows that the convergence rate and tracking ability of the proposed algorithm is less sensitive to the sparseness degree of the acoustic impulse response and less affected in terms of MSD estimate.

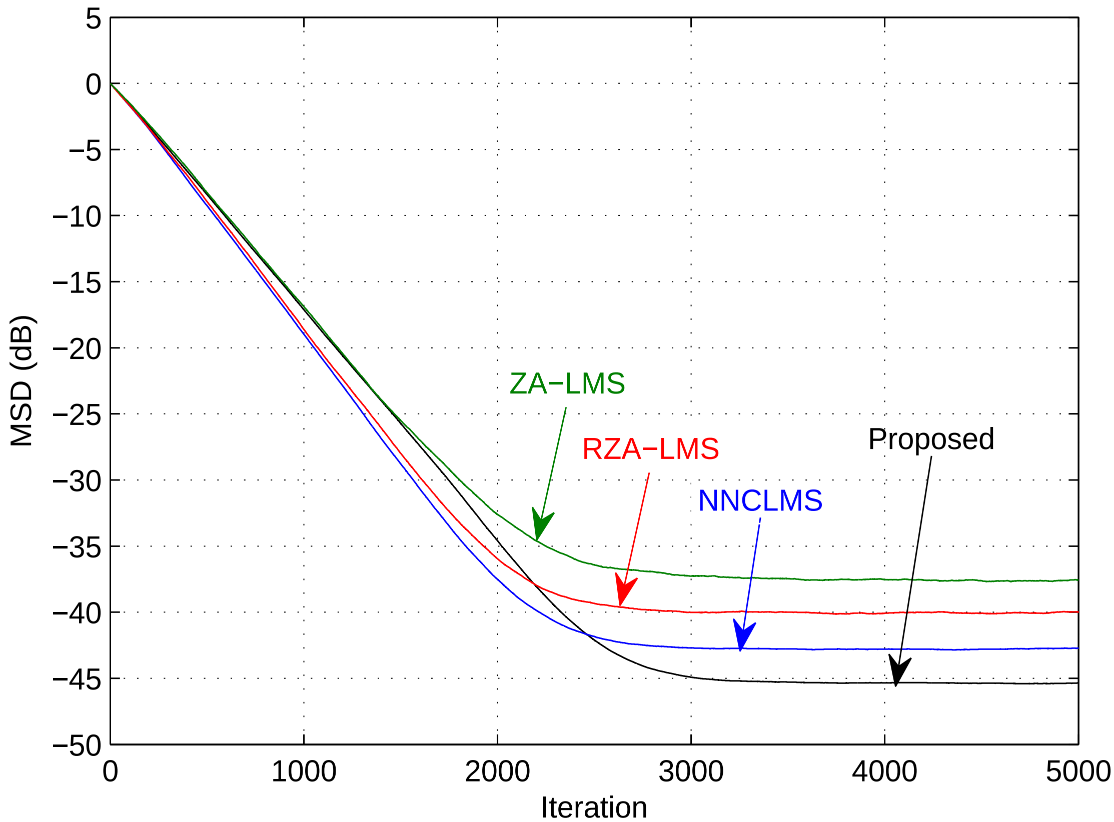

Figure 5.8. Tracking and steady-state behaviors of a 256 taps adaptive filters with 75% sparsity ratio in AWGN.

5.3.2. Additive Correlated Gaussian Noise

The same experiment of Section 5.3.2 is repeated with only replacing the AWGN by an ACGN process. The ACGN process is created as in Section 5.2.5 and with the echo path depicted in Fig. 5.1. Simulations were done with the parameters provided in Table 5.3. The average MSD estimate between the adaptive filter coefficients and the unknown system coefficients of all algorithms are shown in Figs. 5.11, 5.12 and 5.13 at 75%, 50%, and 25% sparsity ratios, respectively. A summary of the obtained results is provided

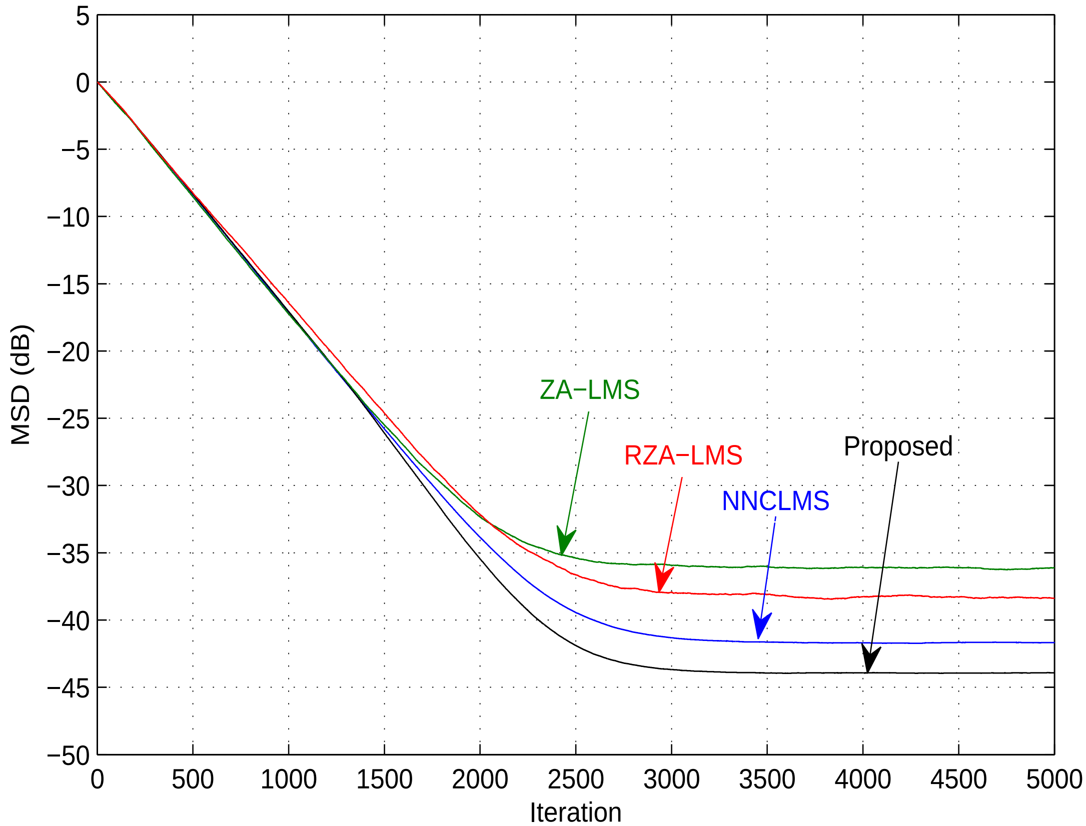

Figure 5.9. Tracking and steady-state behaviors of a 256 taps adaptive filters with 50% sparsity ratio in AWGN.

Table 5.6. Results for experiment in Section 5.3.1.

Algorithm	75% sparse		50% sparse		25% sparse	
	Convergence	MSD(dB)	Convergence	MSD(dB)	Convergence	MSD(dB)
Proposed	3000	-45.1	3000	-44	3000	-43
NNCLMS	2800	-42.5	3000	-42	3000	-38
RZA-LMS	2800	-40	3000	-40.5	3250	-34
ZA-LMS	2800	-37.5	3000	-36	2250	-32

in Table 5.7. We notice that, at 75% sparsity ratio, the proposed algorithm converges faster than the rest algorithms by almost 100 iterations and achieves a better MSD estimate (2.1 dB, 3.1 dB, and 3.6 dB) than NNCLMS, RZA-LMS and ZA-LMS, respectively. This shows the ability of the proposed algorithm in suppressing a correlated noise in sparse system. When the unknown system is set to 50% sparsity ratio,

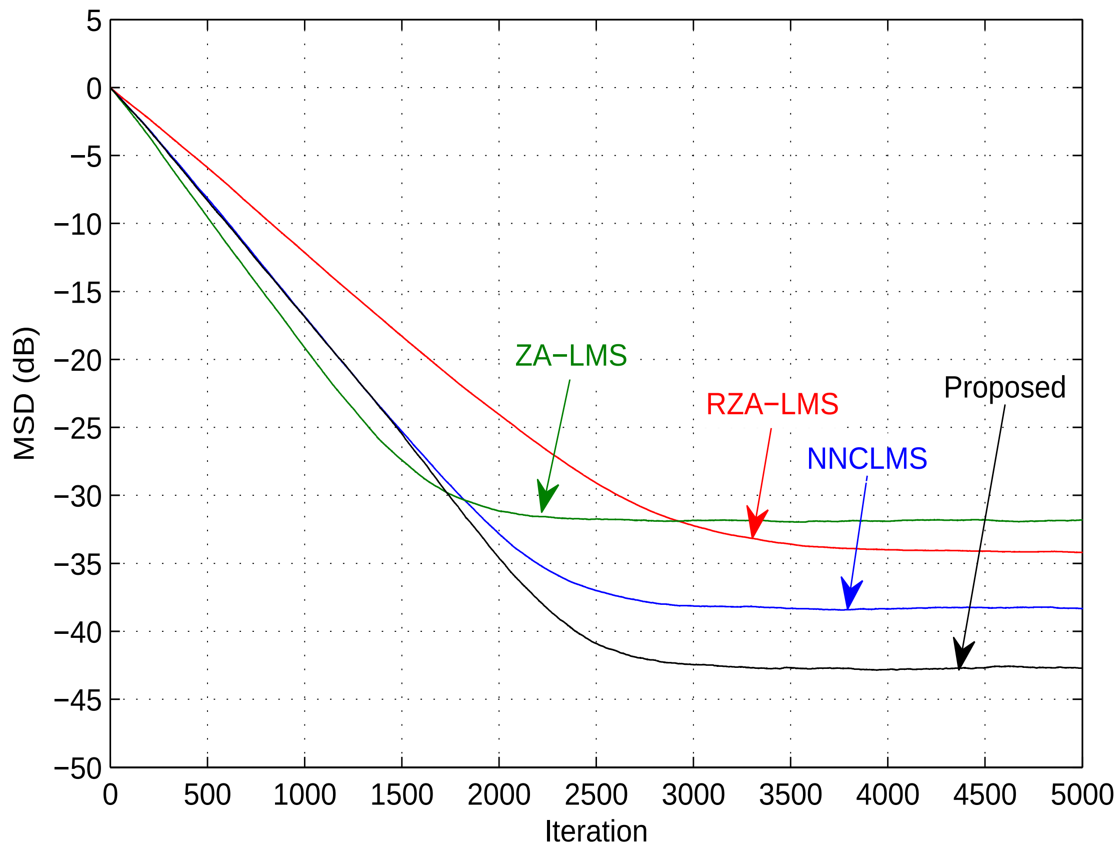

Figure 5.10. Tracking and steady-state behaviors of a 256 taps adaptive filters with 25% sparsity ratio in AWGN.

we still notice almost similar behavior except that the RZA-LMS shows a slower convergence by about 500 iterations and achieves a little higher MSD of almost 0.5 dB than when the unknown system is 75% sparse. But when the unknown system is at 25% sparsity ratio, we notice deterioration in the performance of the NNCLMS, RZA-LMS and ZA-LMS algorithms, with the ZA-LMS algorithm worse among all of them. However, the proposed algorithm still retains its superior performance. This is because the proposed algorithm achieves a lower MSD improvement of almost 2.5 dB, 5 dB and 9 dB compared to NNCLMS, RZA-LMS and ZA-LMS, respectively.

In this experiment, it can be noted that the convergence behavior of the proposed algorithm remains the same in all three sparsity conditions of the unknown system. The proposed algorithm also gives a better MSD estimates in all given sparsity settings. This is due to the ability of the p-norm constraint to efficiently exploit the unknown system's sparsity and the virtue of the variable step-size to track the changes of the unknown system. Even though, the ACGN process causes a little higher MSD estimate, the proposed algorithm has been observed to be more robust among the others in correlated noise environments. We also notice from the result that, unlike when an AWGN is used, all algorithms achieve higher MSD estimate in correlated environment. However, despite this challenge we still observed that the proposed algorithm shows the same convergence rate across every sparsity condition of the system and simultaneously outperform all algorithms in terms of MSD estimates.

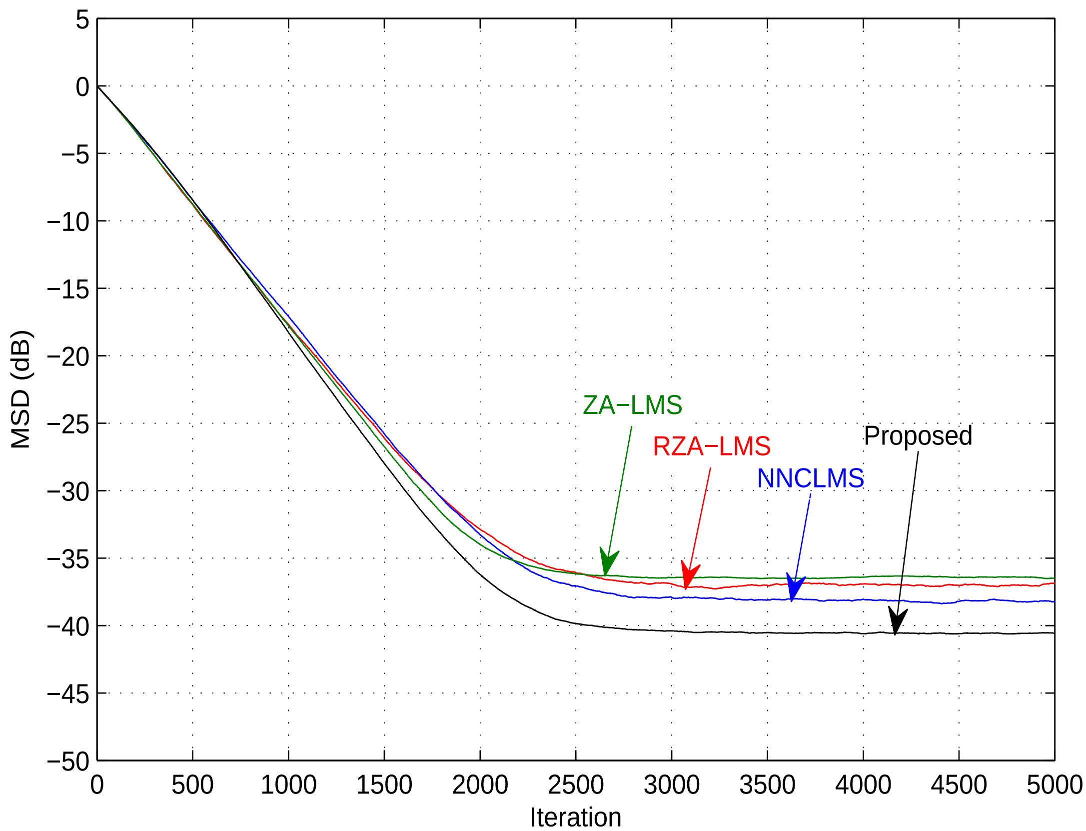

Figure 5.11. Tracking and steady-state behaviors of a 256 taps adaptive filters with 75% sparsity ratio in ACGN.

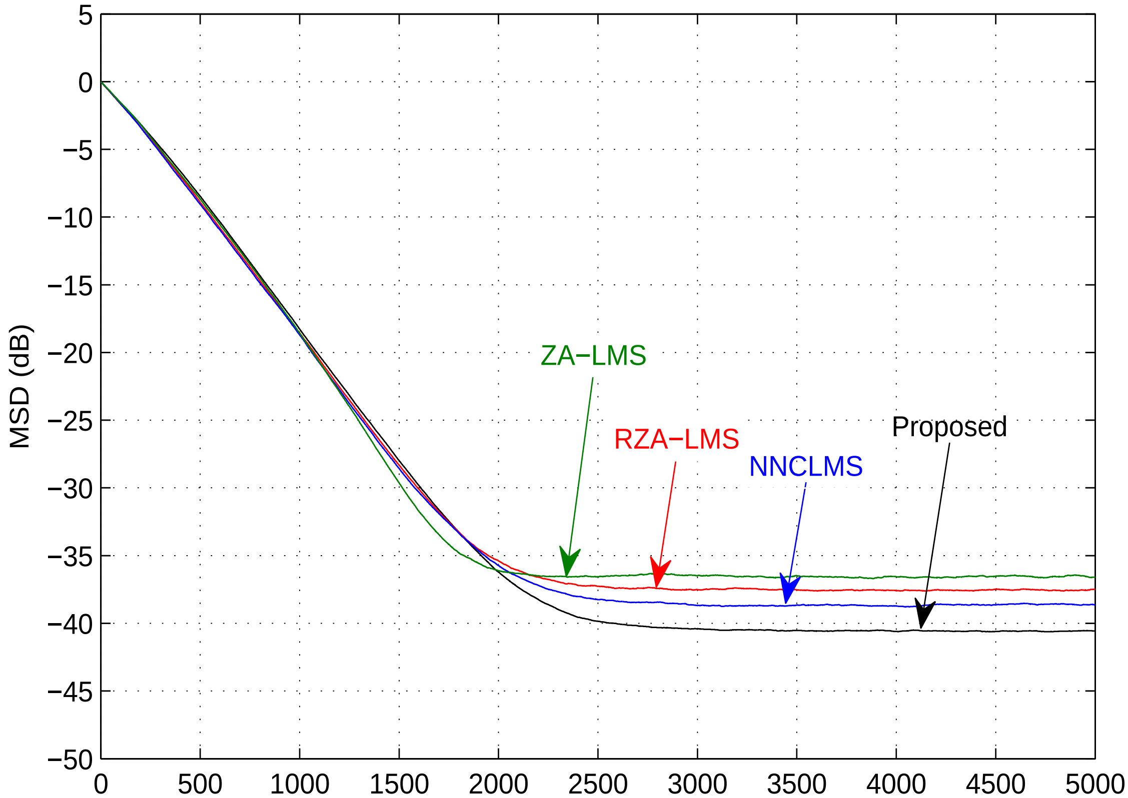

Figure 5.12. Tracking and steady-state behaviors of a 256 taps adaptive filters with 50% sparsity ratio in ACGN.

Table 5.7. Results for experiment in Section 5.3.2.

Algorithm	75% sparse		50% sparse		25% sparse	
	Convergence	MSD(dB)	Convergence	MSD(dB)	Convergence	MSD(dB)
Proposed	2500	-40.1	2500	-40.1	2500	-40
NNCLMS	2600	-38	2500	-38	2500	-37.5
RZA-LMS	2600	-37	2500	-37	2500	-35
ZA-LMS	2600	-36.5	2000	-36	1700	-31

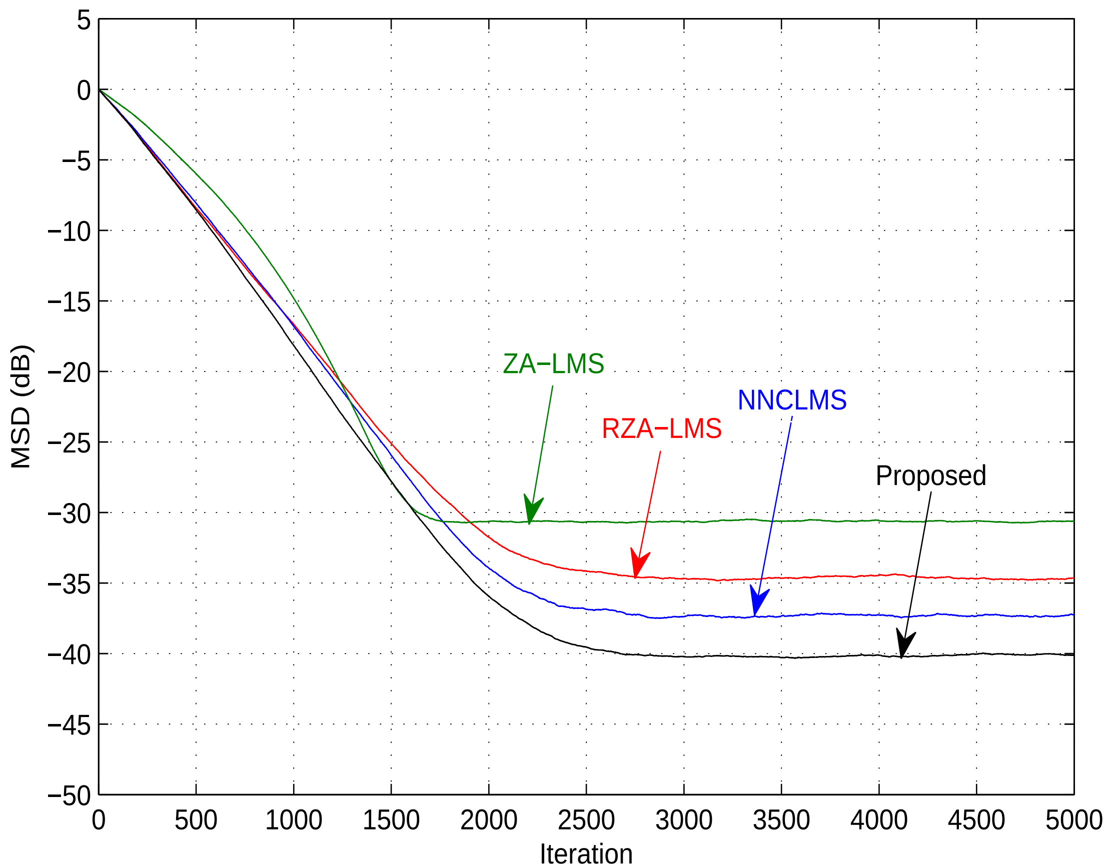

Figure 5.13. Tracking and steady-state behaviors of a 256 taps adaptive filters with 25% sparsity ratio in ACGN.

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

6.1. Conclusion

In this thesis, Some of the challenges caused by undesired acoustic echoes that occur in communication devices are addressed. The research concentrates on the development of a new adaptive filtering algorithm that enables us to identify the sparse echo path of the acoustic room system. Some of the available sparse adaptive algorithms have been reviewed. A new p-norm constraint adaptive algorithm has been proposed. The convergence analysis of the proposed algorithm has been presented and its stability condition is derived. The performances of the proposed algorithm have been investigated through extensive simulation experiments and evaluated in terms of the convergence rate and MSD estimate.

An acoustic echo path of fixed sparsity was simulated in AWGN and a better MSD estimate of the proposed algorithm compared to the best performer among the NLMS, PNLMS, IPNLMS, NNCLMS, ZA-LMS and RZA-LMS algorithms is obtained. Where as in the ACGN, it has been noticed that, even with highly correlated Gaussian noise, the proposed algorithm still much better than the other algorithms in terms of convergence rate and/or MSD. The NLMS, PNLMS and IPNLMS algorithms failed to provide good performance compared to NNCLMS, RZA-LMS and ZA-LMS algorithms. This is due to their lack of available parameters to effectively track sparse impulse responses.

The proposed algorithm was furtherly investigated in identifying an unknown system having a variety of sparsity ratios (ranging from 75%, 50% and 25% sparsity ratios). It has been shown to be performing more robust than the NNCLMS, RZA-LMS and ZA-LMS algorithms in both AWGN and ACGN environments. This is due to the virtue of the variable step-size parameters in addition to p-norm constraint associated with the proposed algorithm.

6.2. Further work

Despite the fact that the results of this work are adequate and satisfactory, there still other works which need to be conducted in the future in order improve the quality of this approach. All our investigations on the performance of the proposed algorithm are limited to only two types of noise environments; AWGN and ACGN. Therefore, one of possible future works could be investigating its performance in other types of noise environments such as additive white impulsive noise, additive correlated impulsive noise, etc. Another area of investigation could be applying the proposed algorithm in other scenarios different from echo cancellation, such as channel equalization, adaptive beam forming, etc. In addition, the performance of the proposed algorithm may also be inspected using a longer length echo paths of about 1024 coefficients and greater.

REFERENCES

[1] Deniz, P. S. R., Adaptive Filtering Algorithms and Practical Implementation, 2008, Third Ed., LLC, NY, Springer.

[2] Haykin, S., Adaptive Filter Theory, 2002, Prentice Hall, Upper Saddle River, NJ.

[3] Hayes, M. H., Statistical Digital Signal Processing and Modeling, 1996, John Wiley & Sons. Inc., New York.

[4] Sondhi, M. M., The History of Echo Cancellation, IEEE Signal Processing Magazine, 2006, 23, 95-102.

[5] Gay, S. L., An Efficient Fast Converging Adaptive Filter for Network Echo Cancellation, Presented at the Thirty-Second Asilomar Conference on Signal, System and Amplifier, California, USA, November 1998, 394-398.

[6] Gilloire, A., Experiments with Sub-Band Acoustic Echo Cancellers for Teleconferencing, IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP1987), April 1987, 2141-2144.

[7] Duttweiler, D. L., Proportionate Normalised Least Mean Square Adaptation in Echo Cancelers, IEEE Transactions on Speech and Audio Processing,2000, 5, 508-518.

[8] Romesburg, E. D., Echo Canceller for Non-Linear Circuits, U.S. Patent, 5, August 1998, 796-819.

[9] Benesty, J., Gansler, T., Morgan, D. R., Sondhi, M. M., Gay, S. L., Advances in Network and Acoustic Echo Cancellation. Berlin, Germany: Springer-Verlag, 2001. DOI: 10.1007/978-3-662-04437-7.

[10] Naylor, P. A., Cui, J., Brookes, M., Adaptive Algorithms for Sparse Echo Cancellation. Signal Processing, June 2006, 6, 1182-1192.

[11] Salman, M. S., Kukrer, O., Hocanin, A., Adaptive Filtering Fundamentals and Applications, 2011, LAP LMBERT, U.S.A.

[12] Sayed, A. H., Adaptive Filters, 2008, John Wiley, Hoboken, New Jersey, 163-167.

[13] Bellanger, M. G., Adaptive Digital Filters, 2001, Second Ed., Marcel Dekker, New York.

[14] Mathews, V. J., Zhenhua X., Stochastic Gradient Adaptive Filters with Gradient Adaptive Step-Sizes, International Conference on Acoustics, Speech, and Signal Processing (ICASSP1990), April 1990, 1385-1388.

[15] Widrow, B., Glover, J. R., McCool, J. M., Kaunitz, J., Williams, C. S., Hearn, R. H., Zeidler, J. R., Eugene Dong, Jr., Goodlin, R. C., Adaptive Noise Cancelling: Principles and Applications, Proceedings of the IEEE, December 1975, 1692-1716.

[16] Li, Y., Gu, Y., Tang, K., Parallel NLMS Filters with Stochastic Active Taps and Step-Sizes for Sparse System Identication, IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP2006), May 2006, 3.

[17] Pelekanakis, K., Chitre, M., Comparison of Sparse Adaptive Filters for Underwater Acoustic Channel Equalization/Estimation, IEEE International Conference on Communication Systems (ICCS2010), November 2010, 17, 395-399.

[18] Gui, G., Peng, W., Adachi, F., Improved Adaptive Sparse Channel Estimation Based on the Least Mean Square Algorithm, IEEE Wireless Communications and Networking Conference (WCNC), Shanghai, China, April 2013, 3105-3109.

[19] Gay, S. L., Douglas, S. C., Normalized Natural Gradient Adaptive Filtering for Sparse and Non-Sparse Systems, IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP2002), Orlando, Florida, March 2002, 2, 1405-1408.

[20] Chen, Y., Gu, Y., Hero, A. O., Sparse LMS for System Identification, IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP2009), Taipei, Taiwan, April 2009, 3125-3128.

[21] Jin, J., Qing, Q., Yuantao, G., Robust Zero-Point Attraction LMS Algorithm on Near Sparse System Identification, IET Signal Processing, 2013, 3, 210-218.

[22] Gu, Y., Jin, J., Mei, S., l_0-Norm Constraint LMS for Sparse System Identification, IEEE Signal Processing Letters, 1985, 9, 774-777.

[23] Christina, B., Control of a Hands-Free Telephone Set, Signal Processing, ScienceDirect, 1997, 61, 131-143.

[24] Elko, G. W., Diethorn, E., Gansler, T., Room Impulse Response Variation Due to Thermal Fluctuation and its Impact on Acoustic Echo Cancellation, International Workshop on Acoustic Echo Noise Control (IWAENC2003), Kyoto, Japan, September 2003, 67-70.

[25] Peterson, P. M., Simulating the Response of Multiple Microphones to a Single Acoustic Source in a Reverberant Room, Journal Of the Acoustical Society of America, Nov. 1986, 5, 1527-1529.

[26] Donoho, D. L., Compressed Sensing, IEEE Transactions on Speech and Audio Processing, 2006, 4, 1289-1306.

[27] Widrow, B., Stearn, S. D., Adaptive Signal Processing, 1985, Printice Hall, New Jersey.

[28] Mandic, D. P., A Generalized Normalized Gradient Descent Algorithm, IEEE Signal Process Letters, February 2004, 11, 115-118.

[29] Cheng Y. F., Etter, D. M., Analysis of an Adaptive Technique for Modeling Sparse Systems, IEEE Transaction on Acoustics, Speech, and Signal Processing, February 1989, 2, 254-264.

[30] Candes, E. J., Wakin, M., An Introduction To Compressive Sampling, IEEE Signal Processing Magazine, March 2008, 2, 21-30.

[31] Salman, M. S., Sparse Leaky-LMS Algorithm for System Identification and its Convergence Analysis, International Journal of Adaptive Control and Signal Processing, 2008, DOI:10.1002/acs. 2428.

[32] Rey, V. L., Rey, H., Benesty, J., Tressens, S., A Family of Robust Algorithms Exploiting Sparsity in Adaptive Filters, IEEE Transactions on Audio, Speech, and Language Processing, May 2009, 4, 572-581.

[33] Etter, D. M., Identification of Sparse Impulse Response System Using an Adaptive Delay Filter, IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP1985), April 1985, 10, 1169-1172.

[34] Khong, A. W. H., Xiang, L., Doroslovacki, M., Naylor, P. A., Frequency Domain Selective Tap Adaptive Algorithm for Sparse System Identification, IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP2008), Las Vegas, Nevada, March 2008, 229-232.

[35] Kawamuri, S., Hatori, M., A Tap Selection Algorithm for Adaptive Filters, IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP1986), April 1986, 11, 2979-2982.

[36] Su, G., Jin, J., Gu, Y., Wang, J., Performance Analysis of l_0-Norm Constraint Least Mean Square Algorithm, IEEE Transactions on Signal Processing, 2011, 6, 2223-2235.

[37] Gu, Y., Jin, J., Mei, S., 2010. A Stochastic Gradient Approach on Compressive Sensing Signal Reconstruction Based on Adaptive Filtering Framework, IEEE Journal of Selected Topics in Signal Processing, 2011, 2, 409-420.

[38] Shi, K., Shi, P., Convergence Analysis of Sparse LMS Algorithms with l_1-norm Penalty Based on White Input Signal, Signal Image and Video Processing, Springer, May 2010, 12, 3289-3293.

[39] Slavakis, K., Kopsinis, Y., Theodoridis, S., Adaptive Algorithm for Sparse System Identification Using Projections onto Weighted l_1-Balls, IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP2012), March 2010, 3742-3745.

[40] Wu, F. Y., Tong, F., Non-Uniform Norm Constraint LMS Algorithm for Sparse System Identification, IEEE Communications Letters 2013, 2, 385-388.

[41] Martin, R. K., Sethares, W. A., Williamson, R. C., Johnson, C. R., Exploiting Sparsity in Adaptive Filters, IEEE Transactions on Signal Processing, August 2002, 8, 1883-1894.

[42] Harris, R. W., Chabries D. M., Bishop, F. A., A Variable Step (VS) Adaptive Filter Algorithm, IEEE Transactions on Acoustics, Speech and Signal Processing 1986, 2, 309-316.

[43] Kwong, R. H., Johnson, E. W., A Variable Step-Size LMS Algorithm, IEEE Transactions on Signal Processing,1992, 7, 1633-1642.

[44] Cui, J., Naylor, P. A. Brown, D. T., An Improved PNLMS Algorithm for Echo Cancellation in Packet-Switched Networks, IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP2004), Toulouse, France, 3, May 2004, 141-144.

[45] Deng, H., Doroslovacki, M., Improving Convergence of the PNLMS Algorithm for Sparse Impulse Response Identication, IEEE Signal Processing Letters,2005, 3, 181-184.

[46] Benesty, J., Morgan, D. R., Sondhi, M. M., A Better Understanding and an Improved Solution to the Specic Problems of Stereophonic Acoustic Echo Cancellation, IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP1998), 2, 156-165.

[47] Salman, M. S., Jahromi, N., Hocanin, A., Kukrer, O., A Zero-Attracting Variable Step-Size LMS Algorithm for Sparse System Identification, IX International Symposium on Telecommunications (BIHTEL2012), Sarajevo, Bosnia and Herzegovina, October 2012, 1-4.

[48] Chartrand, R., Exact Reconstruction of Sparse Signals Via Non-Convex Minimization, IEEE Signal Processing Letters, 2007, 10, 707-710.

[49] Rao, B. D., Delgado, K. K., An Affine Scaling Methodology for Best Basis Selection, IEEE Transaction on Signal Processing, January 1999, 1, 187-200.

[50] Rao, B. D., Bongyong S., Adaptive Filtering Algorithms for Promoting Sparsity, IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP2003), June 2003, 6, 361-364.

[51] Aliyu, M. L., Alkassim, M. A., Salman, M.S., A p-Norm Variable Step-Size LMS Algorithm for Sparse System Identification, Signal Image and Video Processing, Springer, DOI: 10.1007/s11760-013-0610-7.

[52] Gwadabe T. R., Aliyu M. L., Alkassim, M.A., Salman M. S., Haddad H., A New Sparse Leaky LMS Type Algorithm. IEEE 22nd Signal Processing and Communications Applications Conference (SIU 2014), Trabzon, Turkey, April 2014.

[53] Loganathan P., Sparseness-Controlled Adaptive Algorithms for Supervised and Unsupervised System Identification, Ph.D. Thesis, 2011, Imperial College, London.

[54] Gui, G., Adachi, F., Improved Adaptive Sparse Channel Estimation Using Least Mean Square Algorithm, EURASIP Journal on Wireless Communications and Networking, March 2013, 1, 1-18.

[55] Kenney, J.F., Keeping, E.S., In Mathematics of Statistics, 1962, Third Ed. Van Nostrand, Princeton.

[56] Evans, J. B., Xue p., Liu, B., Analysis and Implementation of Variable Step-Size Adaptive Algorithms, IEEE Transaction on Signal Processing, 8, 2517-2535.

[57] Mader, A., Puder, H., Schmidt, G. U., Step-Size Control for Acoustic Echo Cancellation Filters-An Overview. Signal Processing, September 2000, 9, 1697-1719.

[58] Reddy, V. U., Shan, T. J., Kailath, T., Application of Modified Least-Squares Algorithms to Adaptive Echo Cancellation, IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP1983), April 1983, 8, 53-56.

[59] Mayyas, K., Aboulnasr, T., Leaky LMS Algorithm: MSE Analysis for Gaussian Data, IEEE Transactions on Signal Processing, April 1997, 4, 927-934.

[60] Candes, E. J., Wakin, M., Boyd. S., Enhancing Sparsity By Reweighted l_1-Minimization, Journal of Fourier Analysis and Applications, October 2008, 14, 877-905.

[61] Claasen, T., Mecklenbrauker, W., Comparison of the Convergence of Two Algorithms for Adaptive FIR Digital Filters, IEEE Transactions on Acoustics, Speech and Signal Processing, 3, June 1981, 670-678.

[62] Hill, S.I., Williamson, R.C., Convergence of Exponentiated Gradient Algorithms, IEEE Transactions Signal Processing, Jun 2001, 6, 1208-1215.